Civil Engineering and Applied Solutions
Research Article
journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Modeling the Diffusion and Advection of Pollutants from Flaring in Iran, Considering Seasonal Atmospheric Variability

Siamak Boudaghpour [a*], Ehsan Qamari [a], Amirali Bahadori [b]

[a] *Department of Environmental Engineering, Faculty of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran*
[b] *Department of Water Resources Management, Faculty of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran*

A R T I C L E I N F O

Article history:
Received 15 April 2026
Revised 12 May 2026
Accepted 16 June 2026
Available online 01 April 2027

Keywords:
Air pollution
Computational fluid dynamics
Flaring
Advection
Diffusion

A B S T R A C T

Gas flaring in oil-producing regions is a significant source of atmospheric pollutants; however, near-field dispersion dynamics governing worker exposure from low-capacity flares under seasonally contrasting atmospheric stability conditions remain insufficiently characterized. This study employed a two-dimensional Eulerian computational fluid dynamics (CFD) framework with RNG k-ε turbulence closure to simulate the dispersion of carbon monoxide (CO), nitrogen oxides (NO_x), and sulfur dioxide (SO_2) from a representative low-capacity gas flare at exit velocities of 6-7 m/s and combustion temperatures of 900-1200 K. The computational domain, 2050 m × 286 m, was discretized into 66,000 quadrilateral control volumes. Simulations were conducted for two seasonally representative scenarios: an unstable summer atmosphere (Monin-Obukhov length L = −1148 m) and a stable winter surface-layer inversion (L = +364 m). Model predictions were validated against field measurements at three downwind distances, yielding R^2 values of 0.997, 0.9996, and 0.969 for CO, NO_x, and SO_2, respectively. CO concentrations at breathing height (2 m) exceeded the regulatory ambient standard by 55.0% in summer and 35.0% in winter; SO_2 exceeded the primary ambient standard by 375% and 362.5%, respectively. NO_x remained below the flare emission standard at the plume-core level in both seasons. Under unstable summer conditions, buoyancy-driven vertical uplift dominated plume transport; under stable winter conditions, a near-surface recirculation zone between 50 and 400 m downwind generated the widest spatial extent of regulatory exceedance. Atmospheric stability and near-source recirculation are identified as the primary controls on occupational exposure intensity, underscoring the need for targeted flare gas recovery systems and meteorology-aware worker management in Iran's oil-producing regions.

How to cite this article: Boudaghpour, S., Qamari, E., Bahadori, A. Modeling the Diffusion and Advection of Pollutants from Flaring in Iran, Considering Seasonal Atmospheric Variability. Civil Engineering and Applied Solutions. 2027; 3(2): 1–15. doi:10.22080/ceas.2026.31549.1091.

1. Introduction

Advection and diffusion are fundamental processes governing the transport and dispersion of pollutants in the atmosphere. Advection refers to the movement of pollutants due to wind, while diffusion accounts for their spread caused by turbulence and molecular motion. Together, these processes are described mathematically by the advection-diffusion equation, which combines conservation laws with constitutive relations, such as Fick's law of diffusion [1, 2]. Key parameters influencing pollutant dispersion include wind profile, eddy diffusivity, and temperature. Wind profiles dictate the horizontal transport of pollutants, with stronger winds generally enhancing dispersion [3]. Eddy diffusivity, a measure of turbulent mixing, varies with atmospheric stability and height, significantly affecting vertical pollutant spread [4, 5]. Temperature impacts atmospheric stability, as well as advection and diffusion rates [6]. Numerical methods such as finite difference schemes and advanced techniques, including the Generalized

* Corresponding author.
E-mail addresses: bodaghpour@kntu.ac.ir (S. Boudaghpour).

https://doi.org/10.22080/ceas.2026.31549.1091
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

Integral Transform Technique (GILTT), are commonly employed to solve the advection-diffusion equation [7, 8]. However, classical analytical approaches such as GILTT and standard finite difference schemes are subject to well-documented limitations: they rely on idealized assumptions of uniform wind fields and flat terrain, and their accuracy deteriorates significantly under spatially varying meteorological conditions, complex source geometries, or highly transient emission scenarios [5, 7]. These constraints render them inadequate for resolving near-field pollutant concentrations at the breathing-zone scale, particularly in the vicinity of industrial flaring sources where turbulence, buoyancy effects, and ground-level accumulation must be captured with high spatial fidelity. Emerging approaches, such as physics-informed neural networks (PINNs), offer promising alternatives for modeling complex pollutant dynamics under varying atmospheric conditions [8], enabling accurate predictions of pollutant concentrations and aiding environmental planning and policy development.

Gas flaring is a significant source of air pollution in Iran, contributing to greenhouse gas emissions, particulate matter, and volatile organic compounds (VOCs). Iran ranks as the third-largest gas-flaring country globally, accounting for approximately 12.1% of global flaring emissions [9]. Flaring activities emit pollutants such as CO_2, NO_x, black carbon, and VOCs, with emission rates influenced by flare efficiency, gas composition, and combustion conditions [10, 11]. The chemical composition of flare gas in Iran often includes methane, ethane, propane, and hydrogen sulfide, which vary by region and extraction site [12]. Inefficient combustion processes exacerbate emissions, producing harmful byproducts such as polycyclic aromatic hydrocarbons (PAHs) [13]. Seasonal variability, including higher winter wind speeds, further reduces combustion efficiency and increases pollutant dispersion [14]. Iran's regulatory framework has struggled to mitigate flaring due to infrastructure limitations and economic constraints. While flare gas recovery systems have been implemented in some refineries, widespread adoption remains limited [15]. Addressing these challenges requires enhanced monitoring, stricter enforcement, and investment in recovery technologies [16].

Explicit and implicit numerical schemes are widely used to model pollutant diffusion and advection, each offering distinct advantages and limitations. Explicit schemes, such as finite difference methods, are computationally simple and easy to implement but require small time steps to remain stable, especially at high Peclet numbers [17, 18]. Implicit schemes, such as Crank-Nicolson, are unconditionally stable and allow larger time steps, making them suitable for long-term simulations, though they are computationally more intensive [19, 20]. Recent innovations, including semi-Lagrangian Crank-Nicolson (SLCN) methods, combine the strengths of both approaches, offering stability at large Courant numbers and improved accuracy for curved streamlines [21]. Additionally, hybrid schemes such as the QUICK method improve mass conservation and computational efficiency, particularly for transient simulations [18]. Among more recent developments in the broader literature, PINNs have been proposed as an alternative framework that integrates machine learning with physical laws to solve advection-diffusion equations, showing promise for pollutant source identification and transport prediction with reduced computational cost [22, 23]. These methods have been explored in complex scenarios such as groundwater contamination and heterogeneous media [23, 24].

Nevertheless, PINNs remain computationally immature for operational near-field industrial dispersion applications at the spatial resolution required for breathing-zone exposure assessment. The present study, therefore, adopts a two-dimensional Eulerian CFD framework that provides spatial fidelity and physical completeness, including buoyancy effects, turbulence closure, and ground-level recirculation, necessary to resolve pollutant concentrations at occupationally relevant scales. Despite these methodological advances in the field, the application of high-resolution CFD frameworks to near-field dispersion from low-capacity industrial flares, particularly under the combined influence of seasonal atmospheric instability and surface-layer inversion, remains underexplored in the literature, especially in Iran's oil and gas-producing regions.

Seasonal variability significantly influences the dispersion and transport of pollutants from flaring activities in Iran. Meteorological parameters such as wind speed, temperature, humidity, and atmospheric stability vary seasonally, producing distinct patterns of pollutant diffusion and advection. Higher winter wind speeds enhance horizontal transport of pollutants, while lower temperatures increase atmospheric stability and suppress vertical mixing [25, 26]. In contrast, summer conditions foster convective instability, promoting vertical uplift but reducing near-surface dilution under low-wind regimes [27]. These seasonally contrasting regimes, surface-layer inversion in winter versus convective instability in summer, govern the spatial footprint and ground-level magnitude of pollutant accumulation near flaring sources, and constitute the primary meteorological forcing examined in this study. Additionally, reduced rainfall and lower planetary boundary layer heights, as observed during certain seasonal episodes, further hinder pollutant dispersion and elevate ground-level concentrations [28, 29].

Empirical studies in Iran have demonstrated the utility of advanced modeling techniques for predicting pollutant concentrations from flaring activities, while also revealing persistent methodological limitations. The CALPUFF model, combined with meteorological data from the Weather Research and Forecasting (WRF) model, was applied to simulate sulfur dioxide (SO_2) dispersion from gas flares on Sirri Island, showing that low-height flares produce higher ground-level concentrations. In contrast, elevated flares affect areas farther from the source [30]. Satellite-based approaches have also been employed to estimate flaring volumes and emissions, with artificial neural networks (ANNs) achieving an R^2 of 0.73 in correlating flare volumes with remote sensing data [31]. Hybrid models such as ARIMA-TSVR have further demonstrated superior forecasting accuracy relative to traditional methods [32]. Key challenges include variability in flare gas composition and the influence of meteorological conditions, both of which complicate modeling efforts. Critically, prior modeling studies of Iranian flaring zones have largely relied on Gaussian plume models or mesoscale meteorological frameworks, which sacrifice near-field spatial resolution and are unable to resolve breathing-zone concentrations or capture the recirculation zones and buoyancy-driven plume dynamics characteristic of low-capacity flares [30, 32]. These methodological gaps directly motivate the present study.

Against this background, the present study addresses three specific gaps in the existing literature. First, no prior study has applied a two-dimensional Eulerian CFD framework to quantify the dispersion of CO, NO_x, and SO_2 from a low-capacity flare under the seasonally contrasting atmospheric conditions characteristic of Iran's flaring zones. Second, existing modeling efforts have not resolved pollutant concentrations at the breathing-zone scale (approximately 1.5 m above ground level) with sufficient spatial detail to assess occupational exposure relative to established air quality standards. Third, the compound influence of summer atmospheric instability and winter temperature inversion on near-field plume behavior has not been systematically compared within a unified high-resolution modeling framework for this region. To address these gaps, this study employs a CFD approach based on the Eulerian advection-diffusion formulation to simulate the dispersion of CO, NO_x, and SO_2 from a representative low-capacity flare under summer and winter meteorological conditions. The central hypothesis is that seasonal differences in atmospheric stability, specifically convective instability in summer versus surface-layer inversion in winter, produce fundamentally distinct near-field dispersion patterns, with winter conditions generating substantially higher ground-level concentrations within the worker exposure zone. Notably, the seasonally contrasting atmospheric regimes examined here are consistent with atmospheric patterns projected to intensify across southwestern Iran under continued warming [27, 33, 34], thereby making the present findings directly relevant to long-term air quality planning in the region. The findings are intended to inform occupational health risk assessment and emission control strategies for Iran's flaring-intensive oil and gas sector.

2. Methodology

This study focuses on a gas flare from an oil exploitation unit with a nominal capacity of 15,000 barrels per day, representative of low-capacity upstream production facilities common in Iran's southwestern oil fields. The flare in this unit is FLARESIM 1.2.0, characterized as hot, low-pressure (LP), and operating without a mixing agent. The detailed geometric and emission specifications of the flare stack are summarized in Table 1. The flare stack has a height of 22 m, a base diameter of 24 inches (0.6096 m), and a tip diameter of 10 inches (0.254 m). The combustion exit temperature ranges from 900 to 1200 K, and the gas exit velocity at the flare outlet ranges from 6 to 7 m/s. Based on the reported outlet density and pollutant mass fractions, the estimated pollutant mass flow rates are 5.49-7.41 g/s for CO, 0.86-1.66 g/s for NO_x, and 0.039-0.047 g/s for SO_2, derived using combustion-based emission factors for LP Iranian flares [10, 14].

Table 1. Flare stack and emission source specifications.

Parameter	Value	Unit
Stack height	22	m
Stack diameter (base)	24 in (0.6096)	m
Tip diameter	10 in (0.254)	m
Exit gas velocity	6-7	m/s
Exit temperature	900-1200	K
CO mass flow rate	5.49-7.41	g/s
NO_x Mass Flow Rate	0.86-1.66	g/s
SO_2 mass flow rate	0.039-0.047	g/s
Computational domain	2050 × 286	m
Number of control volumes	66,000	—
Turbulence model	RNG k-ε	—

The concentration of air pollutants can be evaluated from two perspectives: the Eulerian and Lagrangian. The main distinction between these approaches lies in their coordinate systems. In the Eulerian system, the coordinate framework remains fixed relative to the movement of air masses, whereas in the Lagrangian system, it moves with the atmospheric masses. The Eulerian framework was selected for this study because it provides superior spatial resolution for near-field concentration fields and is better suited to resolving ground-level pollutant accumulation near a fixed emission source; Lagrangian particle-tracking methods, while effective for far-field transport, are computationally more demanding at the spatial scales required for breathing-zone exposure assessment [35, 36]. The fundamental equations governing fluid flow include the conservation of mass and momentum, which, together with the thermodynamic energy equation, form the cornerstone of atmospheric modeling. The computational domain was discretized into small elements using Gmsh. Subsequently, partial differential equations representing the flow, derived from the Navier-Stokes equations, were applied across the entire grid. This process yielded a large set of nonlinear equations, which were solved simultaneously using ANSYS Fluent. Given the turbulent nature of the flow in this problem, specific turbulence models were incorporated into the analysis, as described in the following sections.

2.1. The governing equations

The advection-diffusion equation governs mass transfer in the atmospheric domain. Because the density variation of air under the simulated meteorological conditions is negligible (Mach number $\ll$ 0.3), the assumption of incompressible momentum conservation equations is adopted throughout, except for the buoyancy term, which is treated via the Boussinesq approximation (see Boundary Conditions section). The incompressible momentum conservation equations in the x (horizontal) and z (vertical) directions, expressed in terms of the Navier-Stokes formulation, are as shown in Eqs. 1 and 2 [35].

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho_a}\frac{\partial p_a}{\partial x}+\frac{1}{\rho_a}\left[\frac{\partial}{\partial x}\left(\rho_a K_{xx}\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial z}\left(\rho_a K_{zx}\frac{\partial u}{\partial z}\right)\right] \quad (1)$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+w\frac{\partial w}{\partial z}=-g-\frac{1}{\rho_a}\frac{\partial p_a}{\partial z}+\frac{1}{\rho_a}\left[\frac{\partial}{\partial x}\left(\rho_a K_{xz}\frac{\partial w}{\partial x}\right)+\frac{\partial}{\partial z}\left(\rho_a K_{zz}\frac{\partial w}{\partial z}\right)\right] \quad (2)$$

where u (m/s) and w (m/s) are the velocity components in the horizontal (x) and vertical (z) directions, respectively; ρ_a (kg/m³) is the air density, treated as constant under the incompressible assumption; p_a (Pa) is the atmospheric pressure; g = 9.81 m/s² is gravitational acceleration; and K_{xx}, K_{zx}, K_{xz}, K_{zz} (m²/s) are the horizontal and vertical components of the turbulent diffusion tensor, whose values are discussed in the context of Table 1 below.

The energy equation for Newtonian fluids with constant density and viscosity is expressed as Eq. 3 [35]:

$$\frac{\partial \theta_V}{\partial t}+(v\cdot\nabla)\theta_v=\frac{1}{\rho_a}(\nabla\cdot\rho_a K_h\nabla)\theta_v+\frac{\theta_V}{c_{p,d}T_V}\sum_{n=1}^{N_ah}\frac{dQ_n}{dt} \quad (3)$$

where θ_V (K) is the virtual potential temperature; K_h (m²/s) is the thermal eddy diffusivity; $c_{p,d}$ (J/kg·K) is the specific heat of dry air at constant pressure; T_V (K) is the virtual temperature; and dQ_n/dt (W/kg) represents the diabatic heating rate from the n-th atmospheric heat source. The summation extends over all Nah diabatic contributions relevant to the boundary layer, including surface heat flux and radiative terms.

The turbulent dispersion tensor K is a 3 × 3 matrix whose components are determined by the RNG k-ε turbulence closure model and meteorological stability parameters, following the theoretical framework of Jacobson [37]. In these equations, u and w are velocities in the x and z directions; ρ_a (kg/m³) is the fluid density of air; and $c_{p,d}$ (J/kg·K) is the specific heat of dry air at constant pressure.

Pollutant transport in the atmosphere is governed simultaneously by advection and diffusion. Each mechanism contributes a distinct transfer factor that describes how a given volume or mass of air transfers the material. The continuity equation governing pollutant transport is expressed as Eq. 4 [35]:

$$\frac{\partial N}{\partial t}+\nabla\cdot(vN)-(\nabla\cdot K_h\nabla)N=0 \quad (4)$$

where N (mol/m³) is the pollutant molar concentration and v (m/s) is the local wind velocity vector. The turbulent mass diffusivity, K_h, specifically its horizontal (K_{xx}, m²/s) and vertical (K_{zz}, m²/s) components, varies with wind speed and atmospheric stability class, following the parameterizations of Jacobson [37] and Akkanen et al. [38]. The values of these coefficients are given in Table 2.

Table 2. Values of turbulent mass diffusion coefficients (m²/s) as a function of wind speed and atmospheric stability class [37, 38].

Wind velocity (m/s)	Good		Cloudy		Rainy	
	K_{xx}	K_{zz}	K_{xx}	K_{zz}	K_{xx}	K_{zz}
0-2	10	300	10	200	10	250
2-3	10	200	10	100	10	150
<3	10	100	10	50	10	50

The substantially larger vertical diffusion coefficients ($K_{zz} \gg K_{xx}$) under low-wind, good-visibility conditions reflect the dominance of convective vertical mixing during unstable summer atmospheric conditions. Under stable winter inversion conditions, the reduced K_{zz} values are consistent with the suppression of vertical mixing characteristic of positive Monin-Obukhov length (L > 0) regimes, as simulated in this study [4, 37].

2.2. Turbulence closure

The high Reynolds number of atmospheric boundary-layer flow in the vicinity of an industrial flare renders molecular viscosity negligible; the dominant influences on pollutant dispersion and plume dynamics arise from vortex turbulence and recirculation. The RNG k-ε model was selected over the standard and realizable k-ε variants for its demonstrated accuracy across a wide range of turbulent atmospheric flow regimes, including buoyancy-affected boundary layer flows and recirculating near-source plumes [39, 40]. The RNG formulation also provides improved predictions of streamline curvature and low-Reynolds-number effects near the ground boundary, which are critical for resolving breathing-zone concentrations in the near-field. The governing transport equations for turbulent kinetic energy k (m²/s²) and its dissipation rate ε (m²/s³) are as shown in Eqs. 5 to 7 [39].

$$\frac{Dk}{Dt}=\nabla\cdot\left(\frac{v_T}{\sigma_k}\nabla k\right)+p-\varepsilon \quad (5)$$

$$\frac{D\varepsilon}{Dt}=\nabla\cdot\left(\frac{v_T}{\sigma_\varepsilon}\nabla\varepsilon\right)+C_{1\varepsilon}\frac{\varepsilon}{k}p-C_{2\varepsilon}\frac{\varepsilon^2}{k} \quad (6)$$

$$v_t=c_\mu\frac{k^2}{\varepsilon} \quad (7)$$

where v_T (m²/s) is the turbulent kinematic viscosity; p (m²/s³) is the turbulent kinetic energy production rate; and σ_k and σ_ε are the turbulent Prandtl numbers for k and ε, respectively**.** The constant c_μ in Eq. 7 is 0.0845. Based on experimental observations, the turbulent kinetic energy production-to-dissipation ratio, p/ε, is set to 1.7 in the atmospheric surface layer [39]. The model constants are $C_{1\varepsilon} = 1.42$, $C_{2\varepsilon} = 1.68$, and $\sigma_k = \sigma_\varepsilon = 0.72$.

The finite volume method was employed to derive discretized algebraic equations for each control volume using the integral form of Gauss's divergence theorem (Eq. 8) [36].

$$\frac{\partial}{\partial t}\left(\int_\forall \rho\phi d\forall\right) + \int_A n\cdot(\rho\phi u)dA = \int_A n\cdot(\Gamma\nabla\phi)dA + \int_\forall S_\phi d\forall \quad (8)$$

where ϕ represents the transported scalar quantity (velocity component, temperature, or pollutant concentration); Γ (kg/m·s) is the effective diffusivity, comprising both molecular and turbulent contributions; and S_ϕ is the volumetric source term. The SIMPLE algorithm was used for pressure-velocity coupling, and second-order upwind discretization was applied to all convective terms to minimize numerical diffusion [40]. In this approach, conservation laws are applied to all variables within the control volume framework, ensuring that no artificial production or destruction of conservative variables occurs.

2.3. Computational domain, grid, and independence study

Based on established standards, the spatial dimensions of the solution domain are defined as 2,050 m in the direction of the prevailing wind and 286 m in height, equivalent to ten times the flare stack height, a domain size selected to ensure that pollutant concentrations at all lateral and upper boundaries remain negligible [36, 41]. The x- and y-axes represent the prevailing wind direction and altitude, respectively, with the coordinate origin at the base of the flare stack at ground level. The computational space was discretized into 66,000 quadrilateral control volumes using Gmsh, with mesh refinement concentrated in the near-source region within approximately 50 m of the flare stack and within the first 20 m above ground level to resolve steep concentration gradients associated with plume rise, buoyancy-driven recirculation, and near-ground pollutant accumulation. A smaller control volume density was employed in regions of high concentration gradient to optimize computational cost while maintaining accuracy. The computational domain and grid layout of the study area are shown in Fig. 1.

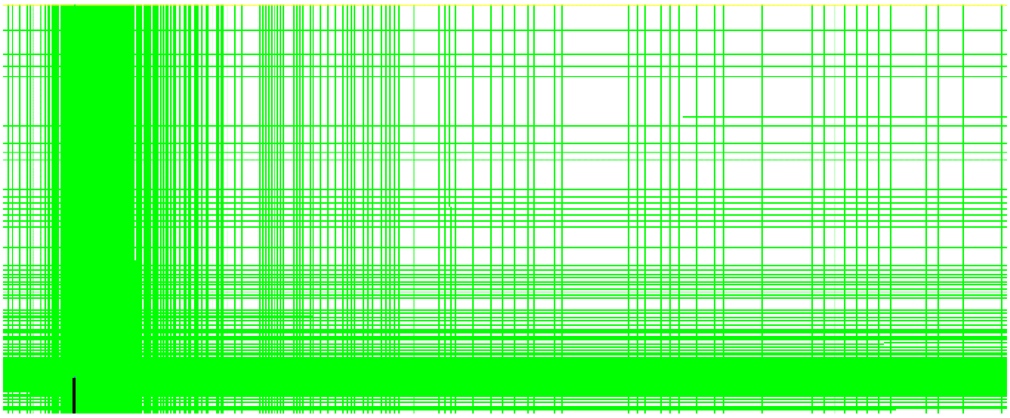


Fig. 1. Schematic of the grid layout of the study area.

To ensure numerical accuracy and verify independence from spatial discretization, a systematic grid sensitivity analysis was performed. Three mesh densities comprising 32,000 (coarse), 66,000 (medium), and 125,000 (fine) control volumes were evaluated. The peak ground-level NO_x concentration at 100 m downwind was monitored as the key indicator of convergence. As shown in Table 3, the relative difference in peak concentration between the medium and fine meshes was found to be only 2.1%, which falls well within the acceptable threshold for grid independence in atmospheric CFD studies. Consequently, the 66,000-cell mesh was selected as the optimal balance between spatial resolution and computational efficiency.

Table 3. Grid independence study results for peak ground-level NO_x concentration.

Mesh level	Number of cells	Peak NO_x concentration (μg/m^3)	Relative error (%)
Coarse	32,000	142.8	—
Medium	66,000	156.4	9.5%
Fine	125,000	159.7	2.1%

2.4. Atmospheric profiles and seasonal boundary conditions

The relationship between wind velocity and height varies significantly within the atmospheric boundary layer. Wind and temperature profiles at the inlet boundary were prescribed using Monin-Obukhov similarity theory [4, 41], which provides a physically rigorous description of the surface layer under both stable and unstable stratification, as shown in Eqs. 9 and 10.

$$\left|\overline{V_h}(z)\right| = \frac{u_*}{k}\left[L_n\left(\frac{z}{z_{0,m}}\right) - \psi_m\right] \quad (9)$$

$$\overline{\theta_v}(z) = \overline{\theta_v}(z_{0,h}) + P_t^r\frac{\theta_*}{k}\left[L_n\frac{z}{z_{0,h}} - \psi_h\right] \quad (10)$$

where u_* (m/s) is the friction velocity; k = 0.41 is the von Kármán constant; $z_{0,m}$ (m) and $z_{0,h}$ (m) are the aerodynamic roughness lengths for momentum and heat, respectively; θ_* (K) is the Monin-Obukhov temperature scale; P_t^r is the turbulent Prandtl number; and ψ_m and ψ_h are the integrated stability correction functions for momentum and heat, respectively [4, 41]. The Monin-Obukhov length L (m) characterizes the prevailing atmospheric stability regime, with $L > 0$ indicating stable (winter inversion) conditions and $L < 0$ indicating unstable (summer convective) conditions.

The meteorological input parameters for each seasonal scenario, derived from surface observations representative of Iran's southwestern oil-producing regions [30], are as follows.

For the winter scenario (stable surface-layer inversion, L = +363.9 m > 0):

$$\overline{\theta_v}(Z_r) = 276.9646, \overline{\theta_v}(z_{0,h}) = 276.767, u_* = 0.16899818, \theta_* = 0.00553947, L = 363.907496$$

$$|V(Z)| = 0.4225\mathrm{Ln}\left(\frac{Z}{0.0003}\right) + 0.006966(Z - 0.0003) \tag{11}$$

$$\overline{\theta_v}(z) = 276.767 + 0.013156\, L_n\left(\frac{z}{3\times10^{-6}}\right) + 0.0003(z - 3\times10^{-6}) \tag{12}$$

The positive value of L confirms stable stratification, with $z/L > 0$ throughout the domain, consistent with the suppression of vertical mixing characteristic of winter temperature inversion in the region.

For the summer scenario (unstable convective boundary layer, L = −1148.0 m < 0):

$$\overline{\theta_v}(z_r) = 299.8646, \overline{\theta_v}(z_{0,h}) = 299.947, u_* = 0.18628209, \theta_* = -0.00231, L = -1147.968$$

$$|\bar{V}(z)| = 0.466\left[\mathrm{Ln}\left(\frac{z}{0.0003}\right) - \mathrm{Ln}\left[\frac{[1+0.0168z]^{\frac{1}{2}}\left[1+(1+0.0168z)^{\frac{1}{4}}\right]^2}{8}\right]\right] + 0.932\tan^{-1}(1+0.0168z)^{\frac{1}{4}} - 0.732 \tag{13}$$

$$\bar{\theta}(z) = 299.947 - 0.00548625\left[\mathrm{Ln}\left(\frac{z}{3\times10^{-6}}\right) - 2\mathrm{Ln}\left[\frac{1+1.0526(1+0.0101048z)^{\frac{1}{2}}}{2.0526}\right]\right] \tag{14}$$

The negative Monin-Obukhov length ($z/L < 0$) confirms unstable atmospheric conditions, with enhanced vertical mixing and convective uplift of the pollutant plume, consistent with summer daytime conditions in southwestern Iran. These profiles were implemented as User-Defined Functions (UDFs) in ANSYS Fluent. The resulting wind velocity and temperature profiles for both seasons are presented in Figs. 2 to 4, confirming physically consistent behavior: a logarithmic-linear (Businger-Dyer corrected) profile under stable winter conditions and an enhanced mixing profile under unstable summer conditions.

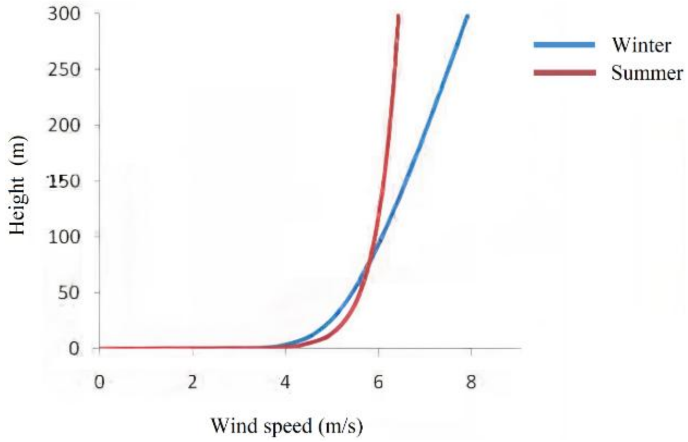


Fig. 2. The height profiles of wind velocity for both winter and summer days.

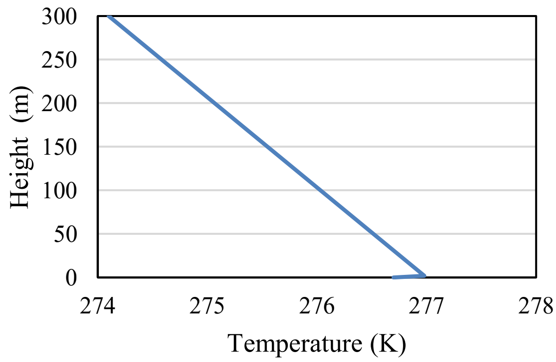


Fig. 3. The temperature profile with respect to height on a winter day.

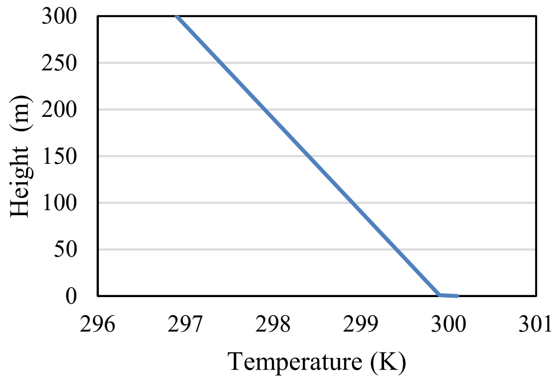


Fig. 4. The temperature profile with respect to height on a summer day.

2.5. Boundary conditions

The boundary conditions applied at each face of the computational domain are summarized as follows and are consistent with standard practice for atmospheric CFD simulations of near-source pollutant dispersion [37, 40, 41].

In this study, the simulation is performed under the assumption of incompressible flow, based on the prevailing wind and temperature profiles imposed at the inlet boundary. This assumption is justified by the relatively low Mach number of atmospheric flows in the context of pollutant dispersion, where density variations are negligible except in buoyancy-driven phenomena. At the inlet (upwind) boundary, spatially varying velocity and temperature profiles prescribed by Eqs. 9 to 14 were imposed as UDFs. At the downstream boundary, an outlet condition maintains the fluid pressure at ambient atmospheric levels, allowing natural outflow without artificial reflections. The upper boundary of the computational domain is treated as a symmetry plane, setting all normal gradients of velocity, pressure, and scalar quantities to zero, effectively ensuring no flux across this boundary. The ground boundary is modeled as a no-slip wall representing the surface of the flare system and the surrounding terrain. The flare stack outlet is defined as a velocity inlet boundary with specified exit velocity (6-7 m/s), exit temperature (900-1200 K), and pollutant mass flow rates (CO: 5.49-7.41 g/s; NO_x: 0.86-1.66 g/s; SO_2: 0.039-0.047 g/s), consistent with the emission specifications in Table 1. The working fluid ambient air is assumed to behave as a Newtonian fluid, where viscous stresses are linearly proportional to the strain rate. The Boussinesq approximation is employed to capture buoyancy effects induced by thermal gradients, allowing density variations to be considered only in the buoyancy term of the momentum equations while treating the flow as incompressible elsewhere, thereby balancing physical realism with computational efficiency [37].

2.6. Model validation

The predictive performance of the CFD framework was assessed by comparing simulated ground-level pollutant concentrations with field measurements at downwind distances of 50, 100, and 200 m from the flare base, in the near-field zone most critical for occupational exposure assessment. The simulated and measured concentrations, together with the corresponding validation metrics, are presented in Table 4.

Table 4. Comparison of simulated and measured ground-level pollutant concentrations at three downwind distances.

Distance (m)	CO_{sim} (mg/m³)	CO_{seam} (mg/m³)	$NO_{x_{sim}}$ (mg/m³)	$NO_{x_{meas}}$ (mg/m³)	$SO_{2_{sim}}$ (mg/m³)	$SO_{2_{meas}}$ (mg/m³)
50	128.0	135.0	106.5	113.0	0.0880	0.0815
100	123.0	128.5	89.5	95.0	0.0850	0.0782
200	105.0	111.0	86.2	92.0	0.0825	0.0768

The predictive performance of the CFD framework was quantitatively assessed by comparing simulated ground-level pollutant concentrations with field measurements at three strategic downwind distances (50, 100, and 200 m) from the flare base. Model accuracy was evaluated using the Root-Mean-Square Error (RMSE) and the coefficient of determination (R^2).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_{sim,i} - C_{meas,i}\right)^2} \tag{15}$$

The calculated R^2 values for CO, NO_x, and SO_2 were 0.9973, 0.9996, and 0.9689, respectively, indicating an exceptional correlation between the numerical predictions and field data within the critical near-field zone. The corresponding RMSE values were 6.20 mg/m³ for CO, 5.95 mg/m³ for NO_x, and 0.00635 mg/m³ for SO_2.

These high correlation coefficients, derived from the available monitoring points, demonstrate the model's robustness in capturing the primary dispersion characteristics and plume grounding effects under varying atmospheric stability regimes. While the model exhibits a minor systematic underprediction of approximately 5-6%, a trend consistent with the known tendency of steady-state RANS formulations to underestimate peak near-source concentrations due to time-averaged turbulent fluctuation effects [38, 39], the consistency of this bias across all pollutant species confirms the physical realism and systematic reliability of the implemented k-ε framework.

3. Results and discussions

The simulations were conducted for two seasonally representative meteorological scenarios: a summer day characterized by unstable atmospheric conditions (Monin-Obukhov length L = −1148 m) and a winter day characterized by stable surface-layer inversion (L = +364 m), using a two-dimensional Eulerian CFD framework to quantify near-field concentrations of CO, NO_x, and SO_2 from a representative low-capacity gas flare. The primary objective was to assess potential exceedances of established air-quality standards within the worker exposure zone, with particular emphasis on the breathing zone at 2 m above ground level. Pollutant concentrations were analyzed at 2 m above ground level, corresponding to the representative human breathing zone, and at the elevated plume core near the flare stack outlet to evaluate vertical and downwind transport of emissions. The simulated flow field and plume-dispersion structure for both seasonal scenarios are presented in Figs. 5 to 8; concentration profiles at breathing height for CO, NO_x, and SO_2 are presented in Figs. 9 to 11. Table 5 summarizes the exceedance analysis relative to the regulatory ambient air-quality and flare-emission standards adopted in this study, and serves as the quantitative basis for the mechanistic interpretation presented in the following sections.

Table 5. Exceedance analysis of simulated pollutant concentrations relative to regulatory ambient air-quality and flare-emission standards (values estimated from simulation profiles).

Pollutant	Condition/height	Max simulated (mg/m³)	Standard (mg/m³)	Standard type	Exceedance
CO	Summer, 2 m	62	40	Ambient, 1-h	+55.0%
CO	Winter, 2 m	54	40	Ambient, 1-h	+35.0%
CO	Summer, elevated plume level	290	160	Flare emission	+81.3%
CO	Winter, elevated plume level	260	160	Flare emission	+62.5%
NO_x	Summer, 2 m	16	0.10	Ambient (screening*)	>100-fold*
NO_x	Winter, 2 m	13	0.10	Ambient (screening*)	>100-fold*
NO_x	Summer, elevated plume level	76	709	Flare emission	−89.3%
NO_x	Winter, elevated plume level	63	709	Flare emission	−91.1%
SO_2	Summer, 2 m	0.38	0.08	Primary ambient	+375%
SO_2	Winter, 2 m	0.37	0.08	Primary ambient	+362.5%
SO_2	Summer, 2 m	0.38	1.30	Secondary ambient	−70.8%
SO_2	Winter, 2 m	0.37	1.30	Secondary ambient	−71.5%
SO_2	Summer, elevated plume level	1.85	2360	Flare emission	−99.92%
SO_2	Winter, elevated plume level	1.85	2360	Flare emission	−99.92%

*NO_x values at 2 m represent near-source occupational screening indicators only. The apparent exceedance reflects the comparison between a near-source model-predicted concentration and a spatially and temporally averaged ambient criterion; see Section 3.3 for detailed interpretation.

3.1. Seasonal dispersion dynamics

The two-dimensional CFD simulations demonstrated a clear and physically consistent seasonal contrast in near-field plume behavior, as illustrated in Figs. 5 and 6. Under unstable summer atmospheric conditions ($L < 0$), buoyancy-driven uplift carried pollutants to elevated atmospheric layers within the near-field region downwind of the flare. At higher elevations, ambient wind speeds of 10-12 m/s rendered horizontal advection dominant, transporting pollutants further downwind before descending to ground level. This behavior is consistent with the convective mixing dynamics of unstable boundary layers, in which enhanced effective vertical turbulent diffusivity promotes rapid dilution [4, 37].

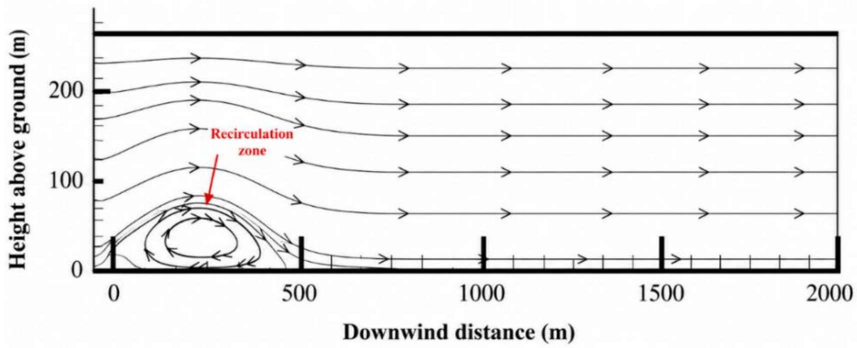
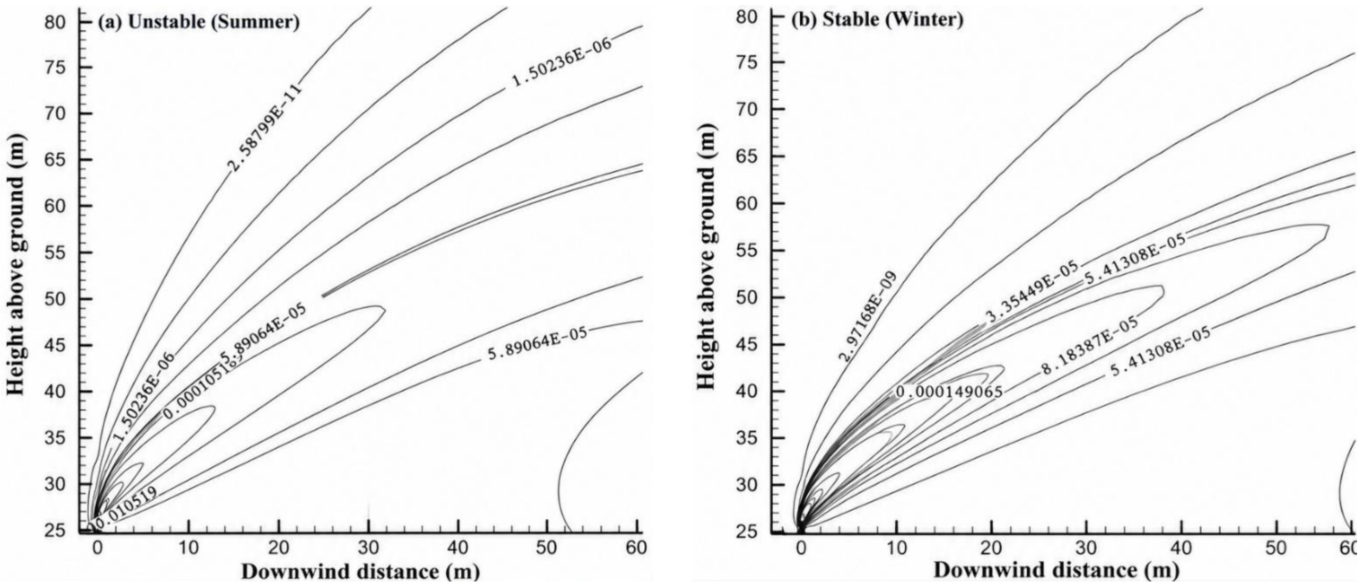


Fig. 5. Streamline pattern and near-surface recirculation structure under stable winter atmospheric conditions (L = +364 m). The simulated flow field shows the formation of a low-velocity recirculation zone within the near-field region downwind of the flare stack, contributing to enhanced near-ground pollutant accumulation.

During stable winter conditions ($L > 0$), the suppression of vertical mixing associated with stable atmospheric stratification confined the plume to lower atmospheric layers. As shown in Fig. 5, the formation of a recirculation zone in the lee of the flare stack, driven by the interaction of the ambient wind field with the vertical stack geometry, generated a region of elevated near-surface concentration between approximately 50 and 400 m downwind. Localized reverse flow and low-velocity recirculation regions were observed near the surface in this zone, further contributing to near-surface pollutant accumulation. These dynamics are physically consistent with the stable boundary layer structure imposed by the winter Monin-Obukhov parameters (u = 0.169 m/s, θ = +0.00554 K, L = +364 m) and with previously reported behavior of industrial plumes under surface-layer inversion conditions [30, 39].

The seasonal contrast between vertical plume rise in summer and near-ground horizontal trapping in winter, as shown comparatively in Fig. 6, constitutes the primary mechanism governing the spatial extent and magnitude of worker-zone exposure in the vicinity of the flare and represents the central physical finding of this study.



(a) (b)

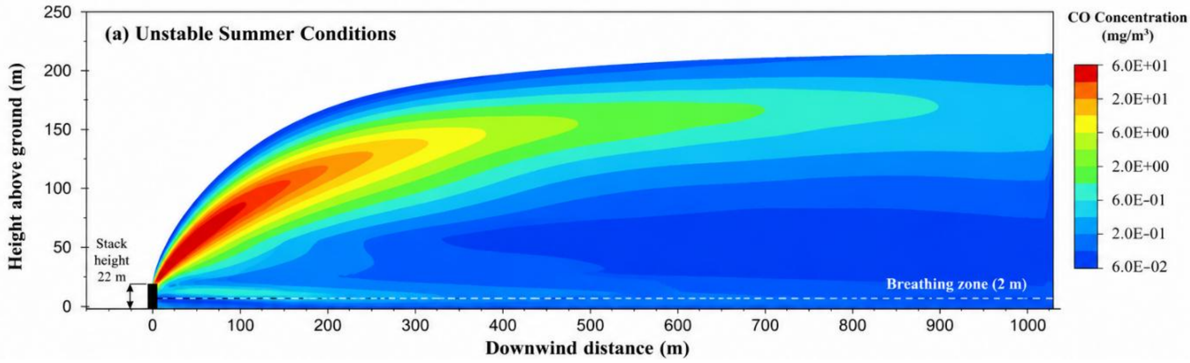
Fig. 6. Near-field plume-dispersion contours predicted by the two-dimensional CFD model under: (a) unstable summer atmospheric conditions; and (b) stable winter atmospheric conditions. The contours illustrate the seasonal contrast between enhanced vertical plume rise during unstable conditions and increased near-surface confinement under stable stratification.

3.2. Carbon monoxide

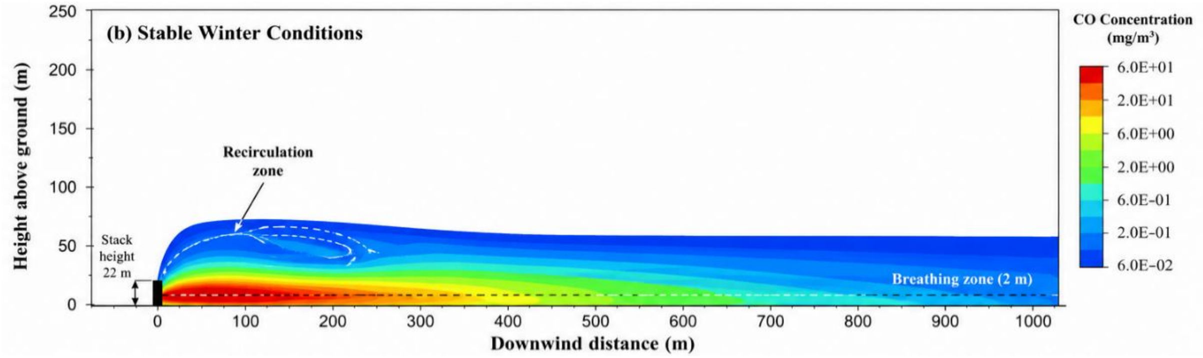
The downwind variation of CO concentration at breathing height (2 m) for both seasonal scenarios, together with the applicable regulatory ambient air-quality standard, is presented in Fig. 8. CO concentrations in the worker exposure zone are substantially higher than at greater distances, with the steepest concentration gradient occurring within a 250 m radius of the flare. At the elevated plume-core level, CO exceeded the flare emission standard over the near-source region, particularly within approximately 3-13 m downwind of the flare stack.

The peak simulated CO concentration at breathing height was 62 mg/m³ in summer and 54 mg/m³ in winter, corresponding to exceedances of the regulatory ambient air-quality standard (40 mg/m³, 1-hour average) of 55.0% and 35.0%, respectively (Table 5). At the elevated plume-core level, peak CO concentrations of 290 mg/m³ in summer and 260 mg/m³ in winter exceeded the applicable flare emission standard of 160 mg/m³ by 81.3% and 62.5%, respectively.

The higher peak breathing-zone CO concentration in summer (62 mg/m³) compared with winter (54 mg/m³) reflects the influence of near-wake recirculation dynamics in the immediate vicinity of the flare stack. Under unstable summer conditions, although overall vertical dilution is enhanced at higher atmospheric levels, the stronger exit momentum of the heated plume drives a portion of the gas downward in the recirculation zone immediately downstream of the stack, producing locally elevated near-ground concentrations within the first 50-100 m. This near-source recirculation effect outweighs the broader dilution advantage of unstable stratification in the immediate worker-exposure zone, indicating that source-exit conditions rather than far-field meteorological dilution are the dominant controls on peak breathing-zone CO exposure near a low-capacity flare. This physical behavior is visually captured in the CO concentration contours presented in Fig. 7.



(a)



(b)

Fig. 7. Simulated CO concentration contours under: (a) unstable summer conditions; and (b) stable winter conditions. The contour fields illustrate the transition from buoyancy-driven plume rise during unstable atmospheric conditions to enhanced near-ground confinement under stable stratification.

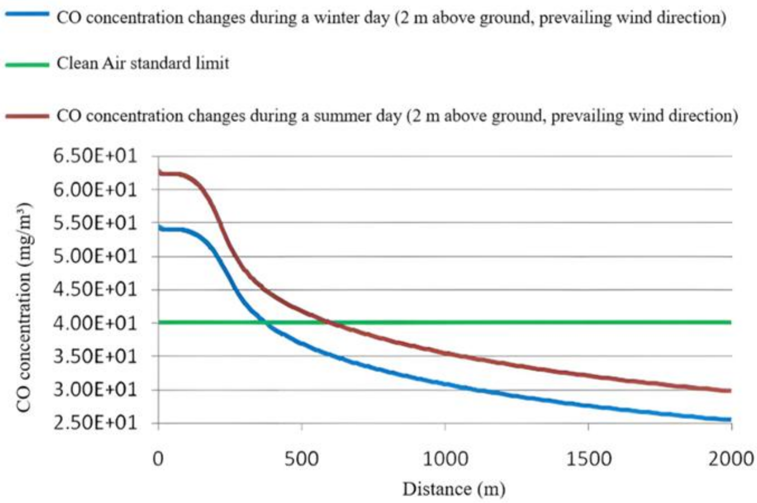


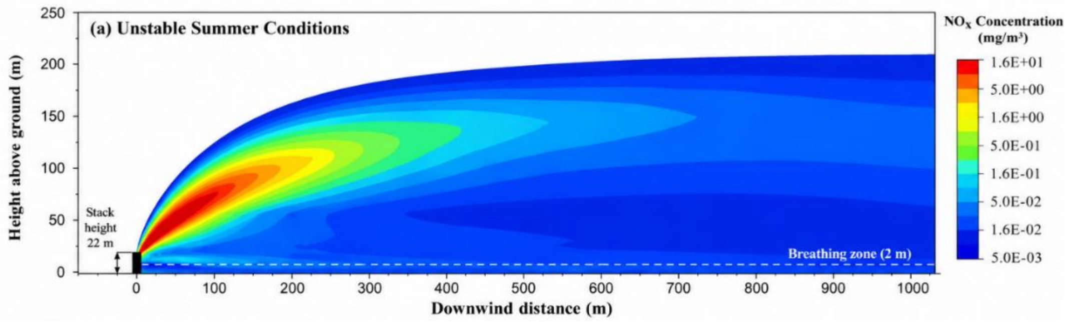
Fig. 8. Downwind variation of CO concentration at breathing height (2 m above ground level) under summer (unstable) and winter (stable) atmospheric conditions, together with the applicable regulatory ambient air-quality standard for CO.

CO concentrations progressively decreased with downwind distance in both seasonal scenarios, approaching regulatory compliance levels beyond the near-source exposure zone. From an occupational safety perspective, these findings underscore the need for targeted emission control measures, particularly within the worker exposure zone immediately surrounding the flare.

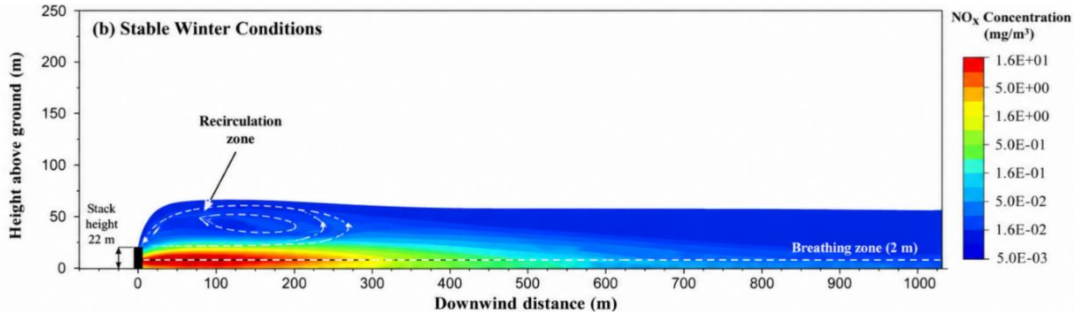
3.3. Nitrogen oxides

The downwind variation of NO_x concentration at breathing height for both seasonal scenarios, together with the applicable regulatory ambient air-quality standard, is presented in Fig. 10. At the elevated plume-core level, predicted NO_x concentrations remained well below the applicable flare emission standard of 709 mg/m³ in both seasons summer: 76 mg/m³ (89.3% below limit); winter: 63 mg/m³ (91.1% below limit) attributable to the high-temperature combustion conditions (900-1200 K) that promote effective plume buoyancy and vertical transport [10, 14].

At breathing height (2 m), simulated NO_x concentrations of 16 mg/m³ in summer and 13 mg/m³ in winter were substantially above the regulatory ambient air-quality standard of 0.10 mg/m³ (Table 5). It should be emphasized that this comparison involves a near-source occupational exposure zone rather than regional ambient air-quality conditions: the simulated values correspond to near-source concentrations in the immediate vicinity of the flare stack, whereas the regulatory ambient standard represents a spatially and temporally averaged criterion. The exceedance values exceeding the ambient standard by more than 100-fold are therefore used here as conservative screening indicators of potential near-field occupational exposure rather than as direct compliance assessments [30, 39].



(a)



(b)

Fig. 9. Simulated NO_x concentration contours under: (a) unstable summer conditions; and (b) stable winter conditions. The contour maps highlight the influence of atmospheric stability on near-field plume transport and localized concentration accumulation.

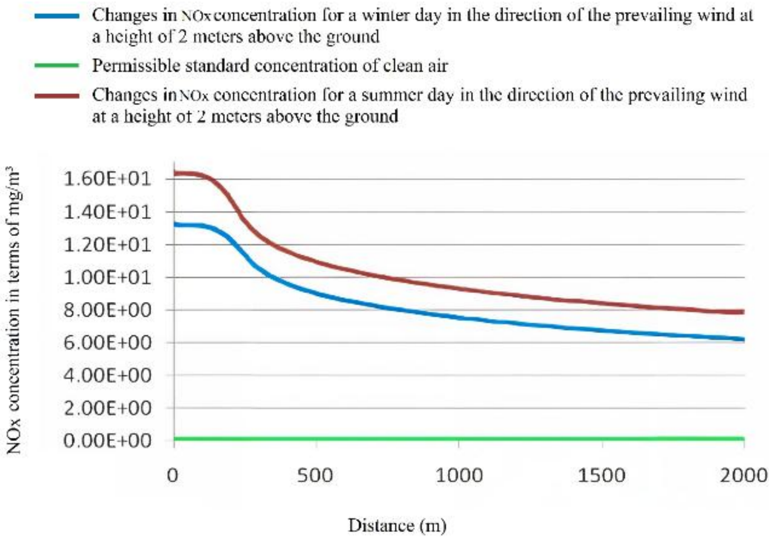


Fig. 10. The changes in NO_x concentration at a height of 2 m above ground level (height of breathing) for both winter and summer days, and the standard concentration for clean air is permitted.

Although unstable summer conditions enhanced overall vertical plume transport, the peak near-source NO_x concentration at breathing height was slightly higher in summer (16 mg/m³) than in winter (13 mg/m³). This pattern is consistent with the near-wake recirculation mechanism described for CO in Section 3.2: source momentum and near-source recirculation controlled the local maximum in the immediate worker exposure zone, independent of the broader seasonal dilution regime. The spatial distribution of NO_x and the resulting near-field concentration hotspots under both seasonal conditions are illustrated in Fig. 9.

3.4. Sulfur dioxide

Fig. 11 presents the downwind variation of SO_2 concentrations at breathing height for both seasonal scenarios, alongside the applicable primary and secondary regulatory ambient air-quality standards and the flare emission standard. Peak SO_2 concentrations at breathing height were 0.38 mg/m³ in summer and 0.37 mg/m³ in winter, exceeding the primary ambient air-quality standard (0.08 mg/m³) by 375% and 362.5%, respectively, while remaining below the secondary ambient standard (1.30 mg/m³) and well below the flare emission limit at the elevated plume-core level, SO_2 concentrations of 1.85 mg/m³ were 99.92% below the applicable flare emission standard in both seasons (Table 5).

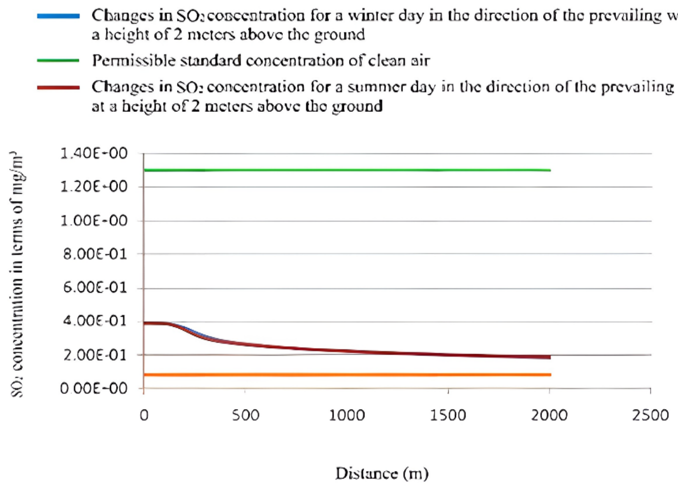


Fig. 11. The change in SO_2 concentration at a height of 2 m above ground level (height of breathing) for both winter and summer days, and authorized concentrations of primary and secondary standards for clean air.

Exceedance of the primary SO_2 standard designed to protect sensitive populations, including children, the elderly, and individuals with pre-existing respiratory conditions, at breathing height in the near-field zone represents a significant occupational health concern. While the secondary standard, intended to protect public welfare, including visibility, crops, and ecosystems, was not exceeded, the persistent exceedance of the primary standard reinforces the need for targeted emission control for sulfur-containing flare gas streams.

SO_2 exhibited notably similar near-ground concentrations between summer (0.38 mg/m³) and winter (0.37 mg/m³) scenarios, suggesting a lower sensitivity of SO_2 near-surface dispersion to the seasonal atmospheric stability variations considered in this study. SO_2 behavior followed a broadly similar spatial pattern to NO_x but at substantially lower absolute concentrations, consistent with the lower SO_2 emission rates applied in the simulations (0.039-0.047 g/s versus 0.86-1.66 g/s for NO_x).

3.5. Sensitivity analysis: seasonal atmospheric stability as a surrogate for meteorological variability

In response to the reviewer's request for a sensitivity analysis with respect to wind speed, this study employed two seasonally contrasting atmospheric scenarios as physically realistic bounding cases: the unstable summer scenario (u^* = 0.186 m/s, $L = -1148$ m) and the stable winter scenario (u^* = 0.169 m/s, L = +364 m). Wind speed and atmospheric stability were not independently varied in the present study in order to preserve physically consistent seasonal boundary-layer conditions derived from Monin-Obukhov similarity theory [4, 41]; independent parametric variation of wind speed while holding stability class constant would require boundary conditions inconsistent with the observed seasonal meteorology of the study region.

The results demonstrate that the transition from stable to unstable stratification produced a fundamentally different plume morphology, vertical uplift, and elevated-level dispersion in summer versus near-ground horizontal trapping and recirculation in winter, as illustrated in Figs. 5 and 6. Despite the stronger overall dilution associated with summer instability, near-source recirculation effects produced higher peak breathing-zone concentrations of CO and NO_x in summer than in winter, highlighting the complexity of near-field exposure dynamics around low-capacity flares and the importance of source exit conditions in governing worker-zone exposure intensity. The two scenarios together bound the range of near-field dispersion outcomes expected across Iran's southwestern oil-producing regions throughout the annual meteorological cycle. A dedicated parametric wind-speed sensitivity study, holding the stability class constant, is identified as a priority for future investigation [41].

These results further suggest that any climate-driven shift in boundary-layer stability could alter the spatial footprint of flare-related pollutant exposure; however, this hypothesis requires dedicated climate-dispersion coupling in future studies to be rigorously evaluated [27, 33].

3.6. Mechanistic interpretation, benchmarking against prior dispersion studies, and practical implications

The present findings are consistent with prior dispersion studies of gas flaring environments and provide additional mechanistic resolution of near-field exposure dynamics. Classical advection-diffusion theory predicts that unstable atmospheric conditions enhance vertical turbulent exchange, whereas stable stratification suppresses vertical mixing and promotes near-surface accumulation [1, 37]. The seasonal contrast resolved in the present simulations follows this expected physical behavior, as captured in the concentration contours of Figs. 7 to 9.

The simulated recirculation zone extending approximately 50-400 m downwind under stable winter conditions indicates that local aerodynamic effects can govern worker-zone exposure independently of emission magnitude alone. Turbulent recirculation of this type has been widely recognized as a controlling mechanism in industrial plume transport because low-velocity vortical regions reduce plume ventilation and prolong pollutant residence time near the surface [38]. In the present simulations, this mechanism explains why substantial near-ground concentrations persisted despite relatively low flare emission rates.

The present CFD results also clarify several limitations of Gaussian and mesoscale dispersion approaches when applied to near-field flare environments. Mirrezaei [30], using the CALPUFF framework to model SO_2 dispersion from gas flares on Sirri Island, reported that lower flare heights resulted in elevated ground-level concentrations. The present simulations support that conclusion while extending it mechanistically by explicitly resolving plume attachment, wake recirculation, and breathing-zone concentration gradients, as well as near-field flow structures that Gaussian formulations cannot reproduce because their underlying parameterizations do not resolve localized vortex formation or stack-induced recirculation [30, 41]. Similarly, regional assessments based on satellite observations and combustion equations have advanced understanding of flare emission burdens across Iranian oil-producing regions [9, 16], but do not resolve the spatial distribution of pollutant concentrations within the worker exposure zone. The present Eulerian CFD framework, therefore, provides complementary information by identifying where pollutant accumulation occurs and which atmospheric mechanisms govern the intensity of localized exposure.

The exceedance analysis presented in Table 5 further differentiates between source-level emission control and near-field occupational exposure. CO exceeded both the ambient air-quality criterion and the flare emission standard, indicating relevance as both a source-performance indicator and a direct occupational exposure concern. SO_2 exceeded the primary ambient criterion at breathing height while remaining below the secondary ambient and flare emission standards, suggesting that its principal risk in this context is localized respiratory exposure rather than regional environmental degradation. The NO_x comparison at breathing height is most appropriately interpreted as a conservative occupational screening indicator, given that the simulated values correspond to near-source instantaneous concentrations rather than the spatially and temporally averaged conditions represented by regulatory ambient standards.

The seasonal dependence of the simulated concentration fields also has implications for climate-sensitive air-quality assessment. Previous studies have demonstrated that climate-driven changes in atmospheric stability, mixing depth, and boundary-layer structure can alter pollutant transport pathways [27, 33, 34, 40]. The present simulations do not constitute a climate projection; however, they demonstrate that changes in atmospheric stability alone can substantially alter the geometry of flare-related worker exposure. Increased instability enhances plume rise and vertical dilution, whereas stable inversion conditions intensify near-ground trapping and expand the horizontal footprint of the worker exposure zone. These findings suggest that future changes in regional boundary-layer dynamics may influence both the intensity and spatial distribution of flare-related exposure in oil-producing regions.

From an operational perspective, the results indicate that mitigation strategies should address both emission magnitude and plume transport dynamics. Flare gas recovery systems remain among the most effective approaches for reducing pollutant release at the source [15, 16]. In addition, the simulated concentration fields indicate that operational controls are necessary to limit occupational exposure under unfavorable meteorological conditions. Increasing effective stack height, improving combustion efficiency, and restricting worker activity within the downwind recirculation zone during stable atmospheric episodes represent practical measures that could substantially reduce near-field exposure intensity.

Several limitations should be acknowledged. First, the two-dimensional CFD framework captures dominant downwind-vertical transport structure but does not resolve crosswind concentration gradients or fully three-dimensional vortex interactions; predicted concentrations should therefore be interpreted as representative of the principal plume plane rather than the complete lateral exposure field. Second, the validation dataset comprised three downwind measurement locations; although agreement between simulated and measured concentrations was strong ($R^2 > 0.97$ for all pollutants), the limited number of field observations reduced statistical robustness. Third, sensitivity analysis was performed using seasonally representative Monin-Obukhov scenarios rather than explicit parametric wind-speed variations. Fourth, steady-state meteorological conditions were assumed, and transient concentration peaks associated with turbulent fluctuations or wind direction shifts were not captured. Future investigations should incorporate expanded field datasets, three-dimensional transient CFD or LES formulations, and explicit meteorological sensitivity analyses to improve characterization of peak occupational exposure conditions.

4. Conclusions

This study applied a two-dimensional Eulerian CFD framework with RNG k-ε turbulence closure to simulate the near-field dispersion of CO, $\mathrm{NO_x}$, and $\mathrm{SO_2}$ from a representative low-capacity gas flare under unstable summer ($L = -1148$ m) and stable winter ($L = +364$ m) atmospheric conditions. The results demonstrate that the atmospheric stability regime, rather than emission magnitude alone, is the dominant control on the spatial distribution and intensity of pollutant exposure in the worker zone. Under unstable summer conditions, buoyancy-driven vertical uplift transported the plume into elevated atmospheric layers; however, near-source exit momentum and wake-induced recirculation produced slightly higher peak breathing-zone concentrations of CO and $\mathrm{NO_x}$ in summer than in winter, confirming that source exit conditions govern near-field occupational exposure intensity. Under stable winter conditions, suppression of vertical mixing generated a recirculation zone between approximately 50 and 400 m downwind, producing a wider near-surface accumulation footprint and the greatest spatial extent of regulatory exceedance.

Quantitatively, CO concentrations at breathing height exceeded the regulatory ambient standard by 55.0% in summer and 35.0% in winter; at the elevated plume-core level, CO exceeded the flare emission standard by 81.3% and 62.5%, respectively. $\mathrm{SO_2}$ exceeded the primary ambient standard by 375% in summer and 362.5% in winter at breathing height, while remaining below the secondary ambient and flare emission limits. $\mathrm{NO_x}$ remained below the flare emission standard at the plume core level in both seasons; near-source breathing-zone concentrations substantially exceeded the ambient criterion, though this should be interpreted as a conservative occupational screening indicator rather than a direct compliance metric. Model validation against field measurements yielded R^2 values of 0.997, 0.9996, and 0.969 for CO, $\mathrm{NO_x}$, and $\mathrm{SO_2}$, respectively, confirming the physical realism of the simulated concentration fields.

From an operational perspective, the findings support three targeted recommendations: implementation of flare gas recovery systems to reduce source-level emissions; increase of effective stack height to elevate the plume core above the near-surface recirculation zone; and meteorology-aware management of worker activity within the 50-400 m downwind zone, with particular caution during stable winter episodes. The seasonal dispersion contrast identified here, vertical uplift in summer versus near-ground horizontal trapping in winter, is physically consistent with projected changes in regional boundary-layer stability, suggesting that shifts in atmospheric stability under continued warming may alter both the intensity and spatial footprint of flare-related occupational exposure. Three-dimensional transient CFD simulations, expanded field measurement campaigns, and dedicated climate-dispersion coupling are identified as priorities for future investigation.

Statements & declarations

Author contributions

Siamak Boudaghpour: Investigation, Formal analysis, Resources, Writing - Original Draft.

Ehsan Qamari: Conceptualization, Methodology, Resources, Writing - Original Draft.

Amirali Bahadori: Formal analysis, Project administration, Writing - Review & Editing.

Funding

The authors did not receive support from any organization for the submitted work.

Data availability

The data presented in this study are available on request from the corresponding author.

Declarations

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

[1] Singh, R. N. Advection diffusion equation models in near-surface geophysical and environmental sciences. Indian Geophysical Union, 2013; 17: 117–127.

[2] Roberts, O. F. T. The theoretical scattering of smoke in a turbulent atmosphere. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 1923; 104: 640–654. doi:10.1098/rspa.1923.0132.

[3] Kafle, J., Adhikari, K., Poudel, E., Pant, R. Mathematical Modeling of Pollutants Dispersion in the Atmosphere. Journal of Nepal Mathematical Society, 2024; 7: 61–70. doi:10.3126/jnms.v7i1.67487.

[4] Monin, A. S., Obukhov, A. M. Basic laws of turbulent mixing in the surface layer of the atmosphere. Trudy Geofiz, Instituta Akademii Nauk, SSSR (Proceedings of Geophysics Institute, National Academy of Science, SSSR), 1954; 24: 163–187.

[5] Buske, D., Vilhena, M., Silva, E., Tirabassi, T., Bodmann, B. Dispersion of pollutants, GILTT, Decomposition method, OLAD; Dispersion of pollutants, GILTT, Decomposition method, OLAD. American Journal of Environmental Engineering, 2016; 6: 9–12. doi:10.5923/s.ajee.201601.03.

[6] Kgabi, N., Mokgwetsi, T. Dilution and dispersion of inhalable particulate matter. WIT Transactions on Ecology and the Environment, 2009; 127: 229–238. doi:10.2495/RAV090201.

[7] Ulfah, S., Awalludin, S. A., Wahidin Advection-diffusion model for the simulation of air pollution distribution from a point source emission. Journal of Physics: Conference Series, 2018; 948: 012067. doi:10.1088/1742-6596/948/1/012067.

[8] Correnty, S., Freitag, M. A., Soodhalter, K. M. Chebyshev HOPGD with sparse grid sampling for parameterized linear systems. Calcolo, 2025; 62: 28. doi:10.1007/s10092-025-00652-1.

[9] Shojaei, S. M., Vahabpour, A., Saifoddin, A. A., Ghasempour, R. Estimation of greenhouse gas emissions from Iran's gas flaring by using satellite data and combustion equations. Integrated Environmental Assessment and Management, 2023; 19: 735–748. doi:10.1002/ieam.4684.

[10] Talebi, A., Fatehifar, E., Alizadeh, R., Kahforoushan, D. The Estimation and Evaluation of New CO, CO2, and NOx Emission Factors for Gas Flares Using Pilot Scale Flare. Energy Sources, 2014; 36: 719–726. doi:10.1080/15567036.2011.584117.

[11] Conrad, B. M., Johnson, M. R. Field Measurements of Black Carbon Yields from Gas Flaring. Environmental Science & Technology, 2017; 51: 1893–1900. doi:10.1021/acs.est.6b03690.

[12] Popovicheva, O., Timofeev, M., Persiantseva, N., Jefferson, M. A., Johnson, M., Rogak, S. N., Baldelli, A. Microstructure and Chemical Composition of Particles from Small-scale Gas Flaring. Aerosol and Air Quality Research, 2019; 19: 2205–2221. doi:10.4209/aaqr.2019.04.0177.

[13] Kindzierski, W. B. Importance of human environmental exposure to hazardous air pollutants from gas flares. Environmental Reviews, 2000; 8: 41–62. doi:10.1139/a00-005.

[14] Leahey, D. M., Preston, K., Strosher, M. Theoretical and observational assessments of flare efficiencies. J Air Waste Manag Assoc, 2001; 51: 1610–6. doi:10.1080/10473289.2001.10464390.

[15] Zadakbar, O., Vatani, A., Karimpour, K. Flare Gas Recovery in Oil and Gas Refineries. Oil & Gas Science and Technology - Revue de l'IFP, 2008; 63: 705–711. doi:10.2516/ogst:2008023.

[16] Shahab-Deljoo, M., Medi, B., Kazi, M.-K., Jafari, M. A techno-economic review of gas flaring in Iran and its human and environmental impacts. Process Safety and Environmental Protection, 2023; 173: 642–665. doi:10.1016/j.psep.2023.03.051.

[17] Ismet Ara, K. N., Rahaman, M., Alam, M. S. Numerical Solution of Advection Diffusion Equation Using Semi-Discretization Scheme. Applied Mathematics, 2021; 12: 1236–1247. doi:10.4236/am.2021.1212079.

[18] Chen, Y., Falconer, R. A. Advection-diffusion modelling using the modified QUICK scheme. International Journal for Numerical Methods in Fluids, 1992; 15: 1171–1196. doi:10.1002/fld.1650151003.

[19] Spiegelman, M., Katz, R. A semi-Lagrangian Crank-Nicolson algorithm for the numerical solution of advection-diffusion problems. Geochemistry Geophysics Geosystems, 2006; 7: 1–12. doi:10.1029/2005GC001073.

[20] Gokieli, M., Szczepańczyk, A. A Numerical Scheme for Evacuation Dynamics. In: Parallel Processing and Applied Mathematics; 2020 Sep 8–11; Bialystok, Poland. p. 277–286. doi:10.1007/978-3-030-43222-5_24.

[21] Hou, Q., Sun, Z., He, L., Karemat, A. Orthogonal grid physics-informed neural networks: A neural network-based simulation tool for advection–diffusion–reaction problems. Physics of Fluids, 2022; 34: 077108. doi:10.1063/5.0095536.

[22] Zio, S., Lamien, B., Beidari, M., Tougri, I. Numerical simulation of pollutants transport in groundwater using deep neural networks informed by physics. Gulf Journal of Mathematics, 2024; 16: 337–352. doi:10.56947/gjom.v16i2.1863.

[23] Berardi, M., Difonzo, F., Icardi, M. Inverse Physics-Informed Neural Networks for Transport Models in Porous Materials. Computer Methods in Applied Mechanics and Engineering, 2025; 435: 117628. doi:10.1016/j.cma.2024.117628.

[24] Huhn, Q. A., Tano, M. E., Ragusa, J. C. Physics-Informed Neural Network with Fourier Features for Radiation Transport in Heterogeneous Media. Nuclear Science and Engineering, 2023; 197: 2484–2497. doi:10.1080/00295639.2023.2184194.

[25] Elham, A., Ghole, V. S., Sen, P. N. Study on the Status of SO2 in Tehran - Iran. Journal of Applied Sciences and Environmental Management, 2009; 10: 75–82. doi:10.4314/jasem.v10i2.43670.

[26] Mobarak Hassan, E., Alizadeh, O. Dust events in southwestern Iran: Estimation of PM10 concentration based on horizontal visibility during dust events. International Journal of Climatology, 2022; 42: 5159–5172. doi:10.1002/JOC.7525.

[27] Jacob, D. J., Winner, D. A. Effect of climate change on air quality. Atmospheric Environment, 2009; 43: 51–63. doi:10.1016/j.atmosenv.2008.09.051.

[28] Broomandi, P., Karaca, F., Nikfal, A., Jahanbakhshi, A., Tamjidi, M., Kim, J. R. Impact of COVID-19 Event on the Air Quality in Iran. Aerosol and Air Quality Research, 2020; 20: 1793–1804. doi:10.4209/aaqr.2020.05.0205.

[29] Ramezani, N. a., Alijani, B., Borna, r. Explaining the Effects of Climate Elements in Tehran's Metropolis Air Quality. Geographical Researches, 2018; 33: 154–169. doi:10.29252/geores.33.3.154.

[30] Mirrezaei, M. A. Impact of Meteorological Parameters on Dispersion Modeling of Sulfur Dioxide from Gas Flares (Case Study: Sirri Island). Iranian (Iranica) Journal of Energy and Environment, 2019; 10: 288–295. doi:10.5829/ijee.2019.10.04.10.

[31] Asadi-Fard, E., Falahatkar, S., Tanha Ziyarati, M., Zhang, X. A new achievement of satellite-based gas flaring volume estimation: decision tree modeling. Earth Science Informatics, 2024; 17: 2887–2901. doi:10.1007/s12145-024-01316-4.

[32] Pakrooh, P., Pishbahar, E. Forecasting Air Pollution Concentrations in Iran, Using a Hybrid Model. Pollution, 2019; 5: 739–747. doi:10.22059/poll.2019.274827.572.

[33] Ebi, K., McGregor, G. Climate Change, Tropospheric Ozone and Particulate Matter, and Health Impacts. Environmental Health Perspectives, 2008; 116: 1449–1455. doi:10.1289/ehp.11463.

[34] Fang, Y., Fiore, A. M., Horowitz, L. W., Gnanadesikan, A., Held, I., Chen, G., Vecchi, G., Levy, H. The impacts of changing transport and precipitation on pollutant distributions in a future climate. Journal of Geophysical Research: Atmospheres, 2011; 116: 1–14. doi:10.1029/2011JD015642.

[35] Rede, G. D., Dongre, S., Kumar, C. R., Kant, R., Ramana, V. V., Rathnamma, M. V. A Novel Approach to Mitigating Air Pollution and Climate Change with Graph Convolutional RNN Models. In: 2025 5th International Conference on Trends in Material Science and Inventive Materials (ICTMIM); 2025 Apr 07–09; Kanyakumari, India. p. 1468–1473. doi:10.1109/ICTMIM65579.2025.10988109.

[36] Zhonghua, W., Mujumdar, A. S. CFD modeling of the gas–particle flow behavior in spouted beds. Powder Technology, 2008; 183: 260–272. doi:10.1016/j.powtec.2007.07.040.

[37] Jacobson, M. Z. Fundamentals of Atmospheric Modeling. 2nd ed. Cambridge (UK): Cambridge University Press; 2005. doi:10.1017/CBO9781139165389.

[38] Akkanen, J., Slootweg, T., Mäenpää, K., Leppänen, M., Agbo, S., Gallambois, C., Kukkonen, J. Bioavailability of Organic Contaminants in Freshwater Environments. In: Guasch, H., Ginebreda, A., Geiszinger, A., editors. Emerging and Priority Pollutants in Rivers. The Handbook of Environmental Chemistry. Berlin: Springer-Verlag Berlin Heidelberg; 2012. p. 25–53. doi:10.1007/978-3-642-25722-3_2.

[39] Pope, S. B. Turbulent Flows. 1st ed. Cambridge (UK): Cambridge University Press; 2000. doi:10.1017/CBO9781316179475.

[40] Lamon, L., von Waldow, H., MacLeod, M., Scheringer, M., Marcomini, A., Hungerbühler, K. Modeling the Global Levels and Distribution of Polychlorinated Biphenyls in Air under a Climate Change Scenario. Environmental Science & Technology, 2009; 43: 5818–5824. doi:10.1021/es900438j.

[41] Zannetti, P. Air pollution modeling: theories, computational methods and available software. 2nd ed. New York (NY): Springer Science & Business Media; 2013. doi:10.1007/978-1-4757-4465-1.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

Investigating the Effect of Silica Fume on the Engineering Behavior and Microstructure of Clay Liners

Sadegh Ghavami [a*], Sina Tahmasebi [b], Narges Delnavaz [a]

[a] *Faculty of Civil Engineering, Sahand University of Technology, Tabriz, Iran*
[b] *School of Civil Engineering, Alaoddoleh Semnani Institute of Higher Education, Garmsar, Iran*

ARTICLE INFO

Article history:
Received 27 April 2026
Revised 17 May 2026
Accepted 16 June 2026
Available online 01 April 2027

Keywords:
Clay liner
Silica fume
Permeability
Strength
Swelling pressure

ABSTRACT

Clay liners are widely used in landfill systems to prevent the migration of contaminants into the surrounding soil and groundwater by providing low permeability and adequate mechanical stability. Various stabilizing agents have been investigated to enhance the engineering performance of these clayey soils, particularly materials that are economical, readily available, and capable of participating in pozzolanic reactions. Limited attention has been given to the application of silica fume, a highly fine amorphous silica produced as a by-product of silicon or ferrosilicon alloy manufacturing, in clay liner systems. In this study, the influence of silica fume contents of 5%, 10%, and 15% (by dry weight of soil) on the behavior of high-plasticity clay was evaluated through standard compaction, unconfined compressive strength, free swell, and permeability tests. The experimental results showed that increasing silica fume content decreased maximum dry density, permeability, swelling percentage, and swelling pressure, while increasing optimum moisture content and unconfined compressive strength. The specimen containing 15% silica fume exhibited more than a 50% reduction in permeability and achieved compressive strength values exceeding the minimum recommended for clay liner materials. In addition, swelling pressure decreased by more than 30% at this dosage. Microstructural observations obtained from scanning electron microscopy confirmed the development of a denser and more compact soil matrix following silica fume treatment.

How to cite this article: Ghavami, S., Tahmasebi, S., Delnavaz, N. Investigating the Effect of Silica Fume on the Engineering Behavior and Microstructure of Clay Liners. Civil Engineering and Applied Solutions. 2027; 3(2): 16–25. doi:10.22080/ceas.2026.31610.1094.

1. Introduction

In many cases, the soils at sites considered for landfill development do not possess the required engineering properties. Because the primary objective of landfill design is to isolate waste from the surrounding environment and groundwater [1], the engineering characteristics of the on-site soil are critically important. Ideally, the soil should have very low permeability, high pollutant adsorption capacity, adequate flexibility to accommodate moisture variations, minimal settlement, and sufficient strength to withstand imposed loads [2]. Clay soils are commonly used as liners and base layers in landfills because of their low permeability. When the clay available at a site does not possess the required properties for use as a liner, and the cost of transporting large volumes of suitable soil is high, a cost-effective alternative is to chemically stabilize the existing soil using appropriate additives [3]. Today, in order to reduce environmental pollution and the consumption of energy and natural resources, there is growing interest in environmentally friendly methods for soil improvement [4-7], including the use of industrial by-products such as silica fume in chemical soil stabilization [8, 9].

According to ACI 116R [10], silica fume is an extremely fine, amorphous silica produced in electric arc furnaces and is a by-

* Corresponding author.
E-mail addresses: ghavamijamal@sut.ac.ir (S. Ghavami).

https://doi.org/10.22080/ceas.2026.31610.1094
ISSN: 3092-7749/

product of the silicon metal and ferrosilicon alloy industries. During the production of 1 ton of ferrosilicon, approximately 200-400 kg of silica fume is generated as a by-product [11]. The silica fume production process is illustrated in Fig. 1.

The amorphous structure and high specific surface area of silica fume particles enhance its reactivity. The use of silica fume as a supplementary or alternative material to Portland cement in concrete production has increased significantly [12]. In recent years, considerable research has focused on the application of silica fume for soil improvement and stabilization. Kalkan [13] demonstrated that the addition of silica fume significantly reduces shrinkage cracks in clay soils, and similar findings were reported by A. Al-Azzawi et al. [14]. Moradi et al. [15] studied marl soil-lime mixtures and found that adding silica fume increased the plastic limit, reduced the plasticity index, enhanced soil strength, and improved bearing capacity. Mostafa et al. [16] also observed that combining lime and silica fume increased the compressive strength, CBR, cohesion, and internal friction angle of the soil while reducing its free swelling. Goodarzi et al. [17] reported that partially replacing Portland cement with silica fume decreased the soil's swelling potential, increased its strength, and accelerated the stabilization process. taghdisi et al. [18] documented similar effects on enhancing soil hardness and strength. Studies by Badv and Emdadi [19] on silty clay soil further confirmed that increasing the silica fume content improves cohesion and strength while reducing soil permeability. In another study, a 50% reduction in collapse potential and a 348% increase in unconfined compressive strength were observed in collapsible soil after the addition of 5% silica fume [20].Türköz et al. [21] indicated that the incorporation of silica fume can enhance the dynamic behavior of clay soils by increasing the initial shear modulus under different effective confining pressures. The positive effect of silica fume on reducing the swelling and swelling pressure of expansive soil was observed by Al-Gharbawi et al. [22]. Waleed and Alshawmar [9] found that the addition of silica fume at an optimum content of 9% improved the shear strength parameters of fine-grained soil, increasing the internal friction angle from 23.8° to 24.51° and cohesion from 15.89 kPa to 23.80 kPa.

Despite numerous studies on the stabilization of clay soils using silica fume, most previous research has mainly focused on general geotechnical applications such as subgrade improvement or soil stabilization for construction purposes. Limited attention has been given to the application of silica fume in clay liner systems used in landfills, where both low permeability and sufficient mechanical strength are required. In addition, the geotechnical behavior of the clayey soils located near the landfill in Varamin city (southeast of Tehran province, Iran) has not been previously investigated. Therefore, this study examines the influence of silica fume on the engineering behavior and microstructure of this local clay soil. Different percentages of silica fume were incorporated into the soil, and its compaction characteristics, strength, swelling potential, and hydraulic properties were evaluated. Furthermore, scanning electron microscopy (SEM) was used to investigate the microstructural changes associated with silica fume addition. The findings contribute to a better understanding of the mechanisms governing the performance of silica fume-stabilized clay liners and their potential use as sustainable landfill barrier materials.

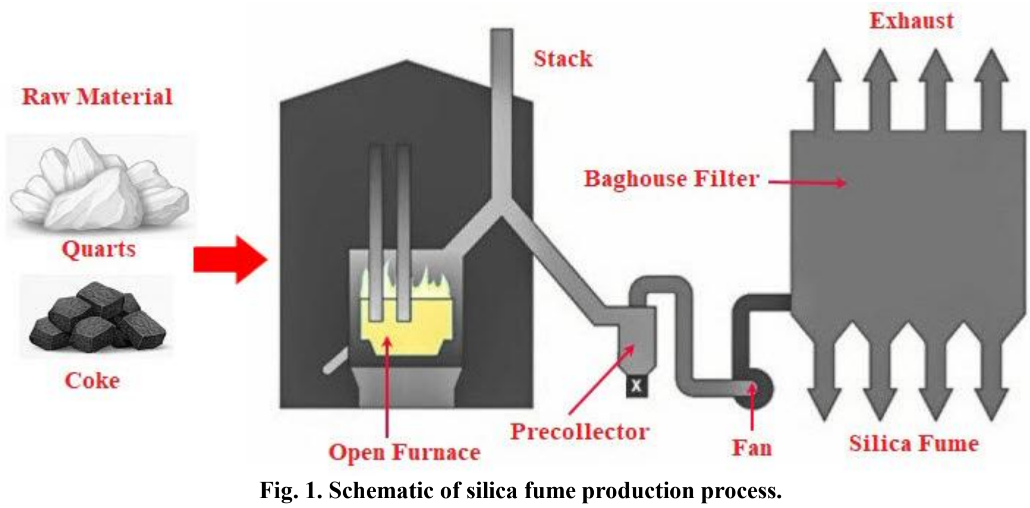


Fig. 1. Schematic of silica fume production process.

2. Materials and methods

The soil investigated in this study was collected from an area in Varamin city, located in the southeast of Tehran province. Bentonite is widely used as an additive to improve on-site soils that do not satisfy the hydraulic conductivity requirements for liner materials [1, 23, 24]. The required bentonite content depends on the type and gradation of the host soil and typically ranges from 10% to 30% by weight [25-27]. Since no standard engineering guideline has yet been established for determining the optimum bentonite content based on soil type, a bentonite content of 20% was selected in this study as a representative value within the commonly reported range for liner materials. The local soil was mixed with 20% bentonite supplied by Semnan Negin Powder Company. Hereafter, the term soil in this article refers to this prepared mixture. To determine the engineering properties of the studied soil, basic soil mechanics tests were conducted in accordance with ASTM standards. The geotechnical properties of the local soil and the local soil mixed with 20% bentonite are presented in Table 1. The grain size distribution curve of the soil (Local Soil+20% Bentonite), determined based on ASTM D422 [28], is presented in Fig. 2. According to the Unified Soil Classification System (USCS) [29], this soil is classified as high-plasticity clay (CH). The chemical compositions of the soils obtained from the X-ray fluorescence (XRF) analysis are presented in Table 2. Silica fume produced by Iran Ferroalloy Industries Company, in compliance with Iranian National Standard No. 13278 [30], was used in this study, and its physical and chemical properties are provided in Table 3. Fig. 3 shows the materials used in this study.

For soil stabilization, the recommended dosage of silica fume is usually between 5% and 15%, depending on the soil type and the targeted engineering properties. For example, Türköz et al. [21] observed the highest shear modulus in a soil mixture containing 15% silica fume. Experiments conducted by Waleed and Alshawmar [9] showed that the optimal amount of silica fume to increase soil strength was 9%. Accordingly, in this study, silica fume contents ranging from 0% to 15% were investigated. To prepare the specimens, silica fume was added to the soil at contents of 5%, 10%, and 15% by dry weight of the soil. After adding water at the target moisture content, the mixture was manually blended until a homogeneous specimen was obtained.

Table 1. Geotechnical properties of soils.

Standard	Local soil	Local soil + 20% bentonite	Standard
Liquid limit, %	37	52	ASTM D4318 [31]
Plastic limit, %	22	28	ASTM D4318 [31]
Plasticity index, %	15	24	ASTM D4318 [31]
Specific gravity (G_s)	2.69	2.67	ASTM D854 [32]
Maximum dry density, kN/m^3	14.8	14.2	ASTM D698 [33]
Optimum moisture content, %	18.2	20.1	ASTM D698 [33]
Unconfined compressive strength, kPa	76.1	86.5	ASTM D2166 [34]
Permeability, cm/s	1.03×10^{-6}	3.1×10^{-7}	ASTM D5084 [35]
Swelling pressure, kPa	91	120	ASTM D4546 [36]

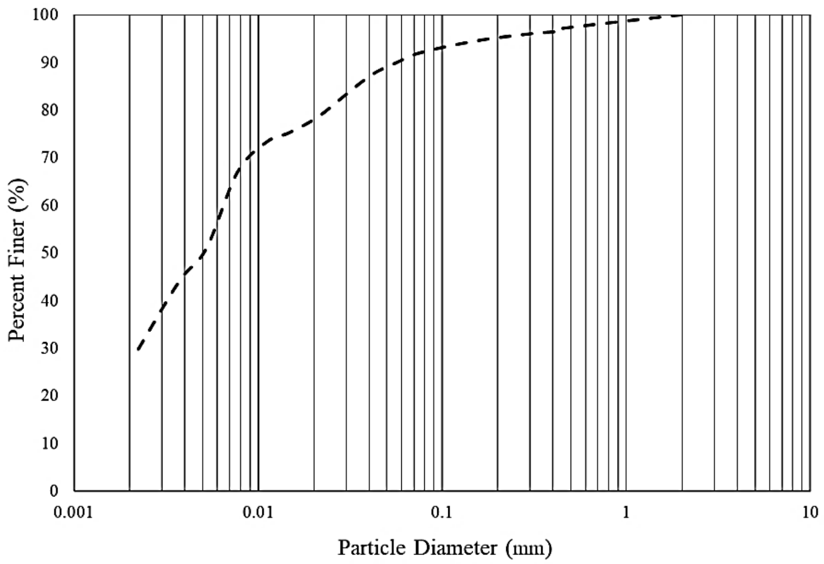


Fig. 2. Grain-size distribution curve of the soil (local soil+20% bentonite).

Table 2. Chemical compositions of the soils.

Compound	Local soil	Bentonite
SiO_2	57.9	60.05
Al_2O_3	16.1	15.48
Fe_2O_3	5.24	4.1
CaO	7.08	3.25
MgO	2.74	2.2
SO_3	1.03	0.41
K_2O	1.35	1.78
Na_2O	1.68	2.84
Other	3.6	5.83
Loss on ignition	3.28	4.06

Table 3. Physical and chemical properties of silica fume.

Chemical properties	Value	Physical properties	Value
90-95	SiO_2 (%)	Color	Grey
0.8-3	C (%)	Diameter (μm)	<1μm
Loss on ignition (%)	1.5-2.5	Specific gravity	2.2
		Surface area (m^2/g)	15-30

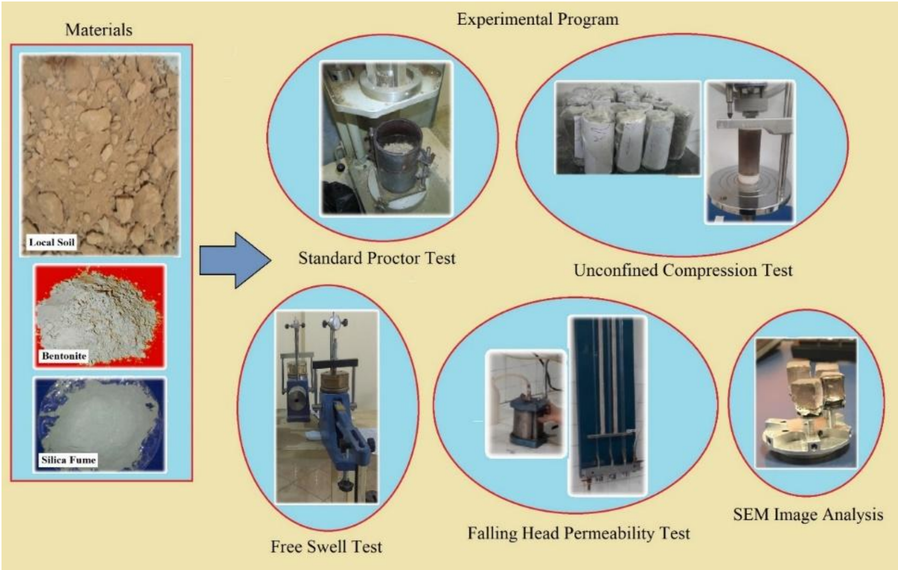


Fig. 3. Materials used and experimental program in this study.

To determine the compaction characteristics of the specimens, including the maximum dry density (MDD) and optimum moisture content (OMC), the Standard Proctor compaction test was conducted. After the addition of silica fume, the soil mixture was stored in a plastic container for 24 hours to allow chemical reactions to take place. The compaction test was then performed on the treated soil. According to ASTM D698 [33], the Standard Proctor test involved compacting the soil in three layers, with 25 blows per layer, corresponding to a compaction energy of 600 kN·m/m³. The compaction was carried out in a mold with a diameter of 101.6 mm using a hammer weighing 24.4 N, dropped from a height of 305 mm.

To perform the unconfined compressive strength (UCS) test in accordance with ASTM D2166 [34], soil and silica fume were mixed at the optimum moisture content, and the specimens were compacted to 100% of the maximum dry density in a mold with a diameter of 38 mm and a height of 76 mm. After removal from the mold, the cylindrical specimens were wrapped in a layer of plastic film, placed in plastic bags, and stored at 23 ± 1.7 °C for 7 days (Fig. 3). Following the curing period, the specimens were subjected to axial loading using a uniaxial testing machine.

To evaluate the swelling characteristics of both untreated and stabilized soils, the free swell test was conducted on the specimens in accordance with ASTM D4546 [36]. Initially, a surcharge pressure was applied to the specimen to achieve the desired void ratio. The surcharge was then removed, and the specimen was inundated to allow free swelling. After the specimen attained full swell, it was reloaded to its initial height, and the swelling pressure was determined accordingly.

Permeability is the ability of water to flow through saturated soil and is commonly measured for fine-grained soils using the falling-head permeability test. In this test, the soil specimen is placed inside a permeameter, and water is allowed to flow through it. The initial hydraulic head difference at time $t = t_1$ is recorded as h_1, and as water flows through the soil, the hydraulic head difference at time $t = t_2$ is recorded as h_2. The coefficient of permeability, k, is calculated using Eq. 1:

$$k = 2.303 \frac{a \, .L}{A \, .\Delta t} log \frac{h_1}{h_2} \tag{1}$$

where h is the hydraulic head difference (cm) at time t, A is the cross-sectional area of the soil specimen (cm²), a is the cross-sectional area of the standpipe (cm²), and L is the length of the soil specimen (cm). The permeability of the specimens was determined using the falling-head permeability test in accordance with ASTM D5084 [35].

To investigate the mechanism of the additive's effect on the soil, scanning electron microscopy (SEM) analysis was conducted using a VEGA\\TESCAN-LMU microscope at 10000x magnification at the Razi Metallurgical Research Center. The selected specimens for SEM analysis were untreated soil and soil treated with 15% silica fume. After curing the stabilized specimen, a small portion was extracted for analysis to examine the pores and evaluate the effect of silica fume on the soil structure. The samples were then analyzed using the scanning electron microscope. The experimental tests conducted in this study are presented in Fig. 3.

3. Results and discussion

Fig. 4 presents the compaction curves for the untreated soil and specimens stabilized with varying percentages of silica fume. For each mix design, five specimens were tested, and the compaction curves were generated based on the measured moisture content and dry density of each specimen. As observed, the addition of silica fume to the soil reduces the MDD while increasing the OMC of the specimens. Specifically, incorporating 15% silica fume decreases the MDD from 14.2 to 13.7 kN/m³ and increases the OMC from 20.1% to 21.7% (Table 4). The reduction in MDD for the silica fume-containing specimens can be attributed to the replacement of soil particles with a higher specific gravity (G_s) by silica fume with a lower density. The increase in OMC is likely due to the presence of silica fume, which increases the specific surface area of the soil, thereby requiring more water for proper compaction of the clayey soil. A similar trend in the variation of compaction parameters with the addition of silica fume to different soils has been reported by Badv and Emdadi [19] and Fattah et al. [37].

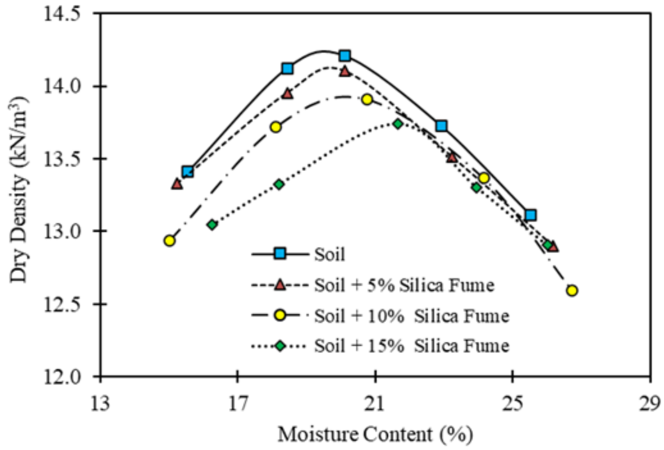


Fig. 4. Compaction curves of the specimens.

Table 4. Changes in the geotechnical properties of the samples.

Properties	Local soil	Local soil + 20% bentonite	Soil + 5% silica fume	Soil + 10% silica fume	Soil + 15% silica fume
MDD, kN/m^3	14.8	14.2	14.1	13.9	13.7
OMC, %	18.2	20.1	20.2	20.8	21.7
UCS, kPa	76.1	86.5	97.5	112.9	121.9
Permeability, cm/s	1.03×10^{-6}	3.1×10^{-7}	2.7×10^{-7}	1.9×10^{-7}	1.4×10^{-7}
Swelling pressure, kPa	91	120	—	—	80

The UCS of the soil was 86.5 kPa, indicating its low strength and limited bearing capacity. As shown in Fig. 5, the addition of silica fume significantly increased the compressive strength. Specifically, specimens containing 5%, 10%, and 15% silica fume exhibited strength increases of approximately 13%, 30%, and 40%, reaching 97.5, 112.9, and 121.9 kPa, respectively. The observed trend suggests that further increases in silica fume content would result in diminishing gains in strength. In the chemical stabilization of clayey soils, the chemical compositions of both the soil and the additive play a crucial role. The soil used in this study contains calcium oxide, which reacts with water molecules to produce Ca^{2+} ions and hydroxyl ions. The active silica present in silica fume reacts with calcium hydroxide to form Calcium Silicate Hydrate (C-S-H) gels. As these gels fill the pore spaces within the soil matrix, a denser structure is formed. Consequently, the soil becomes stiffer, and its strength increases. Other studies have also reported the production of cement products in the presence of silica fume [20, 38, 39]. Another contributing mechanism is the reduction in particle spacing caused by the presence of silica fume, which further increases interparticle friction. The UK Environment Agency [40] stated that a clay liner material should have a minimum remolded undrained shear strength greater than 50 kPa (equivalent to an unconfined compressive strength of 100 kPa). The results of the present study showed that the soils stabilized with 10% and 15% silica fume exhibited strength values higher than this recommended criterion. As seen in the stress-strain curves in Fig. 5, the behavior of the specimens becomes more brittle with increasing silica fume content, and the maximum stress is achieved at lower strains. The strain at peak stress for the untreated soil and soils stabilized with 5%, 10%, and 15% silica fume were 0.043, 0.039, 0.032, and 0.029, respectively. The more brittle behavior observed in the silica fume-treated samples can be attributed to the fact that when the initial soil porosity is high, strength increases due to cementation; however, with increasing stress, the sudden breakdown of these cementitious products leads to a rapid reduction in strength [41].

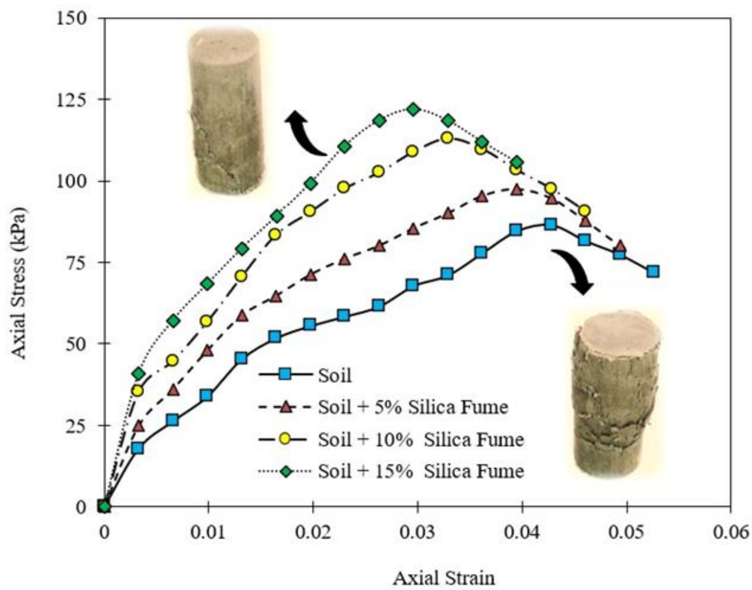
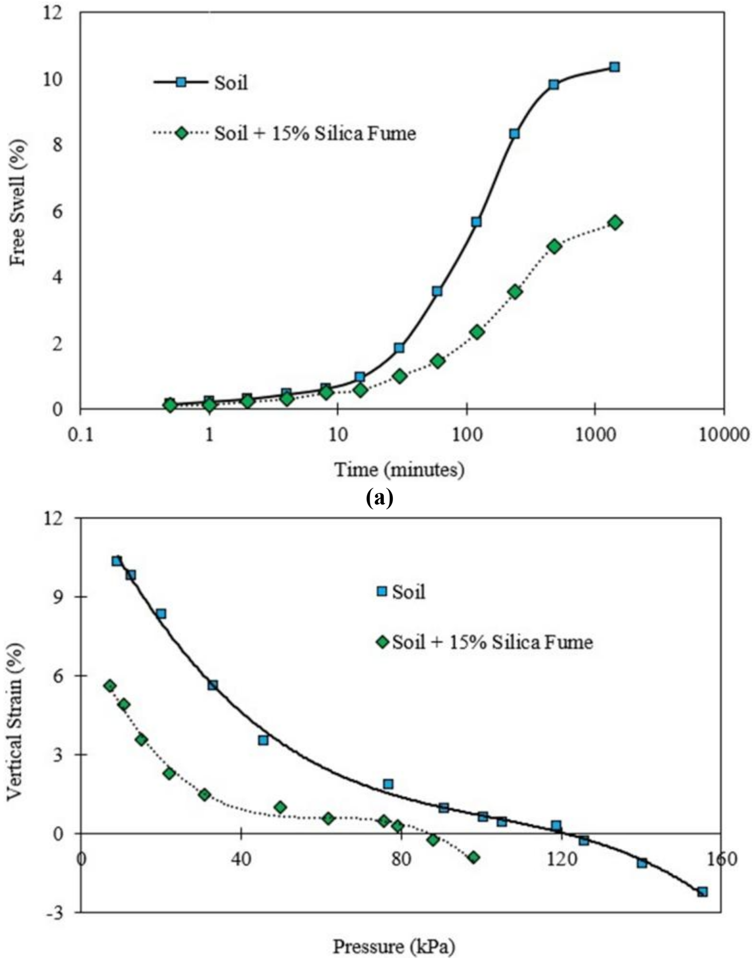


Fig. 5. Stress-strain curves obtained from the UCS test − untreated soil and soil with 15% silica fume, shown after failure.

Fig. 5 also shows that, after loading, the untreated soil and the soil containing 15% silica fume exhibit different failure patterns. The brittle behavior of the silica fume-stabilized specimen results in sudden failure at higher strength, with fewer cracks compared to the untreated soil. Therefore, besides enhancing strength, the incorporation of silica fume is expected to decrease soil permeability by restricting crack development.

The free swell and swelling pressure of the untreated soil and the soil stabilized with 15% silica fume are presented in Fig. 6. The addition of 15% silica fume reduced the swelling percentage from 10.35% to 5.64%, representing a 45% reduction in swelling. The swelling pressure of the untreated soil was measured as 120 kPa, while the soil stabilized with 15% silica fume exhibited a 33% reduction in swelling pressure, decreasing to 80 kPa. The improvement in the swelling behavior of soils due to the addition of silica fume has also been confirmed by other researchers [20, 22, 42]. The formation of bonds between soil particles, as previously discussed, restricts the volumetric expansion of the soil and consequently reduces swelling. Moreover, changes in soil fabric and particle size distribution due to the addition of silica fume reduce the plasticity index of the soil. The reduction in soil plasticity decreases water absorption capacity, which in turn leads to a reduction in percent swell. Several researchers have reported that plasticity is a reliable indicator of the swelling behavior of soils [43].



(a)

(b)

Fig. 6. Effect of silica fume on: (a) swelling percentage; and (b) swelling pressure of the soil.

The permeability of the specimens with varying silica fume contents, measured using the falling-head test, is presented in Fig. 7. Permeability decreased with increasing silica fume content. Specifically, the addition of 15% silica fume reduced soil permeability by 55%, lowering it from 3.1×10^{-7} to 1.4×10^{-7} cm/s. Permeability is one of the most critical parameters in the design of landfill liner systems. According to commonly accepted environmental regulations and design guidelines, the permeability of compacted clay liners is generally required to be equal to or lower than 10^{-7} cm/s to effectively control contaminant migration [44, 45]. The results of the present study showed that increasing silica fume content significantly reduced the permeability of the clay liner specimens. This reduction in permeability can be attributed to the filler effect of silica fume and the formation of bonds between soil particles. These findings indicate that silica fume can be considered a suitable additive for improving the sealing efficiency and long-term performance of clay liner systems.

Fig. 8 presents SEM micrographs of the untreated soil and the soil stabilized with 15% silica fume. As shown, the untreated soil exhibits numerous pores and voids, whereas the silica fume-stabilized soil displays a denser structure. The clay particles and plates are coated with a layer of silica fume. Comparison of the two images indicates a significant reduction in pore size in the stabilized soil. Specifically, the pore size decreased from approximately 10 µm in the untreated soil to less than 2 µm in the soil containing silica fume. This behavior can also be attributed to pozzolanic reactions between the reactive silica in silica fume and calcium compounds present in the soil, leading to the formation of cementitious calcium silicate hydrate (C-S-H) products. The generated gels filled the voids between soil particles and enhanced interparticle bonding, resulting in a more compact and integrated structure.

The reduction in pore size and pore connectivity contributed directly to the decrease in hydraulic conductivity observed in the treated specimens. In addition, the improved bonding and cementation between particles increased soil stiffness and load-bearing capacity, which explains the enhancement in unconfined compressive strength. Therefore, the microstructural observations are consistent with the macroscopic improvements in both hydraulic and mechanical behavior of the silica-fume-treated soil. The microstructural changes observed in this study are in agreement with the results reported by Waleed and Alshawmar [9] and Tiwari et al. [42].

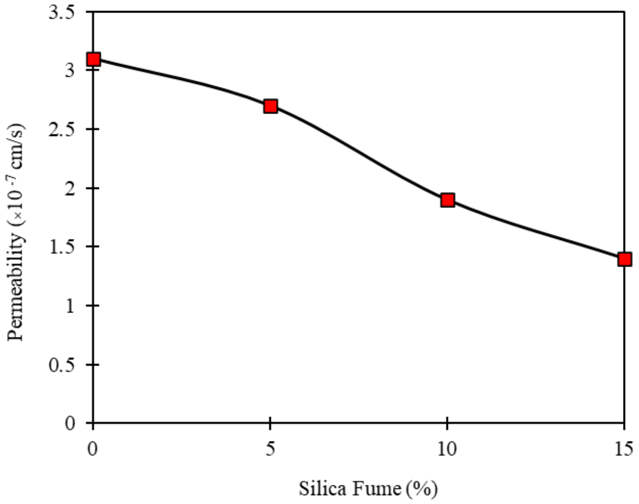


Fig. 7. Permeability variation of specimens with increasing silica fume content.

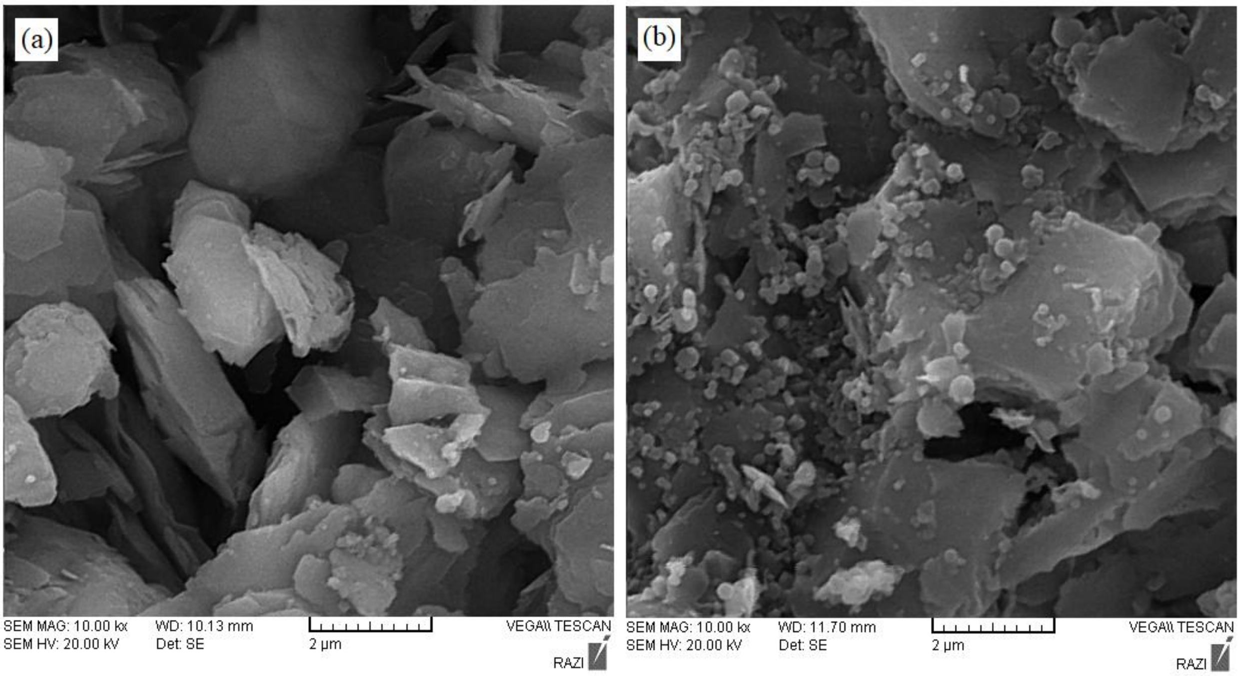


Fig. 8. SEM micrographs of: (a) soil; and (b) soil stabilized with 15% silica fume.

4. Conclusions

This study investigated the effectiveness of silica fume as an environmentally friendly stabilizing additive for improving the geotechnical and microstructural properties of a high-plasticity clay soil intended for landfill liner applications. Based on the experimental results, the following conclusions can be drawn:

- The addition of silica fume led to a reduction in maximum dry density and an increase in optimum moisture content, primarily due to the lower specific gravity and high specific surface area of silica fume particles.
- Unconfined compressive strength increased significantly with increasing silica fume content. The maximum improvement was observed at 15% silica fume, where the strength reached 121.9 kPa, representing an approximately 40% increase compared to the untreated soil and satisfying the minimum strength requirement for landfill liners based on relevant standards. However, the rate of strength gain decreased at higher silica fume contents, indicating diminishing returns. The incorporation of silica fume altered the stress-strain behavior of the soil, making it more brittle and causing peak strength to occur at lower strain levels.
- Swelling characteristics were substantially improved by silica fume stabilization. The free swell percentage and swelling pressure decreased by 45% and 33%, respectively, with the addition of 15% silica fume, mainly due to reduced plasticity and restricted volumetric expansion.

- Permeability decreased markedly with increasing silica fume content. A reduction of approximately 55% was achieved at 15% silica fume, which is attributed to the filler effect and the formation of interparticle bonds. The permeability of the sample containing 15% silica fume was approximately close to the value recommended for clay liner systems (10^{-7} cm/s). Further reduction in permeability may require the use of higher amounts of bentonite or silica fume, while considering the corresponding strength and swelling parameters. The optimum additive contents should be selected by considering both geotechnical performance and practical or economic limitations.
- SEM observations confirmed the densification of the soil structure after stabilization, showing a significant reduction in pore size and a more compact fabric caused by silica fume coating and bonding of clay particles, which is attributed to the formation of C-S-H gels resulting from the reaction between active silica in silica fume and calcium compounds in the soil.

Statements & declarations

Author contributions

Sadegh Ghavami: Investigation, Formal analysis, Supervision, Writing - Review & Editing.

Sina Tahmasebi: Investigation, Methodology, Writing - Original Draft.

Narges Delnavaz: Investigation, Visualization, Writing - Original Draft.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Data availability

The data presented in this study will be available on interested request from the corresponding author.

Declarations

The authors declare no conflict of interest.

References

[1] Proia, R., Salvatore, E., Croce, P., Modoni, G. Compacted sand–bentonite mixtures for the confinement of waste landfills. Acta Geotechnica, 2024; 19: 8007–8022. doi:10.1007/s11440-024-02335-7.

[2] C. M, A., Sunil, B. M. Improving landfill liner performance with bentonite-slag blend permeated with ammonia for a Municipal solid waste landfill. Journal of Environmental Management, 2024; 367: 122013. doi:10.1016/j.jenvman.2024.122013.

[3] Nikbakht, M., Sarand, F. B., Dabiri, R., Hajialilue Bonab, M. Investigation of the Leachate Effect on Permeability and Geotechnical Characteristics of Fine-Grained Soil Modified Using Nanoclay–Nanofiber Composites. Water, 2023; 15: 294. doi:10.3390/w15020294.

[4] Asgari, A., Arjomand, M. A., Bagheri, M., Ebadi-Jamkhaneh, M., Mostafaei, Y. Assessment of Experimental Data and Analytical Method of Helical Pile Capacity Under Tension and Compressive Loading in Dense Sand. Buildings, 2025; 15: 2683. doi:10.3390/buildings15152683.

[5] Ghavami, S., Naseri, H. Prediction of Compressive Strength in Fine-Grained Soils Stabilized with a Combination of Various Stabilization Agents and Nano-SiO2 Using Machine Learning Algorithms. Mathematical and Computational Applications, 2025; 30: 137. doi:10.3390/mca30060137.

[6] Ghavami, S., Naseri, H. Advanced Machine Learning Techniques for Predicting Compaction and Strength Characteristics of Cement Kiln Dust-Stabilized Soils. Transportation Infrastructure Geotechnology, 2026; 13: 84. doi:10.1007/s40515-026-00838-2.

[7] Mostafaei, Y., Arjomand, M. A., Asgari, A., Soltani, A., Ramezanpour, M. A. Sustainable Stabilization of Collapsible Clay Soils Using Eco-Friendly Additives and Sarooj Mortar: Experimental Assessment of Strength and Collapse Behavior. Buildings, 2026; 16: 1561. doi:10.3390/buildings16081561.

[8] Ghavami, S., Farahani, B., Jahanbakhsh, H., Moghadas Nejad, F. Effects of Silica Fume and Nano-silica on the Engineering Properties of Kaolinite Clay. AUT Journal of Civil Engineering, 2018; 2: 135–142. doi:10.22060/ajce.2018.14203.5462.

[9] Waleed, M., Alshawmar, F. Enhancing mechanical properties of low plasticity soil through coal and silica fume stabilization. Scientific Reports, 2025; 15: 9990. doi:10.1038/s41598-025-94149-0.

[10] American Concrete Institute (ACI). ACI 116R-00: Cement and Concrete Terminology. Farmington Hills (MI): ACI; 2000.

[11] Amini, S. H. Ultra clean coal production using dense medium separation for the silicon market [Master's Thesis]. Lexington (KY): University of Kentucky; 2014.

[12] Khan, M. I., Siddique, R. Utilization of silica fume in concrete: Review of durability properties. Resources, Conservation and Recycling, 2011; 57: 30–35. doi:10.1016/j.resconrec.2011.09.016.

[13] Kalkan, E. Influence of silica fume on the desiccation cracks of compacted clayey soils. Applied Clay Science, 2009; 43: 296–302. doi:10.1016/j.clay.2008.09.002.

[14] A. Al-Azzawi, A., A. Daud, K., A. Abdul Sattar, M. Effect of Silica Fume Addition on the Behavior of Silty-Clayey Soils. Journal of Engineering and Sustainable Development, 2012; 16: 92–105.

[15] Moradi, G., Alizade, M., Pashayan, M. Effect of Microsilica on the Strength and Swelling of Lime-Stabilized Marl Soil. In: 9th International Congress on Civil Engineering; 2012 May 8–10; Isfahan, Iran.

[16] Mostafa, A., Ouf, M., Ibrahim, M. Stabilization of Subgrade Pavement Layer Using Silica Fume and Nano Silica. International Journal of Scientific & Engineering Research, 2016; 7: 573–581.

[17] Goodarzi, A. R., Akbari, H. R., Salimi, M. Enhanced stabilization of highly expansive clays by mixing cement and silica fume. Applied Clay Science, 2016; 132-133: 675–684. doi:10.1016/j.clay.2016.08.023.

[18] taghdisi, b., Esmaeiili, K., Khodashenas, S. R. Evaluation of velocity and angle of jet flow for soil whit micro silica lining of water conveyance channels. Iranian Journal of Soil and Water Research, 2018; 49: 101–110. doi:10.22059/ijswr.2018.221484.667582.

[19] Badv, K., Emdadi, H. Utilization of micro silica fume to improve geotechnical properties of Nazloo clay. Sharif Journal of Civil Engineering, 2016; 32.2: 91–100.

[20] Asadoullahtabar, R., Asgari, A., Mohammad Rezapour Tabari, M. Assessment, identifying, and presenting a plan for the stabilization of loessic soils exposed to scouring in the path of gas pipelines, case study: Maraveh-Tappeh city. Engineering Geology, 2024; 342: 107747. doi:10.1016/j.enggeo.2024.107747.

[21] Türköz, M., Umu, S. U., Öztürk, O. Effect of Silica Fume as a Waste Material for Sustainable Environment on the Stabilization and Dynamic Behavior of Dispersive Soil. Sustainability, 2021; 13: 4321. doi:10.3390/su13084321.

[22] Al-Gharbawi, A. S. A., Najemalden, A. M., Fattah, M. Y. Expansive Soil Stabilization with Lime, Cement, and Silica Fume. Applied Sciences, 2023; 13: 436. doi:10.3390/app13010436.

[23] Daniel, D. E. Case histories of compacted clay liners and covers for waste disposal facilities. In: Third International Conference on Case Histories in Geotechnical Engineering; 1993 Jun 01; St. Louis, Missouri.

[24] Wright, W. G. Reuse of Solid Mining Waste Technical and Regulatory Guidance. 1st ed. Washington, DC (US): The Interstate Technology & Regulatory Council (ITRC) Reuse of Solid Mining Waste Team; 2025.

[25] Akgün, H., Ada, M., Koçkar, M. K. Performance assessment of a bentonite–sand mixture for nuclear waste isolation at the potential Akkuyu Nuclear Waste Disposal Site, southern Turkey. Environmental Earth Sciences, 2015; 73: 6101–6116. doi:10.1007/s12665-014-3837-x.

[26] Devapriya, A. S., Thyagaraj, T. Evaluation of red soil-bentonite mixtures for compacted clay liners. Journal of Rock Mechanics and Geotechnical Engineering, 2024; 16: 697–710. doi:10.1016/j.jrmge.2023.04.006.

[27] Sivapullaiah, P. V., Lakshmikantha, H. Chemical Compatibility of Lime Stabilized Indian Red Earth as Liner Material. Soil and Sediment Contamination: An International Journal, 2005; 14: 515–526. doi:10.1080/15320380500263717.

[28] ASTM International. ASTM D422: Standard Test Method for Particle-Size Analysis of Soils. West Conshohochen (PA): ASTM; 2007. doi:10.1520/D0422-63R07.

[29] ASTM International. ASTM D2487: Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System). West Conshohocken (PA): ASTM; 2011. doi:10.1520/D2487-11.

[30] Iranian National Standardization Organization (INSO). INSO 13278 Silica fume (Microsilica) Used in Cementitious Mixtures-Specification. Tehran (IR): INSO; 2017.

[31] ASTM International. ASTM D4318: Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils. West Conshohocken (PA): ASTM; 2005. doi:10.1520/D4318-05.

[32] ASTM International. ASTM D854: Standard Test Methods for Specific Gravity of Soil Solids by Water Pycnometer. West Conshohochen (PA): ASTM; 2002. doi:10.1520/D0854-02.

[33] ASTM International. ASTM D698: Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort. West Conshohocken (PA): ASTM; 2000. doi:10.1520/D0698-00AE01.

[34] ASTM International. ASTM D2166: Standard Test Method for Unconfined Compressive Strength of Cohesive Soil. West Conshohochen (PA): ASTM; 2013. doi:10.1520/D2166_D2166M-13.

[35] ASTM International. ASTM D5084: Standard Test Methods for Measurement of Hydraulic Conductivity of Saturated Porous Materials Using a Flexible Wall Permeameter. West Conshohocken (PA): ASTM; 2016. doi:10.1520/D5084-16.

[36] ASTM International. ASTM D4546: Standard Test Methods for One-Dimensional Swell or Collapse of Soils. West Conshohochen (PA): ASTM; 2014. doi:10.1520/D4546-21.

[37] Fattah, M., Al-Saidi, A., Murad Jebur, M. Characteristics of Clays Stabilized with Lime-Silica Fume Mix. Italian Journal of Geosciences, 2015; 134: 104–113. doi:10.3301/IJG.2014.36.

[38] Alshami, A. W., Ismael, B. H., Aswad, M. F., Majdi, A., Alshijlawi, M., Aljumaily, M. M., AlOmar, M. K., Aidan, I. A., Hameed, M. M. Compaction Curves and Strength of Clayey Soil Modified with Micro and Nano Silica. Materials, 2022; 15: 7148. doi:10.3390/ma15207148.

[39] Tiwari, N., Satyam, N. Experimental Study on the Influence of Polypropylene Fiber on the Swelling Pressure Expansion Attributes of Silica Fume Stabilized Clayey Soil. Geosciences, 2019; 9: 377. doi:10.3390/geosciences9090377.

[40] LFE4 - Earthworks in landfill engineering. ed. Bristol (UK): Environment Agency; 2014.

[41] Ghavami, S., Jahanbakhsh, H., Moghaddasnezhad, F. Laboratory study on stabilization of kaolinite clay with cement and cement kiln dust. Amirkabir Journal of Civil Engineering, 2020; 52: 935–948. doi:10.22060/ceej.2018.15100.5829.

[42] Tiwari, N., Satyam, N., Singh, K. Effect of Curing on Micro-Physical Performance of Polypropylene Fiber Reinforced and Silica Fume stabilized Expansive Soil Under Freezing Thawing Cycles. Scientific Reports, 2020; 10: 7624. doi:10.1038/s41598-020-64658-1.

[43] Saride, S., Puppala, A., Chikyala, S. Swell-shrink and strength behaviors of lime and cement stabilized expansive organic clays. Applied Clay Science, 2013; 85: 39–45. doi:10.1016/j.clay.2013.09.008.

[44] Designing and installing liners. Guide for Industrial Waste Management- Protecting Land, Ground Water, Surface Water, Air. ed. Washington, United States: US Environmental Protection Agency; 2003.

[45] Chaudhary, E., Swami, D., Joshi, N., Reddy, K. Sustainable Landfill Liner Using Local Soils and Wastes Amended with Bentonite: Hydraulic Conductivity and Stochastic Leachate Transport Modeling. Journal of Geotechnical and Geoenvironmental Engineering, 2024; 150: 04023126. doi:10.1061/JGGEFK.GTENG-11470.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Comparison of Numerical Performance of a Modified MPS Method with a Least Squares Method Based on a Mixed Formulation for Elasticity Problems

Gholamreza Shobeyri [a*]

[a] *Faculty of Civil, Water & Environmental Engineering, Shahid Beheshti University, Tehran, Iran*

ARTICLE INFO

Article history:

Received 28 April 2026

Revised 30 May 2026

Accepted 16 June 2026

Available online 01 April 2027

Keywords:

Meshless methods

Improved MPS gradient model

Mixed formulation

Least squares method

Elasticity problems

ABSTRACT

High-precision stress analysis of structures under external loading represents a challenge in civil engineering. Although meshless numerical methods show great potential for solving such problems, their early versions often failed to provide acceptable accuracy. A conventional approach to addressing this is the use of a mixed formulation, which transforms second-order governing differential equations into a system of first-order equations. In this approach, both displacements and structural stresses are treated as unknowns. In contrast, in the original second-order formulation, only displacements are unknowns, and stresses are typically derived indirectly from displacement fields, often resulting in low accuracy. This research compares the numerical performance of an improved MPS gradient model by Shobeyri (2023), a well-known meshless method, for stress calculation via second-order equations with a mixed formulation coupled with a least-squares method. The comparison is conducted on two benchmark elasticity problems. The numerical findings reveal that the improved MPS method provides superior stress accuracy. These results suggest that solving the original second-order differential equations, followed by high-precision stress computation, offers a more effective alternative to the mixed formulation approach.

How to cite this article: Shobeyri, G. Comparison of Numerical Performance of a Modified MPS Method with a Least Squares Method Based on a Mixed Formulation for Elasticity Problems. Civil Engineering and Applied Solutions. 2027; 3(2): 26–36. doi:10.22080/ceas.2026.31619.1095.

1. Introduction

In recent years, numerical methods have emerged as a significant tool for solving complex and practical differential equations within engineering fields. Broadly speaking, these methods are categorized into two main groups: mesh-based and meshless methods. Mesh-based methods, which possess a longer historical background, involve the discretization of the computational domain and the approximation of functions and their derivatives through meshing. However, since meshing is generally a computationally expensive and time-consuming process, meshless methods have been developed. These methods enable the solution of differential equations without the need for any underlying mesh, a characteristic that has led to their widespread adoption and significant development.

Although mesh-based numerical methods are still widely used [1-8], efficient meshless numerical methods have also been developed to solve various engineering problems. Among these methods, the Moving Particle Semi-implicit (MPS) method has gained significant attention, leading to the introduction of various modified versions. In this approach, functions and their derivatives are approximated via interpolation functions based on a series of scattered nodes used for domain discretization. This characteristic facilitates the analysis of problems involving complex geometries.

Various problems in solid and fluid mechanics - such as free surface flows [9-11], convective heat transfer [12], multiphase flows [13], fluid-structure interaction [14], the behavior of a molten jet impinging onto a solid plate [15], progressive water waves

* Corresponding author.
E-mail addresses: g_shobeyri@sbu.ac.ir (G. Shobeyri).

https://doi.org/10.22080/ceas.2026.31619.1095
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

[16], dynamic pressure in dam-break flows [17], two-dimensional hydroelastic slamming [18], fluid lubrication [19], floating structures in waves [20], solid particle dynamics [21], and dense granular flows [22] - have been successfully simulated using the MPS method and its variants. Given the method's inherent capabilities, several techniques have been developed to enhance its performance, most notably those presented in [23-29]. Furthermore, due to the significance of elasticity problems in civil engineering, various meshless numerical methods, such as the Finite Point Method [30], MLPG [31], DLSM [32, 33], and SPH [34] have been employed in numerous studies to analyze such phenomena.

In elasticity problems, once the second-order differential equations - treating structural displacements as unknowns - are solved, the corresponding stresses are typically calculated using numerical methods based on the related formulas. Numerical studies have shown that this approach may lead to reduced accuracy in stress calculation. As an efficient alternative, the second-order differential equations of elasticity can be reformulated as a system of first-order equations that include both structural displacements and stresses; this is known as the mixed formulation [35]. The primary advantage of this formulation is the direct treatment of stresses as unknowns, which yields higher accuracy. However, a larger coefficient matrix is produced due to the increased number of unknowns, leading to a significant increase in computational cost. Despite this, due to its acceptable performance, the mixed formulation has also been applied to various other problems, such as quadratic partial differential equations and free surface problems [36, 37].

This study evaluates the numerical accuracy of a modified MPS gradient model [38] in computing structural stresses via second-order differential equations, comparing it with a mixed formulation coupled with a least-squares method. The investigation is conducted using two benchmark elasticity problems. The results demonstrate that the indirect stress calculation approach, utilizing high-precision numerical methods such as the proposed MPS model, yields superior accuracy compared to the direct stress computation employed in the mixed formulation.

2. The governing equations of elasticity problems

The governing equations for two-dimensional elasticity problems are expressed by Eqs. 1 and 2 [39].

$$(\lambda+2\mu)\frac{\partial^2 u_x}{\partial x^2}+\mu\frac{\partial^2 u_x}{\partial y^2}+(\lambda+\mu)\frac{\partial^2 u_y}{\partial x\partial y}=-fr_x \tag{1}$$

$$\mu\frac{\partial^2 u_y}{\partial x^2}+(\lambda+2\mu)\frac{\partial^2 u_y}{\partial y^2}+(\lambda+\mu)\frac{\partial^2 u_x}{\partial x\partial y}=-fr_y \tag{2}$$

where u_x and u_y represent the horizontal and vertical displacements, respectively, and fr_x and fr_y denote the external forces in the x and y directions. ∂^2 is also the second-order partial derivative operator. The Lamé constants, λ and μ, are defined by Eq. 3.

$$\mu=\frac{E}{2(1+\nu)}>0,\ \ \lambda=\frac{E\nu}{(1-2\nu)(1+\nu)}>0 \tag{3}$$

where ν is the Poisson's ratio and E is the Young's modulus.

In the mixed formulation, a system of first-order equations comprising displacements and stresses (Eqs. 4a to 4e) is employed as an alternative to the second-order differential Eqs. 1 and 2 [35].

$$\frac{\partial\sigma_{xx}}{\partial x}+\frac{\partial\tau_{xy}}{\partial y}=-fr_x \tag{4a}$$

$$\frac{\partial\tau_{xy}}{\partial x}+\frac{\partial\sigma_{yy}}{\partial y}=-fr_y \tag{4b}$$

$$\sigma_{xx}=(\lambda+2\mu)\frac{\partial u_x}{\partial x}+\lambda\frac{\partial u_y}{\partial y} \tag{4c}$$

$$\sigma_{yy}=\lambda\frac{\partial u_x}{\partial x}+(\lambda+2\mu)\frac{\partial u_y}{\partial y} \tag{4d}$$

$$\tau_{xy}=\mu(\frac{\partial u_x}{\partial y}+\frac{\partial u_y}{\partial x}) \tag{4e}$$

where σ_{xx} and σ_{yy} denote the normal stresses and τ_{xy} represents the shear (tangential) stress. The compact form of these equations is given by Eq. 5.

$$L(\varphi)+F=0 \tag{5}$$

where $L(.)$ is a first-order differential operator, defined by Eq. 6.

$$L(.)=L_1(.)_x+L_2(.)_y+L_3(.) \tag{6}$$

The vector of unknowns, φ, is given by Eq. 7.

$$\varphi=\left[u_x,u_y,\sigma_{xx},\sigma_{yy},\tau_{xy}\right]^T \tag{7}$$

The vector F is similarly defined by Eq. 8.

$$F = [0,0,0,\text{-}fr_x,\text{-}fr_y]^T \tag{8}$$

The operators $L_1(.)_x$, $L_2(.)_y$ and $L_3(.)$ can be expressed by Eq. 9.

$$L_1 = \begin{bmatrix} \lambda+2\mu & 0 & 0 & 0 & 0 \\ \lambda & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},\ L_2 = \begin{bmatrix} 0 & \lambda & 0 & 0 & 0 \\ 0 & \lambda+2\mu & 0 & 0 & 0 \\ \mu & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},\ L_3 = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

3. Numerical methods

3.1. Improved MPS gradient model

In the MPS method, the gradient and Laplacian operators of an arbitrary function f are approximated in two dimensions using a weighted averaging approach [9] as shown in Eqs. 10 and 11.

$$\nabla f_i = \sum_{j\neq i} \frac{2}{n_j} \frac{f_j - f_i}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) W(|\mathbf{r}_j - \mathbf{r}_i|, R_s) \tag{10}$$

$$\nabla^2 f_i = \frac{4}{\lambda} \sum_{j\neq i} \frac{1}{n_j} (f_j - f_i) W(|\mathbf{r}_j - \mathbf{r}_i|, R_s) \tag{11}$$

where i denotes the reference node, while j represents its neighboring nodes. Furthermore, r and $W(|\mathbf{r}_j - \mathbf{r}_i|, R_s)$ or $W(|\mathbf{r}_j - \mathbf{r}_i|)$ denote the relative position vector and the interpolation function, respectively. The computational parameters n_j and λ are defined by Eqs. 12 and 13.

$$n_i = \sum_{j\neq i} W(|\mathbf{r}_j - \mathbf{r}_i|, R_s) \tag{12}$$

$$\lambda = \frac{\int_\Omega W(r)\, r^2\, d\Omega}{\int_\Omega W(r) d\Omega} \tag{13}$$

where Ω is the interaction area of node i and $r = |\mathbf{r}_j - \mathbf{r}_i|$.

One of the most efficient interpolation functions is expressed by Eq. 14.

$$W(r, R_s) = \frac{8}{\pi R_s^2} \left[\frac{3}{4} \left(\frac{r}{R_s}\right)^2 - \frac{3}{2}\left(\frac{r}{R_s}\right) + \frac{3}{4} \right] \tag{14}$$

where R_s denotes the size of the interaction area for the reference node.

An improved MPS gradient model, characterized by remarkable numerical performance, is proposed in Eq. 15.

$$\nabla f_i = \left[\frac{1}{\lambda} \sum_{j\neq i} \frac{1}{n_j} W(|\mathbf{r}_j - \mathbf{r}_i|) (\mathbf{r}_j - \mathbf{r}_i) \otimes (\mathbf{r}_j - \mathbf{r}_i)^T \right]^{-1} \left[\frac{1}{\lambda} \sum_{j\neq i} \frac{1}{n_j} W(|\mathbf{r}_j - \mathbf{r}_i|)(f_j - f_i)(\mathbf{r}_j - \mathbf{r}_i) \right] \tag{15}$$

This model has been successfully applied to achieve non-oscillatory stress calculations in elasticity problems. Detailed formulations of this gradient model, including the derivation of second-order derivatives required for Eqs. 1 and 2 via the MPS method, are provided in reference [38].

3.2. The least square scheme in the mixed formulation

The first-order Taylor series expansion for a function f around a central node with position (x_0, y_0) is given by Eq. 16.

$$f(x, y) \approx f_0 + \frac{\partial f_0}{\partial x} \Delta x + \frac{\partial f_0}{\partial y} \Delta y \tag{16}$$

where $\Delta x = x - x_0$ and $\Delta y = y - y_0$.

By applying this expansion to all computational nodes within the interaction area of the reference node, we obtain Eq. 17.

$$\begin{bmatrix} 1 & \Delta x_1 & \Delta y_1 \\ 1 & \Delta x_2 & \Delta y_2 \\ \cdots & \cdots & \cdots \\ 1 & \Delta x_N & \Delta y_N \end{bmatrix} \begin{bmatrix} f_0 \\ \partial f_0/\partial x \\ \partial f_0/\partial y \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \cdots \\ f_N \end{bmatrix} \tag{17}$$

where N denotes the total number of surrounding nodes.

Solving this system allows for the determination of the approximated function values at the computational nodes $(f_1, f_2, \ldots, f_N)$. The total approximation error, $\|E\|$, is expressed by Eq. 18.

$$\|E\| = \sum_{j=1}^{N}\left[f_0 - f_j + \frac{\partial f_0}{\partial x}\Delta x_j + \frac{\partial f_0}{\partial y}\Delta y_j\right]^2 W_j^2 \quad (18)$$

where W_j is the value of the interpolation function at node j.

By minimizing this error function using a weighted least-squares scheme, the vector of unknown parameters can be calculated as shown in Eqs. 19 and 20.

$$\frac{\partial\|E\|}{\partial F} = 0 \quad (19)$$

$$\begin{bmatrix} f_0 \\ \partial f_0/\partial x \\ \partial f_0/\partial y \end{bmatrix} = \begin{bmatrix} \sum W_j^2 & \sum W_j^2\Delta x_j & \sum W_j^2\Delta y_j \\ \sum W_j^2\Delta x_j & \sum W_j^2\Delta x_j^2 & \sum W_j^2\Delta x_j\Delta y_j \\ \sum W_j^2\Delta y_j & \sum W_j^2\Delta x_j\Delta y_j & \sum W_j^2\Delta y_j^2 \end{bmatrix}^{-1} \begin{pmatrix} \sum f_j W_j^2 \\ \sum f_j W_j^2\Delta x_j \\ \sum f_j W_j^2\Delta y_j \end{pmatrix} \quad (20)$$

where $F = [f_0, \partial f_0/\partial x, \partial f_0/\partial y]$ contains the unknown values of the function and its derivatives at the reference node. The interpolation function for this method is defined by Eq. 21.

$$W(r) = \begin{cases} \frac{1}{r^3+\eta^3} & r \le R_s \\ 0 & r > R_s \end{cases} \quad (21)$$

In this study, the average nodal distance (defined later as l_0) is chosen for the parameter η (as a small value) in Eq. 21 to prevent a zero denominator. A similar least-squares method, based on a second-order Taylor expansion, has been employed to solve free-surface and elasticity problems, as well as elliptic PDEs [40-42].

By approximating the function and its first-order derivatives at all computational nodes, the system of first-order equations (Eqs. 4a and 4e) can be solved to obtain the unknown variables of the elasticity problem (displacements and stresses) in the mixed formulation. This system of equations is solvable by applying the boundary conditions using the known analytical results. One advantage of the mixed formulation is that stresses can be calculated directly with lower errors. Conversely, stresses can also be calculated after determining the displacements from the second-order differential equations (Eqs. 1 and 2) by using the Eqs. 4c and 4e; however, this approach may lead to larger errors.

4. Numerical examples

4.1. Cantilever beam

To compare the performance of these two numerical methods, a cantilever beam subjected to a load at its free end is examined in the first numerical example, with the initial geometry illustrated in Fig. 1. The analytical solutions for the displacements and stresses ($u_x, u_y, \sigma_{xx}, \sigma_{yy}, \tau_{xy}$) are provided in [39]. Applying the governing differential equations to the internal nodes results in a non-square matrix. However, by incorporating the analytical solutions of the boundary nodes into the right-hand side (RHS) vector, a square system of equations is obtained, the solution of which yields the unknowns at the internal nodes. It is worth noting that the final system of equations resulting from the discretization of the differential equations was solved using Gaussian elimination. The computational constants $P = 1, E = 10000, c = 1, L = 16$ and $\nu = 0.25$ (in SI units) are used for this test problem.

The computational nodes are initially placed at the vertices of a square grid with a spacing of l_0 and are subsequently moved randomly with a maximum displacement of kl_0, where k is the randomness coefficient. To solve the problem, this coefficient is set to 0.25, and three irregular node distributions with $l_0 = 1/2, 1/4, 1/6$ m are applied (see Fig. 2 with $l_0 = 1/6$ m).

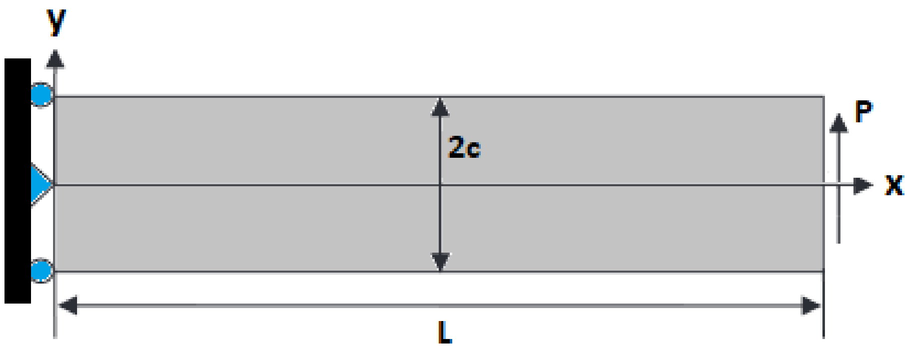


Fig. 1. The Initial geometry of the cantilever beam problem.

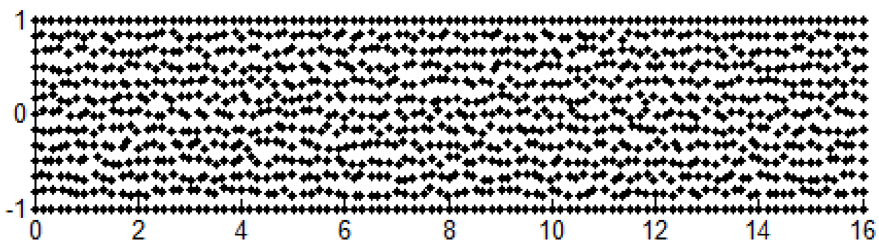


Fig. 2. The nodal distribution with $l_0 = 1/6$ m for the cantilever beam problem.

In this research, Eq. 22 is employed to evaluate the error of the numerical methods in solving this problem.

$$Er_1 = \frac{\Sigma(f_i - \bar{f}_i)^2}{Nb(min(\bar{f}))^2} \tag{22}$$

In Eq. 22, f_i and $\bar{f}_i$ represent the numerical and analytical values of the normal stress (σ_{xx}) at the nodes on the upper face of the beam, respectively. Additionally, $min(\bar{f})$ and Nb denote the minimum value and the total number of computational nodes on this face, respectively.

The computational errors for both numerical methods and the utilized nodal distributions are shown in Fig. 3, according to the defined error indicator. As observed, the improved MPS model offers higher accuracy than the mixed method.

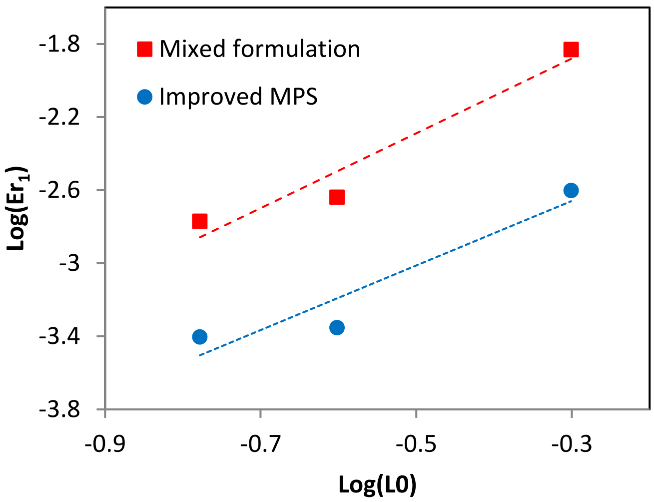
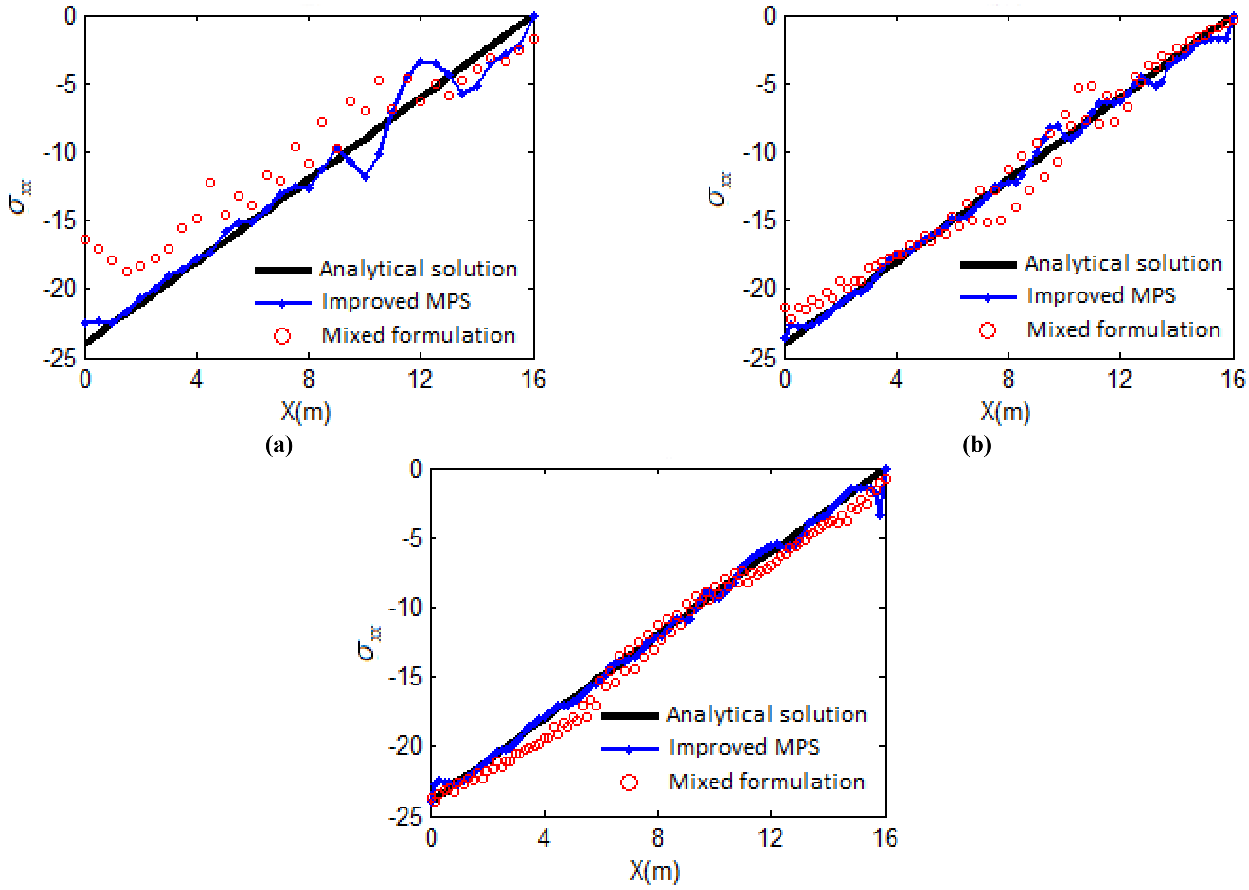


Fig. 3. Error analysis of the numerical methods for the cantilever beam problem.

Fig. 4 illustrates a comparison between the calculated normal stress (σ_{xx}) along the beam's upper edge and the corresponding analytical results. While the MPS method yields results of acceptable accuracy, the mixed formulation exhibits significant fluctuations, thereby highlighting the superiority of the MPS approach. Notably, despite the mixed formulation's ability to solve for stresses directly as primary unknowns, the MPS method provides more consistent and promising results.



(a) **(b)**

(c)

Fig. 4. Comparison of computational and analytical results of stress at the top of the beam for: (a) $l_0 = 1/2$ m; (b) $l_0 = 1/4$ m; and (c) $l_0 = 1/6$ m.

The results of meshless numerical methods are highly dependent on the size of the interaction area. In this study, the optimal value of this parameter is selected such that the defined numerical error (Eq. 22) is minimized. It is evident that the optimal size of R_s differs for the two numerical methods, as presented in Fig. 5. As evident, the MPS method requires a significantly larger R_s for the finest node distribution compared to the mixed method to achieve minimum error, thereby increasing the computational cost (the processor is Intel Core i5, CPU@2.00 GHz-2.6 GHz, and RAM 4.00 GB). Conversely, the larger system of equations in the mixed method also incurs high computational expenses.

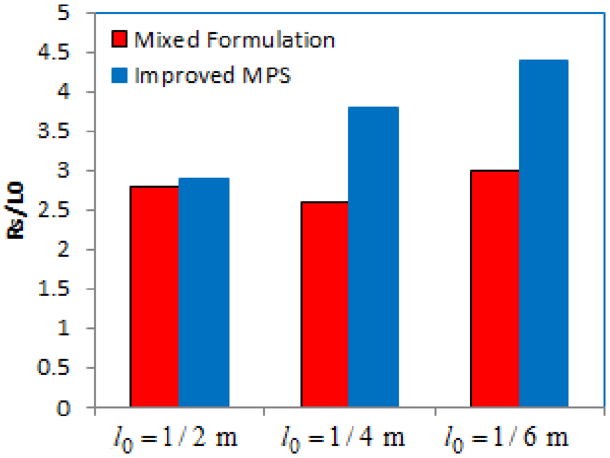


Fig. 5. Optimal values of R_s for the cantilever beam problem.

Under these conditions, the MPS method is only moderately more time-consuming for this specific nodal distribution. For the other two nodal distributions, both methods require approximately the same computational time (see Fig. 6). Fig. 7 illustrates the error variations for different R_s values and the nodal distributions. It is evident that the improved MPS method exhibits less sensitivity to changes in this parameter compared to the mixed method. In other words, reducing the R_s value in the MPS scheme does not lead to a significant decrease in accuracy, but rather considerably reduces the computational cost. Considering both the accuracy of the results and the computational cost simultaneously, the MPS method can be regarded as more efficient.

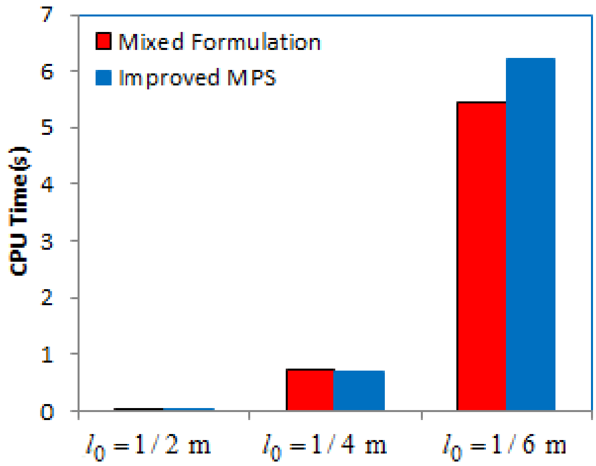


Fig. 6. Computational times of the numerical methods for the cantilever beam problem.

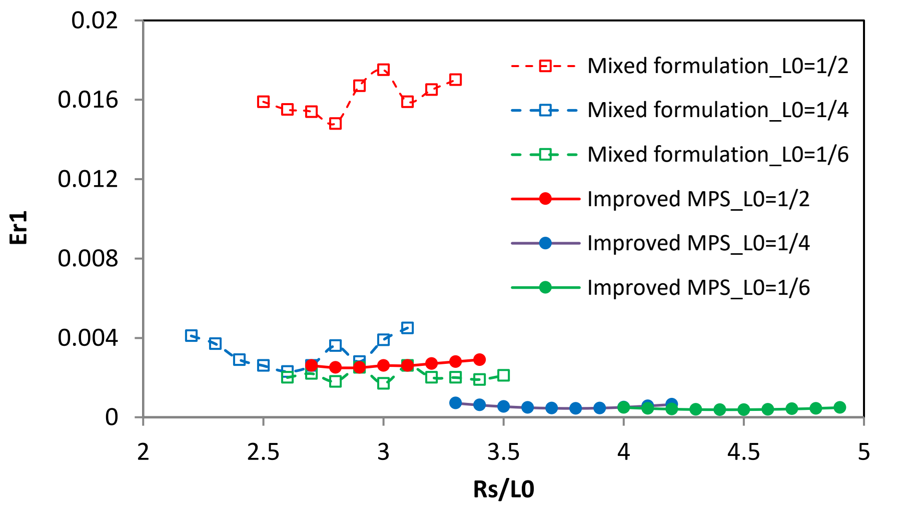


Fig. 7. Variation of numerical error for different R_s values in the cantilever beam problem.

To further evaluate the numerical performance of these two methods under more complex conditions, a highly irregular nodal distribution ($k = 0.5$) with $l_0 = 1/6$ m was employed. As illustrated in Fig. 8, the improved MPS method exhibits significantly lower fluctuations in the calculated normal stress (σ_{xx}) at the upper part of the beam compared to the mixed formulation. This demonstrates the robustness of the proposed method, even when applied to highly irregular nodal distributions. To reduce computational costs for a fixed number of nodes, these nodes can be distributed such that a higher density is placed in regions with higher errors.

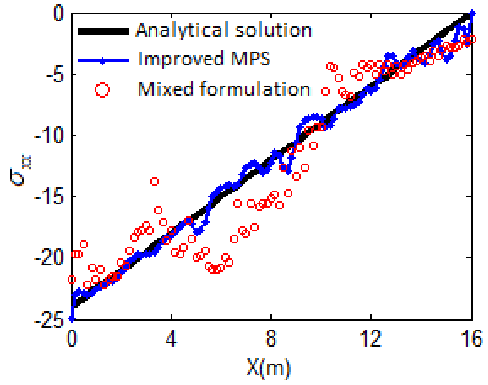
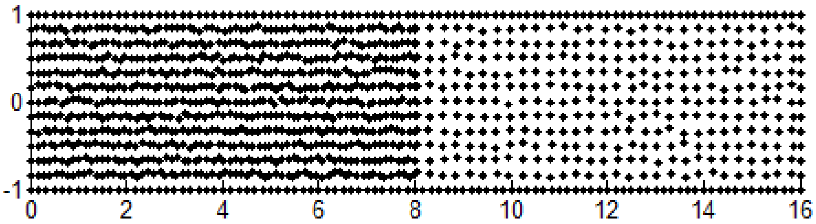
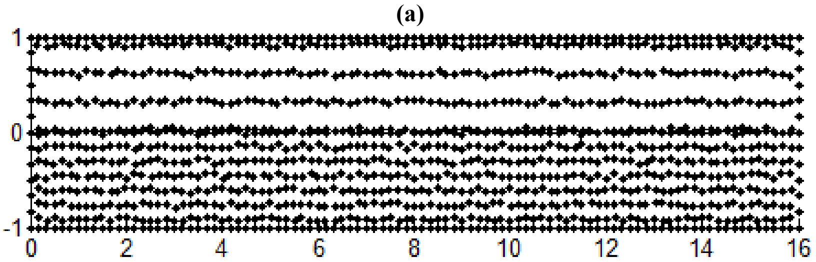


Fig. 8. Comparison of the computational and analytical results of stress at the top of the beam for highly irregular nodal distribution with $k = 0.5$ using $l_0 = 1/6$ m.

This process results in a refined nodal distribution, which is essential for achieving high accuracy in complex conditions. As shown in Fig. 9, two different refined nodal patterns, corresponding to $l_0 = 1/6$ m and $k = 0.15$, are employed to compare the numerical methods investigated in this study. Fig. 10 illustrates the calculated normal stress along the upper edge of the beam for the distributions shown in Fig. 9; these results reaffirm the superior accuracy of the improved MPS model compared to the mixed method. It should be noted that the nodal distribution in Fig. 9b was intentionally chosen with high irregularity to further demonstrate the remarkable capabilities of the MPS method. Consequently, the mixed method exhibits significant oscillations under such conditions. However, when a distribution with lower irregularity is employed, such as the one shown in Fig. 9a, the mixed method yields results with considerably fewer oscillations.

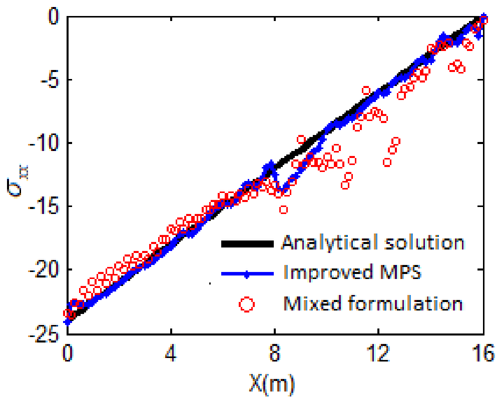


(a)

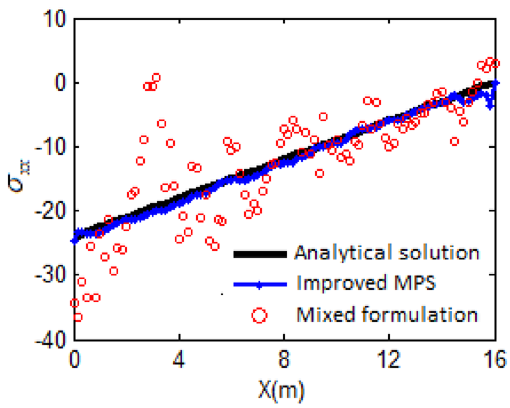


(b)

Fig. 9. Two refined nodal distributions with different patterns with $l_0 = 1/6$ m and $k = 0.15$ for the solution of the cantilever beam problem: (a) first pattern; and (b) second pattern.



(a)



(b)

Fig. 10 Comparison of the computational and analytical results of stress at the top of the beam corresponding to the refined nodal distributions shown in Fig. 9: (a) results corresponding to Fig. 9a; and (b) results corresponding to Fig. 9b.

4.2. Infinite plate with a circular hole

The second benchmark problem in this study is an infinite plate containing a circular hole under uniaxial tension, with the geometry shown in Fig. 11. The analytical solutions for the displacements and stresses of this problem are provided in [39]. Due to the symmetry of the problem, a square domain in the first quadrant with a side length of $6a_p$ (where $a_p = 1$ m is the radius of the hole) is considered as the computational domain. The following constants (in SI units) are applied to solve this test problem: $P = 1$, $E = 1000$ and $\nu = 0.3$.

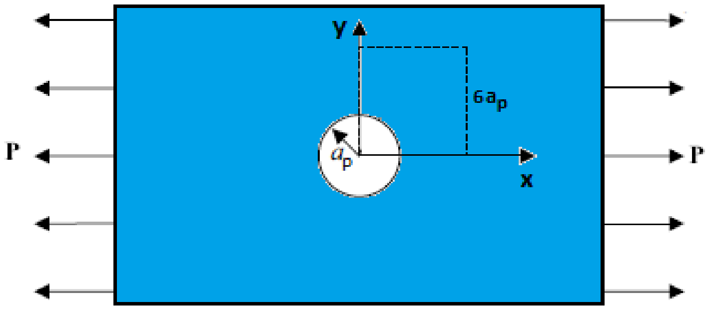


Fig. 11. The geometry of an infinite plate with a circular hole subjected to a uniaxial load.

The error indicator for this problem is given by Eq. 23.

$$Er_2 = \frac{\sum(g_i - \bar{g}_i)^2}{Nh\left(max(\bar{g})\right)^2} \tag{23}$$

In Eq. 23, g_i and $\bar{g}_i$ represent the numerical and analytical values of the normal stress σ_{xx} at the nodes along the left edge of the plate, respectively. Additionally, $max(\bar{g})$ and Nh denote the maximum value and the total number of computational nodes on the edge, respectively.

An irregular nodal distribution with $l_0 = 1/6$ m and $k = 0.3$ as shown in Fig. 12 is used to solve this problem. In Fig. 13, the calculated normal stresses (σ_{xx}) along the left edge of the plate ($x = 0$ m) are compared with the analytical solution, demonstrating the superior accuracy of the improved MPS model. It is worth noting that the best results are obtained based on the error indicator defined in Eq. 23. Additionally, the performance of the numerical methods for the refined nodal distribution shown in Fig. 14 ($l_0 = 1/6$ m, $k = 0.3$) is presented in Fig. 15. As is evident, the MPS method once again proves its ability to produce results with high accuracy, whereas the mixed method yields fluctuating results.

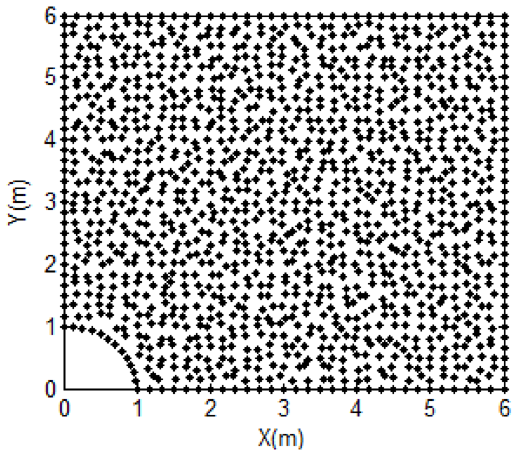


Fig. 12. The nodal distribution with $l_0 = 1/6$ m and $k = 0.3$ for the plate problem.

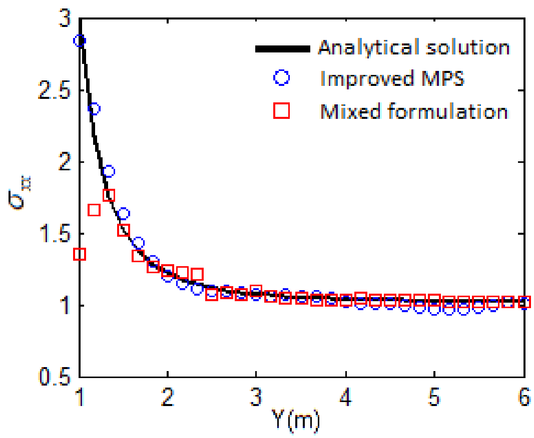


Fig. 13. Comparison of the computational and analytical results of stress along the left edge of the plate.

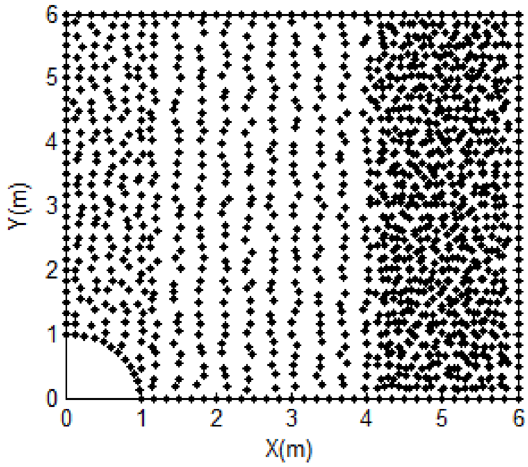


Fig. 14. The refined nodal distribution with $l_0 = 1/6$ m and $k = 0.3$ for the solution of the plate problem.

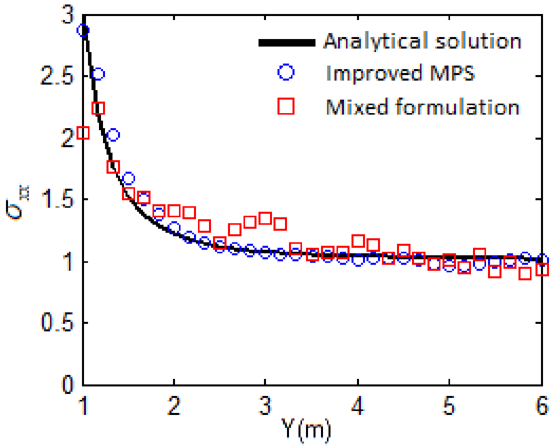


Fig. 15. Comparison of the computational and analytical results of stress at the left edge of the plate for the refined nodal distributions of Fig. 14.

5. Conclusions

This study evaluated the accuracy of stress estimation using numerical methods for two benchmark elasticity problems. In the first approach, the second-order governing differential equations, with structural displacements as the primary unknowns, were solved using the MPS method. The corresponding stresses were subsequently derived through the existing displacement-stress relationships and a modified MPS gradient model [38]. In the second approach, the second-order equations were reformulated into a mixed formulation comprising a system of first-order differential equations, where both displacements and stresses were treated as unknowns and solved via a least-squares method. Based on the obtained numerical results, the following key observations can be drawn:

- The indirect determination of stresses using the modified MPS gradient model (Method 1) yields results with lower error compared to the mixed method (Method 2).
- The results of Method 1 exhibit less sensitivity to variations in the size of the interaction area.
- The accuracy of Method 1 is relatively insensitive to the increasing irregularity of the computational node distribution, whereas Method 2 exhibits greater sensitivity to this factor.
- Method 1 exhibits superior numerical performance for the refined node distributions.
- Considering both accuracy and computational cost simultaneously, it can be concluded that Method 1 is more efficient.

These findings serve as a guide for selecting appropriate numerical methods to achieve high-precision stress analysis in complex, real-world structural problems.

Statements & declarations

Author contributions

Gholamreza Shobeyri: Conceptualization, Methodology, Numerical analysis, Writing - Original Draft, Writing - Review & Editing.

Funding

The author received no financial support for the research, authorship, and/or publication of this manuscript.

Data availability

The data presented in this study will be available on interested request from the corresponding author.

Declarations

The author declares no conflict of interest.

References

[1] Jebalbarezi Sarbijan, M., Shirini, B., Rooholamini, H. Investigation of Crack Growth Behavior in Heterogeneous Asphalt Concrete Using FEM Modeling Based on Random Aggregate Generation and Distribution Algorithms. Civil Engineering and Applied Solutions, 2026; 2: 46–57. doi:10.22080/ceas.2025.29529.1020.

[2] Sadeghi, V., Bagheri, M., Abasi Hamidi, J. Earthquake-Induced Deformation of Road Embankments: A Finite Difference Method. Civil Engineering and Applied Solutions, 2026; 2: 28–41. doi:10.22080/ceas.2026.30800.1066.

[3] Ghaseminejad, V., Moosavi, S. M., Zargar Herijani, T. Parametric Study of Settlement in Shallow Foundations of Adjacent Buildings on Sandy Soil: A Finite Element Parametric Study. Civil Engineering and Applied Solutions, 2026; 2: 73–84. doi:10.22080/ceas.2026.30450.1053.

[4] Ansari, B., Moazzami, A. Soil Freezing Model in the Presence of Underground Inclusion Using Hybrid Boundary-Finite Element Method. Numerical Methods in Civil Engineering, 2026; 10: 25–41. doi:10.61882/NMCE.2511.1107.

[5] Shafiee, A. H., Oulapour, M., Shlaka Drifesh, Z. Uplift Capacity of Granular Pile Anchors in Clay Using Finite Element Limit Analysis. Numerical Methods in Civil Engineering, 2025; 9: 1–8. doi:10.61186/NMCE.2409.1074.

[6] Asadi, R., Zamani Aliabadi, Z. Comparison of numerical methods for the solution of Richards' equation in layered porous media. Numerical Methods in Civil Engineering, 2025; 9: 1–10. doi:10.61186/NMCE.2302.1007.

[7] Daliri, M., Ghohani Arab, H., Miri, M., Khatibi, S. H. Prediction of Shear Strength in SCS Panels with One-End Welded BP Shear Connector Using Numerical Modeling and Gene Expression Programming (GEP). AUT Journal of Civil Engineering, 2025; 9: 185–204. doi:10.22060/ajce.2025.23945.5907.

[8] Kiani Fordoei, M. A., Khodaparast, M., Ghadamgahi, A. Numerical investigation of behavior of energy piles in saturated fine-grained soil. AUT Journal of Civil Engineering, 2024; 8: 145–160. doi:10.22060/ajce.2025.23636.5891.

[9] Koshizuka, S., Oka, Y. Moving-Particle Semi-Implicit Method for Fragmentation of Incompressible Fluid. Nuclear Science and Engineering, 1996; 123: 421–434. doi:10.13182/NSE96-A24205.

[10] Ataie-Ashtiani, B., Farhadi, L. A stable moving-particle semi-implicit method for free surface flows. Fluid Dynamics Research, 2006; 38: 241. doi:10.1016/j.fluiddyn.2005.12.002.

[11] Geng, C., Wang, W.-h., Heng, M.-y., Zhao, Y., Yang, H., Huang, Y. A new free surface identification method for 3D MPS method. Scientific Reports, 2026; 16: 13829. doi:10.1038/s41598-026-44218-9.

[12] Zhang, S., Morita, K., Fukuda, K., Shirakawa, N. An improved MPS method for numerical simulations of convective heat transfer problems. International Journal for Numerical Methods in Fluids, 2006; 51: 31–47. doi:10.1002/fld.1106.

[13] Khayyer, A., Gotoh, H. Enhancement of performance and stability of MPS mesh-free particle method for multiphase flows characterized by high density ratios. Journal of Computational Physics, 2013; 242: 211–233. doi:10.1016/j.jcp.2013.02.002.

[14] Sun, Z., Djidjeli, K., Xing, J. T., Cheng, F. Modified MPS method for the 2D fluid structure interaction problem with free surface. Computers & Fluids, 2015; 122: 47–65. doi:10.1016/j.compfluid.2015.08.017.

[15] Li, G., Liu, M., Duan, G., Chong, D., Yan, J. Numerical investigation of erosion and heat transfer characteristics of molten jet impinging onto solid plate with MPS–LES method. International Journal of Heat and Mass Transfer, 2016; 99: 44–52. doi:10.1016/j.ijheatmasstransfer.2016.03.090.

[16] Wang, L., Jiang, Q., Zhang, C. Improvement of moving particle semi-implicit method for simulation of progressive water waves. International Journal for Numerical Methods in Fluids, 2017; 85: 69–89. doi:10.1002/fld.4373.

[17] Wang, L., Jiang, Q., Nie, S., Zhang, J., Iddy, I. Improvement on MPS Method for Simulation of Dynamic Pressure in Dam Break Flows. Journal of Coastal Research, 2018; 85: 971–975. doi:10.2112/SI85-195.1.

[18] Zha, R., Peng, H., Qiu, W. An improved higher-order moving particle semi-implicit method for simulations of two-dimensional hydroelastic slamming. Physics of Fluids, 2021; 33: 037104. doi:10.1063/5.0033491.

[19] Yamada, D., Imatani, T., Shibata, K., Maniwa, K., Obara, S., Negishi, H. Application of improved multiresolution technique for the MPS method to fluid lubrication. Computational Particle Mechanics, 2022; 9: 421–441. doi:10.1007/s40571-021-00420-2.

[20] Zhang, G., Zhao, W., Wan, D. Moving Particle Semi-implicit method coupled with Finite Element Method for hydroelastic responses of floating structures in waves. European Journal of Mechanics - B/Fluids, 2022; 95: 63–82. doi:10.1016/j.euromechflu.2022.04.005.

[21] Kim, K. S. Numerical Modeling of Solid Particle Dynamics Using the Moving Particle Semi-Implicit (MPS) Method in Seabed Penetration Scenarios. Journal of Coastal Research, 2024; 116: 548–552. doi:10.2112/JCR-SI116-111.1.

[22] Mumeen, A., Rufai, O., Jin, Y.-C. MPS simulation of dense granular flows using a gradient-expansion nonlocal μ(I) rheology. Particuology, 2026; 114: 186–206. doi:10.1016/j.partic.2026.04.013.

[23] Khayyer, A., Gotoh, H. A higher order Laplacian model for enhancement and stabilization of pressure calculation by the MPS method. Applied Ocean Research, 2010; 32: 124–131. doi:10.1016/j.apor.2010.01.001.

[24] Tamai, T., Koshizuka, S. Least squares moving particle semi-implicit method. Computational Particle Mechanics, 2014; 1: 277–305. doi:10.1007/s40571-014-0027-2.

[25] Shobeyri, G., Madadi, H. An improvement in MPS method using Voronoi diagram and a new kernel function. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2018; 40: 172. doi:10.1007/s40430-018-1121-9.

[26] Duan, G., Yamaji, A., Koshizuka, S., Chen, B. The truncation and stabilization error in multiphase moving particle semi-implicit method based on corrective matrix: Which is dominant? Computers & Fluids, 2019; 190: 254–273. doi:10.1016/j.compfluid.2019.06.023.

[27] Li, D., Zhang, H., Qin, G. A Modified MPS Method with a Split-Pressure Poisson Equation and a Virtual Particle for Simulating Free Surface Flows. Journal of Marine Science and Engineering, 2023; 11: 215. doi:10.3390/jmse11010215.

[28] Shobeyri, G. Improved MPS models for simulating free surface flows. Mathematics and Computers in Simulation, 2024; 218: 79–97. doi:10.1016/j.matcom.2023.11.015.

[29] Shobeyri, G. Novel SPH and MPS Laplacian Models Improved by MLS Method for Solving Poisson equations. Numerical Methods in Civil Engineering, 2024; 9: 29–39. doi:10.61186/NMCE.2406.1061.

[30] Oñate, E., Perazzo, F., Miquel, J. A finite point method for elasticity problems. Computers & Structures, 2001; 79: 2151–2163. doi:10.1016/S0045-7949(01)00067-0.

[31] Atluri, S. N., Liu, H. T., Han, Z. D. Meshless Local Petrov-Galerkin (MLPG) Mixed Collocation Method For Elasticity Problems. Computer Modeling in Engineering \& Sciences, 2006; 14: 141–152. doi:10.3970/cmes.2006.014.141.

[32] Afshar, M. H., Amani, J., Naisipour, M. A node enrichment adaptive refinement in Discrete Least Squares Meshless method for solution of elasticity problems. Engineering Analysis with Boundary Elements, 2012; 36: 385–393. doi:10.1016/j.enganabound.2011.08.012.

[33] Nikravesh Kazeroni, S., Afshar, M. H., Faraji, S. RPIM and RPIM-MLS based MDLSM method for the solution of elasticity problems. Scientia Iranica, 2016; 23: 2458–2468. doi:10.24200/sci.2016.2305.

[34] Shobeyri, G. Mixed Smoothed Particle Hydrodynamics Method for Planar Elasticity Problems. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2023; 47: 491–504. doi:10.1007/s40996-022-00883-0.

[35] Amani, J., Afshar, M. H., Naisipour, M. Mixed discrete least squares meshless method for planar elasticity problems using regular and irregular nodal distributions. Engineering Analysis with Boundary Elements, 2012; 36: 894–902. doi:10.1016/j.enganabound.2011.09.012.

[36] Faraji, S., Afshar, M. H., Amani, J. Mixed discrete least square meshless method for solution of quadratic partial differential equations. Scientia Iranica, 2014; 21: 492–504.

[37] Eini, N., Afshar, M. H., Faraji Gargari, S., Shobeyri, G., Afshar, A. A fully Lagrangian mixed discrete least squares meshfree method for simulating the free surface flow problems. Engineering with Computers, 2022; 38: 331–351. doi:10.1007/s00366-020-01157-x.

[38] Shobeyri, G. Improved MPS Gradient Models for Elasticity Problems. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2023; 47: 1831–1843. doi:10.1007/s40996-022-01013-6.

[39] Timoshenko, S., Goodier, J. N. Theory of Elasticity. 3rd ed. New York (NY): McGraw Hill; 1987.

[40] Koh, C. G., Gao, M., Luo, C. A new particle method for simulation of incompressible free surface flow problems. International Journal for Numerical Methods in Engineering, 2012; 89: 1582–1604. doi:10.1002/nme.3303.

[41] Shobeyri, G., Najafabadi, S. H. G., Abed, M. A Comparative Study on Two Mixed Least Squares Meshless Models with Improved SPH, MPS and CPM Methods to Solve Elasticity Problems. Iranian Journal of Science and Technology, Transactions of Mechanical Engineering, 2024; 48: 1565–1580. doi:10.1007/s40997-023-00742-x.

[42] Shobeyri, G., Ghoreishi Najafabadi, S. H., Abed, M. Computational efficiency of mixed least squares meshless models over SPH method for elliptic PDEs. AUT Journal of Civil Engineering, 2024; 8: 17–32. doi:10.22060/ajce.2024.22777.5847.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

A Comparative Analysis of Surrogate Safety Indicators for Road Hazard Assessment Using Smartphone GPS Data

Reza Jalalkamali [a*], Shahriar Afandizadeh Zargari [b], Mohammad Hossain Jalal Kamali [c]

[a] *Department of Civil Engineering, Faculty of Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*
[b] *Faculty of Civil Engineering, Iran University of Science and Technology, Tehran, Iran*
[c] *Department of Civil Engineering, Faculty of Engineering, University of Gonabad, Gonabad, Iran*

ARTICLE INFO

Article history:

Received 23 April 2026

Revised 25 May 2026

Accepted 16 June 2026

Available online 01 April 2027

Keywords:

Road safety

Crash

Surrogate safety indicator (SSI)

Deceleration number of vehicles

Smartphones

Real-time data

ABSTRACT

Road Safety Assessment (RSA) can be conducted through two primary approaches: (1) the reactive approach, which relies on historical crash data; and (2) the proactive approach, which employs Surrogate Safety Indicators (SSIs). In the reactive approach, hazard potential is evaluated directly by measuring accident frequency, whereas in the proactive approach, SSIs must be validated against actual safety measures. However, few studies have systematically compared multiple SSIs to determine which indicators provide the most reliable correlation with actual crash occurrence, particularly in rural contexts where data collection is challenging. The purpose of this research is to compare 39 SSIs within the proactive RSA framework to identify potential hazards on rural road segments. An observational study was conducted employing Pearson correlation analysis, using vehicle velocity data collected via GPS-enabled smartphones from 100 vehicles across 125 km of rural roads. The analysis divided the road network into uniform 5-km segments and evaluated various crash severity levels, including fatal, injury, and total crashes. Among 39 proposed indicators, the deceleration frequency indicator (DNM1, counting all decelerations during braking without predefined thresholds) demonstrates the strongest correlation with crash data ($r = 0.803$ for property-damage crashes), establishing it as a highly robust proactive safety measure for rural road networks. Additionally, results revealed that traditional metrics like average vehicle velocity alone were poor predictors of safety compared to comprehensive deceleration-based measures, emphasizing the necessity of granular kinematic data. The findings enable transportation agencies to implement cost-effective, smartphone-based safety monitoring systems for real-time identification of high-risk rural road segments.

How to cite this article: Jalalkamali, R., Afandizadeh Zargari, S., Jalal Kamali, M. H. A Comparative Analysis of Surrogate Safety Indicators for Road Hazard Assessment Using Smartphone GPS Data. Civil Engineering and Applied Solutions. 2027; 3(2): 37–48. doi:10.22080/ceas.2026.31594.1092.

1. Introduction

Road traffic accidents remain a major global public health concern, resulting in approximately 1.19 million fatalities and 20-50 million injuries annually worldwide. They currently stand as a leading cause of death, particularly among younger age groups. In Iran, traffic-related injuries similarly constitute a primary cause of mortality. The recording of over 16,700 fatalities in 2022 (Iranian calendar year 1401), which significantly exceeds both global and regional averages, highlights a severe crisis in road safety [1].

In recent years, Iran has experienced a simultaneous increase in the number of vehicles, travel speeds, population, and overall volume of road trips. However, the development and quality of road infrastructure have not improved proportionately. For instance,

* Corresponding author.
E-mail addresses: r.jalalkamali@uk.ac.ir (R. Jalalkamali).

https://doi.org/10.22080/ceas.2026.31594.1092
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

studies of two-lane rural highways in Iran show that geometric design elements, such as intersection angles and the existence of spiral curves, significantly influence accident ratios [2]. Furthermore, vehicle safety standards remain sub-optimal, and critically, there is a lack of adequate public education to foster a culture of safe driving. Consequently, the frequency, severity, and associated socio-economic costs of crashes have surged, underscoring the urgent need for effective preventive and mitigative measures [3].

Smart monitoring of road conditions and crash events by tracking vehicles via GPS and transmitting data to emergency services when necessary is an efficient approach to reducing traffic fatalities [4]. Some studies recommend adding modules equipped with various sensors that can record critical road parameters and pavement conditions while also detecting crash events [5].

The measure of road safety is often deduced from crash statistics. However, because crashes are accidental in nature, crash data are typically considered "rare events" in statistical terms. Consequently, at least three years of data are required to draw reliable conclusions about crash patterns. This represents a key limitation of this approach [6].

There are several methodologies that consider historical crash data in order to identify potentially high-risk locations [7]; however, the use of crash data in road safety analysis is sometimes hindered by logical and statistical problems such as small sample sizes, a lack of sufficient details to improve understanding of crash failure mechanisms [8], and unavailability [9]. To address these limitations, researchers have attempted to employ other parameters as surrogate measures for the identification of hazardous road locations, including safety inspections [10], traffic simulations [11, 12], traffic conflict techniques [13-15], and studies based on speed profile variation [15, 16]. These proactive approaches rely on observable traffic behavior and road characteristics rather than historical crash records.

Methods based on vehicle speed variations can be divided into three main categories: a) comparing the speed before a crash with the prevailing speed (event-based) [17]; b) comparing the speeds of drivers involved in a crash with those of drivers not involved in a crash (driver-based) [18]; c) comparing speed variations across different road facilities (facility-based) [19].

To proactively assess road safety, these approaches employ surrogate safety measures rather than relying solely on historical crash statistics. Among these measures, speed parameters have been widely adopted as reliable indicators. Following this rationale, the present study introduces a proactive safety evaluation framework based on continuous speed data collected via smartphone GPS.

2. Literature review

The review of prior investigations shows that most drivers exhibit strong reactions when encountering a hazardous situation. These reactions typically involve braking, steering, accelerating, or a combination thereof. Due to the correlation between crashes and near-crash events, researchers have attempted to prevent crashes through near-crash analysis [20]. In fact, classical naturalistic driving studies have firmly established near-crashes as valid crash surrogates for evaluating driving behavior [21]. Since braking constitutes the majority of near-crash events, many researchers have attempted to identify crash risks by measuring speed reduction using GPS technology.

As highlighted by recent reviews, relying solely on historical crash data has critical limitations, as it is fundamentally a reactive approach. In contrast, Surrogate Safety Measures (SSMs) provide sustainable, proactive indicators that allow researchers to investigate high-risk areas before crashes actually occur [22]. Surrogate Safety Measures can be broadly categorized into two main groups. The first category includes time-based, deceleration-based, and energy-based SSMs, such as Time-to-Collision (TTC), Post Encroachment Time (PET), Time-to-Crash/Accident (TC/TA), and Deceleration Rate to Avoid the Crash (DRAC). These measures are widely used across road safety literature and are typically based on predefined thresholds for traffic conflicts. The second category aims to directly associate each traffic conflict with either a crash or non-crash outcome by estimating its crash probability [23].

Early research employing GPS technology focused on extracting velocity profiles to establish a link between driving behavior and road crash occurrences. Pioneering work in this area by Jun [18] revealed that driving behavior, particularly the speed pattern, is a significant factor in crash potential by comparing drivers with and without recent crash histories [18]. This approach was further developed by Boonsiripant et al. [24], who formalized the technique as the “speed profile” to demonstrate the relationship between its extracted parameters and crash numbers [19]. In their subsequent studies, Boonsiripant et al. [9], Boonsiripant et al. [24] refined this model by identifying “acceleration noise” and “frequency of stops”, both correlated with deceleration, as key predictive parameters [9, 24]. In a similar application of this concept, Camacho-Torregrosa et al. [25] also utilized the speed profile from GPS data to analyze estimated crash rates, specifically for determining average operating speeds and evaluating design consistency parameters [25].

Fontaine and Chun [26] also used GPS probes to estimate the speed differential between adjacent TMC (Traffic Message Channel) segments on freeways. They introduced deceleration events as a surrogate safety measurement for crash data. These events were defined as the number of 15-minute periods where the speed differential between adjacent TMCs exceeded specific thresholds (10, 15, 20, and 25 mph) [26].

Historically, SSM data collection relied heavily on roadside observation techniques, which were labor-intensive and limited in spatial coverage. However, the advent of GPS-enabled devices and smartphone sensors has revolutionized this field by enabling widespread, continuous, and cost-effective data collection under real-world driving conditions [8].

With smartphone GPS accuracy having improved to positioning errors typically within 5 meters [27], their application in safety research has expanded. These devices can determine the geographic coordinates of a vehicle in real time, from which instantaneous speed can be calculated and reported [28]. The built-in sensors of smartphones, particularly GPS and accelerometers, have emerged as cost-effective and reliable tools across various smart transportation domains, from pavement condition monitoring to driver behavior analysis [29]. Furthermore, the rapid advancement of sensor technology has significantly increased reliance on these built-in sensors in naturalistic driving studies. They are now widely utilized to extract critical kinematic SSMs, specifically harsh braking and harsh acceleration events, directly from real-world driving environments [22, 30]. Unlike traditional crash records, which require extensive periods for reliable sample collection, smartphones enable widespread, ongoing, and cost-effective data gathering under real-world driving conditions, providing more timely insights into traffic dynamics. Leveraging this capability, Strauss et al. [31] used smartphone-derived data to examine bicyclists' deceleration rates as a surrogate measure for injury crashes. The correlation coefficient between injury crash frequency and deceleration rate in road sections was found to be r = 0.57 [31].

Stipancic et al. [32] also reviewed the utilization of vehicle manoeuvres as a surrogate measure for crashes. Using smartphone data, they observed that hard braking and hard acceleration events showed a good correlation with the frequency of crashes [32]. The broader literature increasingly supports this positive correlation between harsh driving behaviors, such as harsh braking and acceleration, captured via SSMs, and actual crash risks. Utilizing these harsh events as proactive variables offers a robust alternative to solely relying on reactive historical crash data, allowing for more pre-emptive safety interventions [30].

Further exploring the potential of smartphone telematics, Ziakopoulos [33] utilized high-resolution data to analyse the occurrence of harsh driving events, such as harsh braking (HB) and harsh acceleration (HA). By employing the XGBoost algorithm and SHAP explainability analysis, the study demonstrated that road network characteristics, specifically the number of lanes and speed limits, do not significantly influence the likelihood of these harsh events. Instead, more dynamic factors, including the vehicle's speed and traffic occupancy, were found to be the primary predictors of HB and HA occurrences [33].

In a complementary line of research, Nikolaou et al. [34] introduced spatial analysis to account for geographical dependencies in smartphone-based Surrogate Safety Measures (SSMs). Their study, which focused on harsh braking events, demonstrated that spatial models, such as the Spatial Zero-Inflated Negative Binomial (SZINB) and Spatial Random Forest (SRF), outperform their non-spatial counterparts. The findings identified variables like the number of trips, segment length, speeding, and mobile phone use as significant predictors [34].

While the aforementioned studies confirm the potential of using smartphone-based data for safety analysis, much of this research relies on identifying safety-critical events through the application of predefined and often arbitrary deceleration thresholds. This approach may overlook the cumulative safety implications of more frequent, lower-magnitude speed adjustments, which are particularly relevant on interurban roads. Furthermore, recent comprehensive reviews indicate a critical research gap: future studies must examine how various surrogate indicators correlate with historical road crashes across different severities and regional contexts, moving beyond single-indicator analysis. The present study addresses this gap by introducing and empirically evaluating 39 distinct SSIs, encompassing deceleration frequency and magnitude (across multiple thresholds), as well as indicators derived from velocity, acceleration, and jerk. The primary objective is to systematically correlate these indicators with historical crash data to identify the most robust measure of road safety.

3. The study area

The study area encompassed a two-way, two-lane undivided rural road spanning a total length of 137 km. This finalized route comprised three distinct segments: the first segment traversed level terrain for 20 km. The second segment predominantly passed through rolling terrain for 50 km, while the third segment navigated through mountainous terrain for 55 km. Fig. 1 and Table 1 display the geographical locations of the start and end points for each segment, as well as the entire route.

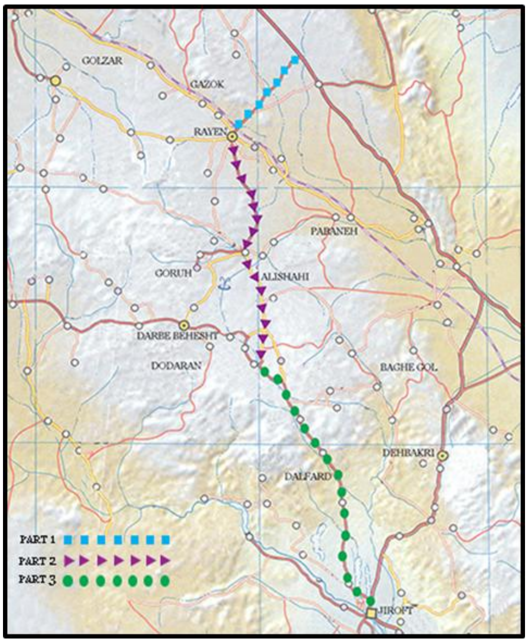


Fig. 1. The study area.

Table 1. Coordinates of the study area.

Part No.	Start point		End point	
	Latitude	Longitude	Latitude	Longitude
Part 1	29.759029°	57.585329°	29.608432°	57.442195°
Part 2	29.585696°	57.442669°	29.138452°	57.499041°
Part 3	29.131210°	57.504835°	28.747257°	57.758172°

4. Data collection

The total number of road crashes for three years was obtained from the Iranian Road Maintenance and Transportation Organization (RMTO). It should be noted that during this specific three-year period (March 21, 2011, to March 21, 2014), no major infrastructural modifications, geometric realignments, or policy changes (e.g., speed limit alterations) occurred along the studied route. This consistency ensures that the historical crash data provides a stable and reliable safety baseline, unskewed by external interventions. Vehicle GPS data were collected from a sample of cars traversing the route. To control for environmental variables and isolate the effect of road geometry, data collection was exclusively conducted under clear weather and dry pavement conditions, which accurately reflects the region's predominantly warm and dry climate with low rainfall. Furthermore, to ensure that the dataset captures a comprehensive representation of naturalistic driving behaviors, the data collection campaigns encompassed both daytime and nighttime driving under these clear conditions.

The AndroSensor mobile application was utilized in this study to record the vehicles' instantaneous speeds and geographic coordinates at one-second intervals with a 5-meter positional accuracy. During the drives, the smartphones were consistently placed on the vehicle dashboards to maintain a clear line of sight to the sky, thereby maximizing GPS signal reception and ensuring data quality. To prevent the Hawthorne effect and ensure naturalistic driving conditions, the data collection was conducted without the drivers' knowledge. The passengers carried out the recording while the drivers were kept completely unaware that their vehicle kinematics were being monitored. This approach ensures that the collected velocity profiles genuinely reflect natural driving behavior in response to the road geometry rather than altered behavior due to observation. A one-second sampling interval (1 Hz) was specifically chosen as it optimally captures macroscopic vehicle kinematics necessary for segment-level safety evaluation. Unlike higher-frequency data (e.g., OBD-II scanners), which are typically suited for microscopic driver behavior analysis and tend to introduce high-frequency sensor noise, a 1 Hz interval aligns with established literature to provide reliable and manageable velocity profiles [27, 35].

Since raw GPS data can be prone to signal fluctuations and sensor noise, which might be falsely interpreted as sudden speed changes, a Kalman filter was applied during the data post-processing phase. This filtering technique effectively smoothed the raw trajectory data and mitigated sensor noise, allowing for the accurate distinction and extraction of true deceleration events (Eq. 1).

$$\mathrm{n} = \frac{z_{1-\alpha/2}^{2} \times p \times (1-p)}{d^2} = \frac{1.96^2 \times 0.5 \times 0.5}{0.1^2} = 96.04 \tag{1}$$

where,

z = 1.96 is the critical value for the 95% confidence interval from the standard normal distribution,

p = 0.5 is the estimated proportion of the population with the characteristic of interest (yielding the maximum variance),

α = 0.05 is the significance level (Type I error rate), and

d = 0.1 is the acceptable margin of error.

A 10% margin of error (d = 0.1) was selected due to the extensive scope of the field study and the inherent logistical constraints of naturalistic data collection. In the context of this study, the calculated minimum sample size of n = 96.04 refers to the number of individual continuous vehicle trips required along the studied route. To account for potential data loss, signal dropouts, and recording errors (dropout rate), data collection was initially conducted for approximately 300 vehicle trips. After a rigorous data screening and quality control process to exclude corrupted, interrupted, or incomplete records, a final set of exactly 100 usable and high-quality vehicle trajectories was successfully extracted. This final sample size effectively accounts for the dropout rate while safely satisfying the statistical requirements for the analysis. The Surrogate Safety Indicators (SSIs) were extracted within 5-km spans, and then the SSIs of all sample vehicles were aggregated in each road segment.

5. Methodology

Following the data collection and sample size determination, this section details the analytical framework employed to evaluate the proposed SSIs. First, a comprehensive set of 39 SSIs is explicitly defined and categorized into two main groups: deceleration-based indicators and other metrics. Next, the procedure for extracting these indicators from the continuous GPS trajectories across the designated 5-km road segments is explained. Finally, the section outlines the statistical approach, utilizing the Pearson Correlation Coefficient (PCC), to establish and quantify the linear associations between these extracted SSIs and historical crash frequencies categorized by severity.

5.1. Suggested indicators

Previous studies on Surrogate Safety Indicators (SSIs) were first reviewed. Subsequently, three road sections were selected according to the study area conditions and road transport regulations. Crash statistics for these three sections were obtained for a 3-year period. Statistical reports on crashes (fatality, injury-causing, damage-causing, total observed, and equivalent property damage only) were extracted for 25 road segments of 5 km length, along with 39 SSIs. GPS-enabled smartphones and appropriate applications, which report the geographic location and the vehicle velocity at one-second intervals, were used to derive SSIs for 100 vehicles on these three road sections. Table 2 indicates the classification of crash types.

Table 2. Crash types (classified by severity).

No.	Crash type	Description
1	Fatality	Crashes leading to the death of at least one person up to 30 hours after their occurrence
2	Injury	Crashes leading to disability, severe injury, light injury, or possible injury
3	Fatality + injury	Fatality + injury crashes (row 1 + row 2)
4	Damaging only	Crashes that lead to financial losses only
5	Total observed	Total fatality, injury, and damaging crashes
6	Equivalent property damage only (EPDO)	In this method, each fatality crash and each injury crash is counted as 9 and 3 damaging crashes, respectively.

The SSIs were divided into two general categories: deceleration-based indicators and indicators.

5.1.1. Deceleration-based indicators

Deceleration-based indicators encompass two groups:

A) Deceleration frequency-based Indicators,

B) Deceleration values-based indicators.

Notably, unlike many urban studies that employ specific deceleration thresholds to isolate safety-critical braking events [18, 32], this study deliberately did not impose a predefined threshold for interurban road segments. In uninterrupted rural traffic flows, drivers typically maintain a steady speed; therefore, even minor but sustained decelerations often reflect essential driving adaptations to roadway geometry, sightline limitations, or potential hazards. By omitting a predefined threshold, the study aimed to prevent bias and empirically evaluate which indicators (including those accounting for all decelerations) yield the highest correlation with actual crash data. Furthermore, applying the Kalman filter during data post-processing ensured that minor signal fluctuations and high-frequency sensor noises were not mistakenly counted as low-magnitude decelerations.

A) Deceleration frequency-based Indicators;

1) The Deceleration Numbers of Mode 1 (DNM1), represent the total count of all decelerations occurring throughout the entire braking duration, recorded at one-second intervals.

2) In contrast to the first mode, the Deceleration Numbers of Mode 2 (DNM2), represent the count of braking events, recording only a single deceleration per event and assigning its spatial coordinates to the onset of braking.

3) An analysis of the deceleration values during braking events revealed that they do not strictly follow an ascending or descending trend, but rather exhibit fluctuations. Consequently, the Deceleration Numbers of Mode 3 (DNM3) represent the count of local maximum decelerations within a braking event, with the spatial coordinates of each local maximum recorded as needed.

4) Similar to DNM, the Deceleration Numbers of Mode 4 (DNM4) records only a single deceleration per braking event; however, its spatial coordinates are assigned to the location of the maximum deceleration rather than the onset of braking.

B) Deceleration values-based indicators;

A pertinent question arises as to why the frequency of decelerations is emphasized over their magnitudes. Recognizing that count-based indicators disregard the numerical value of deceleration (unless classified by a minimum threshold), incorporating the magnitude of each deceleration across the 4 aforementioned modes could potentially improve analytical outcomes. Consequently, this approach may yield an indicator with superior performance in assessing road safety compared to relying solely on historical crash statistics. To investigate this hypothesis, 4 additional indicators are proposed as follows:

5) Corresponding to DNM1, the Value of Deceleration Mode 1 (VDM1) utilizes the sum of deceleration magnitudes rather than their frequency of occurrence. Essentially, this indicator represents the approximate area under the acceleration-time curve for each braking event.

6) The sixth indicator, the Value of Deceleration Mode 2 (VDM2), considers the initial deceleration magnitude during each braking event. This indicator is of particular significance as it may reflect the driver's perception of hazard; consequently, it is investigated as a potentially viable surrogate safety indicator.

7) Corresponding to DNM3, the seventh indicator, the Value of Deceleration Mode 3 (VDM3), utilizes the sum of the

magnitudes of the local maximum decelerations during a braking event, rather than their frequency. This indicator is introduced as it may provide a more accurate representation of the hazard level, making it a potentially more suitable surrogate safety indicator.

8) Determining the maximum deceleration magnitude during each braking event is critical, as it may reflect the severity of the hazard rather than the severity of a crash. Therefore, the eighth indicator, the Value of Deceleration Mode 4 (VDM4), is defined as the sum of the maximum deceleration magnitudes occurring across all braking events. This measure is proposed as a potential candidate for a surrogate safety indicator to replace historical crash statistics.

Upon determining the values of the 8 aforementioned indicators for each braking event, the values were summed across all 90 vehicles for every 5-km road segment. The resulting aggregate served as the representative value for the respective indicator within that segment.

5.1.2. Other indicators

Another pertinent question is why only the frequency or magnitude of decelerations (aggregated within the aforementioned categories) are considered appropriate SSIs in RSA studies. While an SSI is often associated with a braking event, it does not necessarily need to be defined by the cumulative sum or count of decelerations. Instead, it could be related to other parameters, such as velocity or variations in acceleration (both positive and negative). Consequently, the following new indicators are proposed:

9) Average velocity; the velocity-crash association has been extensively investigated. Here we aimed to find whether there is an association between average velocity (Vave) and the number of crashes on each road segment of 5 km length.

10) Average velocity to the second power (Vave2); as with some references, here the association between Vave2 and the number of light injury-causing crashes was investigated.

11) Average velocity to the third power (Vave3); as with some references, here the association between Vave3 and the number of light injury-causing crashes was investigated.

12) Statistics on road crashes in Iran do not separate light and severe injury-causing crashes, and only three general groups of crashes (fatality, injury-causing and damaging) are mentioned. We therefore defined the average velocity to the second plus third power (Vave2+ Vave3), aiming to examine the association between this indicator and crashes.

13) Fatal crashes are a critical component of RSA studies, as they result in significant property damage, severe bodily injuries, and substantial compensation costs, in addition to the loss of life among vehicle occupants. Among crashes of varying severities, fatal crashes occur with the lowest frequency. This low frequency increases the likelihood of random occurrences, making it significantly more difficult to identify a proper SSI that exhibits a strong correlation with these events. Furthermore, fatal crash outcomes are highly dependent on numerous exogenous factors. For instance, the presence of vehicle safety features (e.g., airbags) and occupant compliance (e.g., seatbelt usage) are critical determinants. Additionally, the efficiency of post-crash emergency medical services plays a vital role. These confounding variables, combined with the highly random nature of such crashes, make it exceedingly difficult to derive an appropriate SSI solely from velocity or acceleration data, whether analyzed individually or in combination. Nevertheless, it is still possible to identify the indicator that demonstrates the highest relative correlation with these crashes. Consistent with the existing literature, this study investigates the association between Vave4 and the frequency of fatal crashes.

14) The velocity variance (Vvar), which represents the difference of average velocity to the second power, can be an appropriate SSI, since the vehicle's average velocity may not be the crash's cause, but it may be influenced by inconsistency and turbulence in the vehicle's movement, which can be better expressed in the difference of average velocity.

15) Standard Deviation (SD) of velocity (Vstdv), which indicates the vehicle's velocity difference from the average. This parameter represents inconsistency and turbulence in the vehicle's movement.

16) The average vehicle acceleration on roads (Aave). Vehicles can move more quickly and evenly on a straight road with light traffic; consequently, the Vave increases, and Aave decreases. In contrast, on a winding or mountainous road, vehicles move at a lower velocity and more acceleration changes inevitably occur. Therefore, negative or positive average acceleration is likely to be an appropriate SSI. Variance and SD were also controlled.

17) The variance of acceleration (Avar) for each road segment.

18) The standard deviation of acceleration (Astdv) for each road segment.

19) Average deceleration (Dave) for each road segment. The sum of deceleration numbers and values was expressed as the SSI for deceleration value and number. However, the average, variance, and/or SD of deceleration values may be a more appropriate SSI.

20) The variance of deceleration (Dvar) for each road segment.

21) The standard deviation of deceleration (Dstdv) for each road segment.

22) Average of a jerk to the second power for each road segment (Jave2). Acceleration is a derivative of velocity with respect to time, while jerk is the time derivative of acceleration. In the former, the driver brakes and velocity decreases gradually

with a steady rate over braking. But in the latter, the velocity is suddenly reduced, and the acceleration rate is not constant over braking, where the crash risk is higher. Therefore, the average of a jerk to the second power is probably a more suitable SSI, because, in the second power, the jerk's positive and negative modes cannot neutralize each other's effect.

23) The variance of a jerk to the second power for each road segment (Jvar2).

24) The standard deviation of a jerk to the second power for each road segment (Jstdv2).

25) The average product of the velocity with the deceleration in the DNM1 (SDM1ave) for each road segment. It should be assessed whether the risk level can be greater when braking occurs at a high rather than low velocity, even under the same deceleration. The product of the deceleration with velocity (SD) is therefore considered a possible SSI. Considering this indicator with all four deceleration number modes would probably lead to appropriate SSIs. In each case, we can assess the average, variance, or standard deviation of the product of the velocity with the deceleration.

26) The average product of the velocity with the deceleration in DNM2 (SDM2ave) for each road segment.

27) The average product of the velocity with the deceleration in DNM3 (SDM3ave) for each road segment.

28) The average product of the velocity with the deceleration in DNM4 (SDM4ave) for each road segment.

29) The variance of the product of the velocity with the deceleration in DNM1 (SDM1var) for each road segment.

30) The variance of the product of the velocity with the deceleration in DNM2 (SDM2var) for each road segment.

31) The variance of the product of the velocity with the deceleration in DNM3 (SDM3var) for each road segment.

32) The variance of the product of the velocity with the deceleration in DNM4 (SDM4var) for each road segment.

33) The standard deviation of the product of the velocity with the deceleration in DNM1 (SDM1stdv) for each road segment.

34) The standard deviation of the product of the velocity with the deceleration in DNM2 (SDM2stdv) for each road segment.

35) The standard deviation of the product of the velocity with the deceleration in DNM3 (SDM3stdv) for each road segment.

36) The standard deviation of the product of the velocity with the deceleration in DNM4 (SDM4stdv) for each road segment.

37) Average difference of velocity to the second power (at one-second intervals) in the deceleration modes for each road segment (ΔSD2ave). The length of the brake line is proportional to the changes of velocity to the second power during the braking time. This could therefore be an appropriate SSI.

38) The variance of the difference of velocity to the second power (at one-second intervals) in the deceleration modes for each road segment (ΔSD2var).

39) The standard deviation of the difference of velocity to the second power (at one-second intervals) in the deceleration modes for each road segment (ΔSD2stdv).

The value of each of the indicators and the statistics of observed crashes (classified by severity) were calculated on each of the road segments of 5 km length. The Pearson correlation coefficient (PCC) was used to measure the linear correlation between SSIs and crash statistics.

6. Results

6.1. Results of deceleration-based indicators

The results of the correlation of indicators related to the number or total of decelerations are presented in Tables 3 to 10. Classification of the acceleration number or value in these values is based on the lower limit of deceleration values with steps of 0.5 m/s^2. For example, NMD1 > 1 means the deceleration number in the first mode, but the deceleration value must be greater than 1 m/s^2. The analysis of Tables 3 to 10 showed that in all modes, the deceleration numbers were consistent with the number of crashes (of any severity). The (corresponding) indicators of total deceleration values showed a better correlation. The best results were obtained in Table 3. Among the eight indicators, the frequency of DNM1 (where all the decelerations (deceleration) at the whole braking time (a deceleration per second) were counted) showed the highest correlation with crashes, especially damaging ones (and thus total crashes).

6.2. Results of other indicators

An SSI may be associated with deceleration, but it may not be associated with the sum of the deceleration amount or number, or even deceleration itself. Here, the indicators were defined first, and the results of the correlation test on these indicators and the number of crashes were then presented in Table 11.

7. Discussion

As shown in Table 11, contrary to common assumption, average speed does not exhibit a significant or direct correlation with total crashes or specific crash severities. This observation holds true for the second, third, and fourth powers of average speed as well. Where a correlation does exist, it tends to be inversely proportional. Furthermore, when speed is combined with another parameter into a composite index, such as the product of speed and deceleration, it diminishes the parameter’s correlation with crash frequencies. This phenomenon likely stems from the topographical characteristics of mountainous segments; due to steep gradients and numerous horizontal curves, the average vehicle speed is inherently lower, yet the frequency of crashes in these areas is higher.

Conversely, the standard deviation of speed exhibits a moderate correlation (0.4 to 0.6) with observed crashes. Deviations from the mean speed within a segment induce traffic turbulence and increase overtaking maneuvers, ultimately leading to collisions.

While average acceleration showed no correlation with crashes, the variance and standard deviation of both acceleration and deceleration demonstrate a moderate correlation. These parameters serve as measures of “acceleration noise” (turbulence), indicating the deviation of vehicle acceleration from a normal state. High acceleration noise signifies greater traffic disruption, which consequently leads to increased crash rates. Notably, the average deceleration indicator yields better results, surpassing the 0.6 correlation threshold.

Jerk represents the rate of change of acceleration over time. A higher jerk implies more abrupt acceleration changes, which may indicate the severity of the hazard perceived by the driver (rather than the crash severity itself). To prevent positive and negative values from cancelling each other out, the correlation of the squared jerk with crashes was examined. However, results indicated that the average squared jerk also has only a moderate correlation with crashes. Additionally, the indicator representing the average difference of squared speeds during deceleration, which correlates with braking distance during a braking event, demonstrated an adequate correlation (approximately0.6) with Equivalent Property Damage Only (EPDO) crashes.

Table 3. The correlation coefficient of the deceleration numbers indicator (first mode-DNM1) with the number of observed crashes (classified by severity).

Crash type	NDM1 > 0	NDM1 > 0.5	NDM1 > 1	NDM1 > 1.5	NDM1 > 2	NDM1 > 2.5	NDM1 > 3	NDM1 > 3.5
Fatality	0.376	0.325	0.280	0.235	0.234	0.137	0.153	0.142
Injury	0.102	0.148	0.172	0.165	0.156	0.185	0.140	0.147
Fatality + injury	0.232	0.253	0.257	0.234	0.226	0.215	0.181	0.183
Damage only	0.803	0.681	0.664	0.628	0.617	0.579	0.555	0.526
Total observed	0.741	0.642	0.628	0.592	0.580	0.546	0.517	0.492
EPDO	0.601	0.535	0.514	0.473	0.464	0.412	0.391	0.374

Table 4. The correlation coefficient of the deceleration numbers indicator (second mode-DNM2) with the number of observed crashes (classified by severity).

Crash type	NDM2 > 0	NDM2 > 0.5	NDM2 > 1	NDM2 > 1.5	NDM2 > 2	NDM2 > 2.5	NDM2 > 3	NDM2 > 3.5
Fatality	0.222	0.309	0.293	0.252	0.244	0.162	0.168	0.140
Injury	0.091	0.137	0.159	0.151	0.122	0.163	0.138	0.146
Fatality + injury	0.164	0.237	0.251	0.228	0.199	0.206	0.185	0.182
Damage only	0.713	0.702	0.683	0.649	0.630	0.608	0.584	0.538
Total observed	0.647	0.656	0.643	0.608	0.586	0.568	0.542	0.502
EPDO	0.482	0.533	0.524	0.486	0.463	0.429	0.411	0.379

Table 5. The correlation coefficient of the deceleration numbers indicator (third mode-DNM3) with the number of observed crashes (classified by severity).

Crash type	NDM3 > 0	NDM3 > 0.5	NDM3 > 1	NDM3 > 1.5	NDM3 > 2	NDM3 > 2.5	NDM3 > 3	NDM3 > 3.5
Fatality	0.300	0.316	0.295	0.253	0.245	0.155	0.168	0.140
Injury	0.100	0.146	0.167	0.158	0.128	0.168	0.137	0.145
Fatality + injury	0.201	0.248	0.258	0.235	0.205	0.207	0.184	0.181
Damage only	0.771	0.700	0.685	0.650	0.632	0.604	0.586	0.543
Total observed	0.706	0.657	0.647	0.611	0.588	0.565	0.544	0.506
EPDO	0.549	0.539	0.530	0.490	0.466	0.426	0.411	0.381

Table 6. The correlation coefficient of the deceleration numbers indicator (fourth mode-DNM4) with the number of observed crashes (classified by severity).

Crash type	NDM4 > 0	NDM4 > 0.5	NDM4 > 1	NDM4 > 1.5	NDM4 > 2	NDM4 > 2.5	NDM4 > 3	NDM4 > 3.5
Fatality	0.222	0.291	0.293	0.252	0.240	0.157	0.173	0.136
Injury	0.093	0.131	0.143	0.140	0.114	0.155	0.124	0.153
Fatality + injury	0.165	0.225	0.237	0.218	0.191	0.196	0.175	0.186
Damage only	0.712	0.691	0.683	0.647	0.628	0.604	0.592	0.531
Total observed	0.647	0.644	0.640	0.604	0.581	0.562	0.547	0.498
EPDO	0.482	0.517	0.519	0.481	0.457	0.422	0.412	0.376

Table 7. The correlation coefficient of the sum of the deceleration values indicator (VDM1-correspondent with first mode) with the number of observed crashes (classified by severity).

Crash type	VDM1 > 0	VDM1 > 0.5	VDM1 > 1	VDM1 > 1.5	VDM1 > 2	VDM1 > 2.5	VDM1 > 3	VDM1 > 3.5
Fatality	0.326	0.301	0.259	0.217	0.206	0.124	0.126	0.098
Injury	0.152	0.159	0.178	0.177	0.179	0.211	0.196	0.230
Fatality + injury	0.257	0.254	0.255	0.238	0.235	0.233	0.221	0.240
Damage only	0.700	0.669	0.649	0.616	0.600	0.562	0.528	0.475
Total observed	0.659	0.632	0.614	0.583	0.568	0.535	0.504	0.463
EPDO	0.546	0.521	0.499	0.463	0.451	0.407	0.386	0.359

Table 8. The correlation coefficient of the sum of the deceleration values indicator (VDM2-correspondent with second mode) with the number of observed crashes (classified by severity).

Crash type	VDM2 > 0	VDM2 > 0.5	VDM2 > 1	VDM2 > 1.5	VDM2 > 2	VDM2 > 2.5	VDM2 > 3	VDM2 > 3.5
Fatality	0.229	0.223	0.222	0.205	0.202	0.162	0.160	0.069
Injury	0.124	0.127	0.118	0.141	0.112	0.176	0.268	0.286
Fatality + injury	0.195	0.195	0.188	0.202	0.175	0.217	0.297	0.279
Damage only	0.650	0.620	0.592	0.573	0.520	0.501	0.466	0.277
Total observed	0.602	0.576	0.550	0.537	0.485	0.480	0.469	0.303
EPDO	0.466	0.450	0.433	0.424	0.386	0.382	0.397	0.270

Table 9. The correlation coefficient of the sum of the deceleration values indicator (VDM3-correspondent with third mode) with the number of observed crashes (classified by severity).

Crash type	VDM3 > 0	VDM3 > 0.5	VDM3 > 1	VDM3 > 1.5	VDM3 > 2	VDM3 > 2.5	VDM3 > 3	VDM3 > 3.5
Fatality	0.299	0.297	0.272	0.233	0.217	0.139	0.137	0.137
Injury	0.150	0.155	0.171	0.169	0.156	0.200	0.200	0.200
Fatality + injury	0.245	0.249	0.254	0.237	0.219	0.229	0.228	0.228
Damage only	0.704	0.686	0.669	0.637	0.617	0.586	0.556	0.556
Total observed	0.659	0.645	0.631	0.600	0.579	0.555	0.529	0.529
EPDO	0.534	0.526	0.513	0.478	0.456	0.421	0.406	0.406

Table 10. The correlation coefficient of the sum of the deceleration values indicator (VDM4-correspondent with fourth mode) with the number of observed crashes (classified by severity).

Crash type	VDM4 > 0	VDM4 > 0.5	VDM4 > 1	VDM4 > 1.5	VDM4 > 2	VDM4 > 2.5	VDM4 > 3	VDM4 > 3.5
Fatality	0.266	0.278	0.270	0.231	0.213	0.141	-0.016	0.090
Injury	0.137	0.140	0.152	0.153	0.144	0.190	0.234	0.245
Fatality + injury	0.221	0.228	0.235	0.222	0.207	0.221	0.200	0.251
Damage only	0.689	0.678	0.666	0.634	0.613	0.585	-0.150	0.478
Total observed	0.640	0.633	0.624	0.594	0.572	0.552	-0.080	0.468
EPDO	0.507	0.508	0.503	0.470	0.448	0.418	0.005	0.363

Table 11. Results of the correlation test on SSIs and the number of crashes (classified by severity).

Crash type	Fatality	Injury	Fatality + injury	Damage only	Total observed	EPDO
V_{ave}	-0.406	-0.141	-0.277	-0.774	-0.727	-0.959
$V_{ave}{}^{2}$	-0.421	-0.15	-0.29	-0.772	-0.728	-0.943
$V_{ave}{}^{3}$	-0.434	-0.154	-0.3	-0.766	-0.725	-0.924
$V_{ave}{}^{2,3}$	-0.434	-0.154	-0.299	-0.766	-0.725	-0.924
$V_{ave}{}^{4}$	-0.445	-0.156	-0.305	-0.757	-0.719	-0.904
V_{var}	0.241	-0.021	0.072	0.198	0.186	0.2
V_{stdv}	0.244	0.283	0.342	0.559	0.559	0.486
A_{ave}	0.06	0.06	0.002	-0.141	-0.12	-0.11
A_{var}	0.189	0.303	0.339	0.524	0.528	0.452
A_{stdv}	0.244	0.283	0.342	0.559	0.559	0.486
D_{ave}	0.329	0.2	0.3	0.601	0.585	0.516
D_{var}	0.178	0.327	0.356	0.476	0.492	0.432
D_{stv}	0.239	0.306	0.36	0.524	0.533	0.476
J^{2}_{ave}	0.156	0.332	0.352	0.504	0.515	0.754
J^{2}_{var}	-0.193	0.461	0.335	-0.09	0.004	-0.017

J^2_{stv}	-0.118	0.505	0.403	0.062	0.15	0.139
$SDM1_{ave}$	0.143	0.225	0.253	0.137	0.178	0.32
$SDM2_{ave}$	-0.369	0.106	-0.044	-0.328	-0.291	-0.178
$SDM3_{ave}$	0.037	0.26	0.243	0.095	0.14	0.286
$SDM4_{ave}$	-0.156	0.251	0.163	-0.007	0.034	0.187
$SDM1_{var}$	-0.129	0.394	0.3	-0.078	0.006	-0.011
$SDM2_{var}$	-0.113	0.202	0.136	-0.277	-0.204	-0.237
$SDM3_{var}$	-0.132	0.405	0.309	-0.061	0.022	0.004
$SDM4_{var}$	-0.154	0.408	0.303	-0.067	0.015	-0.004
$SDM1_{stv}$	-0.085	0.354	0.281	-0.054	0.021	0.010
$SDM2_{stv}$	-0.153	0.177	0.099	-0.315	-0.245	-0.273
$SDM3_{stv}$	-0.088	0.363	0.288	-0.037	0.038	0.027
$SDM4_{stv}$	-0.122	0.362	0.275	-0.052	0.022	0.014
$\Delta S_D{}^2{}_{ave}$	-0.060	0.069	0.038	0.369	0.324	0.595
$\Delta S_D{}^2{}_{var}$	-0.122	0.066	0.013	0.317	0.273	0.498
$\Delta S_D{}^2{}_{stv}$	-0.084	0.125	0.079	0.346	0.314	0.567

8. Conclusions

This study evaluated various Surrogate Safety Indicators (SSIs) derived from smartphone GPS data to assess road hazards, particularly focusing on mountainous terrains. The comparative analysis of 39 indicators across different modes of calculation yielded several critical insights:

First, simplistic indicators such as average speed are inadequate for safety assessment in mountainous topographies. The inherent geometric constraints (steep gradients and sharp curves) naturally lower average speeds in high-risk zones, leading to an inverse or insignificant correlation with actual crash frequencies. Instead, measures of traffic turbulence, such as the standard deviation of speed and acceleration noise, serve as more reliable indicators of crash potential.

Second, indicators associated with deceleration proved to be substantially more reliable than those related to acceleration or composite speed-deceleration indices. Among the 31 indicators categorized in the second group (Table 11), average deceleration (Dave) exhibited the highest correlation with crash occurrences (over 0.6). Furthermore, the average difference of squared velocities during deceleration (ΔSDave2) showed promising results, particularly for EPDO crashes, highlighting the importance of braking dynamics in hazard evaluation.

Ultimately, when comparing the most effective indicators from all groups, the frequency-based indicators significantly outperformed value-based or variance-based metrics. The frequency of type-1 deceleration (DNM1), which assigns one count per second of deceleration across the entire braking duration, demonstrated the strongest overall correlation with both damaging and total observed crashes. Therefore, DNM1 is recommended as the most appropriate and robust surrogate safety indicator for proactive road hazard assessment using generalized GPS data.

9. Limitations and future directions

While this study successfully utilized the Pearson Correlation Coefficient (PCC) as a straightforward metric to initially screen, comparatively evaluate, and rank the proposed surrogate safety indicators, it is important to acknowledge certain methodological limitations. First, the primary scope of this research was the comparative ranking of indicators rather than developing predictive crash count models. For predictive modelling purposes, future studies should employ advanced count-data regression models, such as Poisson or Negative Binomial (NB) models, which are specifically designed to handle the nature of crash frequencies. Second, the current analysis did not account for potential spatial autocorrelation among the adjacent 5-km road segments. Addressing spatial dependencies using spatial econometric models in future research would be a highly valuable extension to refine the accuracy and reliability of the predictive safety models based on these surrogate indicators.

Furthermore, as this study focused on establishing statistical correlations, the identified indicators were not validated on an independent dataset. Future studies should focus on developing predictive models using these indicators and validating their predictive power on external datasets.

Additionally, while this study aimed to identify a generalized indicator across diverse topographies, future studies could perform segregated analyses to develop terrain-specific safety indicators and thresholds.

Statements & declarations

Author contributions

Reza Jalalkamali: Conceptualization, Investigation, Methodology, Formal analysis, Resources, Writing - Review & Editing.

Shahriar Afandizadeh Zargari: Conceptualization, Investigation, Methodology, Project administration.

Mohammad Hossain Jalal Kamali: Investigation, Writing - Review & Editing.

Funding

Funding information is not available.

Data availability

Data available on request due to restrictions, e.g., privacy or ethics: The data presented in this study are available on request from the corresponding author. The data are not publicly available because the researchers intend to retain the dataset for future comparative analyses and extended research purposes.

Declarations

The authors declare no conflict of interest.

References

[1] Khedri, E., Amin, R., Khodaee, A. Mortality trend due to traffic accident in Iran in Years 2018-2021. Civil and Project, 2023; 5: 40–56. doi:10.22034/cpj.2023.400882.1206.

[2] Jalal Kamali, M. H., Monajjem, M. S., Ayubirad, M. S. Studying the effect of spiral curves and intersection angle, on the accident ratios in two-lane rural highways in Iran. Promet - Traffic&Transportation, 2013; 25: 343–348. doi:10.7307/ptt.v25i4.332.

[3] Shafabakhsh, G. A., Famili, A., Akbari, M. Spatial analysis of data frequency and severity of rural accidents. Transportation Letters, 2023; 15: 1243–1250. doi:10.1080/19427867.2016.1138605.

[4] Butkar, U. D., Gandhewar, N. Accident Detection and Alert System (Current Location) Using Global Positioning System. International Journal of Mechanical Engineering, 2022; 7.

[5] Jahan, I. A., Jamil, I. A., Fahim, M. S. H., Huq, A. S., Faisal, F., Nishat, M. M. Accident detection and road condition monitoring using blackbox module. In: 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN); 2022 Jul 05–08; Barcelona, Spain. p. 247–252. doi:10.1109/ICUFN55119.2022.9829589.

[6] Lord, D., Mannering, F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice, 2010; 44: 291–305. doi:10.1016/j.tra.2010.02.001.

[7] Highway Safety Manual (HSM). 1st ed. Washington, D.C. (US): American Association of State Highway Transportation Officials (AASHTO); 2010.

[8] Tarko, A. P. Surrogate measures of safety. In: Lord, D., Washington, S., editors. Safe mobility: challenges, methodology and solutions. Leeds: Emerald Group Publishing Ltd.; 2018. p. 383–405. doi:10.1108/S2044-994120180000011019.

[9] Boonsiripant, S., Rodgers, M. O., Hunter, M. P. Speed profile variation as a road network screening tool. Transportation research record, 2011; 2236: 83–91. doi:10.3141/2236-10.

[10] Li, K., Sun, J., Ni, Y., Zhou, S. e. Evaluation of safety factors at Chinese intersections. Proceedings of the ICE - Transport, 2012; 165: 195–204. doi:10.1680/tran.10.00021.

[11] Abdel-Aty, M. Crash Simulation and Animation: "A New Approach for Traffic Safety Analysis". Orlando (FL): University of Central Florida; 2001. Report No.: 01.

[12] Young, W., Archer, J. The measurement and modelling of proximal safety measures. Proceedings of the ICE - Transport, 2010; 163: 191–201. doi:10.1680/tran.2010.163.4.191.

[13] Essa, M., Sayed, T. Comparison between Surrogate Safety Assessment Model and Real-Time Safety Models in Predicting Field-Measured Conflicts at Signalized Intersections. Transportation Research Record: Journal of the Transportation Research Board, 2020; 2674: 036119812090787. doi:10.1177/0361198120907874.

[14] Ewadh, H. A., Neham, S. S. Conflict to study safety at four leg-signalised intersections. Proceedings of the Institution of Civil Engineers - Transport, 2011; 164: 221–230. doi:10.1680/tran.7.00055.

[15] Saad, M., Abdel-Aty, M., Lee, J., Wang, L. Access Design Safety Analysis for Managed Lanes Including Accessibility Level and Weaving Length. In: Transportation Research Board 97th Annual Meeting; 2018 Jan 07–11; Washington, D.C., United States. p. 18–00860.

[16] Castro, M., Iglesias, L., Rodriguez-Solano, R., Sánchez, J. A. Highway safety analysis using geographic information systems. Proceedings of the Institution of Civil Engineers - Transport, 2008; 161: 91–97. doi:10.1680/tran.2008.161.2.91.

[17] Kockelman, K., Ma, J. Freeway Speeds and Speed Variations Preceding Crashes, Within and Across Lanes. Journal of the Transportation Research Forum, 2007; 46: 43–61. doi:10.5399/osu/jtrf.46.1.976.

[18] Jun, J. Potential Crash Measures Based on GPS-Observed Driving Behavior Activity Metrics. Atlanta (GA): Georgia Institute of Technology; 2006.

[19] Boonsiripant, S. Speed profile variation as a surrogate measure of road safety based on GPS-equipped vehicle data. 1st ed. Atlanta (GA): Georgia Institute of Technology; 2009.

[20] Guo, F., Klauer, S. G., McGill, M. T., Dingus, T. A. Evaluating the relationship between near-crashes and crashes: Can near-crashes serve as a surrogate safety metric for crashes? Washington, D.C. (US): U.S. Department of Transportation National Highway Traffic Safety Administration; 2010. Report No.: DOT HS 811 382.

[21] Guo, F., Klauer, S., Hankey, J., Dingus, T. Near Crashes as Crash Surrogate for Naturalistic Driving Studies. Transportation Research Record: Journal of the Transportation Research Board, 2010; 2147: 66–74. doi:10.3141/2147-09.

[22] Nikolaou, D., Ziakopoulos, A., Yannis, G. A Review of Surrogate Safety Measures Uses in Historical Crash Investigations. Sustainability, 2023; 15: 7580. doi:10.3390/su15097580.

[23] Bonela, S. R., Kadali, R. Review of traffic safety evaluation at T-intersections using surrogate safety measures in developing countries context. IATSS Research, 2022; 46: 307–321. doi:10.1016/j.iatssr.2022.03.001.

[24] Boonsiripant, S., Hunter, M., Dixon, K., Rodgers, M. Measurement and Comparison of Acceleration and Deceleration Zones at Traffic Control Intersections. Transportation Research Record: Journal of the Transportation Research Board, 2010; 2171: 1–10. doi:10.3141/2171-01.

[25] Camacho-Torregrosa, F. J., Pérez-Zuriaga, A. M., Campoy-Ungría, J. M., García-García, A. New geometric design consistency model based on operating speed profiles for road safety evaluation. Accident Analysis & Prevention, 2013; 61: 33–42. doi:10.1016/j.aap.2012.10.001.

[26] Fontaine, M., Chun, P. Feasibility of Using Segment Speed Deceleration Events as Safety Surrogate Measures at a Series of Freeway Interchanges. In: Transportation Research Board 95th Annual Meeting; 2016 Jan 10–14; Washington, D.C., United States. p. 1–16.

[27] Wolf, J. L., Bachman, W. H., Oliveira, M. S., Auld, J. A., Mohammadian, A., Vovsha, P. S., Zmud, J. Applying GPS data to understand travel behavior. ed. Washington, D.C. (US): Transportation Research Board; 2014.

[28] Chugh, G., Bansal, D., Sofat, S. Road condition detection using smartphone sensors: A survey. International Journal of Electronic and Electrical Engineering, 2014; 7: 595–602.

[29] Jalalkamali, R., Kamali Sarvestani, P. Investigating the Situation of Urban Pathways Pavements Surface in Future Smart Cities Using Accelerometer Sensors and GPS Tools of Smartphones and GIS Application. Modares Civil Engineering journal, 2022; 22: 21–36. doi:10.22034/22.5.21.

[30] Sharma, S. N., Singh, D., Dehalwar, K. Surrogate safety analysis-leveraging advanced technologies for safer roads. Suranaree Journal of Science and Technology, 2024; 31: 10.55766. doi:10.55766/sujst-2024-04-e03837.

[31] Strauss, J., Zangenehpour, S., Miranda-Moreno, L. F., Saunier, N. Cyclist deceleration rate as surrogate safety measure in Montreal using smartphone GPS data. Accident Analysis & Prevention, 2017; 99: 287–296. doi:10.1016/j.aap.2016.11.019.

[32] Stipancic, J., Miranda-Moreno, L., Saunier, N. Vehicle manoeuvers as surrogate safety measures: Extracting data from the GPS-enabled smartphones of regular drivers. Accid Anal Prev, 2018; 115: 160–169. doi:10.1016/j.aap.2018.03.005.

[33] Ziakopoulos, A. Analysis of harsh braking and harsh acceleration occurrence via explainable imbalanced machine learning using high-resolution smartphone telematics and traffic data. Accident Analysis & Prevention, 2024; 207: 107743. doi:10.1016/j.aap.2024.107743.

[34] Nikolaou, D., Ziakopoulos, A., Kontaxi, A., Theofilatos, A. A., Yannis, G. Spatial analysis of telematics-based surrogate safety measures. Journal of Safety Research, 2025; 92: 98–108. doi:10.1016/j.jsr.2024.09.012.

[35] Ahmadinejad, M., Zargari, S., Jalalkamali, R. Are deceleration numbers a suitable index for road safety? Proceedings of the Institution of Civil Engineers - Transport, 2017; 171: 247–252. doi:10.1680/jtran.16.00117.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Performance-Based Seismic Assessment of RC Frames Strengthened with Steel Angle and Strap Systems

Hananeh Dashtaki [a], Javad Shayanfar [b*], Morteza Jamshidi [a]

[a] *Department of Civil Engineering, Cha. C., Islamic Azad University, Chalous, Iran*

[b] *Department of Civil Engineering, University of Minho, Azurém, Guimarães 4800-058, Portugal*

A R T I C L E I N F O

Article history:

Received 25 February 2026

Revised 10 April 2026

Accepted 16 June 2026

Available online 01 April 2027

Keywords:

Reinforced concrete frames

Seismic retrofit

Steel angles and straps

Nonlinear analysis

A B S T R A C T

The seismic vulnerability of reinforced concrete (RC) structures designed according to outdated standards lacking seismic provisions has been widely documented through experimental investigations and post-earthquake observations. Consequently, reliable assessment of the seismic performance of such structures and the development of effective strengthening strategies remain critical objectives in earthquake engineering. This study investigates the seismic behavior of RC frames strengthened using steel angle and strap systems through analytical and numerical approaches. A novel sectional moment-curvature model is developed, incorporating confinement effects from stirrups and steel straps as well as local buckling of compressive steel angles. Sectional curvature is converted to member rotation to derive plastic hinge properties for nonlinear structural analysis. The proposed model is validated against experimental data from existing studies, demonstrating strong agreement in predicting flexural and shear behavior. Subsequently, the seismic performance of strengthened RC members and structural systems is evaluated through nonlinear pushover analysis, with emphasis on performance levels, plastic hinge development, and failure mechanisms. Results indicate that the steel angle and strap system substantially enhances load-carrying capacity and ductility, leading to improved seismic performance and reduced structural vulnerability. These findings confirm the effectiveness of the strengthening strategy and provide an analytical framework for performance-based seismic assessment and retrofit design of existing RC structures.

How to cite this article: Dashtaki, H., Shayanfar, J., Jamshidi, M. Performance-Based Seismic Assessment of RC Frames Strengthened with Steel Angle and Strap Systems. Civil Engineering and Applied Solutions. 2027; 3(2): 49–68. doi:10.22080/ceas.2026.31395.1082.

Nomenclature

Parameter	Description	Parameter	Description
A_g	Total area of the section	t_1	Thickness of steel angles
A_{con}	Areas of confined regions	t_2	Thickness of steel straps
A_{unc}	Areas of unconfined regions	v	Poisson's ratio
A_{s1}	Area of steel bars at tension	V_{ang}	Shear strength due to steel angles
A_{s2}	Area of steel bars at compression	V_{bat}	Shear strength due to steel straps
b	Smaller width of the section	v_c	Shear stress capacity in compression
c	Neutral axis	v_t	Shear stress capacity in tension
C_{con}	Concrete compressive force	v_n	Shear capacity due to concrete
d_b	Diameter of steel bars	V_n	Total shear strength

* Corresponding author.

E-mail addresses: arch3d.ir@gmail.com (J. Shayanfar).

https://doi.org/10.22080/ceas.2026.31395.1082

ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

E_c	Unconfined concrete elastic modulus	x	Axial strain ratio
E_{sec}	Modulus of unconfined concrete at the peak	α_1	Empirical factor for confinement-induced gain
E_s	Steel elastic modulus	α_2	Empirical factor for non-circularity effects
f_c	Concrete axial stress	φ	Curvature
f_c'	Peak strength of unconfined concrete	φ_y	Yield curvature
$f_{cc,sj}'$	Peak strength of confined concrete	φ_u	Ultimate curvature
f_j	Confining stress of steel straps	θ_i	Rotation
f_l	Minimum confinement pressure	θ_{bb}	Ultimate rotation capacity
F_l	Maximum confinement pressure	θ_{p_bb}	Plastic rotation threshold
f_s	Confining stress of steel stirrups	ε_0	Axial strain at the peak stage
F_{sa}	Axial stress at steel angles	ε_c	Axial strain
f_t	Concrete tensile strength	$\varepsilon_{cc,sj}$	Axial strain at the peak stage
f_{ya}	Yield stress of the steel angles	ε_l	Lateral/hoop strain
f_{ys}	Yield stress of steel stirrups	ε_s	Tensile strain in steel bars
h	Larger width of the section	ε_{sa}	Axial strain at steel angles
K_e	Confinement effectiveness factor	ε_{sb}	Yield strain of steel straps
$K_{e,j}$	Confinement effectiveness factor	ε_{sy}	Yield strain of steel stirrups
L_1	Width of steel angles	ε_{ya}	Yield strain of steel angles
L_2	Width of steel straps	ρ_{jx}	Strap reinforcement ratio in x direction
L_{eff}	Effective column length	ρ_{jy}	Strap reinforcement ratio in y direction
L_p	Plastic hinge length	ρ_{sx}	Stirrup reinforcement ratio in x direction
M_f	Bending capacity	ρ_{sy}	Stirrup reinforcement ratio in y direction
M_n	Equivalent bending capacity	ρ_t	Total longitudinal reinforcement ratio
N	Applied axial load	σ	Normal stress
N_{max}	Maximum axial capacity	σ_c	Principal compressive stress
n_c	Local buckling-induced reduction	σ_t	Principal tensile stress
s'	Steel stirrup spacing	t_1	Thickness of steel angles
s_b	Steel strap spacing	t_2	Thickness of steel straps

1. Introduction

Evidence from past earthquakes, as well as numerous experimental and analytical studies [1-3] , clearly demonstrates the seismic vulnerability of reinforced concrete (RC) structures that were inadequately designed according to outdated codes. It is noteworthy that a significant portion of existing RC structures worldwide were constructed prior to the development and enforcement of modern seismic design codes; consequently, the seismic provisions embedded in contemporary codes were not considered in their original design [4, 5]. Therefore, in order to enhance the seismic performance of existing structures and align them with current performance-based requirements, it is first essential to identify structural weaknesses and accurately assess seismic behavior, and subsequently, if necessary, to implement an appropriate strengthening strategy.

In recent years, one of the retrofitting methods that has gained considerable attention from researchers and practicing engineers is the use of steel angle and strap (or batten) systems. This system not only significantly improves the seismic performance of members and the overall structure, but also offers advantages such as cost-effectiveness, ease of material availability, and relatively rapid and straightforward implementation [6-12]. Experimental studies by Salman and Al-Sherrawi [6] on RC columns strengthened with steel angle and strap systems under lateral loading demonstrated that this retrofitting method enhances the flexural behavior of members and substantially increases energy dissipation capacity. Montuori and Piluso [13] and Nagaprasad et al. [14] further evaluated the performance of the system through tests on eccentrically loaded strengthened columns. Their results indicated that the steel angle and strap system increases flexural strength through the participation of steel angles in resisting tensile and compressive forces, while simultaneously improving member ductility by enhancing confinement pressure provided by the straps. Campione et al. [15] focused on analytical modeling and proposed a model to estimate the flexural and shear capacity of RC columns strengthened with steel angle and strap systems. In this model, the compatibility of strains between the angles and concrete was neglected to account for shear stresses, resulting in a more conservative and safer prediction of structural behavior. Moreover, Campione [16], through a numerical study, examined the effect of steel angles and straps on the strength and ductility of RC columns, demonstrating that reducing the spacing between straps not only increases strength within the axial load-moment interaction domain but also significantly improves the ductility of strengthened members due to enhanced lateral confinement, as also confirmed by Shhatha et al. [9], Hu et al. [10], Campione [17], Badalamenti et al. [18].

The novelty of this study lies in the development of an integrated nonlinear analytical model that explicitly captures the coupled behavior of flexure, shear, confinement, and instability mechanisms in RC members strengthened with steel angle and strap systems.

At the material level, the model extends conventional confined concrete formulations by incorporating the combined and direction-dependent confinement effects of stirrups and external steel straps through an iterative procedure for determining effective lateral pressure. At the sectional level, it introduces a unified fiber-based moment-curvature framework in which steel angles are fully integrated into strain compatibility and force equilibrium, while their compressive behavior includes local buckling effects, enabling realistic prediction of post-peak response. Furthermore, the model explicitly couples flexural deformation with shear capacity degradation by updating shear resistance as a function of member rotation and accounting for contributions from concrete, stirrups, steel straps, and angles. The bar buckling formulation is also extended to include the influence of external confinement. Importantly, unlike existing studies that are generally limited to member-level or uncoupled formulations, this work provides a consistent framework that captures coupled flexure-shear-confinement behavior and applies it from the sectional level to full structural analysis. To the best of the authors' knowledge, such a comprehensive and unified modeling approach for steel angle and strap strengthening systems, validated at both member and structural levels, has not been previously reported.

The paper first describes the strengthening system and the underlying modeling assumptions, followed by the development of the nonlinear analytical framework, including material constitutive relationships and sectional analysis procedures. The proposed model is then validated against available experimental results, and subsequently employed to assess the seismic response of multi-story RC frames before and after strengthening. Finally, the key findings are discussed, and the main conclusions are drawn.

2. Steel angle and strap strengthening method

The strengthening system investigated in this study consists of steel angles installed at the corners of columns, which are interconnected by steel straps (or batten). This system, by establishing a combined confinement and structural participation mechanism, can effectively enhance the seismic performance of RC members.

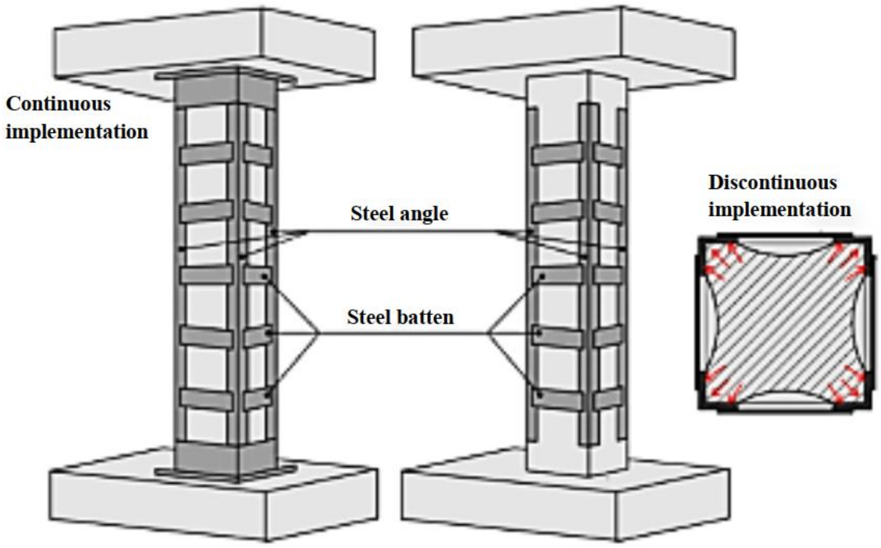


Fig. 1. Configuration of the steel angle and strap strengthening system.

The main advantages of the steel angle and strap system can be summarized in three key aspects:

1. Increased lateral confinement pressure on the concrete core, leading to a significant improvement in member ductility;
2. Enhanced shear capacity and, when properly designed, the potential to shift the failure from brittle shear to ductile flexure;
3. Provision of axial resistance (tension and compression) through the steel angles, which subsequently increases the flexural capacity of the column section.

The effectiveness of the steel angle and strap system is directly dependent on the execution and quality of component connections. When the system is implemented in a discontinuous manner (as shown in Fig. 1), the steel straps primarily serve to provide lateral confinement, and the improvement in member behavior is reflected in enhanced strength and ductility [5]. In this case, due to the lack of effective force transfer, the steel angles do not actively participate in resisting tensile and compressive forces. In contrast, with continuous implementation of the strengthening system (Fig. 1), the steel angles actively contribute to the flexural behavior of the member by resisting both tensile and compressive forces, in addition to the ductility enhancement provided by concrete confinement. Therefore, the selection of appropriate detailing and ensuring proper continuity between the steel components and the RC member play a decisive role in the effectiveness of this strengthening method.

3. Steel angle and batten strengthening method

In this section, the methods used to determine the nonlinear behavior of RC members strengthened with the steel angle and strap system are described. Based on the discrete plastic hinge approach [19], the properties of RC members in the moment-rotation relationship can be determined through sectional moment-curvature analysis and the subsequent conversion of curvature to rotation. To perform this analysis, it is essential to define the nonlinear behavior of the materials used in the member; therefore, the method for calculating material properties is presented below.

3.1. Confined concrete with stirrups and steel batten

The steel angle and strap strengthening system, in addition to the transverse reinforcement in the section, can enhance concrete behavior by increasing the confinement pressure [6, 8, 20]. The elevated confinement pressure results in improved strength and ductility of the reinforced concrete members, with detailed information of its mechanism that can be found in Shayanfar et al. [21], Shayanfar et al. [22], Shayanfar et al. [23]. In Fig. 2, the regions of an RC member section subjected to confinement pressure from the steel angle and strap strengthening system, as well as transverse reinforcement, are illustrated.

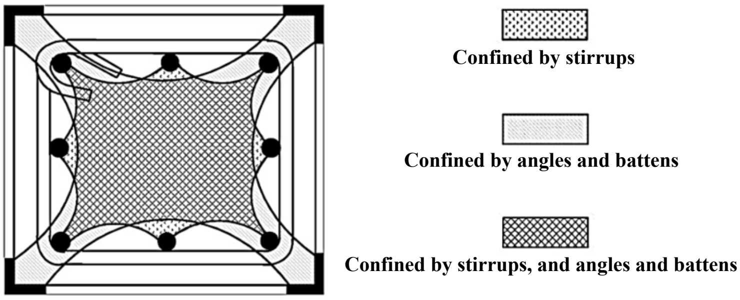


Fig. 2. Sectional regions exposed to confinement provided by angle and batten system as well as stirrups.

As can be observed in Fig.2, the cross-section of a reinforced concrete member strengthened with a steel angle and steel strip system can be divided into two regions based on the confinement pressure [24-26]: (I) Regions subjected to confinement pressure from the steel angles and steel strips; (II) Regions subjected to confinement pressure from the steel angles and steel strips in combination with the transverse reinforcement. However, in this paper, for the sake of conservatism and simplicity, the confined regions located outside the core of the section, which are confined by the aforementioned strengthening system, have been neglected. The details of the strengthening system under investigation are illustrated in Fig. 3. The parameters t_1 and t_2 denote the thicknesses of the steel angles and steel strips, respectively. The remaining parameters are shown and can be identified in Fig. 3.

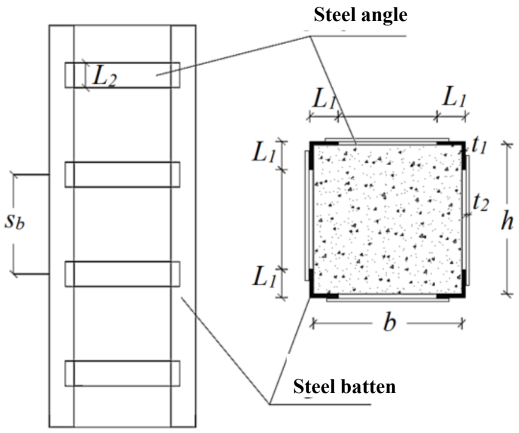


Fig. 3. Specimen of a reinforced concrete column strengthened with a steel angle and batten system.

In this study, the axial stress-strain model for concrete proposed by Mander et al. [27], with some modifications, has been adopted to account for the confinement effect by the steel angle and batten system. The formulation can be expressed as Eq. 1.

$$f_c = \frac{f'_{cc,sj} x r}{r-1+x^r} \tag{1}$$

where:

$$x = \frac{\varepsilon_c}{\varepsilon_{cc,sj}} \tag{2}$$

$$r = \frac{E_c}{E_c - E_{sec}} \tag{3}$$

$$E_{sec} = \frac{f'_{cc,sj}}{\varepsilon_{cc,sj}} \tag{4}$$

$$\varepsilon_{cc,sj} = \varepsilon_0 \left[1 + 5\left(\frac{f'_{cc,sj}}{f'_c} - 1\right)\right] \tag{5}$$

Here, E_c is the concrete elastic modulus; f'_c and ε_0 are the axial stress and strain at the peak stage of unconfined concrete. The maximum compressive strength ($f'_{cc,sj}$) of concrete confined by the steel angle and strap strengthening system, as well as transverse reinforcement, corresponds to $\varepsilon_{cc,sj}$ and can be calculated by Eq. 6.

$$f'_{cc,sj} = \alpha_1 \alpha_2 f'_c \tag{6}$$

where:

$$\alpha_1 = 2.254\sqrt{1+\frac{7.94F_l}{f'_c}} - 2\frac{F_l}{f'_c} - 1.254 \quad (7)$$

$$\alpha_2 = \left(1.4\frac{f_l}{F_l} - 0.6\left(\frac{f_l}{F_l}\right)^2 - 0.8\right)\sqrt{\frac{F_l}{f_l}} + 1 \quad (8)$$

In Eq. 1, α_2 is used to account for the effect of varying lateral confinement in the vertical and horizontal directions of the section, following the study by Wang and Restrepo [28]. In this equation, f_l and F_l represent the minimum and maximum effective lateral confinement pressures, respectively, which can be calculated by Eqs. 9 and 10.

$$f_l = min(f_{l,sx} + f_{l,jx}, f_{l,sy} + f_{l,jy}) \quad (9)$$

$$F_l = max(f_{l,sx} + f_{l,jx}, f_{l,sy} + f_{l,jy}) \quad (10)$$

where:

$$f_{l,sx} = K_e \rho_{sx} f_s \quad (11)$$

$$f_{l,sy} = K_e \rho_{sy} f_s \quad (12)$$

$$f_{l,jx} = K_{e,j} \rho_{jx} f_j \quad (13)$$

$$f_{l,jy} = K_{e,j} \rho_{jy} f_j \quad (14)$$

In the above equation, ρ_{sx} and ρ_{sy} denote the ratios of transverse reinforcement in the horizontal and vertical directions of the section, respectively. Similarly, ρ_{jx} and ρ_{jy} represent the ratios of steel straps in the horizontal and vertical directions, respectively. The parameter f_j and f_s corresponds to the lateral confining stress induced by the steel straps and transverse reinforcement. In addition, the coefficients K_e and $K_{e,j}$ represent the confinement effectiveness factors associated with the transverse reinforcement and the steel angle and strap strengthening system, respectively, whose values are determined by Eqs. 15 and 16.

$$K_e = \frac{\left(1-\sum_{i=1}^{n}\frac{{w_i}^2}{6bh}\right)\left(1-\frac{s'}{2b}\right)\left(1-\frac{s'}{2h}\right)}{1-\rho_t} \quad (15)$$

$$K_{e,j} = \frac{(1-\rho_t-\{(h-2L_1)^2+(b.-2L_1)^2\})}{3A_g(1-\rho_t)} \times \left(1-\frac{s_b-L_2}{2h}\right)\left(1-\frac{s_b-L_2}{2b}\right) \quad (16)$$

The parameter w_i denotes the clear spacing between adjacent longitudinal reinforcing bars. In addition, A_{sv} and A_g represent the total area of transverse reinforcement and the gross cross-sectional area, respectively. The ratio ρ_t corresponds to the longitudinal reinforcement ratio with respect to the gross sectional area. To determine the lateral confining pressure corresponding to the maximum compressive strength of concrete, a novel procedure is proposed, which is described in the following section.

1. Assume an initial value for the axial strain of concrete ε_a.
2. Convert the axial concrete strain to the corresponding lateral strain (parallel to the member cross-section) in the steel straps and transverse reinforcement, ε_l, assuming a Poisson's ratio equal to 0.5 [29, 30].
3. Calculate the stress developed in the steel batten.
4. Calculate the stress developed in the transverse reinforcement.
5. Determine the lateral confinement pressure induced by the steel angle and batten system using Eqs. 13 and 14.
6. Determine the lateral confinement pressure provided by the transverse reinforcement using Eqs. 11 and 12.
7. Calculate the maximum compressive strength of concrete based on Eq. 6.
8. Determine the axial strain corresponding to the maximum compressive strength of concrete using Eq. 5.
9. Compare the assumed axial strain in Step (1) with the axial strain obtained in Step (8).

$$\varepsilon_l = 0.5 \times \varepsilon_a \quad (17)$$

$$f_j = E_j \times \varepsilon_l \qquad \varepsilon_l \le \varepsilon_{yb} \quad (18)$$

$$f_j = f_{yj} \qquad \varepsilon_l > \varepsilon_{yb} \quad (19)$$

$$f_{s,j} = E_s \times \varepsilon_l \qquad \varepsilon_l \le \varepsilon_{sy} \quad (20)$$

$$f_{s,j} = f_{ys} \qquad \varepsilon_l > \varepsilon_{sy} \quad (21)$$

If the axial strain assumed in Step (1) shows adequate convergence with the value obtained in Step (8), the concrete stress-strain behavior can be evaluated using Eq. 1, based on the maximum compressive strength determined in Step (7) and the corresponding axial strain obtained in Step (8). Otherwise, a new value for the axial strain is assumed in Step (1), and the calculation procedure is repeated until convergence is achieved.

3.2. Local buckling of compressed steel angles

Based on previous studies [13, 14, 18, 20, 31], the effects of local buckling must be considered in the moment-curvature analysis of steel angles under compression used in strengthening systems. In the present study, the analytical model proposed by Campione [17] is adopted, which is based on the findings of Badalamenti et al. [18] and accounts for the influence of local buckling on the nonlinear behavior of compressed steel angles. According to this model, the stress-strain relationship of the compressed steel angles is defined by Eqs. 22 and 23.

$$min\left(E_s\varepsilon_s, n_c\, f_{ya}\right) \qquad \varepsilon_s \leq \varepsilon_{ya} \tag{22}$$

$$min\left(f_{ya}, n_c\, f_{ya}\right) \qquad \varepsilon_s > \varepsilon_{ya} \tag{23}$$

where:

$$n_c = \frac{1}{\sqrt{2\varepsilon_s - \varepsilon_s^2}}\left\{0.5\frac{L_1 t_1}{s_b t_2}\left(1 - 0.5\frac{R}{f_{ya}}\right) - C\right\} \tag{24}$$

$$R = min\left(\varepsilon_s E_s, f_{ya}\right) \tag{25}$$

$$C = 0.47\frac{E_s\varepsilon_s}{t_2 b f_{ya}}\left(\frac{e^{\left[-1.5\left(\frac{s_b}{b}\right)\right]}}{\left[\frac{L_1}{(s_b t_1)}\right]+\left[\frac{(0.5b-L_1)}{(L_2 t_2)}\right]}\right) \tag{26}$$

where f_{ya} denotes the yield stress of the compressed steel angles corresponding to the compressive strain ε_{ya}. By modifying the stress-strain relationship of the compressed angles, when the local buckling mechanism initiates ($0 < n_c < 1$), a decrease in n_c leads to a reduction in the longitudinal stress and, consequently, the compressive force developed in the steel angles. This reduction in longitudinal stress directly affects the strain compatibility and force equilibrium relationships.

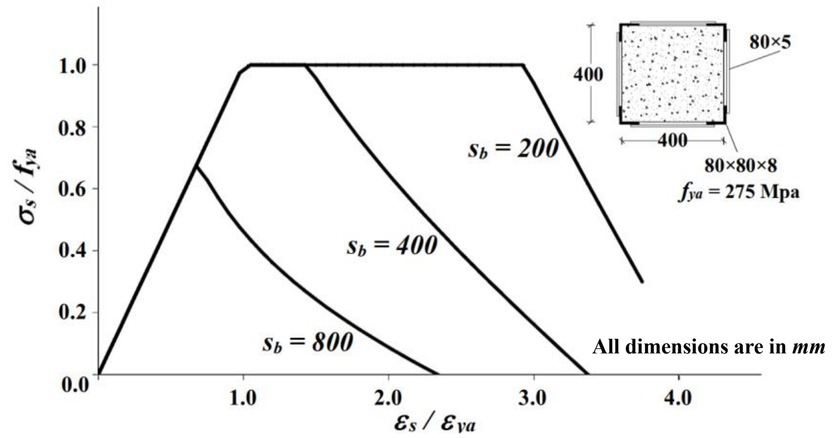


Fig. 4. Longitudinal stress-longitudinal strain relationship of steel angles under compression.

3.3. Moment-curvature analysis

In this section, the theoretical framework and analytical procedure for performing moment-curvature analysis of an RC member strengthened with a steel angle and steel strip system are presented. Moment-curvature analysis is a sectional nonlinear analysis method used to characterize the flexural response of structural members. It establishes the relationship between the applied bending moment and the resulting curvature of the cross-section, thereby enabling evaluation of stiffness degradation, yielding progression, confinement effects, and ultimate flexural capacity.

The method is particularly suitable for composite or strengthened sections, where multiple materials with distinct constitutive laws interact. In the present study, by following Shayanfar et al. [32], Shayanfar et al. [33], Shayanfar et al. [34], a fiber-based sectional modeling approach is employed. In this technique, the cross-section is discretized into numerous small fibers, each assigned an appropriate uniaxial stress-strain relationship according to its material type (confined concrete, unconfined concrete, longitudinal reinforcement, and steel angles). Under the assumptions of strain compatibility and plane sections remaining plane, the strain distribution across the depth of the section is determined from the imposed curvature. The internal forces are then obtained by integrating the stresses over the sectional area, and equilibrium with the applied axial load is enforced to compute the corresponding bending moment [35]. Fig. 5 schematically illustrates the fiber-section model of a reinforced concrete member strengthened with the steel angle and steel strip system. In Fig. 5, F_{sa} and f_{sa} denote the longitudinal force and longitudinal stress in the steel angles at the section level, respectively. The remaining variables, including sectional geometry, strain distribution, and internal force components, are defined within Fig. 5.

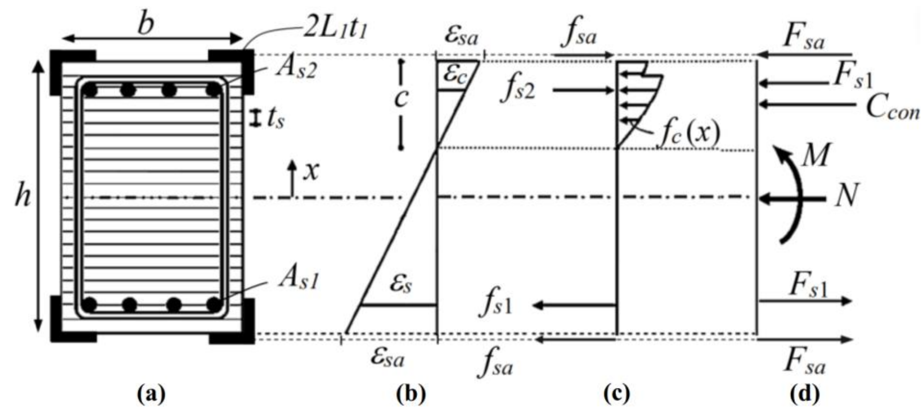


Fig. 5. Fiber analysis of a reinforced concrete member strengthened with a steel angle and steel strip system: (a) cross-section; (b) distribution of longitudinal strains; (c) distribution of longitudinal stresses; and (d) distribution of compressive and tensile forces within the section.

The moment-curvature analysis procedure employed in this study can be summarized as follows:

1. Assume a value for the concrete strain at the extreme fiber of the section.
2. Assume a value for the neutral axis position, c.
3. Calculate the total compressive forces in the concrete, the tensile and compressive forces in the longitudinal reinforcement, and the tensile and compressive forces in the steel angles (Eqs. 27 to 29).

$$C_{con} = \sum f_c(x) b\, t_s \tag{27}$$

$$F_{si} = \sum f_{si} A_{si} \tag{28}$$

$$F_{sa} = \sum 2 f_{sa} L_1 t_1 \tag{29}$$

4. Check the force equilibrium in the section ($C_c + F_{si} + F_{sa} - N \ = 0$). If the sum $C_c + F_{si} + F_{sa}$ is close to the applied axial force, the assumed neutral axis position is correct. Otherwise, a new value for the neutral axis position must be assumed.
5. Calculate the bending moment (Eq. 30), in which d_i and d_{ai} are the distances of the steel bar and steel angles to the top of the section.

$$M_f = \sum f_c(x) b\, t_s x_i + \sum F_{si}\left(\frac{h}{2} - d_i\right) + \sum F_{sa}\left(\frac{h}{2} - d_{ai}\right) \tag{30}$$

6. Curvature Calculation (Eq. 31)

$$\varphi_i = \frac{\varepsilon_c}{c} \tag{31}$$

7. Assume a new value for the concrete strain at the extreme fiber of the section, and repeat the process until the strain in a fiber of the confined region reaches its maximum value.
8. Determine the moment-curvature relationship.

Using the above procedure, the moment-curvature relationship of the RC member strengthened with the steel angle and strap system can be obtained. By converting this relationship into a moment-rotation relationship based on the plastic hinge approach (Fig. 6), the properties of plastic hinges can be derived. The rotation of the RC member corresponding to each curvature level can be calculated by Eqs. 32 and 33).

$$\theta_i = \frac{\varphi_i \times L_{eff}}{2} \qquad \varphi_i < \varphi_y \tag{32}$$

$$\theta_i = \frac{\varphi_y \times L_{eff}}{2} + (\varphi_i - \varphi_y) \times L_p \qquad \varphi_i \geq \varphi_y \tag{33}$$

where:

$$L_{eff} = L + 0.022 f_s d_b \qquad f_s \leq f_y \tag{34}$$

$$L_p = 0.08L + +0.022 f_y d_b \geq 0.044 f_y d_b \tag{35}$$

Here, φ_y and φ_u represent the yield and ultimate curvature of the transverse reinforcement, respectively. The plastic hinge length of the member, L_p, is calculated based on the relationship proposed by Paulay and Priestly [36]. L is the distance between the maximum and zero moment, produced in the member.

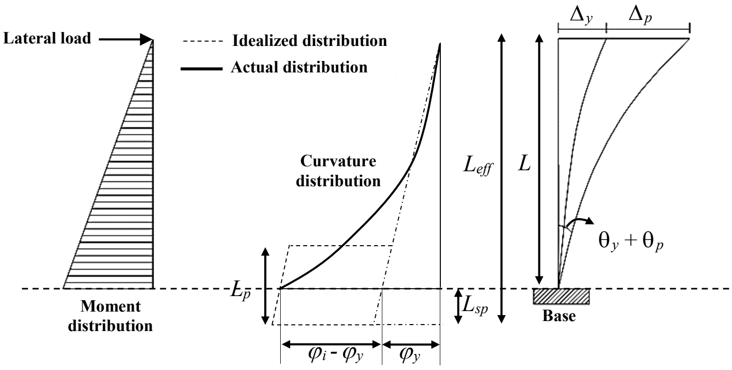


Fig. 6. Plastic hinge approach.

3.4. Shear behavior of RC member with steel angle and batten strengthening method

Insufficient shear capacity in RC members such as beams and columns can lead to nonlinear behavior and a brittle shear failure mechanism [37-39]. The formation of diagonal shear cracks in RC columns and the subsequent reduction in axial load-carrying capacity may cause axial failure of these members, thereby increasing the vulnerability of the structure to seismic actions. Conversely, increased flexural deformations and rotations of the member, accompanied by the development of diagonal tensile cracks, can reduce shear capacity. Under seismic loading, as flexural deformations and rotations increase, after yielding of the longitudinal tensile reinforcement, the width of flexural cracks grows and tensile crushing and cracking of concrete occur in compressive zones of the section [1, 40-42]. Moreover, with the increase in flexural deformations, the height of the neutral axis decreases, which can further reduce the shear capacity of the RC member. Therefore, simulation of the shear mechanism in nonlinear analysis of structural members as a function of flexural deformations is necessary and constitutes one of the objectives of this paper. For a member strengthened with a steel angle and steel strip system, in addition to stirrups and concrete, the steel strips and angles also contribute to the shear capacity of the member. Accordingly, in this study, to determine the contribution of the angles and strips to shear resistance, the model proposed by Campione et al. [15] has been employed, which is expressed in the following equations.

$$V_{angles} = c_1 \frac{{L_1}^2 \times t_1}{2 \times s_b} f_{ya} \tag{36}$$

$$V_{batten} = c_2 \frac{2L_2 \times t_2 \times d}{s_b} f_{yb} \tag{37}$$

where:

$$c_1 = \frac{1}{1 + \left[\left(\frac{L_2}{L_1}\right)^3 \times \frac{t_2 s_b}{t_1(b - 2L_1)}\right]} \tag{38}$$

$$c_2 = 1 - \frac{N}{N_{max}} \tag{39}$$

In the above equations, f_{yb} represent the yield stresses of the steel angles and steel strips, respectively. N_{max} denotes the maximum axial capacity of the member strengthened with the steel angle and steel strip system, which in this study is calculated by Eq. 40.

$$N_{max} = f'_{cc}\, A_{con} + 0.85 f'_c\, A_{unc} + f_y A_{st} + 8 n_a f_{ya} L_1 t_1 \tag{40}$$

In which, based on studies by Campione [15-17]:

$$0 \le n_a = \frac{1}{2 f_{ya} L_1 t_1} \sqrt{4 L_1 f_{ya} \left({L_1}^2 t_1 f_{ya} - 2\sqrt{2} \frac{L_2 t_2 s_b f_{yb}}{3 L_1} e^{\left(-1.5 \frac{s_b}{b}\right)}\right)} \le 1 \tag{41}$$

In the above relations, A_{con} and A_{unc} denote the areas of confined and unconfined regions, respectively. The coefficient n_a accounts for the effects of angle buckling on the compressive force contributed by the angles. To determine the concrete contribution to shear capacity, various models have been proposed in the technical literature [20, 43-47]. However, since significant regions of the member cross-section are affected by confinement pressure from the steel strips and their interaction with stirrups as functions of eccentricity, the aforementioned models do not appear fully efficient.

In the present study, the shear model proposed by Shayanfar and Akbarzadeh Bengar [19] which determines the concrete shear capacity based on fiber analysis, has been extended to compute the concrete contribution to shear capacity in an RC member strengthened with a steel angle and steel strip system. By adopting Mohr's circle theory, the principal compressive (σ_c) and tensile (σ_t) stresses in an element under normal (axial) stress σ and shear stress ν of the RC member cross-section (Fig. 7) can be determined by Eqs. 42 and 43.

$$\sigma_c = \frac{\sigma}{2} + \sqrt{\frac{\sigma^2}{4} + \nu^2} \le f_{cc} \tag{42}$$

$$\sigma_t = \frac{\sigma}{2} - \sqrt{\frac{\sigma^2}{4} + \nu^2} \le f_t \tag{43}$$

Here, f_{cc} and f_t represent the compressive and tensile failure surfaces, respectively, which have been defined based on the model proposed by Park et al. [48] and the Rankine failure criterion. The tensile strength f_t has been taken equal to $f_t = 0.292\sqrt{f'_c}$ according to the study conducted by Shayanfar and Akbarzadeh Bengar [19].

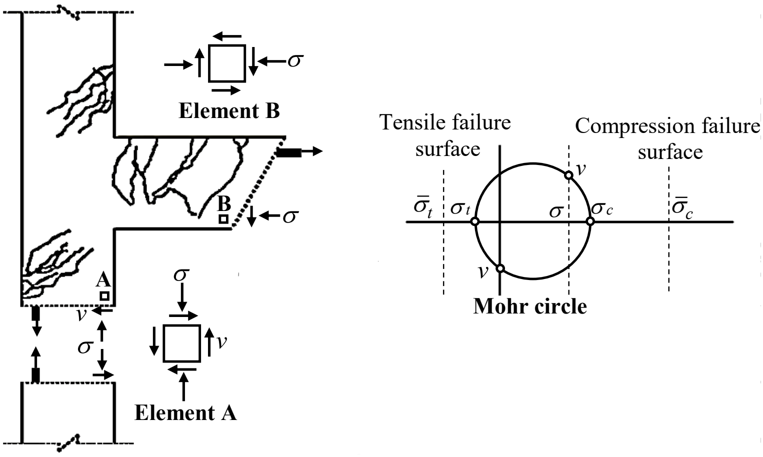


Fig. 7. Rankine failure criteria adopted for concrete.

By rewriting Eqs. 42 and 43, the shear stress capacity of an element located in the compression zone of the section can be expressed as shown in Eqs. 44 to 46.

$$v_c(y_i) = \sqrt{\sigma_c[\sigma_c - \sigma(y_i)]} \qquad \text{Contributed by compression} \tag{44}$$

$$v_t(y_i) = \sqrt{\sigma_t[\sigma_t + \sigma(y_i)]} \qquad \text{Contributed by tension} \tag{45}$$

$$v_n = \sum v(y_i) = \sum min\left(\sqrt{\sigma_t[\sigma_t + \sigma(y_i)]}, \sqrt{\sigma_c[\sigma_c - \sigma(y_i)]}\right) \tag{46}$$

where y_i is the distance of each element within the compression zone from the neutral axis c, which is determined through fiber section analysis. The shear stress capacity contribution of concrete in each strip of the compression zone, $v(y_i)$, is taken as the smaller value between $v_{c,i}(y_i)$ and $v_{t,i}(y_i)$. Accordingly, for each assumed neutral axis depth, the total shear strength of the RC member (considering the contributions of concrete, steel angles, steel batten, and transverse reinforcement) can be determined as shown in Eq. 47.

$$V_n = \sum v(y_i) \times b \times t_y\Big]_{Concrete} + c_1 \frac{L_1^2 \times t_1}{2 \times s_b} f_{ya}\Big]_{Angle} + c_2 \frac{2L_2 \times t_2 \times d}{s_b} f_{yb}\Big]_{Batten} + \frac{A_{sv} f_{sy} d_s}{s}\Big]_{Stirrup} \tag{47}$$

where t_y denotes the thickness of each fiber element in the vertical direction.

To define the plastic hinge properties of RC members while accounting for shear mechanism effects, the shear contribution is incorporated into the flexural moment-rotation relationship by multiplying the shear capacity by the shear span, L (Eq. 48).

$$M_n = V_n \times L \tag{48}$$

Using this approach, the influence of shear can be introduced into the member's flexural moment-rotation response. Fig. 8 illustrates the interaction between nonlinear flexural and shear behaviors in RC members, both strengthened and unstrengthened. As shown in the figure, the shear capacity of the member is a function of its rotation; specifically, with increasing member rotation, the shear capacity decreases. Furthermore, three distinct nonlinear behavioral modes (flexural failure, flexural-shear failure, and shear failure) can be determined by comparing the shear and flexural capacities of the member. Therefore, through the aforementioned procedure, the effects of the shear mechanism in RC members can be simulated by means of an equivalent flexural plastic hinge or a rotational spring model.

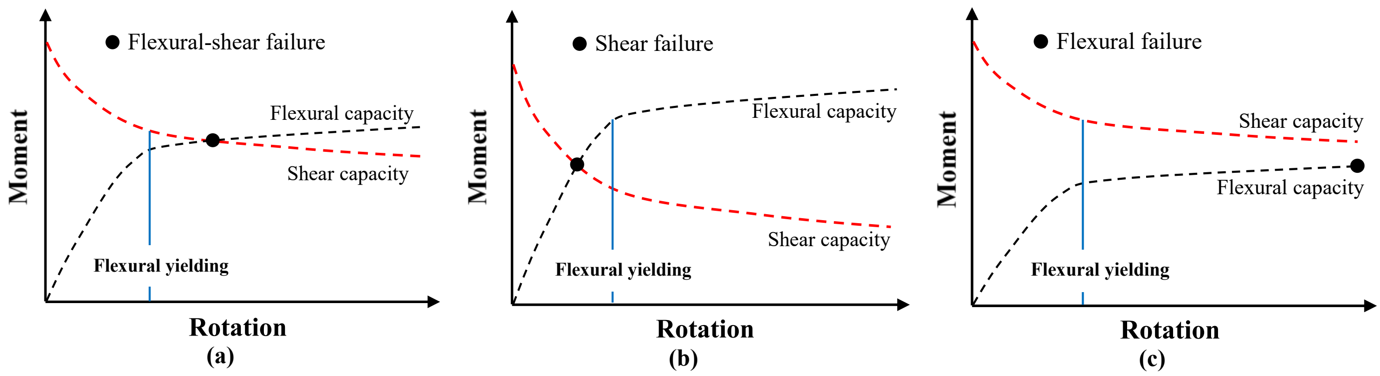


Fig. 8. Interaction between flexural and shear behaviors: (a) shear failure; (b) flexural-shear failure; and (c) flexural failure.

3.5. Longitudinal reinforcement buckling

In this section, a procedure is presented to simulate the effects of longitudinal reinforcement buckling on the seismic response of an RC member strengthened with a steel angle and steel strip system. In this paper, the model proposed by Berry and Eberhard [49] has been adopted to simulate longitudinal bar buckling. According to this model, the maximum rotation capacity of an RC member is governed by the onset of longitudinal reinforcement buckling. The steel angle and strip strengthening system, similar to transverse reinforcement (stirrups), can act as a lateral restraint for the longitudinal bars, thereby delaying the buckling mechanism [50]. In this study, the Berry and Eberhard [49] buckling model has been extended to account for the influence of steel angles and strips on the buckling phenomenon. Accordingly, the plastic rotation at the onset of longitudinal bar buckling can be determined by Eq. 49.

$$\theta_{p\_bb} = C_0\left(1 + C_1\rho_{eff,f}\right)\left(1 + C_2\frac{N}{A_g f'_c}\right)^{-1}\left(1 + C_3\frac{L}{h} + C_4\frac{f_y d_b}{h}\right) \tag{49}$$

where:

$$\rho_{eff,j} = \frac{f_{yv}\rho_s + 0.85 f_{yb}\rho_{s,b}}{f_c} \tag{50}$$

In these expressions, C_0, C_1, C_2, C_3, and C_4 are constants proposed by Berry and Eberhard [49], equal to 0.019, 1.65, 1.797, 0.012, and 0.072, respectively. The parameters ρ_s and $\rho_{s,b}$ denote the volumetric ratios of stirrups and steel strips, respectively. The coefficient 0.85 in Eq. 50 reflects the relatively lower effectiveness of the external steel angle-strip strengthening system compared with internal stirrups in restraining longitudinal bar buckling, due to its installation outside the section core. Finally, the ultimate rotation of the strengthened member corresponding to longitudinal reinforcement buckling can be calculated by Eq. 51.

$$\theta_{bb} = \frac{\varphi_y L}{2} + \theta_{p\_bb} \tag{51}$$

Therefore, the ultimate rotation capacity of the RC member can be controlled by the aforementioned buckling criterion.

3.6. Ultimate conditions of nonlinear analysis

In this section, the criteria considered in this paper for defining the ultimate conditions of the nonlinear analysis of an RC member are described. To perform the moment-curvature analysis in accordance with the proposed model, it is necessary to define termination criteria at which the nonlinear analysis is stopped. In this paper, the stopping criteria are defined as follows:

1. When the longitudinal strain of concrete at the extreme fiber of the confined compression zone reaches the ultimate longitudinal strain of confined concrete, the moment-curvature analysis is terminated.
2. When the longitudinal strain in the tensile longitudinal reinforcement reaches the maximum steel strain, the moment-curvature analysis is terminated.
3. When the longitudinal strain in the tensile steel angles reaches the maximum steel strain, the moment-curvature analysis is terminated.
4. In the case of shear failure, based on the studies by Williams and Sexsmith [51] and Park et al. [48], the maximum member rotation corresponds to 80% of the maximum member strength at the intersection point of the shear and flexural capacity curves.
5. If a reduction greater than 20% occurs in the flexural capacity of the member, the moment-curvature analysis is terminated.
6. When the member rotation exceeds the longitudinal reinforcement buckling limit, the moment-curvature analysis is terminated.

Thus, by employing the above criteria, the ultimate conditions for terminating the moment-curvature analysis can be defined.

4. Validation of the proposed model for nonlinear behavior of strengthened columns

For validation purposes, the experimental study conducted by Li et al. [52] on RC columns strengthened with a steel angle and steel strip system under constant axial load was utilized. The tested columns had a concrete compressive strength of 32.7 MPa. The experimental program included specimens strengthened using a steel jacket composed of steel angles at the corners connected by transverse steel strips. The applied axial loads ranged from 180 kN to 420 kN, corresponding to axial load ratios between 0.13 and 0.29. Two types of steel angles were used in the strengthening scheme: Angle 1 (40 × 40 × 4 mm, yield strength 350 MPa) and Angle 2 (30×30×3 mm, yield strength 338.5 MPa). Angle 2 was used only in specimen B223, while the remaining strengthened specimens utilized Angle 1. The steel strips had a cross-section of 30 × 3 mm with a yield strength of 533.3 MPa. The longitudinal reinforcement consisted of 14-mm diameter bars with a yield strength of 384.8 MPa, while the transverse reinforcement consisted of 8-mm diameter stirrups with a yield strength of 326.9 MPa. All tested specimens exhibited flexural failure modes. Further details regarding specimen geometry and test configuration are provided in Li et al. [52]. To validate the proposed nonlinear model, the predicted moment-axial force interaction diagrams were compared with the experimental results reported by Li et al. [52] in Fig. 9. The comparison indicates good agreement between analytical predictions and experimental data for different axial load levels. This consistency confirms the reliability of the proposed modeling approach and supports the assumptions adopted in the nonlinear analysis of RC members strengthened with the steel angle and strip system.

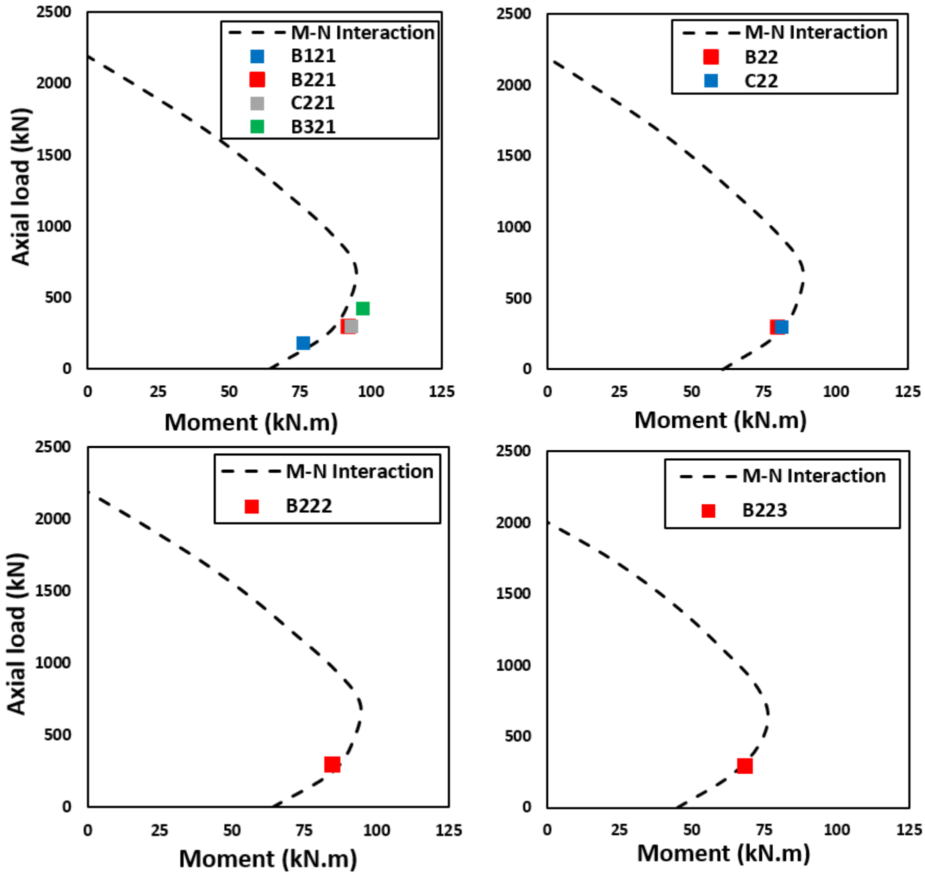


Fig. 9. Results of the nonlinear model for the column specimens tested by Li et al. [52].

Nagaprasad et al. [14] conducted experimental tests on steel-jacketed RC columns subjected to constant axial load and lateral loading. The specimens were strengthened using a steel angle and strip configuration, and all exhibited flexural failure. The concrete compressive strength was reported as 37.7 MPa. The experimental program included measurements of material properties for longitudinal reinforcement, transverse reinforcement, and the steel strengthening components, which were used to characterize the mechanical behavior of the specimens. Further experimental details are available in Nagaprasad et al. [14]. To validate the proposed nonlinear model, analytical predictions were compared with the experimental results.

As shown in Fig. 10, the moment-axial force interaction diagrams generated by the model showed satisfactory agreement with the test data across different loading conditions. This consistency confirms the capability of the model to capture the nonlinear response of strengthened RC columns and supports the assumptions adopted in the analytical formulation. The validation results demonstrate the reliability of the proposed modeling approach for predicting the behavior of members strengthened with steel angle and strip systems under combined axial and lateral loading.

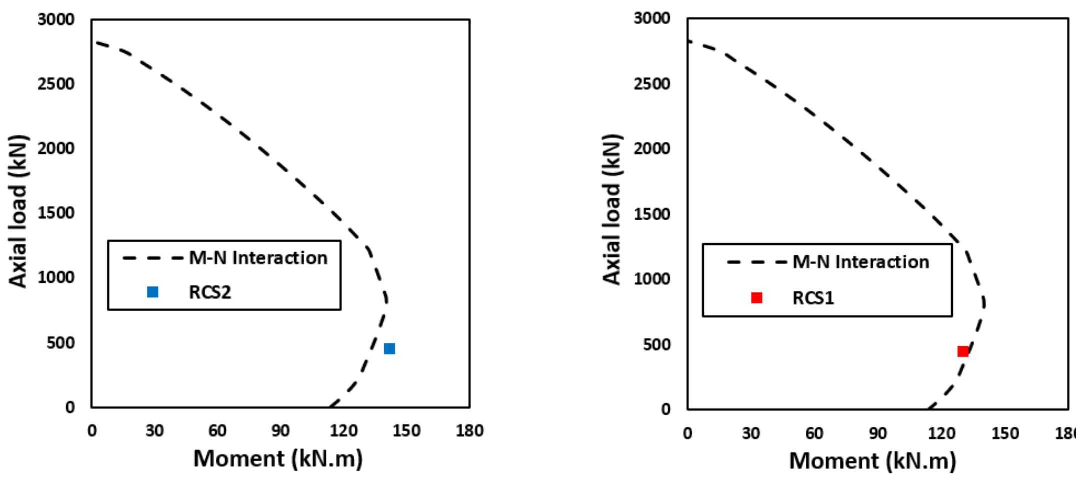


Fig. 10. Comparison between nonlinear analysis results and experimental data reported by Nagaprasad et al. [14].

Germano et al. [53] conducted experimental tests on RC columns subjected to constant axial load and lateral displacement. The specimens differed in transverse reinforcement detailing (types A and C) to investigate the influence of stirrup configuration on structural response. The columns had a height of 1800 mm and a shear span of 1565 mm, and all specimens exhibited flexural failure. Experimental load-displacement responses were reported and used for model validation. Based on Fig. 11, the comparison between analytical predictions and test results demonstrates satisfactory agreement in terms of lateral stiffness, peak resistance, and displacement capacity. The nonlinear model successfully captures the overall load-displacement behavior and seismic response characteristics of the tested columns. Although minor deviations were observed in the prediction of initial stiffness for one specimen, the model provides reliable estimations of ultimate strength and ductility. These results confirm the applicability of the proposed analytical approach for assessing the nonlinear behavior of RC columns under lateral loading conditions.

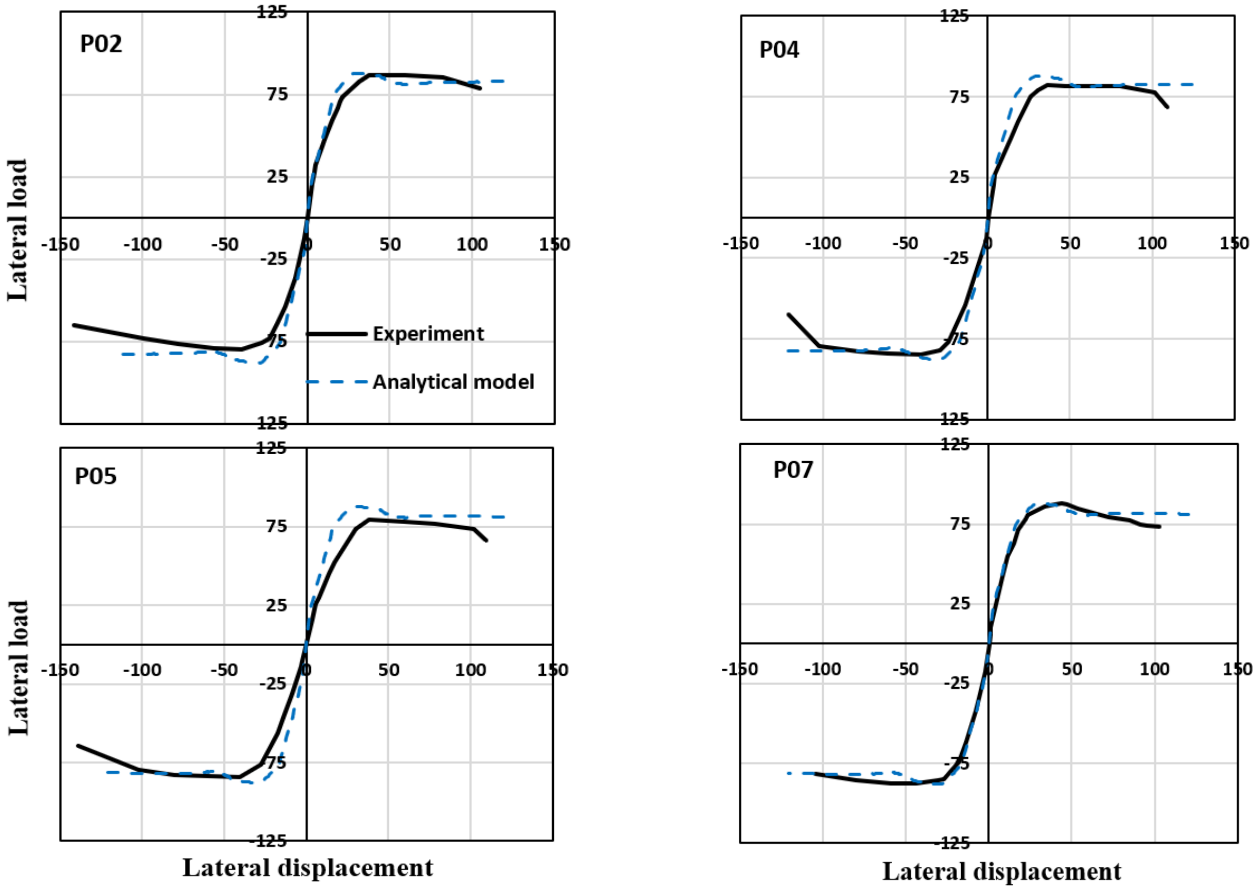


Fig. 11. Load-displacement results from nonlinear analysis compared with experimental data for column specimens tested by Germano et al. [53].

A single-span, two-storey reinforced concrete frame with a fixed base was experimentally studied by Duong et al. [54]. Lateral displacement was imposed at the second-storey beam to evaluate seismic behavior. The concrete compressive strength was measured as 42.9 MPa, and further experimental details are available in the original publication. For validation of the analytical procedure, nonlinear simulations were performed using two modeling approaches: (I) a model in which shear effects were neglected, and member response was assumed to be governed solely by flexure, and (II) a model incorporating shear effects according to the procedure adopted in this study. Based on Fig. 12, the comparison of numerical and experimental load-displacement responses indicates satisfactory agreement when shear effects are included, with the model accurately capturing stiffness degradation and ultimate displacement capacity. Conversely, exclusion of shear effects resulted in significant discrepancies in the predicted structural response. Both strength and ductility were inadequately represented, and the experimentally observed failure mechanisms – shear failure in the beams and flexural cracking in the columns – could not be reproduced with sufficient accuracy. These results demonstrate that consideration of shear effects is essential for realistic simulation of nonlinear behavior and failure modes in RC frames.

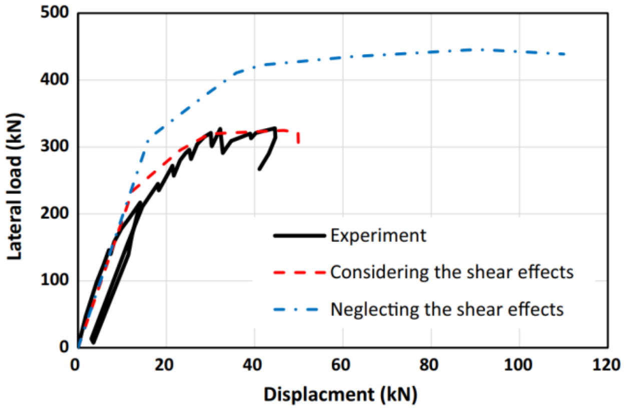


Fig. 12. Comparison of force-displacement responses of the frame: experimental results (Duong et al. [54]) and numerical analysis.

5. Evaluation of the steel angle and strip strengthening system at the structural level

In this section, the influence of the steel angle and strip strengthening system on the seismic performance of reinforced concrete structures is investigated. Three RC frame structures with four, six, and eight stories and three spans were designed, each with a story height of 3 m and a span width of 5 m. The configurations of the four-, six-, and eight-story frames (denoted as frames A, B, and C, respectively) are illustrated schematically in Fig. 13. The structural design was carried out in accordance with the Iranian seismic code, adopting standard equivalent static procedures for seismic load determination, with a lateral force distribution consistent with an inverted triangular pattern along the height of the structures. In the design process, the total dead load and a representative portion of the live load were considered in combination with the seismic loads, and member stiffness properties were defined in line with standard reinforced concrete design provisions (e.g., ACI-based practice). A notable characteristic of these structures is the relatively large spacing between transverse reinforcement (stirrups), which results in reduced confinement of the concrete core. The dimensions and reinforcement details for the beam and column sections (sections A-F) are summarized in Table 1. The notations $A_{s,top}$ and A_s, bot represent the areas of reinforcement at the top and bottom of the beam sections, respectively. The longitudinal reinforcement arrangement is specified as 16 bars of 25 mm diameter (denoted as n16φ25). The concrete compressive strength and the yield and ultimate strengths of the longitudinal and transverse reinforcement were taken as 25 MPa, 420 MPa, and 650 MPa, respectively. The diameter of the stirrups was 10 mm, consistent with the reinforcement details provided in Table 1. In all stories, three layers of stirrups were provided in the beam-column joint regions to enhance confinement and mitigate diagonal cracking within the joint core. Based on previous studies [3, 55], this reinforcement detailing is sufficient to improve shear resistance and reduce joint damage, thereby justifying the assumption of rigid joint behavior in the analysis. The seismic design of the frames was based on a design ground acceleration of 0.3 g and soil type III, corresponding to a high seismic hazard zone. The base shear was determined in accordance with the seismic design provisions. Dead and live loads were considered as 30 kN/m and 10 kN/m, respectively, in the structural design. This structural configuration enables evaluation of the effectiveness of the steel angle and strip strengthening system in enhancing seismic performance and mitigating damage under lateral loading conditions.

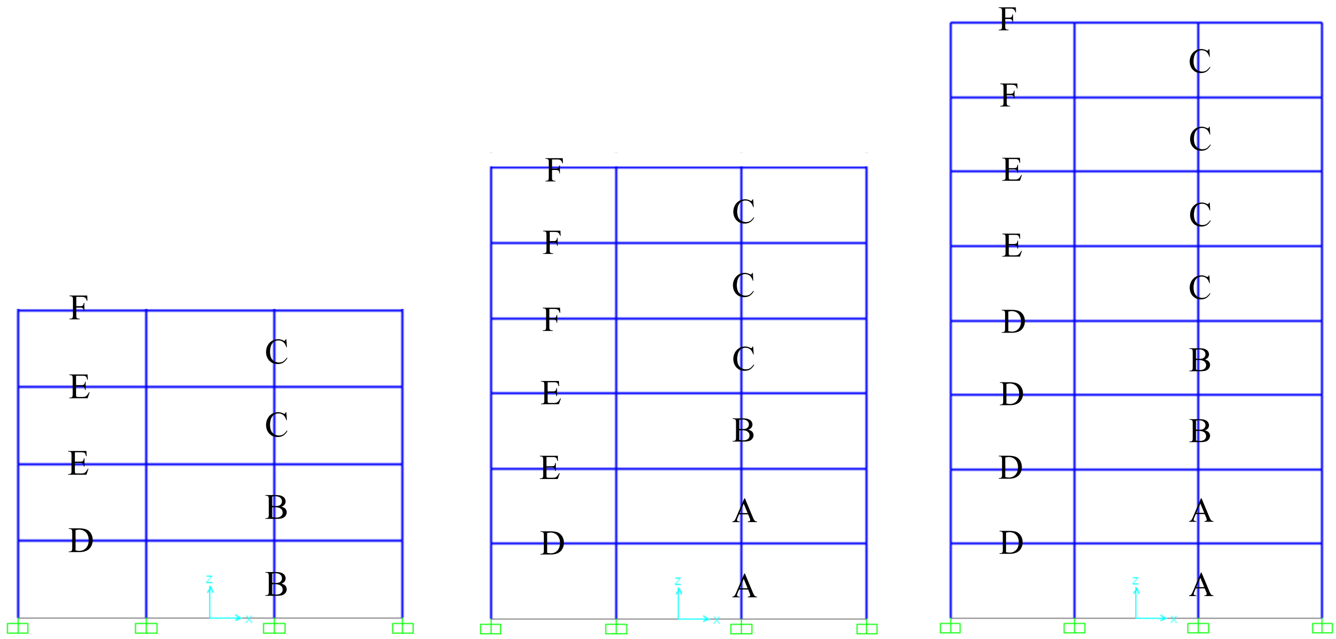


Fig. 13. Geometric configuration of the studied structures: (a) four-story frame (frame A); (b) six-story frame (frame B); and (c) eight-story frame (frame C).

In the nonlinear analysis performed in SAP2000 Version 20.2.0 [56], the plastic hinges assigned to beams and columns were defined as user-defined hinges based on the proposed analytical model. Specifically, M3 hinges were assigned to beam elements and P-M3 interaction hinges were assigned to column elements. Although these hinges are classified as flexural hinges, their defining properties were derived from the proposed model, which inherently incorporates the influence of shear behavior, confinement effects, and material nonlinearities in the moment-curvature response used to define hinge properties. The hinges were located at a distance equal to 50% of the plastic hinge length from the element ends, which is consistent with common modeling practice to represent the development of inelastic rotations away from the critical section. The hinge behavior, including yield and post-yield characteristics, was directly informed by the analytical formulation presented in the study, ensuring consistency between the sectional modeling and the global structural response. In the studied frames, the strengthening strategy is focused on improving the seismic performance of the structure and enhancing nonlinear behavior at the locations of plastic hinge formation. Tables 2 to 4 present the details of the steel angles and steel strips designed for the four-, six-, and eight-story frames.

Fig. 14 illustrates an example of the steel angle and strip configuration in frame F, first story, at an interior beam-column joint. It should be noted that in all members, the spacing of steel strips within the plastic hinge region was designed as 200 mm, while outside this region it was increased to 350 mm. The strengthening system was designed as a continuous configuration in all frames, allowing the steel angles to contribute to both tensile and compressive load transfer. This continuity enhances the flexural capacity of the members and improves overall seismic performance.

Table 1. Dimensions and reinforcement details of sections (A to F).

Section	Type	b (mm)	h (mm)	A_{st}	$A_{s,top}$	$A_{s,bot}$	Stirrup
A-A	Column	600	600	n16φ25	–	–	d10@450
B-B	Column	–	–	n16φ18	–	–	–
C-C	Column	500	500	n16φ16	–	–	–
D-D	Beam	–	–	–	n4φ25	n6φ25	d10@140
E-E	Beam	–	–	–	n4φ22	n6φ22	d10@175
F-F	Beam	–	–	–	n3φ18	n6φ18	d10@250

Table 2. Steel angle and strip details for frame D.

Story	External angle (beam)	External angle (column)	External strip (beam)	External strip (column)	Internal angle (beam)	Internal angle (column)	Internal strip (beam)	Internal strip (column)
1	4 × 40 × 40	6 × 60 × 60	4 × 40	5 × 50	4 × 40 × 40	6 × 60 × 60	4 × 40	5 × 50
2	–	8 × 80 × 80	–	7 × 70	–	8 × 80 × 80	–	7 × 70
3	–	3 × 30 × 30	–	3 × 30	–	5 × 50 × 50	–	5 × 50
4	–	3 × 30 × 30	–	3 × 30	–	3 × 30 × 30	–	3 × 30

Table 3. Steel angle and strip details for frame E.

Story	External angle (beam)	External angle (column)	External strip (beam)	External strip (column)	Internal angle (beam)	Internal angle (column)	Internal strip (beam)	Internal strip (column)
1	–	5 × 50 × 50	–	4 × 40	–	5 × 50 × 50	–	4 × 40
2	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30	4 × 40 × 40	5 × 50 × 50	4 × 40	4 × 40
3	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30
4	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30
5	–	4 × 40 × 40	–	3 × 30	–	4 × 40 × 40	4 × 40	3 × 30
6	–	–	–	–	–	–	–	–

Table 4. Steel angle and strip details for frame F.

Story	External angle (beam)	External angle (column)	External strip (beam)	External strip (column)	Internal angle (beam)	Internal angle (column)	Internal strip (beam)	Internal strip (column)
1	4 × 40 × 40	7 × 70 × 70	4 × 40	6 × 60	4 × 40 × 40	5 × 50 × 50	5 × 50	4 × 40
2	5 × 50 × 50	4 × 40 × 40	5 × 50	3 × 30	5 × 50 × 50	4 × 40 × 40	4 × 40	3 × 30
3	5 × 50 × 50	4 × 40 × 40	5 × 50	3 × 30	5 × 50 × 50	4 × 40 × 40	4 × 40	3 × 30
4	5 × 50 × 50	4 × 40 × 40	5 × 50	3 × 30	5 × 50 × 50	4 × 40 × 40	4 × 40	3 × 30
5	4 × 40 × 40	4 × 40 ×40	4 × 40	3 × 30	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30
6	4 × 40 × 40	4 × 40 × 40	4 × 40	3 × 30	4 × 40 × 40	5 × 50 × 50	5 × 50	4 × 40
7	–	–	–	–	–	–	–	–
8	–	–	–	–	–	–	–	–

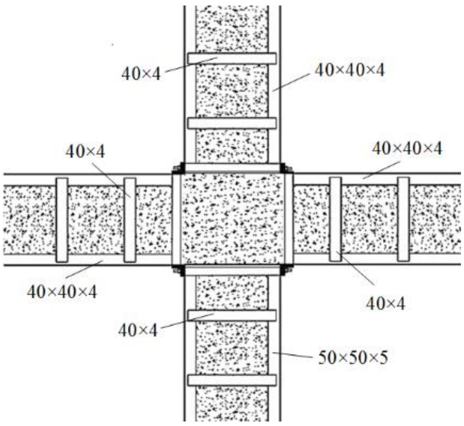


Fig. 14. Example of the steel angle and strip strengthening configuration in frame F, first story, interior beam-column joint.

In this sense, the geometry and mechanical properties of the strengthening elements are not arbitrary inputs, but are the direct outcome of a performance-driven design process. It should be emphasized that there is no unique or closed-form solution for selecting the strengthening layout, as the response depends on load distribution, members' capacity, and the interaction between confinement, flexural resistance, and shear capacity at both member and structural levels. Therefore, the adopted iterative approach represents a rational and standard methodology in performance-based seismic design, ensuring that the final configuration is both efficient and consistent with the targeted performance objective.

5.1. Seismic behavior of structures A to F

In this section, the nonlinear response and seismic performance of the RC structures (Frames A to F) are evaluated before and after strengthening with the steel angle and shear strip system. The parameters of interest include the load-displacement behavior, seismic performance, and plastic hinge formation patterns at the performance point and ultimate deformation state. Nonlinear static analysis was conducted using SAP2000 Version 20.2.0 [56], and plastic hinges were defined in accordance with the procedures described in this study. A triangular lateral load distribution was applied. Fig. 15 presents the results of nonlinear analysis in terms of lateral load-displacement response and the distribution of plastic hinges at the ultimate state. For the four-story frame, the results indicate that frame A exhibits improved strength and ductility after strengthening.

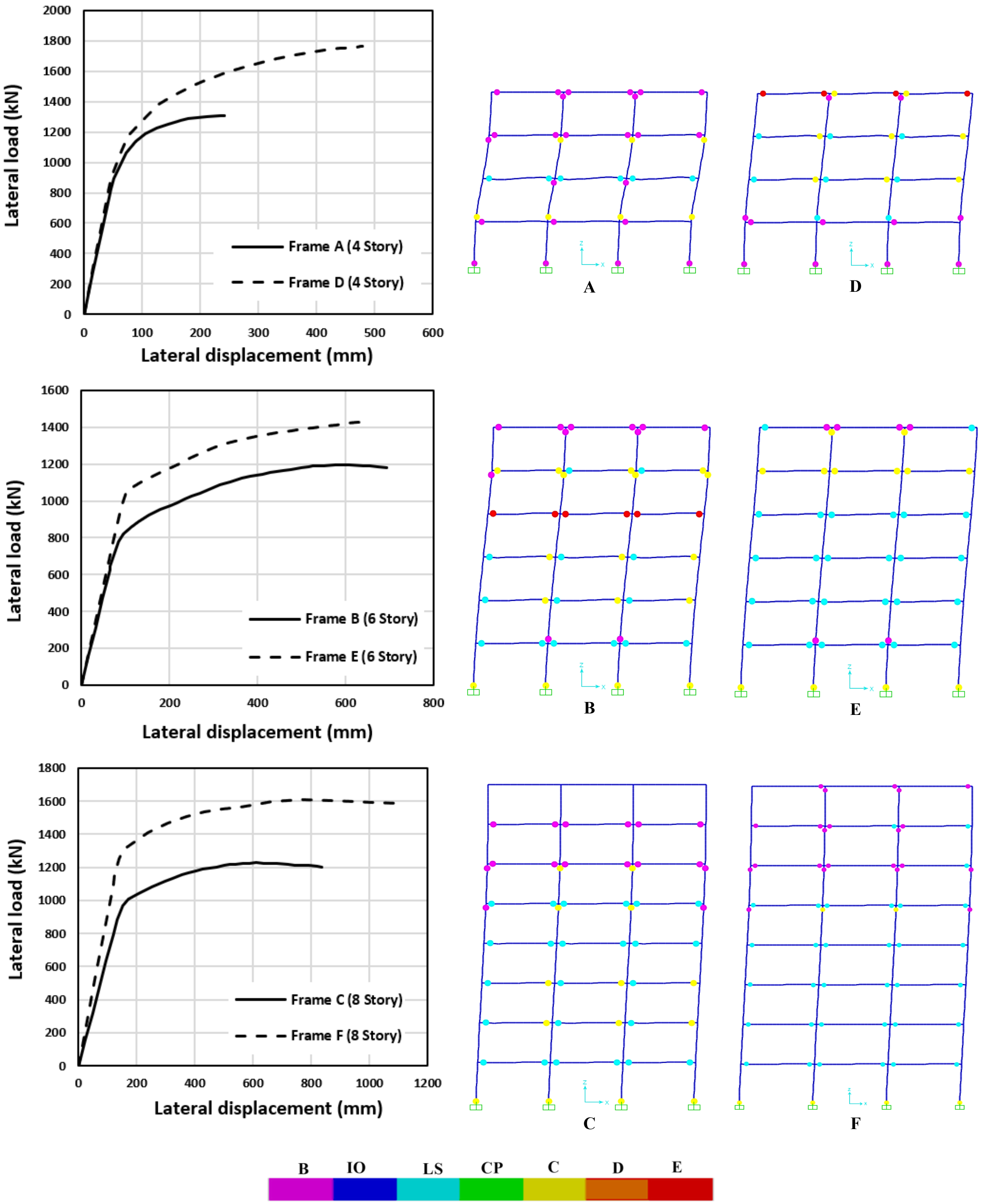


Fig. 15. Load-displacement response and corresponding damage pattern at the ultimate state.

The damage pattern shows that plastic hinges form in all columns prior to strengthening, with rotations in the second and third stories being more pronounced. This suggests the possibility of a soft-story mechanism in these regions. Following strengthening, plastic hinges predominantly form in the beams, while hinge rotations in the columns of the second and fourth stories are reduced and remain well below their ultimate limits.

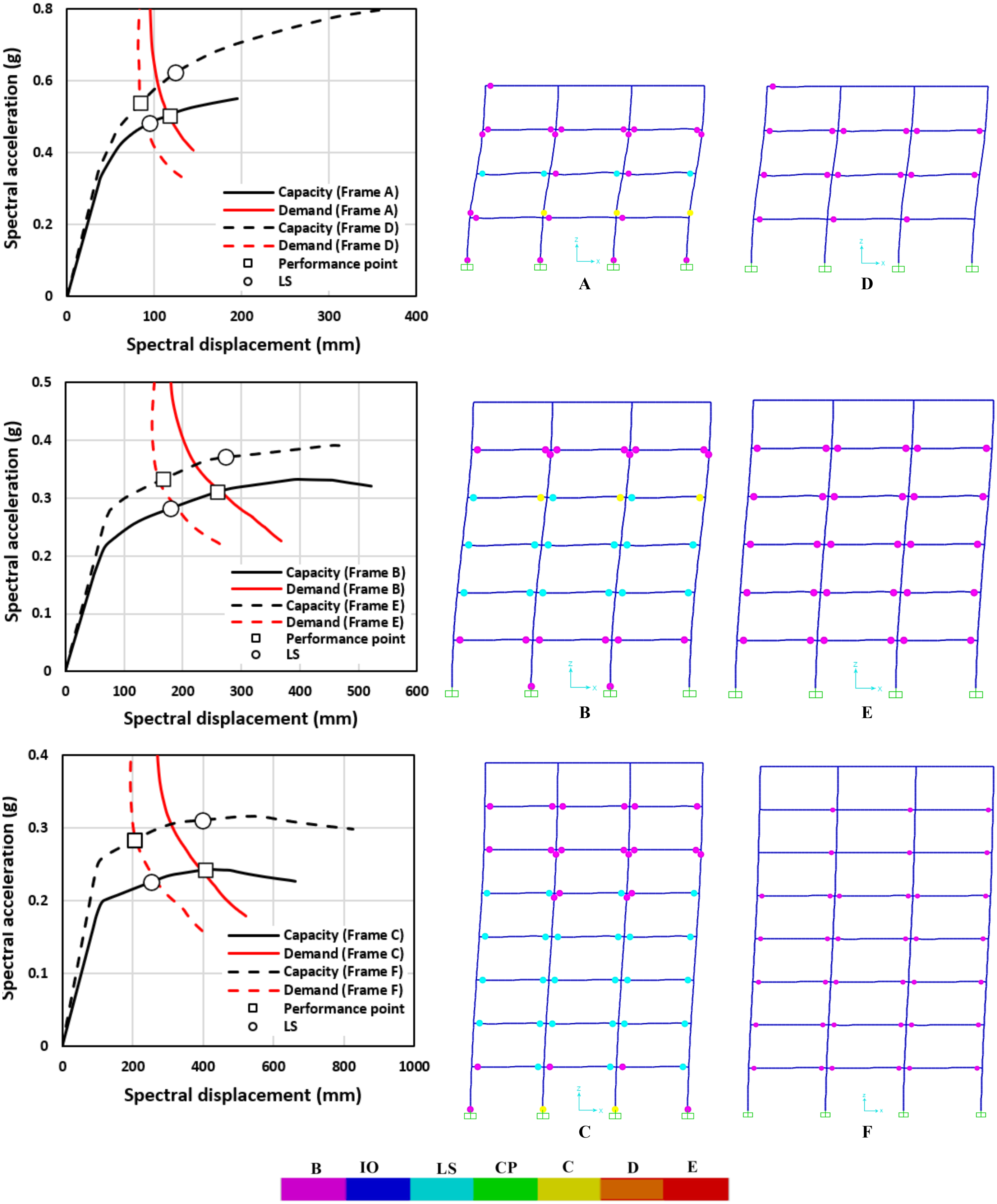


Fig. 16. Spectral acceleration-spectral displacement results and the damage pattern corresponding to the performance point.

Although the maximum lateral displacement of the strengthened frame exceeds that of the unstrengthened configuration, the post-strengthening hinge distribution reflects a more favorable failure mechanism. In the six-story frame, strengthening improved lateral stiffness and initial strength. However, the maximum lateral displacement in the strengthened configuration was slightly lower than in the original frame. The damage pattern indicates that plastic hinge rotations in the fifth-story columns approached their ultimate limits, increasing the likelihood of a soft-story mechanism. To address this, the strengthening strategy prioritized column enhancement while avoiding excessive stiffening of adjacent beams, thereby encouraging plastic hinge formation in beams rather than columns. After strengthening, plastic hinges predominantly developed in beams, while column hinge rotations were limited to yield-level deformations in the second story and remained moderate in the top story where strengthening was not applied. For the eight-story frame, strengthening resulted in improved strength, ductility, and initial stiffness. The damage pattern revealed that plastic hinge rotations in the fifth and sixth stories approached ultimate limits, indicating a potential soft-story mechanism in these regions. After strengthening, hinge formation was primarily observed in the beams, while column rotations in the critical stories were reduced. Although the maximum lateral displacement of the strengthened frame remained greater than that of the unstrengthened frame, the plastic hinge distribution and failure pattern were significantly improved.

Fig. 16 illustrates the nonlinear analysis results with respect to seismic performance and plastic hinge distribution at the performance point. Seismic performance is assessed by comparing the performance point with the life safety (LS) performance level, as defined in FEMA 356 [57]. At the LS level, structural and nonstructural components experience moderate damage and deformation. Lateral stiffness and the ability to resist additional lateral loads are reduced, and while collapse is unlikely, the structure may no longer be suitable for immediate occupancy. Repair may be possible but is often not economically practical. For the four-story frame, frame A remained beyond the LS performance level, indicating limited seismic performance in its original configuration. The hinge distribution confirmed the likelihood of a soft-story mechanism in the second story. After strengthening, plastic hinges formed primarily in the beams and remained within yield-level rotations, well below the LS limit. No significant hinge formation occurred at the base of the structure. The results in terms of load-displacement response, seismic performance, and hinge distribution at the performance point confirm the effectiveness of the steel angle and strip strengthening system in improving seismic behavior. In the four-story configuration, frame B remained beyond the life safety (LS) performance level, indicating inadequate seismic performance under design-level earthquakes. The plastic hinge distribution suggests the possibility of a soft-story mechanism in the fifth story, while hinge rotations in the fourth-story beams approached their ultimate limits.

After strengthening, plastic hinges were primarily confined to the beams and remained within yield-level rotations, with considerable margin from the LS threshold. The load-displacement response, seismic performance, and plastic hinge distribution at the performance point and ultimate deformation confirm the effectiveness of the steel angle and strip strengthening system. Similarly, in the four-story configuration, frame C also remained beyond the LS performance level, reflecting limited seismic performance in its unstrengthened state. The damage pattern indicated potential soft-story behavior in the fifth and sixth stories, and column hinge rotations at the base of the structure approached their ultimate limits. Following strengthening, plastic hinges developed predominantly in the beams, with rotations remaining within the yield range and far from the LS limit. The results in terms of structural response, seismic performance, and hinge formation at the performance point validate the effectiveness of the strengthening strategy using the steel angle and strip system.

It is noteworthy that the proposed model is developed within a nonlinear static (monotonic) framework; however, its formulation is based on sectional constitutive relationships that are inherently suitable for extension to cyclic and dynamic analyses. For application under cyclic loading, the model would require the incorporation of appropriate hysteretic rules to capture unloading-reloading behavior, stiffness and strength degradation, and energy dissipation. Similarly, for dynamic analysis, the model can be implemented within a time-history framework by integrating the proposed nonlinear hinge properties into the structural model along with suitable damping definitions. These extensions are straightforward and consistent with the fiber-based formulation adopted in this study.

6. Conclusions

This study focused on evaluating the nonlinear behavior and seismic performance of reinforced concrete (RC) structures and their strengthening using a steel angle and strip system. Analytical procedures were developed to capture nonlinear mechanisms before and after strengthening, enabling a rigorous assessment of seismic response. A modeling framework was proposed to incorporate the effects of the strengthening system on seismic behavior. The strengthening strategy influences nonlinear response through longitudinal strain development in the steel angles, allowing them to resist tensile and compressive forces and contribute to load transfer. Possible buckling of steel angles in compression zones was also included in the nonlinear formulation to represent realistic structural behavior. Furthermore, the combined action of steel strips and angles provides confinement to the concrete core, and this confinement effect was considered in the moment-curvature analysis of strengthened members. Insufficient shear capacity in RC members can lead to brittle shear failure and increased seismic vulnerability. Diagonal shear cracking in columns may reduce axial load-carrying capacity, while large flexural deformations and diagonal tensile cracking can further diminish shear resistance. Nonlinear analysis that neglects shear mechanisms may produce unconservative predictions of strength and deformation capacity, highlighting the necessity of incorporating shear effects in performance-based assessment. For strengthened members, a shear model was developed that considers contributions from stirrups, concrete, steel angles, and strips to overall shear capacity. Four-, six-, and eight-story RC frames were designed and strengthened to evaluate seismic performance. The results demonstrated close agreement between analytical predictions and experimental data, confirming the capability of the proposed model to simulate flexural and shear behavior under different loading conditions. Strengthening with the steel angle and strip system significantly improved initial stiffness, load-carrying capacity, and ductility. Load-displacement results of strengthened frames showed that the strengthening system enhances initial stiffness and seismic resistance when properly designed. Additionally, plastic hinge formation was redirected from columns to beams, mitigating undesirable strong-beam-weak-column mechanisms and improving overall structural performance. Seismic performance evaluation indicated that strengthening can shift the performance point toward a more favorable state and reduce damage potential at critical locations. The findings confirm that the steel angle and strip strengthening system is an effective strategy for improving nonlinear response and seismic performance of RC structures when appropriately designed and implemented.

Statements & declarations

Author contributions

Hananeh Dashtaki: Investigation, Formal analysis, Data curation, Software, Validation, Writing - Original Draft.

Javad Shayanfar: Conceptualization, Methodology, Software, Writing - Review & Editing.

Morteza Jamshidi: Project administration, Resources, Supervision, Writing - Review & Editing.

Acknowledgments

The authors declare that there are no acknowledgments to report for this work.

Funding

Funding information is not available.

Data availability

All data are available in the manuscript.

Declarations

The authors declare no conflict of interest.

References

[1] Sadeghpour Haji, M., Niknam, R., Shayanfar, J. ANN-Based Modeling of Shear Behavior of Reinforced Concrete Columns under Constant Axial Loads. Civil Engineering and Applied Solutions, 2026; 2: 28–48. doi:10.22080/ceas.2025.30228.1050.

[2] Abedi, M., Fangueiro, R., Correia, A. G., Shayanfar, J. Smart Geosynthetics and Prospects for Civil Infrastructure Monitoring: A Comprehensive and Critical Review. Sustainability, 2023; 15: 9258. doi:10.3390/su15129258.

[3] Shayanfar, J., Hemmati, A., Bengar, H. A. A simplified numerical model to simulate RC beam–column joints collapse. Bulletin of Earthquake Engineering, 2019; 17: 803–844. doi:10.1007/s10518-018-0472-z.

[4] Mohammadi, A., Shayanfar, J., Barros, J. A. O. Reliability-based design model for compressive strength of axially loaded FRP-confined concrete columns. Composite Structures, 2026; 384: 120159. doi:10.1016/j.compstruct.2026.120159.

[5] Shayanfar, J., Baharloo, S., Barros, J. A. O., Mohammadi, A. Strength Model for FRP-Confined Concrete: Comprehensive Database and Reliability-Based Partial Factor Calibration. Journal of Composites for Construction, 2026; 30: 04026006. doi:10.1061/JCCOF2.CCENG-5448.

[6] Salman, H. M., Al-Sherrawi, M. H. M-N Interaction Diagrams of RC Columns Strengthened with Steel C-Sections and Battens. Civil Engineering Journal, 2024; 10: 1974–1986. doi:10.28991/CEJ-2024-010-06-016.

[7] Abedi, M., Shayanfar, J., Al-Jabri, K. Infrastructure damage assessment via machine learning approaches: a systematic review. Asian Journal of Civil Engineering, 2023; 24: 3823–3852. doi:10.1007/s42107-023-00748-5.

[8] Erange, Y. G. P., Jenothan, M., Jayasinghe, J. A. S. C., Bandara, C. S., Dammika, A. J. Enhancement of the Lateral Performance of RC Columns Using Steel Jacketing: A Numerical Approach. Engineer: Journal of the Institution of Engineers, Sri Lanka, 2023; 56: 51–64. doi:10.4038/engineer.v56i2.7576.

[9] Shhatha, M. A., Mahdi, W. H., Alalikhan, A. A. Assessment for behavior of axially loaded reinforced concrete columns strengthened by different patterns of steel-framed jacket. Open Engineering, 2023; 13: 20220414. doi:10.1515/eng-2022-0414.

[10] Hu, Q., Jia, S., Wu, H., Wang, T. Experimental Research of Axial Compression Bearing Capacity on RC Columns Strengthened with Angle Steel and Self-Compacting Concrete. International Journal of Steel Structures, 2022; 22: 253–262. doi:10.1007/s13296-021-00571-1.

[11] Salah, A., Elsanadedy, H., Abbas, H., Almusallam, T., Al-Salloum, Y. Behavior of axially loaded L-shaped RC columns strengthened using steel jacketing. Journal of Building Engineering, 2022; 47: 103870. doi:10.1016/j.jobe.2021.103870.

[12] Akbarzadeh Bengar, H., Abdollahtabar, M., Shayanfar, J. Predicting the Ductility of RC Beams Using Nonlinear Regression and ANN. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2016; 40: 297–310. doi:10.1007/s40996-016-0033-0.

[13] Montuori, R., Piluso, V. Reinforced concrete columns strengthened with angles and battens subjected to eccentric load. Engineering Structures, 2009; 31: 539–550. doi:10.1016/j.engstruct.2008.10.005.

[14] Nagaprasad, P., Sahoo, D. R., Rai, D. Seismic strengthening of RC columns using external steel cage. Earthquake Engineering & Structural Dynamics, 2009; 38: 1563–1586. doi:10.1002/eqe.917.

[15] Campione, G., Monaco, A., Papia, M. Simplified analytical model for moment–axial force domain in the presence of shear in R.C. members externally strengthened with steel cages. Materials and Structures, 2016; 49: 3143–3157. doi:10.1617/s11527-015-0710-7.

[16] Campione, G. Strength and ductility of R.C. columns strengthened with steel angles and battens. Construction and Building Materials, 2012; 35: 800–807. doi:10.1016/j.conbuildmat.2012.04.090.

[17] Campione, G. RC Columns Strengthened with Steel Angles and Battens: Experimental Results and Design Procedure. Practice Periodical on Structural Design and Construction, 2013; 18: 1–11. doi:10.1061/(ASCE)SC.1943-5576.0000125.

[18] Badalamenti, V., Campione, G., Mangiavillano, M. L. Simplified Model for Compressive Behavior of Concrete Columns Strengthened by Steel Angles and Strips. Journal of Engineering Mechanics, 2010; 136: 230–238. doi:10.1061/(ASCE)EM.1943-7889.0000069.

[19] Shayanfar, J., Akbarzadeh Bengar, H. Nonlinear analysis of RC frames considering shear behaviour of members under varying axial load. Bulletin of Earthquake Engineering, 2017; 15: 2055–2078. doi:10.1007/s10518-016-0060-z.

[20] Khant, K., Shayanfar, J., Nguyen, H., Rezazadeh, M. Multi-Layer Perceptron for Prediction of Axial Strength of Fully FRP-Confined Circular Concrete. In: 12th International Conference on FRP Composites in Civil Engineering (CICE 2025); 2025 Jul 14–16; Lisbon, Portugal. p. 1041–1048. doi:10.1007/978-3-032-09387-5_101.

[21] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Analysis-oriented model for partially FRP-and-steel-confined circular RC columns under compression. Engineering Structures, 2023; 276: 115330. doi:10.1016/j.engstruct.2022.115330.

[22] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Cross-sectional and confining system unification on peak compressive strength of FRP confined concrete. Structural Concrete, 2023; 24: 1531–1545. doi:10.1002/suco.202200105.

[23] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Analytical model to predict axial stress-strain behavior of heat-damaged unreinforced concrete columns wrapped by FRP jacket. Engineering Structures, 2023; 289: 116244. doi:10.1016/j.engstruct.2023.116244.

[24] Pereira, J. P. C., Shayanfar, J., Barros, J. A. O. Effect of Different Passive FRP Confinement Arrangements on Stress-Strain Behavior of Axially-Loaded Fire-Damaged Concrete Elements. In: 12th International Conference on FRP Composites in Civil Engineering (CICE 2025); 2025 Jul 14–16; Lisbon, Portugal. p. 1033–1040. doi:10.1007/978-3-032-09399-8_97.

[25] Shayanfar, J., Barros, J. A. O. Design-Oriented Model of Unified Character to Determine Softening–Hardening Stress–Strain Behavior of FRP-Confined Concrete Columns of General Cross Section. Journal of Composites for Construction, 2024; 28: 04024059. doi:10.1061/JCCOF2.CCENG-4772.

[26] Shayanfar, J., Omidalizadeh, M., Nematzadeh, M. Analysis-oriented Model for Seismic Assessment of RC Jacket Retrofitted Columns. Steel and Composite Structures, 2020; 37: 371–390. doi:10.12989/scs.2020.37.3.371.

[27] Mander, J. B., Priestley, M. J. N., Park, R. Theoretical Stress-Strain Model for Confined Concrete. Journal of Structural Engineering, 1988; 114: 1804–1826. doi:10.1061/(ASCE)0733-9445(1988)114:8(1804).

[28] Wang, Y. C., Restrepo, J. I. Investigation of concentrically loaded reinforced concrete columns confined with glass fiber-reinforced polymer jackets. Structural Journal, 2001; 98: 377–385.

[29] Nematzadeh, M., Mousavimehr, M., Shayanfar, J., Omidalizadeh, M. Eccentric compressive behavior of steel fiber-reinforced RC columns strengthened with CFRP wraps: Experimental investigation and analytical modeling. Engineering Structures, 2021; 226: 111389. doi:10.1016/j.engstruct.2020.111389.

[30] Shayanfar, J., Rezazadeh, M., Barros, J. A. Analytical Model to Predict Dilation Behavior of FRP Confined Circular Concrete Columns Subjected to Axial Compressive Loading. Journal of Composites for Construction, 2020; 24: 04020071. doi:10.1061/(ASCE)CC.1943-5614.0001087.

[31] Wang, J., Xiao, H., Lu, L., Yang, J., Lu, S., Shayanfar, J. Axial stress-strain model for concrete in partially FRP wrapped reinforced concrete columns. Construction and Building Materials, 2024; 416: 135028. doi:10.1016/j.conbuildmat.2024.135028.

[32] Shayanfar, J., Barros, J. A. O., Pereira, J. P. C. A versatile model with a design framework for axially-loaded FRP-confined concrete with/without a stress reduction-recovery behavior. Construction and Building Materials, 2024; 448: 138097. doi:10.1016/j.conbuildmat.2024.138097.

[33] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Design-oriented stress–strain model for RC columns with dual FRP- steel confinement mechanism. Composite Structures, 2024; 330: 117821. doi:10.1016/j.compstruct.2023.117821.

[34] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Stress–Strain Model for FRP-Confined Circular Concrete Columns Developing Structural Softening Behavior. Journal of Composites for Construction, 2024; 28: 04023065. doi:10.1061/JCCOF2.CCENG-4364.

[35] Shayanfar, J., Barros, J. A. O., Rezazadeh, M., Kafshgarkolaei, H. J. Enhancing the performance of heat-damaged rectangular RC columns using prestressed FRP confinement. Construction and Building Materials, 2025; 501: 144346. doi:10.1016/j.conbuildmat.2025.144346.

[36] Paulay, T., Priestly, M. J. N. Seismic Design of Reinforced Concrete and Masonry Buildings. First ed. New York (NY): John Wiley & Sons, Inc; 1992. doi:10.1002/9780470172841.

[37] Shayanfar, J., Barros, J. A., Rezazadeh, M. Unified model for fully and partially FRP confined circular and square concrete columns subjected to axial compression. Engineering Structures, 2022; 251: 113355. doi:10.1016/j.engstruct.2021.113355.

[38] Shayanfar, J., Barros, J. A. O., Rezazadeh, M. Generalized Analysis-oriented model of FRP confined concrete circular columns. Composite Structures, 2021; 270: 114026. doi:10.1016/j.compstruct.2021.114026.

[39] Shayanfar, J., Akbarzadeh Bengar, H., Niroomandi, A. A proposed model for predicting nonlinear behavior of RC joints under seismic loads. Materials & Design, 2016; 95: 563–579. doi:10.1016/j.matdes.2016.01.098.

[40] Shayanfar, J., Pereira, J. P. C., Barros, J. A. O. Passive confinement scenarios for strengthening fire-damaged concrete columns: Experimental and analytical research. Structures, 2025; 73: 108375. doi:10.1016/j.istruc.2025.108375.

[41] Abedi, M., Roshan, M. J., Gulisano, F., Shayanfar, J., Adresi, M., Fangueiro, R., Correia, A. G. An advanced cement-based geocomposite with autonomous sensing and heating capabilities for enhanced intelligent transportation infrastructure. Construction and Building Materials, 2024; 411: 134577. doi:10.1016/j.conbuildmat.2023.134577.

[42] Akbarzadeh Bengar, H., Assadikiadehi, M., Shayanfar, J., Nazari, M. Effective flexural rigidities for RC beams and columns with steel fiber. Steel and Composite Structures, 2019; 34: 453–465. doi:10.12989/scs.2020.34.3.453.

[43] Tariq, M., Khan, A., Ullah, A., Shayanfar, J., Niaz, M. Improved Shear Strength Prediction Model of Steel Fiber Reinforced Concrete Beams by Adopting Gene Expression Programming. Materials (Basel), 2022; 15: 3758. doi:10.3390/ma15113758.

[44] Tariq, M., Khan, A., Shayanfar, J., Hanif, M. U., Ullah, A. A regression model for predicting the shear strength of RC knee joint subjected to opening and closing moment. Journal of Building Engineering, 2021; 41: 102727. doi:10.1016/j.jobe.2021.102727.

[45] Shayanfar, J., Akbarzadeh Bengar, H. A practical model for simulating nonlinear behaviour of FRP strengthened RC beam-column joints Steel and Composite Structures, 2018; 27: 49–74. doi:10.12989/scs.2018.27.1.049

[46] Pan, Z., Li, B. Truss-Arch Model for Shear Strength of Shear-Critical Reinforced Concrete Columns. Journal of Structural Engineering, 2013; 139: 548–560. doi:10.1061/(ASCE)ST.1943-541X.0000677.

[47] Lynn, A. C. Seismic evaluation of existing reinforced concrete building columns [PhD Thesis]. Berkeley (CA): University of California, Berkeley; 2001.

[48] Park, H.-G., Yu, E.-J., Choi, K.-K. Shear-strength degradation model for RC columns subjected to cyclic loading. Engineering Structures, 2012; 34: 187–197. doi:10.1016/j.engstruct.2011.08.041.

[49] Berry, M. P., Eberhard, M. O. Practical Performance Model for Bar Buckling. Journal of Structural Engineering, 2005; 131: 1060–1070. doi:10.1061/(ASCE)0733-9445(2005)131:7(1060).

[50] Shayanfar, J., Akbarzadeh, B. Numerical model to simulate shear behaviour of RC joints and columns. Computers and Concrete, 2016; 18: 877–901. doi:10.12989/cac.2016.18.4.877.

[51] Williams, M. S., Sexsmith, R. G. 8. Seismic Damage Indices for Concrete Structures: A State-of-the-Art Review. Earthquake Spectra, 1995; 11: 319–349. doi:10.1193/1.1585817.

[52] Li, J., Gong, J., Wang, L. Seismic behavior of corrosion-damaged reinforced concrete columns strengthened using combined carbon fiber-reinforced polymer and steel jacket. Construction and Building Materials, 2009; 23: 2653–2663. doi:10.1016/j.conbuildmat.2009.01.003.

[53] Germano, F., Plizzari, G. A., Tiberti, G. Experimental study on the behavior of SFRC columns under seismic loads. In: VIII International Conference on Fracture Mechanics of Concrete and Concrete Structures FraMCoS-8; 2013 Mar 10–13; Toledo, Spain. p. 1–12.

[54] Duong, V., Sheikh, S., Vecchio, F. Seismic behavior of shear-critical reinforced concrete frame: Experimental investigation. ACI Structural Journal, 2007; 104: 304–313. doi:10.14359/18620.

[55] Shayanfar, J., Bengar, H. A., Parvin, A. Analytical prediction of seismic behavior of RC joints and columns under varying axial load. Engineering Structures, 2018; 174: 792–813. doi:10.1016/j.engstruct.2018.07.103.

[56] Computers and Structures, Inc. (CSI). SAP2000 Integrated Software for Structural Analysis and Design. Version 20.2.0. Berkeley (CA): Computers and Structures, Inc. (CSI); 2018.

[57] Federal Emergency Management Agency (FEMA). FEMA 356: Prestandard and Commentary for the Seismic Rehabilitation of Buildings. Washington, D.C. (US): ASCE; 2000.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Determinants of Ride-Hailing Adoption in Developing Urban Contexts: Evidence from Western Districts of Mashhad, Iran

Mahdi Nosrati [a], Amir Abbas Rassafi [a*], Hossein Rahimi Farahani [a]

[a] *Department of Transportation Planning, Faculty of Engineering, Imam Khomeini International University, Qazvin, Iran*

ARTICLE INFO

Article history:

Received 25 May 2026

Revised 24 June 2026

Accepted 01 August 2026

Available online 01 April 2027

Keywords:

Ride-hailing adoption

Travel behavior

Mode choice

Pandemic-related impacts

Discrete choice modeling

Public transportation

ABSTRACT

Ride-Hailing (RH) services have emerged as a preferred non-shared mobility option, contributing to a notable increase in their volume and modal share within urban transportation systems in Iranian cities. Despite this trend, limited empirical research exists on RH adoption in Iran and other developing countries. This study investigates the factors influencing RH usage in Iran, focusing on Mashhad, the nation's second-largest metropolis. Data were collected in 2022 through 619 valid questionnaires encompassing four categories of variables: demographic, travel-related, environmental, and latent psychological factors. Exploratory factor analysis was first used to derive factor scores, followed by the development of an ordered probit discrete choice model. Findings reveal that variables such as marital status, employment, physical disability, household size, and social media use showed no significant effect. However, travel purpose, limited public transit access, proximity to traffic-restricted area, and trip characteristics such as duration and time of day significantly influenced usage. Among latent variables, privacy concerns and dissatisfaction with public transportation positively affected RH preference. Whereas pandemic-related apprehensions and economic constraints diminished it. The study offers insights into urban mobility behavior in developing contexts and highlights the multifaceted factors shaping RH demand.

How to cite this article: Nosrati, M., Rassafi, Amir A., Rahimi Farahani, H. Determinants of Ride-Hailing Adoption in Developing Urban Contexts: Evidence from Western Districts of Mashhad, Iran. Civil Engineering and Applied Solutions. 2027; 3(2): 69–86. doi:10.22080/ceas.2026.31910.1103.

1. Introduction

The rise of platform businesses has sparked a clash between traditional economic players and new, innovative actors [1]. App-based mobility encompasses a range of transportation services facilitated via smartphone applications. Among these, ride-hailing remains the most prevalent service. Despite being introduced over a decade ago, ride-hailing continues to generate considerable debate regarding its effects on urban transportation systems [2]. RH has altered traffic patterns, potentially affecting related externalities such as congestion, traffic accidents, and environmental impacts [3]. Transportation Network Companies (TNCs) such as Uber, Lyft, and Didi have fundamentally transformed the conventional taxi industry on a global scale. For instance, Uber's entry into New York in 2011 reduced hourly traditional taxi usage by 25% [4], similarly Shenzhen, China, saw a 20% decline in taxi revenues [5]. Such shifts highlight ride-hailing's dual role as both a complement to and a substitute for public transit, leading to substantial changes in urban mobility patterns [6, 7]. In Iran, the 2014 launch of RH services like Snapp facilitated by Information and Communication Technology (ICT) advancements and internet penetration revolutionized transportation. Snapp enables users to book rides, view driver profiles, estimate fares, and choose payment methods via smartphones, amassing over 40 million users and 2 million daily trips by 2020 [8].

Focusing on Mashhad, Iran's second-largest city, this research combines Exploratory Factor Analysis (EFA) and ordered probit

* Corresponding author.
E-mail addresses: rasafi@eng.ikiu.ac.ir (A. A. Rassafi).

https://doi.org/10.22080/ceas.2026.31910.1103
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

modeling to analyze 619 survey responses in 2022. It identifies key demographic, travel-related, environmental, and latent psychological predictors of RH usage while highlighting the pandemic's unique impact [9]. The findings offer policymakers and urban planner insights into sustainable mobility solutions tailored to developing economies.

This study contributes to the existing literature in several ways. First, it investigates ride-hailing adoption in Mashhad, one of the largest metropolitan areas in Iran, where empirical evidence on ride-hailing behavior remains limited. Second, the study jointly examines demographic, travel-related, environmental, and latent psychological factors affecting ride-hailing usage. Third, it explicitly incorporates pandemic-related perceptions and concerns, as well as privacy concerns associated with public transportation, factors that have received limited attention in previous studies conducted in developing countries. Finally, by combining exploratory factor analysis with an ordered probit model, the study provides a comprehensive framework for analyzing both latent attitudes and observed travel behavior.

2. Research background

2.1. Analysis approaches

Recent advancements integrate machine learning with macro-level data (e.g., land use, urban density) and micro-level variables (e.g., demographics, real-time weather, traffic congestion) to model RH demand [10]. RH platforms operate as two-sided markets, where supply (drivers) and demand (passengers) interact dynamically [11, 12]. This framework emphasizes the interdependence of driver incentives (e.g., wage structures, surge pricing) and passenger demand in shaping service accessibility, pricing, and transaction volumes. For instance, driver availability directly affects wait times and user satisfaction, while demand surges influence pricing algorithms. Balancing these dynamics is essential for platform sustainability, as mismatches between supply and demand can lead to service inefficiencies or driver attrition [13].

Mode choice modeling, examines the behavioral and contextual factors influencing individuals' preferences for RH over alternatives like public transit or private vehicles. By accounting for the interconnected factors on mode choice (service attributes, passenger characteristics, and external conditions), the framework ensures a holistic analysis of transportation systems, balancing user needs, operational efficiency, and environmental variables [14]. Structural Equation Models empower the analysis by considering interactions between latent factors such as passenger attitudes and perceptions and observable variables like socioeconomic traits [4]. For instance, multigroup MIMIC models have been used to investigate the pandemic-driven shifts in RH intentions [9]. This study employs an ordered probit model to assess both observed variables and latent constructs. In ordered probit, a normal distribution of the error term is assumed, which can be more appropriate for modeling processes [15, 16].

2.2. SH demand determinants

RH services have revolutionized urban transportation, yet their demand dynamics remain context-dependent, particularly in understudied developing economies [3, 17]. This review synthesizes global evidence on socio-demographic, attitudinal, and environmental determinants of RH adoption, highlighting critical research gaps in Iran and similar contexts. By integrating latent variables and the pandemic-era behavioral shifts, this study advances methodological and empirical insights for sustainable mobility planning. The rapid proliferation of RH platforms like Uber, Lyft, and Snapp has redefined urban mobility, offering flexible alternatives to private vehicles and public transit [7, 8, 17, 18]. This review examines the determinants of RH demand, critiques methodological approaches, and identifies gaps in Iran's mobility literature.

2.2.1. Socio-demographics

Demographic factors such as age, income, and educational attainment serve as reliable indicators of RH adoption, with younger, affluent, and highly educated groups typically exhibiting higher usage rates [9, 19, 20]. However, regional variations challenge this trend; For instance, in China, users frequently lack higher education qualifications [21]. A counterintuitive pattern also emerges in low-income New York communities, which show unexpectedly high RH demand [22]. Gender remains a subject of debate in adoption research; while some studies in underdeveloped nations suggest higher female adoption rates of ride-hailing services [23, 24], recent evidence from countries such as Pakistan indicates lower female usage due to stronger safety concerns and contextual constraints. However, other findings suggest no significant gender or racial influences [25, 26]. Similarly, car ownership is generally linked to reduced demand in most studies [7, 23, 27], though this trend does not hold in Chile, where research shows no such relationship [18].

2.2.2. Travel and built environment

Environmental factors demonstrate a stronger association with usage frequency compared to socio-demographic characteristics [28]. Neighborhood walkability, urban compactness, pedestrian-oriented design, transit accessibility, educational institutions, and accessibility to town centers are associated with individuals' use of ride-hailing services [29, 30]. Demand patterns are notably shaped by trip purposes (such as work, education, leisure, or business) and urban density [17, 31]. Additionally, individuals residing near metro systems, commercial centers, or in densely populated urban areas show a higher likelihood of adopting ride-sourcing services [19, 32]. Similarly, limited parking availability and traffic congestion further drive reliance on RH platforms [33].

2.2.3. Latent factors

Psychological factors, including perceived safety, service quality, travel time efficiency, reliability, cost-effectiveness, flexibility, and environmental citizenship serve as primary motivators for adopting RH services [25, 31, 34-37]. Features like real-time tracking and driver verification significantly increase usage in regions with high crime rates [38], whereas dissatisfaction with public transportation drives transitions to RH platforms [39].

2.3. Iran's RH landscape

In Iran, RH services constitute a vital component of the urban mobility economy [40], distinguished by high public trust, perceived utility, and service convenience [41]. Iranian users prioritize flexibility and distrust in public transit [41]. Women, students, and formal-sector employees constitute primary adopters, while vehicle ownership reduces usage [24, 42]. Despite its societal embeddedness, empirical studies on Iran's RH ecosystem remain scarce, reflecting a broader neglect of developing contexts.

2.4. Research gaps and contributions

This study addresses three critical gaps. First, while latent variables (e.g., attitudes, perceptions) are recognized predictors of transportation demand, their role in RH adoption remains rarely understudied. We integrate attitudinal questions with observed variables in a Discrete Choice Model (DCM) to capture these nuances. Second, by examining the pandemic-era data, we reveal how health and economic shocks reshape mobility preferences. Third, although technology-enabled services such as ride-hailing comprise a substantial portion of the modal split in many developing countries, including Iran, scholarly research on this topic remains limited [3]. In particular, there is a notable gap in studies exploring the factors that influence the frequency of ride-hailing usage. Iran, as a developing country with a population of more than 90 million and large usage of RH services, is a worthy case study and opens up new insights on travelers' attitudes in developing countries.

3. Methodology

This study employs EFA and DCM to analyze factors influencing RH demand. Fig. 1 summarizes the overall research framework adopted in this study. The process begins with questionnaire design and data collection, followed by data preprocessing and exploratory factor analysis (EFA) to identify latent constructs. The extracted factor scores are then combined with demographic, travel-related, and environmental variables and incorporated into an ordered probit model to estimate the determinants of ride-hailing usage. Finally, marginal effects are calculated to facilitate the interpretation of the model results.

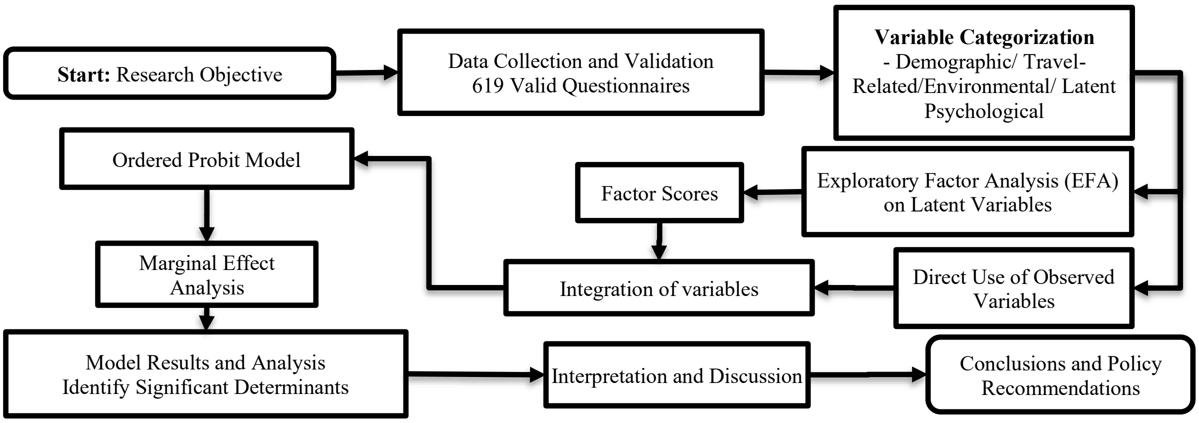


Fig. 1. Framework for estimating RH demand.

3.1. DCM framework and ordered probit model

DCMs analyze decisions among mutually exclusive alternatives, assuming individuals choose the alternative maximizing their utility [16]. The utility function U_{in} for individual i choosing alternative n is expressed as Eq. 1 [43].

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

where:

U_{in}: The utility of alternative n for the decision maker i (i denotes the individual respondent and n denotes the alternative)

V_{in}: The observable or measurable component in the utility that the analyst calculates or estimates

ε_{in}: The portion of utility that is unobservable or error remains unknown to the analyst

The Ordered Probit Model is particularly suited for analyzing ordinal survey data (e.g., Likert-scale responses), where categories have inherent order but lack numerical meaning. The model preserves the ordinal nature of responses while accommodating unequal spacing between categories, making it a statistically rigorous method for interpreting ordered categorical variables [44].

Let i denote a respondent ($i = 1, \dots, N$), where N represents the sample size. The observed ordinal response of individual i to a survey question is denoted y_i, which can take integer values $(1, 2, 3, \dots, J)$. The model posits a latent variable y^*_i, representing the unobserved propensity of respondent i to agree with a statement, bounded between $(-\infty < y^*_i < +\infty)$. This latent variable y^*_i is linked linearly to a vector of explanatory features x_i by Eq. 2.

$$y^*_i = x_i\beta + u_i \qquad i = 1, \dots, n \tag{2}$$

Here, β parameter analogously to slope parameters in linear regression. While y^*_i is unobserved, its relationship to the observed y is defined as Eq. 3.

$$\begin{aligned} &y = 1 \ \ if \ \ -\infty < y^* < \kappa_1 \\ &y = 2 \ \ if \ \ \kappa_1 < y^* < \kappa_2 \\ &y = J \ \ if \ \ \kappa_{J-1} < y^* < \infty \end{aligned} \tag{3}$$

The parameters $\kappa_j (j = 1, \dots, J-1)$ are referred to as cut points or threshold parameters. For example, the probability density function of y^*_i for $J = 4$ is illustrated in Fig. 2 [44].

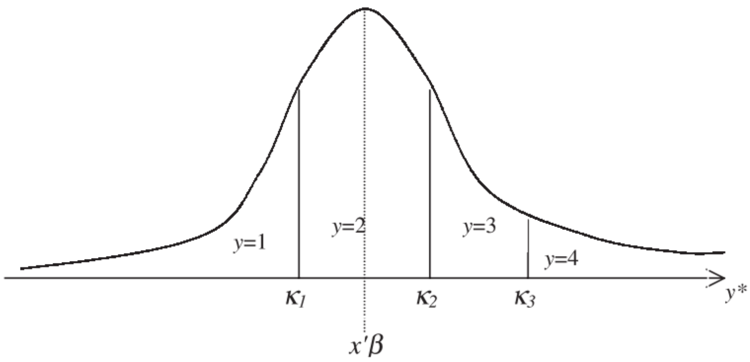


Fig. 2. Density function of y^*_i.

Let $P_i(y)$ represent the probability that respondent *i* chooses category *y*. This probability is expressed as Eq. 4.

$$P_i(y) = P(\kappa_{y-1} < y^* < \kappa_y) = \Phi(\kappa_y - x_i\beta) - \Phi(\kappa_{y-1} - x_i\beta) \ \ y = 1, \dots, J \tag{4}$$

where Φ is the standard normal cumulative distribution function. The log-likelihood for the sample is expressed as Eq. 5.

$$LogL = \sum_i ln[P_i(y_i)] = \sum_i ln[\Phi(\kappa_{yi} - x_i\beta) - \Phi(\kappa_{yi-1} - x_i\beta)] \tag{5}$$

DCMs, such as logit and probit models, require systematic validation to evaluate their accuracy. Common statistical methods used for this purpose include hypothesis testing, diagnostic assessments, and goodness-of-fit measures such as McFadden's pseudo-R^2 and the Akaike/Bayesian Information Criteria (AIC/BIC).

3.2. Exploratory factor analysis

Factor analysis is a versatile multivariate method extensively utilized in psychology, education, health, and social sciences to refine and validate measurement tools and theoretical concepts. EFA is a statistical method used to identify latent structures within a set of observed variables. Basically, it includes 5 steps, namely: Data suitability, extraction criteria, factor extraction, rotation, and interpretation [45]. In the current study, principal component analysis and varimax were used for factor extraction and rotation approaches, respectively [45, 46].

To validate the factor analysis framework, three statistical criteria were applied. First, the correlation matrix was examined to ensure moderate-to-strong correlations (≥ 0.3) among variables within proposed factors, a prerequisite for justifying shared latent constructs [46]. Second, the Kaiser-Meyer-Olkin (KMO) test assessed sampling adequacy, with values ≥ 0.6 confirming the data's suitability for EFA [45]. Finally, Bartlett's test of sphericity; ($p < 0.05$) rejected the null hypothesis of random noise, thereby supporting the use of EFA [46, 47]. Together, these tests ensure methodological rigor in identifying latent structures within the dataset.

4. Case study and data

The empirical survey underlying this research was conducted between January and March 2022, when the COVID-19 pandemic continued to affect everyday urban mobility patterns in Mashhad. During this phase, several public health measures and mobility restrictions were still enforced, shaping individuals' travel decisions and their reliance on ride-hailing services. Therefore, this study employing an online survey, investigated the determinants influencing the adoption of RH services by incorporating an analysis of both latent attitudinal and observed variables. Given the well-established presence and widespread utilization of RH services in

Mashhad, it was presumed that the respondents had substantial familiarity with this mode of transportation. Additionally, precautionary measures were implemented at the survey's outset, and comprehensive explanations were provided to participants to ensure clarity and accuracy in their responses. The internal consistency of the questionnaire was assessed using Cronbach’s alpha, which yielded an acceptable level of reliability.

Participants were recruited through an online survey using a convenience sampling method, as the pandemic-related restrictions prevented in-person data collection. The questionnaire link was distributed via social media platforms and local digital communities of ride-hailing users in Mashhad. The study focused on residents of recently developed neighborhoods in the city’s western districts, representing a more modern urban population. Eligible participants had completed at least secondary school education, resided in the selected areas, and were familiar with ride-hailing services. Individuals not meeting these criteria or providing incomplete responses were excluded from the analysis.

Mashhad, spanning an area of over 300 square kilometers, is situated in Khorasan Razavi Province in northeastern Iran, with a population of approximately 3.5 million and over 30 million visitors annually [48, 49]. Mashhad serves as Iran’s leading tourist hub, drawing international visitors from countries such as Iraq, Bahrain, Kuwait, Lebanon, Pakistan, Afghanistan, Azerbaijan, and Qatar. The city is structured into 13 districts, with the Samen district being of particular religious and governmental importance (Fig. 3). Like many other major metropolitan areas, Mashhad grapples with substantial challenges associated with the widespread use of motor vehicles.

Before administering the full-scale survey, a pilot study was conducted. The pilot sample consisted of 45 transportation experts and students, each of whom provided individualized feedback on survey questions to ensure clarity and eliminate ambiguity. Following the pilot phase, and preliminary modeling, unclear or irrelevant questions were modified to enhance response accuracy and participation rates. The final data collection took place between January and March 2022. Traditional communities are primarily concentrated in the eastern districts of Mashhad. In contrast, the current study centers on more modern populations and neighborhoods characterized by recent urban development. The survey was distributed among individuals in Mashhad’s Districts 9, 10, and 11, resulting in 619 completed questionnaires (District 12 is a developing area that currently lacks a significant residential population). The RH Questionnaire was systematically structured into two sections, comprising 46 questions in total. The first section focused on observed variables, employing multiple-choice and fill-in-the-blank formats, while the second section assessed latent variables using a five-point Likert scale.

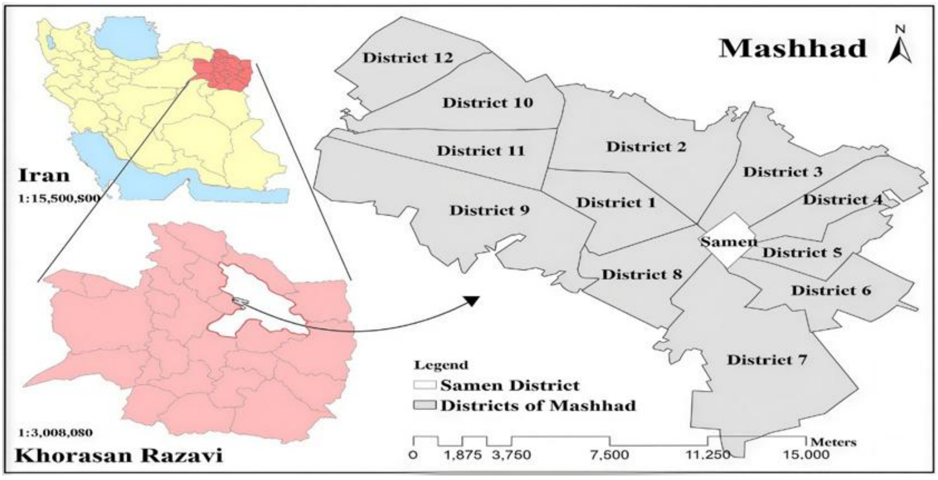


Fig. 3. The location and urban zoning of Mashhad [50].

5. Results

5.1. Demographic, travel-related and environmental characteristics

Table 1 summarizes the key demographic characteristics of the study sample. According to the table, more than 50% of the respondents are under 25 years old, possess educational qualifications below a postgraduate degree, and are predominantly single. This trend suggests that a significant portion of the respondents are students or pupils from universities and schools. Women comprise approximately 53% of the respondents, while men account for 47%. This gender disparity may be attributed to higher response rates among women, despite the survey being distributed nearly equally among both genders. Approximately 60% of the respondents are either students or part-time employed, aligning with the prevalence of lower income levels below 30 million Rials (MR) and between 30 and 80 MR. Additionally, the most common household structure consists of four family members with two licensed drivers.

The data reveals that 53% of the respondents reside within a five-minute distance from public transit stations, while 30.4% are between five and ten minutes away, indicating widespread accessibility to transit options. Additionally, the data suggests that most individuals have access to both private and public transportation options. Approximately half of the respondents prefer to delegate parcel deliveries to RH drivers, rather than making trips themselves.

Regarding RH travel behavior, 39.7% of the respondents did not specify a particular trip purpose. Among those who did, the most frequently cited trip purposes were returning home, work, and personal errands. Notably, the afternoon emerged as the most preferred time for RH usage, followed by morning trips and travel during the pandemic restrictions.

Table 1. Studied variables and their specifications.

Variable	Unit	Type*	Category	Freq. (%)
Age	Year	O	Under 18	19.4
			18 to 24	32.5
			24 to 30	6.8
			30 to 40	16.2
			40 to 50	15.3
			Over 50	9.9
Educational attainment	-	O	Less than high school diploma	23.6
			High school diploma or associate degree	31.2
			Bachelor's degree	28.9
			Postgraduate degree	16.3
Marital status	-	N	Single	57.4
			Married	42.6
Gender	-	N	Female	52.5
			Male	47.5
Use of RH for parcel delivery	-	N	Yes	51.2
			No	48.6
Walking time to the nearest public transit station	min	O	0 to 5	53
			5 to 10	30.4
			10 to 15	8.7
			>15	7.9
Trip purpose	-	N	Work	16.3
			Education	4.4
			Returning home	18.4
			Leisure	4.2
			Personal errands	16.2
			Various destinations	39.7
Preferred trip duration for RH use	min	N	>20	14.2
			≤20	24.7
			No preference	60.4
Daily social media usage	h	O	≤1	9.4
			1 to 2	27.3
			>2	63.2
Monthly income	MR	N	≤30	27.1
			30-80	27.1
			80-120	11.6
			>120	2.4
			Prefer not to answer	31.5
Employment status	-	N	Full-time	29.6
			Part-time	12.9
			Student	46
			Retired	2.7
			Housekeeper	3.9
			Unemployed	4.8
Household size	Persons	O	1	1
			2	7.3
			3	19.2
			4 or more	72.5

Number of licensed drivers in household	Persons	O	0 or 1	16.6
			2	48
			3 or more	35.2
Physical or intellectual disability	-	N	Yes	2.9
			No	96.8
Type of vehicle accessible to the respondent	-	N	Motorcycle	3.9
			Car	66.7
			both	5
			No vehicle	24.4
Preferred time of day for RH usage	-	N	During the morning period	12.1
			During the afternoon period	34.9
			During the pandemic restrictions	4.8
			At different times throughout the day	47.5
Frequency of RH usage in the past month	Times/month	O	0	25.4
			1 to 8	51.9
			> 8	22.8
RH usage before the pandemic	Times/month	O	0	17.1
			0 to 8	59.1
			> 8	23.4
Workplace/study location in CBD	-	N	Yes	25.5
			No	74.5
Residence in CBD zone	-	N	Yes	16.6
			No	83.4

* Type = Measurement type (O: Ordinal / N: Nominal)

The study also found high engagement with social media, as about two-thirds of the respondents reported spending more than two hours per day on social networking platforms. This aligns with the high participation rate in the online questionnaire, underscoring the respondents' familiarity with mobile technology and digital platforms.

In terms of geographical distribution, 74.5% of the respondents' workplaces or educational institutions and 83.4% of their residences were located outside the Mashhad traffic-restricted area. This finding is consistent with the urban layout, where the traffic-restricted area primarily accommodates administrative, commercial, and religious centers, as well as hotels, with few permanent residents. Given that this study primarily focuses on Mashhad's developed residential population, only a small proportion of the respondents resided within the traffic-restricted area.

The willingness to use RH services has declined in the aftermath of the COVID-19 pandemic and is shown in Fig. 4. Before the pandemic, 17.1% of the respondents reported not using RH services at all during the month, whereas this number increased to 25.4% during the pandemic. Notably, the overall 8.3% decline in RH usage primarily stemmed from a reduction among moderate users, while the proportion of frequent users remained relatively stable throughout the same period.

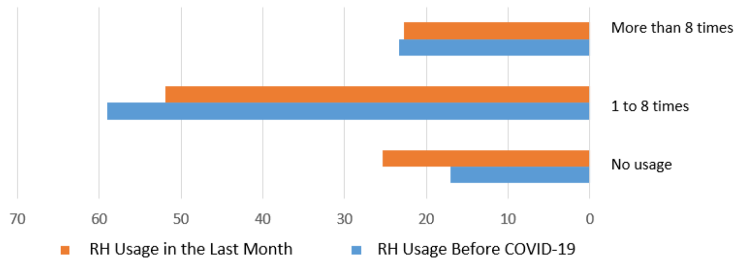


Fig. 4. Frequency of internet taxi usage.

5.2. Psychological variables

Fig. 5 portrays a summary of the level-of-importance questions made by the respondents. A substantial proportion of the respondents expressed significant concerns regarding parking availability during urban travel, highlighting the urgent need for adequate parking infrastructure across all districts of Mashhad. Furthermore, over 50% of the respondents agreed with statements emphasizing the importance of general safety during travel, additionally, more than 75% highlighted the necessity of always being aware of their location throughout a trip. Also, the majority of respondents reported difficulty in trusting strangers, reflecting security concerns. Additionally, the majority of participants are technologically proficient and remain informed about modern advancements. Knowing the cost of travel beforehand is a priority for most respondents. However, within this group, only 35.9% strongly agreed, and 26.7% agreed that they always opt for the most economical mode of transportation.

Most of the respondents preferred using pedestrian bridges when available, and 70% considered motorcycles to be a hazardous mode of transportation. The majority of the respondents are reluctant to share rides with other passengers, even if it would reduce travel costs, emphasizing the importance placed on privacy during travel. Less than half of the respondents perceived the facilities at stations and the fleets' conditions as inadequate. Additionally, 78.1% found buses and subways to be overcrowded and unpleasant.

Data collection for this study occurred during the COVID-19 pandemic and the cold seasons of 2021–2022. As expected, a significant proportion of the respondents avoided crowded places, gatherings, and similar activities to minimize the risk of contracting the virus. Most of the respondents exhibited environmentally conscious attitudes. Specifically, 72.2% believed that environmental protection efforts should persist; similarly, 69.9% expressed a willingness to pay more for environmentally friendly products. 75% of the respondents identified traffic congestion as a major cause of inconvenience. Congestion and pollution in central urban districts are significant concerns, as the majority of the respondents rely more on private vehicles than on public transportation.

Fig. 5. Aggregate shares of responses to the level-of-importance questions from the respondents.

5.3. EFA

The inclusion of numerous correlated attitudinal variables in behavioral models introduces complexity. To address this, exploratory factor analysis is employed to reduce observed variables into fewer uncorrelated latent constructs (factors). This section applies principal components analysis to distill the respondents' attitudes into interpretable dimensions, generating factor scores for integration into DCMs. Table 2 presents the factors identified by loadings exceeding 0.4. After pruning variables with weak loadings (<0.4), six factors were explored, namely (1) Parking and congestion concerns, (2) Safety and security concerns, (3) Concerns about the pandemic, (4) Privacy concerns regarding public transit, (5) Economic concerns, and (6) Variety-seeking perceptions.

Table 2. Results of the EFA.

Component items	Mean	Standard deviation	Communality	1	2	3	4	5	6
I often worry about finding a parking spot during urban trips.	3.85	1.064	0.560	0.704					
Constructing parking lots in city centers should be a priority.	4.43	0.834	0.543	0.601					
Traffic congestion is a major problem for me, and it bothers me.	4.09	1.001	0.489	0.570					
Time spent on daily commutes is wasted time.	3.68	1.204	0.485	0.559					
If a pedestrian bridge exists, I usually use it to cross the street.	4.24	1.060	0.446		0.595				
Being able to track my location at any moment during a trip is important to me.	4.25	0.947	0.462		0.569				
I find it hard to trust strangers.	4.16	0.915	0.447		0.555				
I always choose the shortest route, even if it is only two or three minutes shorter.	3.90	1.062	0.498		0.444				
I believe motorcycles are dangerous vehicles.	3.93	1.225	0.426		0.433				
Due to not contracting the pandemic, I have almost completely stopped visiting relatives, going to cafés, restaurants, etc.	3.77	1.201	0.653			0.795			
Avoiding the pandemic due to sensitivity, weakness, or illness of my relatives is very important to me.	4.40	0.809	0.626			0.746			
Buses and subways are usually crowded, and this is bothersome.	4.15	0.994	0.369			0.407			
Using public transportation with a friend, spouse, or children is difficult.	3.30	1.407	0.569				0.654		
I prefer living in a large house, even if it is far from public transportation and places I frequently visit.	3.26	1.344	0.456				0.636		
I am not interested in sharing travel with other passengers, even if it reduces travel costs.	3.58	1.298	0.542				0.622		
I try to choose the cheapest mode of transportation.	3.73	1.249	0.700					0.827	
Knowing the cost of travel before starting a trip is important to me.	4.51	0.747	0.584					0.688	
I believe having new experiences is important, and I am always trying new things.	3.79	1.009	0.560						0.740
I am so busy that I do not get to complete many of my daily tasks.	3.18	1.208	0.516						0.647

Table 3 presents the eigenvalues for the factors. As observed, in the final EFA, a total of six latent factors with eigenvalues greater than 1 have been identified. Additionally, the percentage of variance explained by each factor and the total variance explained for each category are provided. The explained variance represents the proportion of total variance accounted for by each extracted factor, indicating its relative contribution to explaining the observed data. Factors with higher explained variance capture a greater share of the information contained in the original variables. The cumulative explained variance reflects the overall proportion of total variance explained by all retained factors and provides an indication of how well the extracted factor structure represents the underlying data.

5.4. Model estimation

The modeling process was conducted in Stata software, evaluating a range of model specifications. Ultimately, the Ordered Probit Model with the highest content validity and statistical significance was selected. Additionally, the average marginal effects quantify the percentage change in the probability of choosing RH associated with each variable. This section introduces the variables used in the modeling process and outlines the steps taken to construct the model in Stata software. Following this, the results of the modeling are presented, along with a detailed discussion of their implications.

The dependent variable in the RH modeling is the frequency of using RH after the pandemic. Since the dependent variable in choice models must be discrete, it was grouped into intervals. The grouping was determined by evaluating the statistical and content significance of the model. The three ordered groups of the dependent variable demonstrated the highest significance in the ordered probit model. The classification thresholds were determined based on an iterative modeling process and the observed distribution of the data, ensuring meaningful behavioral segmentation of users. The independent variables encompass observed variables, including demographic attributes (e.g., age, gender, income), travel-related variables (e.g., access to personal vehicles, trip purpose), and environmental conditions (e.g., proximity to public transit, traffic congestion), as well as factors for attitudinal factors. A detailed description of the estimated coefficients is provided in Table 4.

The pseudo R^2 indicates an acceptable level of explanatory power for behavioral modeling; however, it also suggests that additional relevant factors may further improve the model's performance in future research. A diagnostic assessment showed that all Variance Inflation Factor (VIF) values were below 2.75, confirming the absence of significant collinearity among the predictors.

Table 3. Eigenvalues for the factors.

Component	Without factor rotation			With orthogonal factor rotation		
	Eigenvalue	Variance percentage	Cumulative variance percentage	Eigenvalue	Variance percentage	Cumulative variance percentage
1	3.646	19.188%	19.188%	2.108	11.094%	11.094%
2	1.443	7.594%	26.782%	1.743	9.171%	20.266%
3	1.385	7.289%	34.070%	1.666	8.767%	29.032%
4	1.205	6.344%	40.414%	1.533	8.066%	37.098%
5	1.145	6.027%	46.441%	1.487	7.824%	44.923%
6	1.108	5.834%	52.275%	1.397	7.353%	52.275%

Table 4. Ordered probit results.

Variable	Coefficient	T-statistic	P-value
AGE (reference group: above 50 years)			
Below 18 years	0.241	0.86	0.390
18–24 years	0.442	2.09	0.037**
24–30 years	0.766	2.91	0.004***
30–40 years	0.648	3.12	0.002***
40–50 years	0.226	1.11	0.266
Gender (reference group: female)			
Male	-0.641	-6.28	0.000***
Educational attainment (reference group: postgraduate degree)			
Less than high school diploma	0.099	0.36	0.723
High school diploma or associate degree	0.386	1.94	0.052*
Bachelor's degree	0.246	1.59	0.111
Monthly income (reference group: less than 30 MR)			
30-80 MR	0.005	0.03	0.974
80-120 MR	-0.124	-0.61	0.545
More than 120 MR	-0.010	-0.03	0.978
Prefer not to answer	-0.271	-2.05	0.041**
Type of vehicle (reference group: car)			
Motorcycle	-0.153	-0.59	0.556
Motorcycle and car	0.336	1.48	0.139
No vehicle	0.565	4.61	0.000***
Trip purpose (reference group: leisure)			
Work	0.893	3.28	0.001***
Education	0.697	2.08	0.037**
Returning home	0.528	2.01	0.045**
Personal errands	0.365	1.36	0.174
Multiple purposes	0.615	2.44	0.015**
Workplace/study location in traffic-restricted zone	0.379	3.25	0.001***
Preferred time of RH use (reference group: during COVID restrictions)			
Morning	0.505	1.89	0.059*
Afternoon	0.634	2.67	0.008***
Various times	0.525	2.24	0.025**
Using RH for package delivery	0.359	3.51	0.000***

Preferred trip duration (reference group: trips longer than 20 minutes)			
Trips shorter than 20 minutes	0.305	1.88	0.060*
No preference	0.358	2.44	0.015**
Walking distance to nearest public transit (reference group: less than 5 minutes)			
5–10 minutes	0.352	3.10	0.002***
10–15 minutes	0.349	1.95	0.051*
More than 15 minutes	0.049	0.26	0.796
Concerns & perceptions			
Parking and congestion concerns	0.083	1.54	0.123
Safety and security concerns	-0.036	-0.74	0.462
Concern about the pandemic	-0.126	-2.47	0.013**
Privacy concerns regarding public transit	0.092	1.84	0.066*
Economic concerns	-0.075	-1.50	0.134
Variety-seeking perceptions	-0.034	-0.68	0.497

Model fit statistics: Pseudo R^2: 0.167; Log likelihood: -515.507; Chi^2 (37): 206.35; p-value (prob > chi^2): 0.000; $p < 0.01$ (highly significant), $p < 0.05$ (statistically significant), $p < 0.10$ (marginally significant)

5.4.1. Marginal effects

The marginal effects were calculated for three groups of the dependent variable, including (1) No use of RH, (2) Moderate use of RH and (3) High use of RH. However, the average marginal effects for the moderate use group were not statistically significant, unlike those for the no use and high use groups. As a result, the marginal effects for the moderate use group are not discussed further in this section. The marginal effects for the no use and high use groups are presented in Table 5. These results provide insights into the factors influencing the likelihood of individuals either not using RH at all or using them frequently, offering valuable information for policymakers and service providers. Table 5 presents the average marginal effects of variables for groups with no usage and high usage of RH services. It shows how different factors (such as age, gender, education, income, vehicle ownership, trip purpose, and travel preferences) influence the likelihood of using RH services either very frequently or not at all.

5.4.2. Attitudinal factors

In the model, two attitudinal factors, namely concern about the pandemic and privacy regarding public transportation, demonstrated strong statistical significance, and at a lower level of confidence, parking and economic concerns were significant. Individuals who are more concerned about the pandemic and prioritize cost-effectiveness are less inclined to use RH. Additionally, individuals who value privacy and have a negative perception of public transportation are more likely to use RH. Two factors, diversity seeking and security, did not significantly influence the frequency of RH usage.

5.4.3. Demographic variables

According to Table 5, individuals aged 18-24, 24-30, and 30-40 are 9.6%, 18.5%, and 15.1% more likely, respectively, to use RH frequently. Gender also influenced RH usage; Being male reduces the probability of frequent RH use by 15.5%. Educational attainment impacted usage as well; Individuals with a high school diploma or associate degree are more inclined (9.2%) to use RH compared to those with a postgraduate degree. Contrary to expectations, monthly income and employment status did not significantly affect RH demand. The only significant coefficient related to income was the unwillingness to report monthly income, which negatively impacted the likelihood of using RH. Vehicle ownership was another significant factor; Not having access to a vehicle increases the probability of frequent RH use by 14.8. Social media usage did not show statistical significance in the model.

5.4.4. Travel-related variables

Trip purpose significantly influenced RH usage. According to Table 5, work, education, and returning home trips increase the likelihood of using RH by 19.6%, 14.3%, and 10.2%, respectively. Time of travel also played a role; Traveling in the afternoon increases the likelihood of using RH by 13.1%, while morning travel increases it by 9.9%. Additionally, trips lasting less than 20 minutes increase the likelihood of RH use by 6.7%. Using RH for transporting goods also increases the likelihood of usage by 8.6%.

5.4.5. Environmental variables

The location of the workplace or educational institution within the traffic-restricted area increases the likelihood of RH usage by 9.6%. However, residing within the traffic-restricted area did not show statistical significance. Regarding proximity to public transportation, walking distances of up to 10 minutes increased the likelihood of using RH, but walking distances exceeding 10 minutes had no significant effect.

Table 5. Average marginal effects of variables on high and no usage of RH services.

Variable	High usage group			No usage group		
	dy/dx	T-statistic	Significance level	dy/dx	T-statistic	Significance level
Age (reference group: above 50)						
Below 18	0.049	0.84	0.401	-0.071	-0.87	0.383
18–24	0.096	2.18	0.029**	-0.124	-2.02	0.043**
24–30	0.185	2.74	0.006***	-0.196	-3.03	0.002***
30–40	0.151	3.18	0.001***	-0.172	-3.02	0.003***
40–50	0.045	1.12	0.262	-0.066	-1.10	0.269
Gender (reference group: female)						
Male	-0.155	-6.36	0.000***	0.168	6.31	0.000***
Educational attainment (reference group: postgraduate degree)						
Less than high school diploma	0.021	0.35	0.724	-0.028	-0.36	0.721
High school diploma or associate degree	0.092	1.99	0.047**	-0.100	-1.90	0.057*
Bachelor’s degree	0.056	1.63	0.103	-0.067	-1.55	0.120
Monthly income (reference group: less than 30 MR)						
30-80 MR	0.001	0.03	0.974	-0.001	-0.03	0.974
80-120 MR	-0.031	-0.61	0.539	0.031	0.60	0.552
More than 120 MR	-0.002	-0.03	0.978	0.002	0.03	0.978
Prefer not to answer	-0.064	-2.05	0.040**	0.071	2.04	0.041**
Vehicle type (reference group: car)						
Motorcycle	-0.032	-0.62	0.533	0.044	0.57	0.566
Motorcycle + Car	0.083	1.36	0.173	-0.086	-1.62	0.105
No vehicle	0.148	4.33	0.000***	-0.134	-5.06	0.000***
Trip purpose (reference group: leisure)						
Work	0.196	3.82	0.000***	-0.241	-3.02	0.003***
Education	0.143	2.03	0.043**	-0.197	-2.11	0.035**
Returning home	0.102	2.28	0.023**	-0.155	-1.93	0.054*
Personal errands	0.066	1.48	0.139	-0.110	-1.33	0.184
Multiple purposes	0.123	2.98	0.003***	-0.177	-2.28	0.022**
Preferred time for RH (reference group: during COVID restrictions)						
Morning	0.099	2.01	0.044**	-0.146	-1.84	0.066*
Afternoon	0.131	3.20	0.001***	-0.178	-2.48	0.013**
Various times	0.104	2.65	0.008***	-0.151	-2.10	0.036**
Trip duration for RH (reference group: trips longer than 20 minutes)						
Trips shorter than 20 minutes	0.067	1.93	0.053*	-0.084	-1.85	0.064*
No preference	0.080	2.64	0.008***	-0.097	-2.33	0.020**
Work/study location in traffic-restricted area	0.096	3.10	0.002***	-0.092	-3.46	0.001***
Use of RH for cargo/packages	0.086	3.52	0.000***	-0.093	-3.50	0.000***
Walking distance to nearest public transit stop (reference group: less than 5 minutes)						
5–10 minutes	0.086	3.02	0.003***	-0.088	-3.20	0.001***
10–15 minutes	0.086	1.82	0.069*	-0.088	-2.10	0.036**
More than 15 minutes	0.011	0.26	0.798	-0.013	-0.26	0.794

Significance levels: p < 0.10 */ p < 0.05 **/ p < 0.01 ***
Interpretation:
Positive coefficients → Increase in probability of using RH frequently – Negative coefficients → Increase in probability of not using RH

6. Discussion

Given that the factors influencing the choice of RH services in developing countries have rarely been analyzed, this study aims to fill this research gap. This section provides a detailed discussion of the results of demand modeling and discusses explanatory variables.

6.1. Latent attitudinal factors

Previous studies have reported that attitudinal variables such as being technology-friendly, environmentally friendly, variety-seeking, concerns about security, cost-effectiveness, and privacy significantly impact the willingness to use RH services [4, 31, 33, 51, 52].

Among previously introduced attitudinal factors, variety-seeking, security, economic concerns, and parking and traffic congestion were statistically insignificant at the 90% confidence level. The insignificance of some psychological factors may reflect contextual and socio-spatial differences compared to previous studies.

Our findings, similar to Baghestani et al. [53] reveal that concerns about the pandemic reduce the likelihood of using RH services. Given the pandemic, it was expected that individuals would prefer to use more RH services. However, since most survey participants were from affluent urban districts of Mashhad (districts 9, 10 and 11) where car ownership is high, they preferred using their private vehicles instead of RH during the pandemic. Car ownership and its negative effects on RH usage have been reported in numerous studies [23, 42, 54].

The consideration of RH for travel is influenced by total costs [55]. In a study conducted before the pandemic, cost-effectiveness was reported as a significant positive factor in the adoption of RH in Tehran, Iran [4]. According to our results, economic concerns reduce the tendency to use RH; This may be due to the economic crises caused by the pandemic and economic instability in Iran, leading cost-conscious respondents to maximize savings across all expenses. Similarly, Azimi et al. stated that individuals concerned about affordability are less likely to use such services [39].

Contrary to the past studies, variety-seeking and safety and security did not significantly impact RH usage in this study [14, 17, 33, 34, 51]. RH services have been available in Iran for a decade, making them a common mode of urban transportation, and the lack of an effect from variety-seeking may stem from the diminished novelty of these services.

In Iran, users of RH services typically rely on public, semi-public, or non-private transportation and do not own personal vehicles [4, 24]. Despite competition between RH and public transportation due to short commute times and most respondents' proximity to public transit hubs like bus stations, concerns about safety and security do not significantly influence the frequency of RH usage. This contrasts with findings from other regions which attitudinal factors such as reliability, safety, and convenience have been shown to reduce frequent use of RH services [56]. In Mexico City, social issues like perceived crime rates and risks of sexual harassment in public transportation strongly shape the adoption of RH platforms [57].

Previous studies indicate that parking unavailability, high parking costs, and privacy concerns increase the likelihood of using RH [4, 17, 33]. Contrary to expectations, concerns about parking and traffic congestion had no effect on RH usage in our study. Given the limited parking spaces, traffic issues in most urban districts, and the stressful nature of driving in congestion, these factors were expected to influence demand.

However, privacy concerns and negative attitudes toward public transportation significantly increased the likelihood of using RH. This may be due to factors such as overcrowding, inadequate facilities, the pandemic [58], and the lack of cleanliness in public transit and traditional taxis.

6.2. Demographic variables

Findings indicate that RH services in Mashhad are predominantly used by young women without private vehicles, holding high school diplomas or bachelor's degrees. However, due to the online data collection method, despite validity checks, potential biases may exist.

Model estimations show that household size, number of licensed family members, marital status, monthly income, employment status, and social media usage were statistically insignificant at the 90% confidence level. About half of the demographic variables were statistically insignificant in the demand model for RH in Mashhad. This represents that, compared to attitudinal, travel-related, and environmental variables, demographic variables have less influence on RH adoption. Similarly, other studies reported that environmental and Attitudinal factors had a greater effect on the frequency of RH use than demographic variables [28, 56].

Contrary to previous studies, income and employment status had no significant impact on RH usage. These differences may reflect contextual variations across countries in terms of socioeconomic conditions, technology adoption, and the maturity of ride-hailing services. In the Iranian context, ride-hailing services have become widely accessible across different social groups, which may reduce the explanatory power of income and employment characteristics. Similarly, the widespread use of smartphones and mobile applications may limit the influence of social media usage as a distinguishing factor. These findings highlight the importance of local context when interpreting ride-hailing adoption behavior. However, employment status was a significant factor in Tehran's RH usage, possibly due to higher service costs in Tehran, leading students, unemployed individuals, or retirees to prefer traditional taxis [42].

Previous studies have shown that mobile phone usage and interest in technology directly influence the adoption of RH services [4, 24, 31]. However, contrary to expectations, the current study found no significant relationship between social media usage and RH adoption. This discrepancy may stem from methodological limitations, such as insufficient demographic categorization in the questionnaire design.

As expected, young individuals (18-40 years old) were more inclined to frequently use RH. This is likely due to their familiarity with mobile applications and RH services, as well as their vehicle ownership rates. Additionally, results indicate that gender is a significant determinant of ride-hailing usage, with women exhibiting a higher likelihood of using these services compared to men. This finding is consistent with previous studies highlighting gender-based differences in urban mobility patterns [58]. In particular, it may reflect the role of contextual factors in shaping women's driving experiences and in contributing to variations in perceived stress levels [23, 24].

Individuals with high school diplomas or associate degrees were more likely to use RH compared to highly educated individuals. This may be due to lower car ownership rates among students and self-employed individuals. Some studies found that educational attainment had no impact on travel mode choice [21, 42], while others reported that highly educated individuals were major users of RH in the US [19, 20, 31].

As expected, the absence of private vehicles significantly increased the likelihood of using RH due to mobility challenges, public transportation limitations, and security concerns, aligning with previous research [7, 23, 35, 42].

6.3. Travel-related variables

Findings indicate that all travel-related variables significantly influence the demand for RH services. Among these, work trips exert the greatest impact on increasing the likelihood of RH usage. In addition to work trips, educational trips and home-returning trips similarly contribute to a higher tendency to use RH services. This outcome may be attributed to the fact that RH services offer significant time savings compared to other transportation modes (such as traditional taxis and public transit), enabling passengers to reach their workplace or educational institution without delay. Moreover, since most administrative, commercial, educational, and healthcare centers were established in the central districts of Mashhad, these areas now experience traffic congestion, parking shortages, and are subject to traffic restrictions, all of which contribute to the high utilization of RH services. Additionally, due to exhaustion from work or study, travelers tend to prioritize getting home as quickly as possible, opting for the fastest and most convenient mode of transportation after private vehicles. However, for leisure trips, passengers often prefer private vehicles. The finding regarding the strong influence of work trips on increased RH usage aligns with previous studies [4, 17, 42, 59].

Contrary to expectations, findings indicate that during the pandemic restriction hours (10 PM to 4 AM), RH usage was at its lowest. This outcome may stem from the fact that individuals engaging in late-night activities such as social gatherings, dining out, or entertainment are generally from wealthier demographics and prefer to use their own comfortable vehicles. During other times of the day, the highest likelihood of RH usage occurs in the afternoons and mornings, respectively. This result may be explained by the tendency of citizens to schedule shopping, social events (e.g., dining out, visiting coffee shops, or seeing relatives), and medical appointments for the afternoon. Factors such as limited parking, traffic-related stress, safety concerns when waiting for transportation at night [7], and overcrowded public transit contribute to the preference for RH during these times. Similarly, other studies have reported that RH is frequently used for social-related travel, such as trips to bars, restaurants, and concerts [7]. Acheampong et al. also found that the majority of RH trips occur in the evening, with 31% between 12 PM and 5 PM and 29% between 5 PM and 11 PM [17].

Prior research reported that RH is often used for short-duration trips (10 to 30 minutes) [17]. Similarly, in the present study, trips lasting shorter than 20 minutes were associated with an increased likelihood of RH usage. This preference may arise from the time-consuming and cumbersome process of retrieving a private vehicle from a parking space, navigating traffic, and searching for parking at the destination. Other studies have also cited time savings and stress reduction as key reasons for choosing RH [7, 25]. Additionally, longer trips increase their costs and may lead to discomfort from prolonged sitting in low-quality vehicles (as RH fleets in Iran typically consist of older, less-maintained cars). In such cases, individuals may prefer public or private transportation.

Results from Table 5 indicate that using RH services for freight and package delivery significantly increases the likelihood of frequent RH usage. This suggests that individuals without access to a private vehicle prefer RH for its convenience. Additionally, RH services allow for package delivery without requiring the sender's presence during the trip, offering time savings and better tracking capabilities compared to alternative methods.

6.4. Environmental variables

Previous studies have identified several environmental factors that influence RH usage, including land use type, private car accessibility, public transportation availability, urban versus suburban residence, and population and employment density [4, 19, 31, 32, 54]. Circella et al. even reported that environmental factors have a greater impact than demographic variables on RH usage [28].

As expected, working or studying in districts subject to traffic restrictions is one of the most influential variables increasing RH usage. In these central urban districts, parking prohibitions, congestion, and travel restrictions make private vehicle use highly impractical for employees and students, leading them to opt for alternative modes. Compared to traditional public transportation, RH services offer greater comfort, shorter travel times, cost-effectiveness, privacy, technology-driven convenience, and enhanced security for these individuals [4, 42].

Some studies have reported that high accessibility to public transit (buses and metro) negatively affects RH usage [21, 23]. Conversely, other studies suggest that proximity to metro stations correlates with increased RH adoption [4, 32]. Findings from the present study indicate that when walking distance to a public transit station exceeds five minutes, passengers are more likely to use RH services. Notably, when walking time exceeds ten minutes, no significant impact on RH adoption is observed. Furthermore, given the high population density in restricted-traffic areas, it was expected that residents in these districts would have a higher tendency to use RH services [54]. However, the model does not indicate a significant relationship between living in a traffic-restricted area and RH usage.

7. Conclusions

RH have become a primary mode of transportation in Iran and developing countries. Despite their growing importance, there is a lack of sufficient research on RH usage in Iran. This study aimed to estimate the demand for RH in Iran (Mashhad) and provide a deeper understanding of the travel behavior of RH users. The findings reveal that urban transportation users are more likely to rely on RH when they do not have access to a personal vehicle. Therefore, identifying and analyzing the factors influencing RH usage is crucial for urban planning and transportation policy. As in previous studies, these factors were categorized into four groups, including demographic, travel-related, environmental characteristics, and latent factors. In this study, a stated preference online questionnaire was designed to gather comprehensive data on Mashhad citizens, covering these factors. A total of 619 questionnaires were collected. Due to pandemic-related restrictions and the online nature of data collection, the sample was mainly concentrated in the western and relatively more affluent districts of Mashhad (9, 10, and 11).

Results highlight the significance of both observed variables and latent factors of RH demand. Among the observed variables, demographic characteristics were found to be less influential compared to travel-related and environmental characteristics. More than half of the demographic variables, such as marital status, employment status, physical disability or illness, social media usage, family size, and the number of licensed drivers in the household, showed no significant impact on RH usage. Additionally, the demographic variables of education and income were statistically significant only at specific levels. Variables such as age, gender, and vehicle ownership were significant, and young women without access to a vehicle showed a high tendency to use RH frequently. Travel-related variables, including trip purpose, preferred travel time, and use of RH for transporting goods, were all significant and influential.

The findings indicate that travel-related characteristics are more important than other factors in determining RH usage. Business trips, educational trips, and returning home increased the probability of using RH, respectively. In contrast to other studies, leisure trips had a negative effect on RH usage. Regarding preferred travel times, traveling in the afternoon and morning increased the probability of using RH, respectively. Surprisingly, the pandemic restriction times (10 PM to 4 AM) had a negative effect on RH usage. Additionally, trips lasting shorter than 20 minutes and the use of RH for transporting goods were significant factors, though they have received less attention in prior literature. Among environmental variables, residing within the Mashhad traffic-restricted area did not significantly affect RH usage. However, being within the traffic-restricted area for work or school and increasing the walking distance to public transportation from 5 to 10 minutes increased the probability of using RH, respectively.

This study identified six latent factors through EFA, including concerns about (1) the pandemic, (2) parking and congestion, (3) safety and security, (4) privacy of public transit, (5) cost-effectiveness, and (6) variety seeking. Among these, concerns about the pandemic and privacy were the most significant ones. The factors of the pandemic and economic concerns had negative effects on RH demand. Conversely, concerns about public transportation, privacy and parking, and congestion increased the likelihood of using RH.

Overall, it should be noted that the study is based on an online, non-probability sample collected during the COVID-19 pandemic and concentrated in the western districts of Mashhad, which are generally characterized by relatively higher socioeconomic status. Consequently, younger, more educated, and technology-oriented individuals may be overrepresented in the sample. In addition, the data were collected during a unique phase of the COVID-19 pandemic, which may have influenced travel behavior and ride-hailing usage patterns. Therefore, the findings should be interpreted in light of both spatial and temporal contexts and may not be fully generalizable to the entire population of Mashhad, other Iranian cities, or different urban environments. Nevertheless, several of the identified behavioral relationships are expected to remain relevant under normal conditions, although their magnitude may vary over time.

Statements & declarations

Author contributions

Mahdi Nosrati: Data curation, Formal analysis, Investigation, Software, Visualization, Writing - Original Draft.

Amir Abbas Rassafi: Conceptualization, Investigation, Methodology, Supervision, Validation, Writing - Review & Editing.

Hossein Rahimi Farahani: Formal analysis, Investigation, Methodology, Validation, Writing - Review & Editing.

Funding

Funding information is not available.

Data availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to privacy and ethical considerations, individual-level responses are not publicly shared.

Declarations

The authors declare no conflict of interest.

References

[1] Soltanzadeh, J., Blind, K., Elyasi, M. Exploring how regulators face platform business issues in the lifecycle stages: Evidence of Iranian ride-hailing platform business. Telecommunications Policy, 2023; 47: 102582. doi:10.1016/j.telpol.2023.102582.

[2] Oviedo, D., Sabogal-Cardona, O., Scholl, L. Transport and social (dis)advantage in ride-hailing. Area Development and Policy, 2025; 10: 295–325. doi:10.1080/23792949.2024.2404024.

[3] Martins, D. L., Meira, L. H., Costa, C. S., Pitombo, C. S. An overview of the impacts of ridesourcing in developing countries: Main topics and questions for future research. Latin American Transport Studies, 2024; 2: 100017. doi:10.1016/j.latran.2024.100017.

[4] Etminani-Ghasrodashti, R., Hamidi, S. Individuals' Demand for Ride-hailing Services: Investigating the Combined Effects of Attitudinal Factors, Land Use, and Travel Attributes on Demand for App-based Taxis in Tehran, Iran. Sustainability, 2019; 11: 5755. doi:10.3390/su11205755.

[5] Nie, Y. How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China. Transportation Research Part C: Emerging Technologies, 2017; 79: 242–256. doi:10.1016/j.trc.2017.03.017.

[6] Tirachini, A. Ride-hailing, travel behaviour and sustainable mobility: an international review. Transportation, 2020; 47: 2011–2047. doi:10.1007/s11116-019-10070-2.

[7] Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transport Policy, 2016; 45: 168–178. doi:10.1016/j.tranpol.2015.10.004.

[8] Snapp. Snapp group annual report. Tehran (IR): Snapp; 2020. Report No.: 2.

[9] Said, M., Soria, J., Stathopoulos, A. Shifting Mobility Behaviors in Unprecedented Times: A Multigroup MIMIC Model Investigating Intentions to Use On-Demand Ride Services During the COVID-19 Pandemic. Transportation Research Record, 2023; 2677: 704–722. doi:10.1177/03611981221147520.

[10] Ke, J., Yang, H., Zheng, H., Chen, X., Jia, Y., Gong, P., Ye, J. Hexagon-Based Convolutional Neural Network for Supply-Demand Forecasting of Ride-Sourcing Services. IEEE Transactions on Intelligent Transportation Systems, 2019; 20: 4160–4173. doi:10.1109/TITS.2018.2882861.

[11] Weyl, E. G. A Price Theory of Multi-sided Platforms. American Economic Review, 2010; 100: 1642–72. doi:10.1257/aer.100.4.1642.

[12] Armstrong, M. Competition in Two-Sided Markets. RAND Journal of Economics, 2006; 37: 668–691. doi:10.1111/j.1756-2171.2006.tb00037.x.

[13] Wang, H., Yang, H. Ridesourcing systems: A framework and review. Transportation Research Part B: Methodological, 2019; 129: 122–155. doi:10.1016/j.trb.2019.07.009.

[14] Alemi, F., Circella, G., Mokhtarian, P., Handy, S. What drives the use of ride-hailing in California? Ordered probit models of the usage frequency of Uber and Lyft. Transportation Research Part C: Emerging Technologies, 2019; 102: 233–248. doi:10.1016/j.trc.2018.12.016.

[15] Greene, W. H. Econometric Analysis. 7th ed. Boston (MA): Pearson; 2012.

[16] Train, K. Discrete choice methods with simulation. 2nd ed. Cambridge (UK): Cambridge University Press; 2009.

[17] Acheampong, R. A., Siiba, A., Okyere, D. K., Tuffour, J. P. Mobility-on-demand: An empirical study of internet-based ride-hailing adoption factors, travel characteristics and mode substitution effects. Transportation Research Part C: Emerging Technologies, 2020; 115: 102638. doi:10.1016/j.trc.2020.102638.

[18] Tirachini, A., del Río, M. Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour. Transport Policy, 2019; 82: 46–57. doi:10.1016/j.tranpol.2019.07.008.

[19] Dias, F. F., Lavieri, P. S., Garikapati, V. M., Astroza, S., Pendyala, R. M., Bhat, C. R. A behavioral choice model of the use of car-sharing and ride-sourcing services. Transportation, 2017; 44: 1307–1323. doi:10.1007/s11116-017-9797-8.

[20] Vinayak, P., Dias, F. F., Astroza, S., Bhat, C. R., Pendyala, R. M., Garikapati, V. M. Accounting for multi-dimensional dependencies among decision-makers within a generalized model framework: An application to understanding shared mobility service usage levels. Transport Policy, 2018; 72: 129–137. doi:10.1016/j.tranpol.2018.09.013.

[21] Tang, B.-J., Li, X.-Y., Yu, B., Wei, Y.-M. How app-based ride-hailing services influence travel behavior: An empirical study from China. International Journal of Sustainable Transportation, 2020; 14: 554–568. doi:10.1080/15568318.2019.1584932.

[22] Atkinson-Palombo, C., Varone, L., Garrick, N. W. Understanding the Surprising and Oversized Use of Ridesourcing Services in Poor Neighborhoods in New York City. Transportation Research Record, 2019; 2673: 185–194. doi:10.1177/0361198119835809.

[23] Nugroho, S. B., Zusman, E., Nakano, R. Explaining the spread of online taxi services in Semarang, Bogor and Bandung, Indonesia; a discrete choice analysis. Travel Behaviour and Society, 2020; 20: 358–369. doi:10.1016/j.tbs.2020.04.008.

[24] Lesteven, G., Samadzad, M. Ride-hailing, a new mode to commute? Evidence from Tehran, Iran. Travel Behaviour and Society, 2021; 22: 175–185. doi:10.1016/j.tbs.2020.09.006.

[25] Smith, A. Shared, Collaborative and On Demand: The New Digital Economy. Washington, D.C. (US): Pew Research Center; 2016. Report No.: 02.

[26] Hussain, I., Hussain, Q., Dias, C., Bargi, W. A., Ali, N., Abdullah, M., Cheng, L. Exploring Commuter's Preferences and Future Intentions to Use Ride-Sharing: A Case Study From a Developing Country. Journal of Advanced Transportation, 2025; 2025: 5516034. doi:10.1155/atr/5516034.

[27] Sabouri, S., Brewer, S., Ewing, R. Exploring the relationship between ride-sourcing services and vehicle ownership, using both inferential and machine learning approaches. Landscape and Urban Planning, 2020; 198: 103797. doi:10.1016/j.landurbplan.2020.103797.

[28] Circella, G., Alemi, F., Tiedeman, K., Handy, S., Mokhtarian, P. L. The Adoption of Shared Mobility in California and Its Relationship with Other Components of Travel Behavior. Davis (CA): National Center for Sustainable Transportation; 2018. Report No.: 03.

[29] Wali, B., Santi, P., Ratti, C. Modeling ride-hailing and carsharing adoption & use patterns: deciphering the substitutive and complementary impacts of built environment, transit accessibility, & active travel. Sustainable Cities and Society, 2025; 120: 105931. doi:10.1016/j.scs.2024.105931.

[30] Li, W., Ma, J., Cai, H., Chen, F., Qin, W. The role of built environment in shaping reserved ride-hailing services: Insights from interpretable machine learning approach. Research in Transportation Business & Management, 2024; 56: 101173. doi:10.1016/j.rtbm.2024.101173.

[31] Alemi, F., Circella, G., Mokhtarian, P., Handy, S. Exploring the latent constructs behind the use of ride-hailing in California. Journal of Choice Modelling, 2018; 29: 47–62. doi:10.1016/j.jocm.2018.08.003.

[32] Wang, S., Noland, R. B. Variation in ride-hailing trips in Chengdu, China. Transportation Research Part D: Transport and Environment, 2021; 90: 102596. doi:10.1016/j.trd.2020.102596.

[33] Sarjana, S., Ramadan, O., Sisiopiku, V. Analysis of Transportation Users' Preferences and Attitudes for Identifying Micro-Level Determinants of Transportation Network Companies' (TNCs) Growth. Journal of Transportation Technologies, 2020; 10: 251–264. doi:10.4236/jtts.2020.103016.

[34] Bhaduri, E., Pal, S., Goswami, A. K. Analysing heterogeneity in factors affecting adoption of ride-hailing services: a stepwise LCCA-MCDM modelling approach. Transportation, 2024; doi:10.1007/s11116-024-10563-9.

[35] Sikder, S. Who Uses Ride-Hailing Services in the United States? Transportation Research Record, 2019; 2673: 40–54. doi:10.1177/0361198119859302.

[36] Nguyen-Phuoc, D. Q., Su, D. N., Tran, P. T. K., Le, D.-T. T., Johnson, L. W. Factors influencing customer's loyalty towards ride-hailing taxi services – A case study of Vietnam. Transportation Research Part A: Policy and Practice, 2020; 134: 96–112. doi:10.1016/j.tra.2020.02.008.

[37] Anser, M., Ahmed, F., Ali, A. Investigating Factors Impacting the Use of Ride-Hailing Service Applications. International Journal of Management Research and Emerging Sciences, 2024; 14: doi:10.56536/ijmres.v14i3.637.

[38] Ghaffar, A., Mitra, S., Hyland, M. Modeling determinants of ridesourcing usage: A census tract-level analysis of Chicago. Transportation Research Part C: Emerging Technologies, 2020; 119: 102769. doi:10.1016/j.trc.2020.102769.

[39] Azimi, G., Rahimi, A., Asgari, H., Jin, X. Role of Attitudes in Transit and Auto Users' Mode Choice of Ridesourcing. Transportation Research Record, 2020; 2674: 1–16. doi:10.1177/0361198120941841.

[40] Tohidimehr, H., Bakhtiarizadeh, H. Money on the Way: Analyzing the Impact of Ride-Sharing on Labor, Income, and Expenditure in Iran. Income, and Expenditure in Iran (October 22, 2024), 2024; doi:10.2139/ssrn.5166284.

[41] Akbari, M., Amiri, N. S., Zúñiga, M. Á., Padash, H., Shakiba, H. Evidence for Acceptance of Ride-Hailing Services in Iran. Transportation Research Record, 2020; 2674: 289–303. doi:10.1177/0361198120942224.

[42] Edrisi, A., Rezaei, H. Assessment of how Ride-Hailing services affect individuals' travel mode choice in urban transportation. International Journal of Human Capital in Urban Management, 2021; 6: 81–96. doi:10.22034/IJHCUM.2021.01.07.

[43] Koppelman, F., Bhat, C. A Self-Instructing Course in Mode Choice Modelling: Multinomial and Nested Logit Models. Washington, D.C. (US): US Federal Transit Administration - Department of Transportation; 2006. Report No.: 04.

[44] Daykin, A. R., Moffatt, P. G. Analyzing Ordered Responses: A Review of the Ordered Probit Model. Understanding Statistics, 2002; 1: 157–166. doi:10.1207/S15328031US0103_02.

[45] Tabachnick, B. G., Fidell, L. S. Using Multivariate Statistics. 6th ed. Boston (MA): Pearson; 2013.

[46] Williams, B., Onsman, A., Brown, T. Exploratory Factor Analysis: A Five-Step Guide for Novices. Australasian Journal of Paramedicine, 2010; 8: 1–13. doi:10.33151/ajp.8.3.93.

[47] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. Multivariate data analysis. 8th ed. Boston (MA): CENGATE; 2019.

[48] Mirkatouli, J., Samadi, R., Hosseini, A. Evaluating and analysis of socio-economic variables on land and housing prices in Mashhad, Iran. Sustainable Cities and Society, 2018; 41: 695–705. doi:10.1016/j.scs.2018.06.022.

[49] Talebian, K., Riza, M. Mashhad, City of Light. Cities, 2020; 101: 102674. doi:10.1016/j.cities.2020.102674.

[50] Soltanifard, H., Roshandel, T., Ghodrati, S. Assessment and ranking of influencing factors in the relationship between spatial patterns of urban green spaces and socioeconomic indices in Mashhad urban districts, Iran. Modeling Earth Systems and Environment, 2020; 6: 1589–1605. doi:10.1007/s40808-020-00776-z.

[51] Jing, P., Chen, Y., Wang, X., Pan, K., Yuan, D. Evaluating the effectiveness of Didi ride-hailing security measures: An integration model. Transportation Research Part F: Traffic Psychology and Behaviour, 2021; 76: 139–166. doi:10.1016/j.trf.2020.11.004.

[52] Sedighi, M., Parsaeiyan, H., Araghi, Y. An Empirical Study of Intention to Continue Using of Digital Ride-hailing Platforms. The Review of Socionetwork Strategies, 2021; 15: 489–515. doi:10.1007/s12626-021-00098-1.

[53] Baghestani, A., Heshami, S., Mahpour, A. The impact of pandemic experiences on future similar situations: analysis of latent variables and trip purposes in ride-hailing choice. Transportation Planning and Technology, 2026; 49: 691–710. doi:10.1080/03081060.2025.2458606.

[54] Conway, M. W., Salon, D., King, D. A. Trends in Taxi Use and the Advent of Ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. Urban Science, 2018; 2: 79. doi:10.3390/urbansci2030079.

[55] Habib, K. N. Mode choice modelling for hailable rides: An investigation of the competition of Uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation. Transportation Research Part A: Policy and Practice, 2019; 129: 205–216. doi:10.1016/j.tra.2019.08.014.

[56] Loa, P., Nurul Habib, K. Examining the influence of attitudinal factors on the use of ride-hailing services in Toronto. Transportation Research Part A: Policy and Practice, 2021; 146: 13–28. doi:10.1016/j.tra.2021.02.002.

[57] Sabogal-Cardona, O., Oviedo, D., Scholl, L., Crotte, A., Bedoya-Maya, F. Not my usual trip: Ride-hailing characterization in Mexico City. Travel Behaviour and Society, 2021; 25: 233–245. doi:10.1016/j.tbs.2021.07.010.

[58] Firoozi Yeganeh, S., Khademi, N., Farahani, H., Besharat, M. A. A qualitative exploration of factors influencing women's intention to use shared taxis: A study on the characteristics of urban commuting behavior in Iran. Transport Policy, 2022; 129: 90–104. doi:10.1016/j.tranpol.2022.10.001.

[59] Zhen, C. Impact of ride-sourcing services on travel habits and transportation planning [Master's Thesis]. Pittsburgh (PA): University of Pittsburgh; 2015.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Enhancing the Performance of High-Speed Railway Slab-Ballast Transition Zones Using Geotextile Reinforcement: A Dynamic-Equivalent 3D Finite Element Investigation

Nima Moslemi Patrudi [a*], Hossein Ghasemzadeh Tehrani [a], Iman Aghayan [a, b], Mohsen Bagheri [c]

[a] *Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran*
[b] *Department of Civil & Environmental Engineering, University of Wisconsin, Milwaukee, WI 53211, USA*
[c] *Faculty of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran*

ARTICLE INFO

Article history:

Received 29 May 2026

Revised 07 July 2026

Accepted 01 August 2026

Available online 01 April 2027

Keywords:

High-speed railway

Transition zone

Geotextile reinforcement

Finite element analysis

Nonlinear dynamic analysis

ABSTRACT

Transition zones between ballast and slab tracks are critical points in high-speed railways due to abrupt stiffness variations. This study investigates the optimization of speed-induced track responses in railway transition zones using geotextile reinforcement through a three-dimensional finite element dynamic-equivalent analysis. To address the inherent tensile weakness of the ballast layer, various geotextile lengths (6, 9, and 12 m) were modeled under heavy loading at speeds of 160 and 500 km/h. The numerical model was validated against experimental data of geosynthetic-stabilized soil under cyclic loading, showing a deviation of less than 10%. Results indicate that increasing train speed significantly intensifies rail deflection and impact coefficients. However, the inclusion of geotextile over cement-stabilized ballast enhances confinement and tensile strength, leading to a 60% reduction in vertical rail deformation compared to the unreinforced state. Among the tested configurations, a 12-meter geotextile length provided the optimal performance in reducing vertical rail displacement and smoothing the stiffness transition. These findings demonstrate that geotextile reinforcement is a highly effective and practical reinforcement strategy for enhancing the structural integrity and longevity of high-speed railway transition zones.

How to cite this article: Moslemi Patrudi, N., Ghasemzadeh Tehrani, H., Aghayan, I., Bagheri, M. Enhancing the Performance of High-Speed Railway Slab-Ballast Transition Zones Using Geotextile Reinforcement: A Dynamic-Equivalent 3D Finite Element Investigation. Civil Engineering and Applied Solutions. 2027; 3(2): 87–101. doi:10.22080/ceas.2026.31946.1105.

1. Introduction

High-speed railway (HSR) systems, characterized by operating speeds exceeding 160 km/h, have revolutionized modern transportation by demanding unprecedented levels of safety, passenger comfort, and structural precision. The operational integrity of these lines is highly sensitive to the dynamic interaction between the rolling stock and the trackbed, where any sudden variation in support conditions can compromise stability. A fundamental challenge in HSR infrastructure involves managing 'transition zones'-critical interfaces where abrupt changes in vertical track stiffness occur, leading to magnified dynamic loads and accelerated structural degradation. Consequently, optimizing the mechanical response at these discontinuities is vital for ensuring the long-term reliability and efficiency of high-speed rail networks. The transition between flexible soil subgrades and rigid bridge abutments represents a critical discontinuity in railway tracks, primarily due to the sharp contrast in structural stiffness. Such disparities are a leading cause of geometric anomalies and uneven track settlement. Engineering practice dictates the use of specialized approach sections to harmonize these stiffness variations, providing a smoother mechanical transition and reducing the impact of differential

* Corresponding author.
E-mail addresses: moslemi_nima@shahroodut.ac.ir (N. Moslemi Patrudi).

https://doi.org/10.22080/ceas.2026.31946.1105
ISSN: 3092-7749/

settlement across the interface.

Shan et al. [1] studied the use of the Finite Element Method (FEM) to investigate the dynamic performance of two prevalent bridge-subgrade transition designs in high-speed railways: the two-part and the inverted trapezoid sections. Comparative analysis of track-subgrade displacement, acceleration, and surface stress reveals that the two-part transition zone exhibits superior dynamic stability. Furthermore, the results highlight a critical 3-meter zone from the bridge abutment where abrupt response changes occur, suggesting a minimum recommended transition length of 3 meters for enhanced structural integrity. Shih et al. [2] investigated the application of time-domain finite element (FE) models for representing moving loads on flexible railway tracks. Their systematic study compared various mesh geometries, boundary conditions, and soil damping models to provide comprehensive guidance for selecting appropriate FE parameters. They demonstrated that while infinite elements provide accuracy for harmonic loads, a sufficiently wide cuboid mesh combined with mass-proportional damping is more computationally efficient for moving load problems. Furthermore, they established that the model length required to achieve steady-state response significantly increases – up to 130 m – as the load approaches critical speeds in homogeneous grounds.

According to Chen and Zhou [3], the stiffness matching between high-rigidity fasteners and layered subgrade is essential for the dynamic stability of ballastless tracks. Using a 3D FEM model of the Beijing-Shanghai line, this study finds that while higher subgrade moduli reduce total displacement, fastener stiffness primarily impacts rail-level response. Notably, the modulus ratio between subgrade layers is identified as the dominant factor governing surface displacement under high-frequency loading. Connolly and Costa [4] identified a novel dynamic load amplification mechanism triggered by the interaction between vehicle axle configuration and the soil-guideway's vibration characteristics. Their numerical modeling, validated by field data, demonstrates that at speeds exceeding the critical velocity, load spacing sensitivity can increase displacements by up to 400%, significantly overshadowing the traditional critical velocity effects (50-100%). The study highlights the necessity of integrated vehicle-track design for future ultra-high-speed systems like Hyperloop and HSR to mitigate massive resonance effects. Zhou et al. [5] reviewed the key issues of subgrade settlement in high-speed railways built on soft soil, emphasizing that high compressibility and low permeability make post-construction settlement difficult to control and pose serious safety risks. They summarized theoretical and experimental foundation treatment methods for ballasted and ballastless tracks, discussed settlement evolution and control performance, and highlighted the limitations of existing prediction methods and the need for improved settlement control theories and technologies. Charoenwong et al. [6] proposed a novel numerical framework to predict track irregularity evolution by integrating empirical settlement laws with 3D finite element theory and multi-body vehicle-track interaction. Their study emphasizes the necessity of updating the track profile after every load passage to capture the evolving dynamic interaction accurately. The findings demonstrate that neglecting these progressive geometry changes leads to significant modeling errors and that subgrade material properties play a decisive role in long-term plastic settlement.

Wang et al. [7] investigated the mechanisms of railway heave in ballastless tracks induced by subgrade swelling in expansive soil regions. By integrating long-term longitudinal deformation monitoring, laboratory oedometer tests, and coupled FLAC3D-PFC3D numerical simulations, the authors identified that heave is a cumulative response to swelling in both subgrade filler and mudstone foundation layers. They established logistic functional relationships between swelling behavior and in-situ water content, while defining Expassoc and Nelder distribution models to quantify heave displacement based on swelling ratios. This research provides critical reference values for assessing swelling potential and predicting long-term deformations in similar geotechnical railway environments. Charoenwong et al. [8] conducted a comparative analysis of train-induced differential settlement between ballasted and concrete slab tracks at speeds of 200 km/h and 300 km/h. By developing a novel numerical model validated for both trackforms, the study demonstrates that ballasted tracks exhibit significantly higher rates of geometry deterioration compared to slab tracks, especially as line speeds increase. Their findings provide a scientific basis for trackform selection, highlighting that while slab tracks have higher initial costs, their superior resistance to differential settlement makes them more suitable for high-speed routes where maintenance windows are limited. Coelho et al. [9] investigated the transition zone near a concrete culvert, concluding that excessive dynamic displacements often stem from poor sleeper support rather than approach slab stiffness. Their field data also highlighted that ground motion and accelerations escalate as train speeds approach critical thresholds, necessitating enhanced support conditions in transition areas.

Sañudo et al. [10] at the University of Cantabria investigated the impact of train direction and pad positioning on the performance of track transition zones. Their study quantified the maximum vertical stresses and displacements within the ballast (beneath the sleepers) and under the main slab in ballastless track systems. By analyzing these peak responses, the research highlighted how the spatial configuration of track components and the direction of loading cycles significantly influence the stress distribution and long-term settlement at transition interfaces. Sadeghi et al. [11] employed nonlinear dynamic modeling to analyze the seismic response of road embankments, focusing on the role of soil liquefaction and PGA in driving permanent displacements. They evaluated how modifications to slope geometry and induced cracking compromise structural integrity, contributing to improved predictive assessments for the seismic durability of transportation infrastructure. Esen et al. [12] investigated the settlement behavior of a Geosynthetically Reinforced Soil with Retaining Wall (GRS-RW) railway system under static and cyclic loading conditions. A full-scale laboratory model including both concrete slab track and ballasted track was constructed on a 1.2 m subgrade with a frost protection layer. Results showed that the slab track exhibited significantly lower elastic and permanent deformations compared to the ballasted track, with about 25% lower rail displacement amplitude and 6-7 times smaller sleeper displacement under cyclic loading. These findings highlight the superior performance of slab track systems on GRS-RW foundations for railway applications. Jain et al. [13] investigated the persistent problem of degradation in railway transition zones by evaluating commonly used mitigation measures and proposing a new transition structure. The study first assessed traditional solutions, including horizontal and inclined approach slabs and transition wedges, using kinematic response, stress analysis, and an energy-based criterion based on total strain energy minimization. Subsequently, a newly proposed Safe Hull-Inspired Energy Limiting Design (SHIELD) was analyzed using the same evaluation criterion. The results of a two-dimensional finite element analysis indicated that the SHIELD concept effectively manages energy flow in transition zones and provides insight into the limitations of conventional transition structures. The study also demonstrated the robustness of the energy-based criterion and highlighted the potential of the SHIELD concept as a promising solution for improving the design of railway transition zones. Jain et al. [14] examined the challenges associated with railway transition zones caused by sudden stiffness variations and differential settlements when tracks pass over rigid structures such as bridges or culverts. The study evaluated the effectiveness of different sleeper configurations and reduced sleeper spacing in mitigating dynamic amplification in transition zones using a criterion based on the total strain energy in track-bed layers. A two-dimensional finite element model of an embankment-bridge transition was employed, and the effect of hanging sleepers resulting from loss of contact between sleepers and ballast was also investigated. The results indicated that although modifications in sleeper configuration and spacing can influence the track response, these measures alone are insufficient to completely eliminate transition zone effects. Li et al. [15] investigated the dynamic behavior of the vehicle-track system at subgrade-bridge transition zones in seasonal frozen regions, considering environmental temperature changes and wheel-rail irregularities. A coupled environment-vehicle-track-foundation model was developed and validated using monitoring data. The results showed that freeze-thaw cycles can cause frost heave and thaw settlement, leading to significant changes in track deformation and dynamic responses. Freeze-thaw effects and differential settlement were found to have a stronger impact on wheel-rail vertical forces than frost heave alone. Rail distortion near bridge abutments was also identified as a key factor increasing lateral vehicle vibrations in transition zones. Cui et al. [16] investigated the nonlinear interaction between vehicle-track dynamics and structural damage in ballastless tracks under uneven subgrade settlement. A nonlinear vehicle-track-subgrade interaction model including interlayer separation and concrete damage plasticity was developed. The results showed that subgrade settlement can significantly amplify vibrations and accelerate structural damage in ballastless tracks. Settlement wavelength and amplitude were identified as key factors influencing deformation and damage evolution. The study highlights the importance of controlling uneven settlement to ensure the safety and durability of high-speed railway track systems. Tutumluer et al. [17] reviewed the use of geosynthetics in transportation infrastructure such as roads, railways, and airfields. The study discusses the role of materials like geogrids, geotextiles, and geocells in improving subgrade restraint and stabilizing unbound aggregates. The findings highlight that geosynthetics enhance stiffness and load distribution through bearing capacity improvement and lateral restraint mechanisms. Overall, the research demonstrates their effectiveness in improving the performance and durability of transportation systems. Hasheminezhad et al. [18] examined the role of geosynthetics in improving the resilience and sustainability of geotechnical infrastructure. The study highlights that geosynthetics enhance load distribution, reduce deformation, and increase the durability of structures such as roads, embankments, and retaining systems. Their applications in drainage improvement, landslide recovery, and flood mitigation are also discussed. In addition, recent advances in geosynthetic technologies and their role in life cycle cost analysis are reviewed. The findings emphasize the importance of geosynthetics in developing resilient and sustainable infrastructure. Mohan and Nandan [19] investigated the use of pine-needle-based natural geotextiles as a sustainable solution for improving road infrastructure performance. The study showed that these geotextiles enhance load distribution, water retention, and subgrade stability while reducing environmental impacts compared to conventional materials. Mechanical, durability, and environmental tests, along with numerical modeling, were conducted to evaluate their performance. The results indicate that pine-needle geotextiles can improve the durability and lifespan of road pavements and provide a sustainable alternative for road construction. Mohan and Nandan [20] demonstrated that alkali-treated pine needle geotextile (TPNG) serves as an effective, sustainable reinforcement for pavement subgrades. Their research shows that TPNG significantly enhances load-bearing capacity, increasing the California Bearing Ratio (CBR) by up to 185% and improving shear strength by raising the internal friction angle from 22° to 35°. Furthermore, the material reduces pavement settlement by up to 50% under cyclic loading, offering a viable, eco-friendly solution for resilient infrastructure. Petriaev and Fedorenko [21] propose a nonlinear elastoplastic finite element (FEM) approach in Plaxis to model geogrid behavior, accounting for detailed geometry and rib thickness. By verifying the model against experimental data, the study demonstrates that linear elastic assumptions are insufficient for accurate analysis. This research effectively captures the geogrid's performance in both unconfined and soil-confined conditions, providing a more precise framework for predicting tensile stress-strain relationships. Ahmadi et al. [22] proposed a hybrid numerical framework combining DEM, FDM, and FEM to simulate the initial development of differential settlement in railway transition zones. The approach models granular ballast behavior, track structure response, and vehicle-track dynamic interaction simultaneously. Results show that the method can capture void formation beneath sleepers, redistribution of contact forces, and progressive track irregularities under repeated axle loads. The study highlights the importance of gradual stiffness transitions to reduce long-term track settlement and maintenance needs. Punetha and Nimbalkar [23] investigated the dynamic behavior of railway track-bridge transition zones under moving train loads using a three-dimensional finite element model validated with field data. The study analyzed mechanisms such as unsupported sleeper formation and stress variations within track substructure layers. Results indicate that train speed, axle load, and wheel motion significantly affect stresses, displacements, and sleeper-ballast gaps. The research also shows that geoinclusion reinforcement can improve performance and reduce differential settlement in critical transition zones. Asghari et al. [24] investigated the dynamic behavior of railway bridge-track transition zones considering the effects of approach slab length, thickness, and train speed using a three-dimensional FEM model. The results indicate that increasing train speed intensifies deformation in the transition zone. Increasing the thickness and length of the approach slab significantly reduces deformation and improves track performance. The study highlights the importance of proper transition zone design to mitigate stiffness changes and dynamic loads. Hassan et al. [25] investigated the dynamic performance of railway transition zones between rigid structures and embankments using a validated three-dimensional finite element model. A parametric analysis was conducted to evaluate the effects of approach slab geometry and wedge-shaped backfill on reducing dynamic impacts and differential settlement. Results indicate that increasing slab thickness and length significantly reduces rail displacement, with thickness having a stronger influence. The study also shows that wedge-shaped backfill improves stiffness transition and overall track performance under high-speed loading.

In recent years, geosynthetic reinforcement has become an effective solution for improving the performance of geotechnical systems subjected to dynamic loading, particularly in transportation infrastructure. Previous studies have employed analytical formulations, laboratory experiments, centrifuge testing, and three-dimensional finite element simulations to investigate the behavior of reinforced embankments, pile-supported systems, reinforced foundations, and load transfer platforms under moving and cyclic loads. These investigations consistently demonstrated that geotextiles and geogrids enhance stress distribution, improve lateral confinement, reduce permanent deformation and settlement, and increase the overall stability and serviceability of reinforced soil systems [26-36].

Despite the considerable body of research on railway transition zones, previous studies have predominantly focused on transition-zone geometry optimization, stiffness variation analysis, settlement prediction, moving-load modeling, and dynamic vehicle-track interaction. Although these investigations have significantly improved the understanding of transition-zone behavior, limited attention has been devoted to the combined application of geotextile reinforcement and cement-stabilized ballast as a practical solution for mitigating stiffness discontinuities in slab-ballast transition zones. Furthermore, the influence of geotextile reinforcement length on the mechanical performance of high-speed railway transitions has not been systematically quantified. Therefore, a clear research gap remains regarding the effectiveness of geotextile confinement mechanisms and their contribution to stiffness enhancement and displacement reduction under high-speed railway loading conditions.

To address this gap, the present study develops a comprehensive three-dimensional finite element model to investigate the behavior of geotextile-reinforced slab-ballast transition zones. The proposed approach incorporates geotextile layers within cement-stabilized ballast and systematically evaluates the performance of different reinforcement lengths (6, 9, and 12 m). The principal novelties of this research include: (I) the development of a detailed 3D finite element framework for geotextile-reinforced railway transition zones; (II) the investigation of geotextile confinement effects on ballast behavior and stress redistribution; (III) the evaluation of the optimal reinforcement length for minimizing rail deflection and track deformation; and (IV) the quantification of track stiffness enhancement achieved through geotextile inclusion. By applying quasi-static loading near the slab-ballast interface, the study provides new insights into the mechanisms governing stiffness transition and offers a practical design framework for improving the long-term performance and structural reliability of high-speed railway infrastructure.

2. Numerical simulation

The numerical analysis was performed using ABAQUS 2017 finite element software. To achieve a balance between computational efficiency and accuracy, the model was initially conceptualized in 3D and then refined to capture the critical dynamic responses of the transition zone.

2.1. Geometry and model components

The geometric configuration of the transition zone was developed in accordance with Instruction 301 (Technical Specifications for Railway Tracks) to ensure structural compliance (Fig. 1). The foundation of the model comprises an embankment modeled as a solid block with a length of 100 m, a height of 2.3 m, and a width of 7 m. Integrated into this system is a 50-m-long slab track section, characterized by a width of 3.6 m and a height of 0.5 m, which includes pre-defined recesses for precise sleeper placement (Table 1). To represent the flexible portion of the track, a ballast layer extending 47.4 m was incorporated with a total thickness of 50 cm. This thickness accounts for the mandatory 35-cm minimum clearance beneath the sleepers plus the sleeper height, while the top width of 3.4 m provides a 40-cm ballast shoulder to maintain lateral stability during high-speed operations.

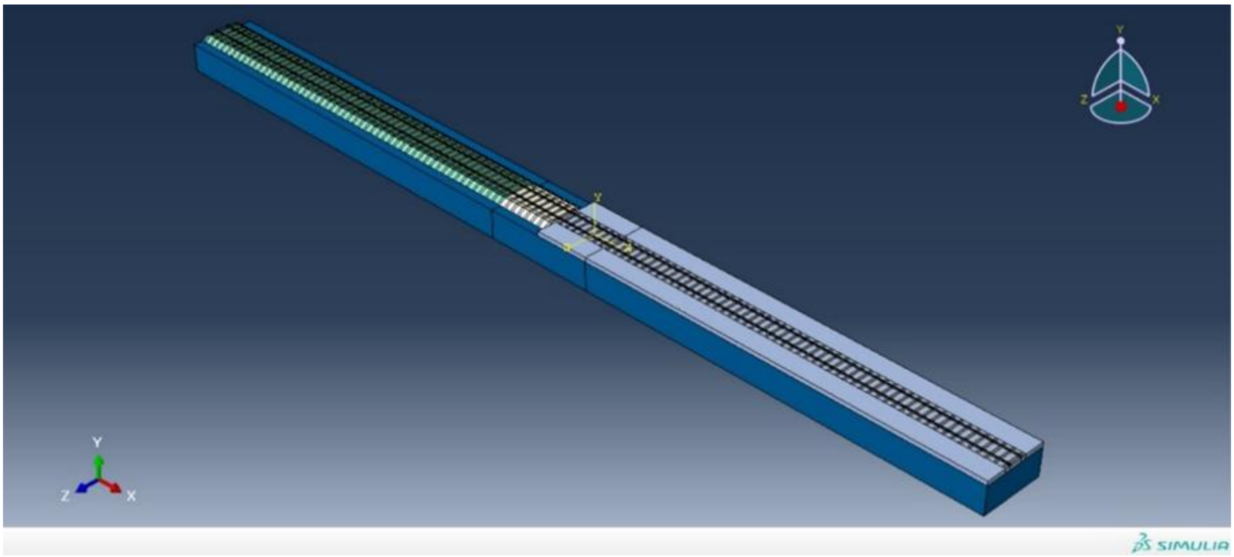


Fig. 1. General view of the 3D finite element model of the railway track system developed in Abaqus.

Table 1. Material properties used in the finite element model.

Model components	Density (kg/m^3)	Poisson's ratio	Modulus of elasticity (MPa)
Rail	7850	0.3	210000
Traverse	2500	0.2	29000
Ballast	1900	0.3	300
Railway embankment	1950	0.33	95
Element designed for cement-stabilized ballast	2500	0.2	25700
Slab track	3000	0.17	45000

The geotextile reinforcement was modeled using shell elements available in ABAQUS. A linear elastic constitutive model was adopted because the primary function of the geotextile in railway applications is to provide tensile reinforcement and lateral confinement, while its bending stiffness is generally negligible compared with the surrounding granular materials. The geotextile was therefore assumed to behave as a tension-resisting membrane with linear elastic properties throughout the analyzed loading range. The material properties were defined according to the manufacturer's technical specifications, including a tensile strength of 70 kN/m and a nominal thickness of 6.2 mm. Based on these parameters, the equivalent axial stiffness (EA) and corresponding elastic modulus were calculated and assigned to the shell elements. This modeling approach has been widely adopted in numerical studies of geosynthetic-reinforced transportation infrastructure and provides a reasonable representation of the reinforcement mechanism while maintaining computational efficiency.

Only one geotextile type was considered in the present study, and its mechanical properties were maintained constant in all simulations to isolate the influence of reinforcement length on transition-zone performance.

In the present study, the ballast layer was modeled as a linear elastic material to provide a consistent basis for comparing the relative effectiveness of different geotextile reinforcement lengths under identical modeling assumptions. This simplification was considered appropriate because the primary objective was to evaluate the comparative improvement in transition-zone performance rather than to reproduce the full nonlinear behavior of ballast materials. Nevertheless, it is recognized that railway ballast exhibits stress-dependent stiffness, irreversible plastic deformation, particle breakage, and rearrangement under repeated loading cycles, which are not explicitly captured in the current model.

The superstructure and reinforcement elements were modeled to capture the mechanical interactions at the interface. Concrete sleepers, with dimensions of 2.6 × 0.25 × 0.22 m (length × width × height), were positioned within the trackbed. The rail was represented by a rectangular section with an equivalent cross-sectional area of 0.0162 m^2, meticulously calibrated to match the moment of inertia and standard properties of a UIC 60 rail profile. Furthermore, the geosynthetic (geotextile) reinforcements were modeled as specialized elements designed to provide axial tensile strength and confinement (Table 2). These elements were specifically characterized by their zero bending stiffness, ensuring they resist only tensile forces to enhance the structural integrity of the ballast layer within the critical transition region.

Table 2. Characteristics of the applied geosynthetic material.

Name	Geotextile (800 g/m²)
Polymer type	PET
Unit weight (gr/m^2)	800
Thickness (mm)	6.2
Elongation in grab test	Over 50%
Wide-width tensile strength (kN/m)	70
Visible pore size (mm)	0.09
Permeability (cm/sec)	0.20

2.2. Interactions and boundary conditions

To simulate the physical connectivity of the track components, the rail-sleeper connections were modeled using linear springs to represent the rail pads and fastening systems. For the interfaces between the ballast/embankment and slab/subgrade, tie constraints were employed assuming no-slip conditions, justified by the quasi-static nature of the loading (Fig. 2). To optimize computational efficiency, the slab and sleeper components were merged, given their negligible relative displacement. Regarding the boundary conditions, the bottom surfaces were fully fixed in all directions, while the longitudinal ends were constrained to prevent any out-of-plane movement.

To minimize the influence of artificial boundary reflections, a total model length of 100 m was adopted based on recommendations reported in previous finite element studies of railway track systems. This length was considered sufficient to ensure that reflected stress waves do not significantly affect the response within the transition zone, which constitutes the primary region of interest. Nevertheless, it is acknowledged that advanced boundary treatments, such as infinite elements, viscous boundaries, or absorbing boundary conditions, may further reduce wave reflections in fully dynamic moving-load simulations. The implementation of such techniques is recommended for future investigations involving transient vehicle-track interaction analyses.

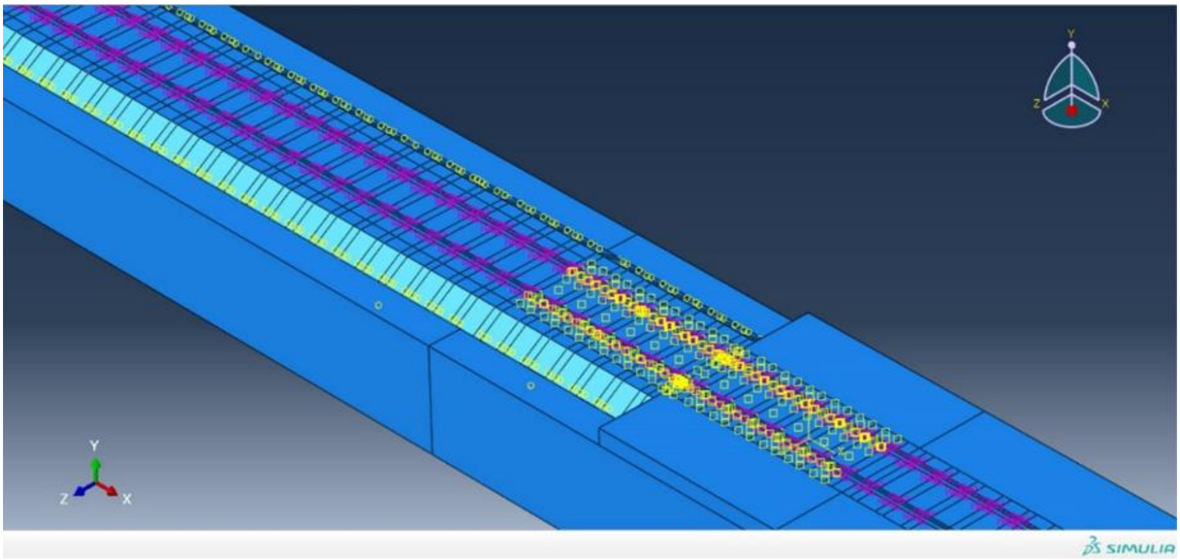


Fig. 2. Illustration of the interaction properties and boundary conditions assigned to the rail-sleeper and sleeper-ballast interfaces.

2.3. Loading and dynamic impact factor

The wheel loads were applied as quasi-static concentrated loads, incorporating dynamic effects at high speeds by multiplying the static wheel load (P_s) by a dynamic impact factor (φ), as Eq. 1.

$$P_d = \varphi . P_s \tag{1}$$

The impact factor was determined using a formula calibrated for the specific conditions of the Iranian railway network (Eq. 2).

$$\varphi = 1 + 4/73\frac{V}{D} \tag{2}$$

where V represents the train speed (km/h) and D is the wheel diameter (840 mm). Two distinct speed scenarios of 160 km/h and 500 km/h were analyzed, with loads strategically positioned 3 meters from the transition interface to accurately capture the peak vertical displacement (Table 3 and Fig. 3).

Table 3. Loading values of different train elements.

Train components	Load (N)
Locomotive body	1115000
Wagon body	940000
Three-axle bogie (locomotive)	180000
Two-axle bogie (wagon)	42000
Locomotive bogie axle	58000
Wagon bogie axle	18000
Total	2353000

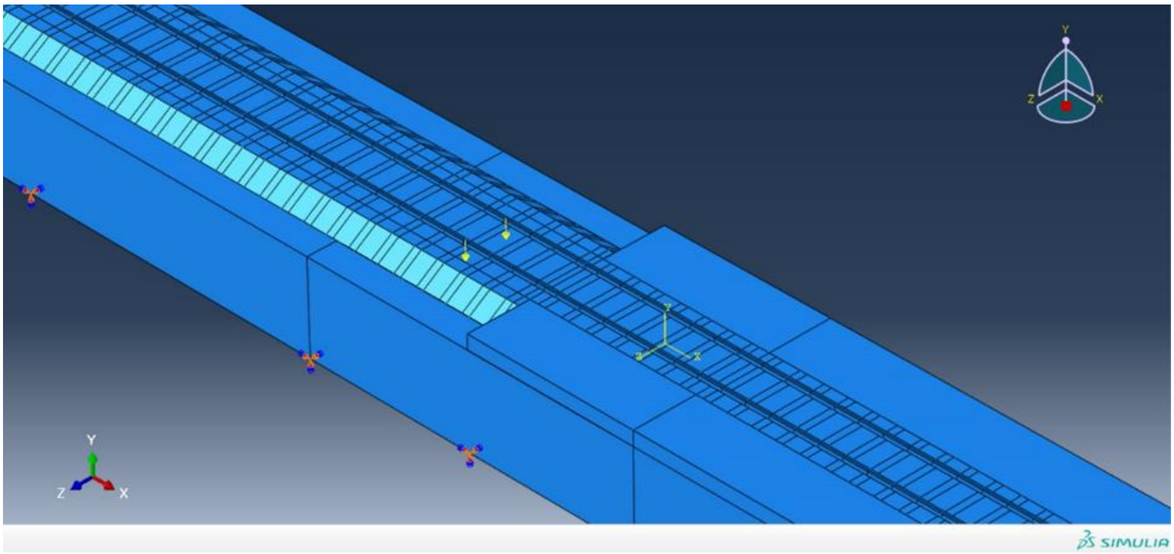


Fig. 3. Application of concentrated wheel loads on the rail surface for dynamic analysis and impact factor evaluation.

The Dynamic Impact Factor (DIF) adopted in this study was selected to account for the amplification of wheel loads induced by train speed and track irregularities. The employed formulation has been widely used in railway engineering practice and provides a simplified yet effective representation of speed-dependent dynamic loading effects. Although the adopted approach does not explicitly simulate moving wheel loads or vehicle-track interaction, it enables the incorporation of dynamic load amplification within a computationally efficient framework suitable for comparative parametric investigations of transition-zone performance. In addition to current high-speed railway operating conditions, a train speed of 500 km/h was considered as an upper-bound design scenario to evaluate the robustness of transition-zone performance under future ultra-high-speed railway systems. The inclusion of this speed does not imply current operational practice; rather, it reflects the growing interest in next-generation transportation technologies and the need to assess infrastructure performance beyond existing high-speed rail limits. As highlighted by Connolly and Costa (2020), the geodynamics of very-high-speed transport systems represent an emerging research area requiring the evaluation of track and subgrade behavior under increasingly demanding operational conditions. Accordingly, the 500 km/h scenario was adopted to investigate the resilience of the proposed geotextile-reinforced transition zone under extreme but plausible future loading conditions. The loading position was selected based on the findings of Shan et al. (2013), who reported that the region approximately 3 m away from the transition interface is where the most significant changes in response and peak displacements are observed. In addition, preliminary numerical analyses performed during the model development stage of this study also indicated that the maximum rail displacement occurs in the vicinity of this location. Therefore, the load was applied at this distance to capture the critical response behavior of the transition zone.

In this study, a dynamic-equivalent loading approach based on the Dynamic-Implicit formulation is adopted. Accordingly, explicit damping definitions such as Rayleigh damping are not introduced, as the response is governed by quasi-static equilibrium with inertia effects rather than full transient wave propagation. Nevertheless, it is recognized that energy dissipation mechanisms, including material and radiation damping, may influence the system response. These effects are therefore suggested to be incorporated in future studies for more comprehensive dynamic simulations. It should be noted that the present modeling framework does not represent a fully coupled vehicle-track dynamic interaction analysis. Instead, dynamic effects associated with train speed are incorporated through a Dynamic Impact Factor (DIF) applied to the static wheel loads. Therefore, the adopted methodology should be interpreted as a dynamic-equivalent analysis rather than a moving-load dynamic simulation. The adopted methodology represents a dynamic-equivalent analysis in which train-induced dynamic effects are incorporated through a Dynamic Impact Factor (DIF) rather than a fully coupled moving-load simulation. This simplified approach is commonly used in railway engineering studies to evaluate the influence of speed-dependent load amplification while maintaining computational efficiency.

The selected transition lengths of 6 m, 9 m, and 12 m were defined based on practical railway construction considerations. These values were chosen to represent short, medium, and extended reinforcement zones commonly implemented in transition-zone improvement practices. This range is consistent with typical design recommendations reported in railway engineering studies and allows for a comparative assessment of different reinforcement extents on system performance.

2.4. Solver and meshing

The Dynamic-Implicit solver was employed to ensure numerical convergence while accurately capturing time-dependent responses. The meshing strategy incorporated C3D8 (8-node linear brick) and C3D10 (10-node quadratic tetrahedron) elements for solid components, while S4 (4-node shell) elements were utilized for planar reinforcements. Furthermore, the mesh density was locally refined at the transition interface and beneath the loading points to enhance the accuracy of gradient calculations in these critical zones (Fig. 4).

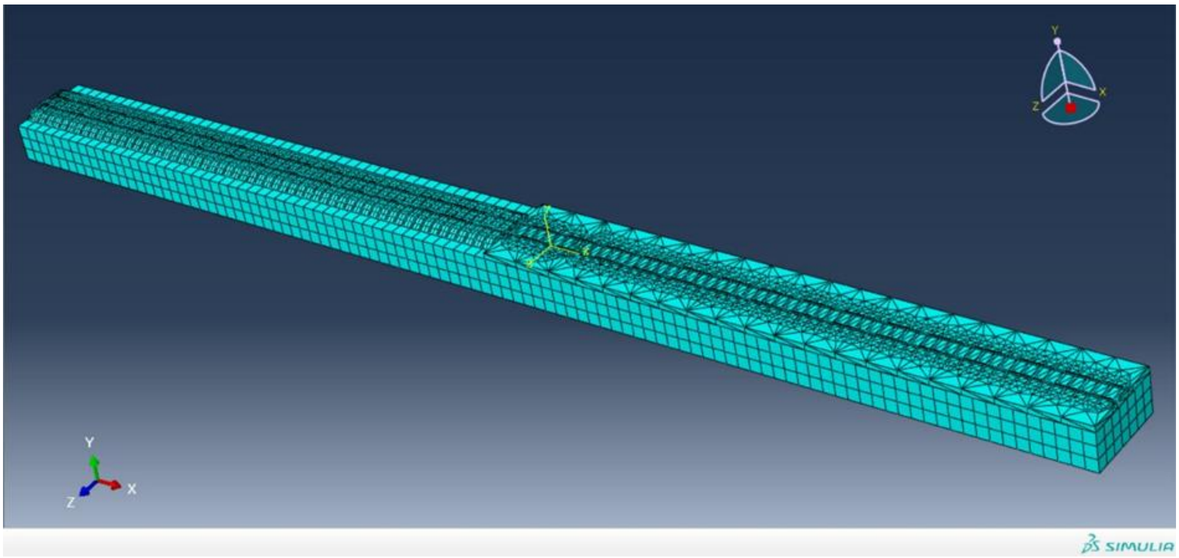


Fig. 4. Discretized finite element mesh of the 3D model with local mesh refinement in the zone of interest.

2.5. Mesh convergence study

To ensure that the numerical results are independent of mesh density, a mesh convergence study was performed. Three different mesh configurations were considered by progressively refining the mesh in the transition zone and beneath the loading area, where the highest stress and displacement gradients occur. The maximum vertical rail displacement was selected as the convergence criterion because it represents the primary response parameter used throughout the study. The results indicate that increasing the mesh density beyond the medium refinement level produced only marginal changes in the predicted rail displacement. The difference between the medium and fine meshes remained below 3%, demonstrating that the numerical solution had reached mesh independence. Therefore, the medium-fine mesh configuration was adopted for all subsequent analyses to achieve a suitable balance between computational efficiency and solution accuracy [37, 38].

2.6. Track bed stiffness

The track bed stiffness (K) is defined as the force required to produce a unit vertical displacement (Y), calculated as Eq. 3.

$$K = \frac{P_S}{Y} = \frac{P_S}{\sum_{m=0}^{m=\infty} Y_{n\pm m}} \tag{3}$$

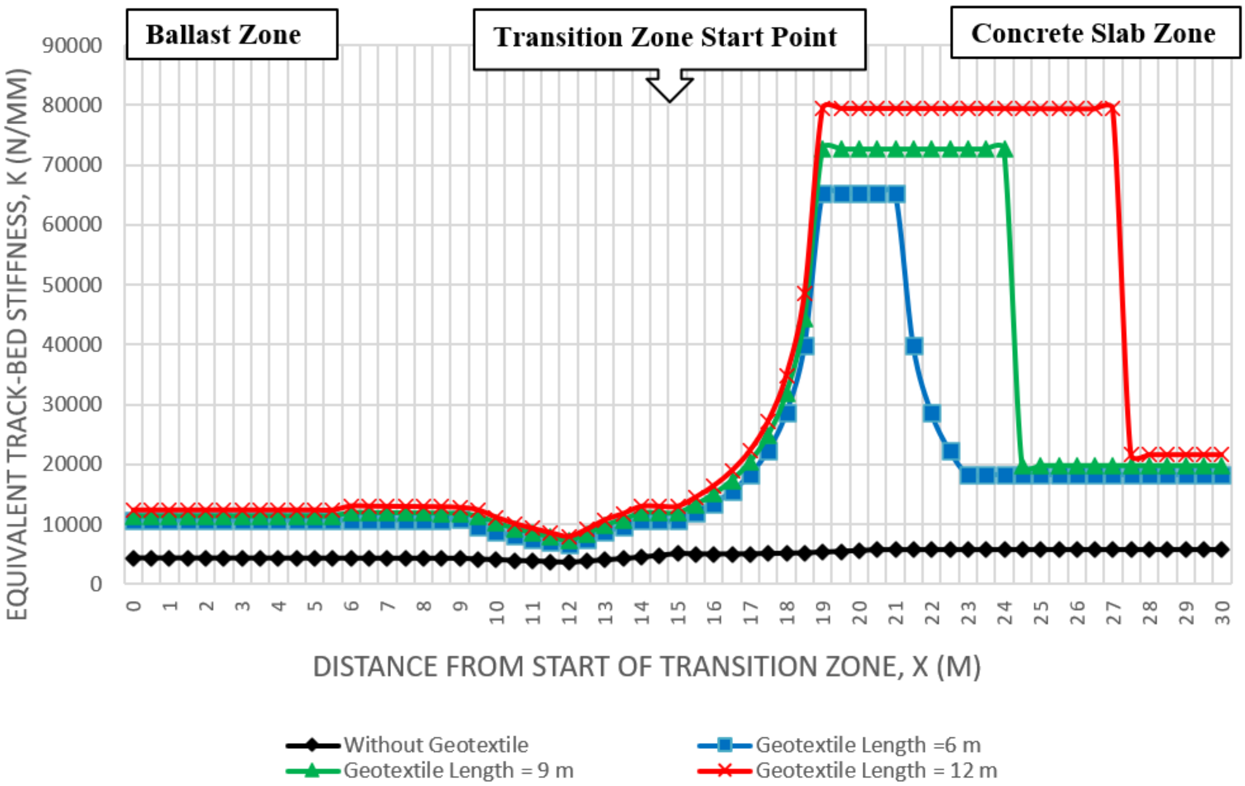
This parameter serves as a key performance indicator (KPI) to evaluate the effectiveness of geotextile reinforcement in smoothing the stiffness transition between the ballasted track and the slab track sections.

The reported stiffness values in this study (Fig. 5) correspond to the equivalent track-bed stiffness, which is calculated using Eq. 3. This parameter accounts for the combined contribution of the track system components and is employed consistently throughout the numerical model to represent the stiffness condition of the railway structure.

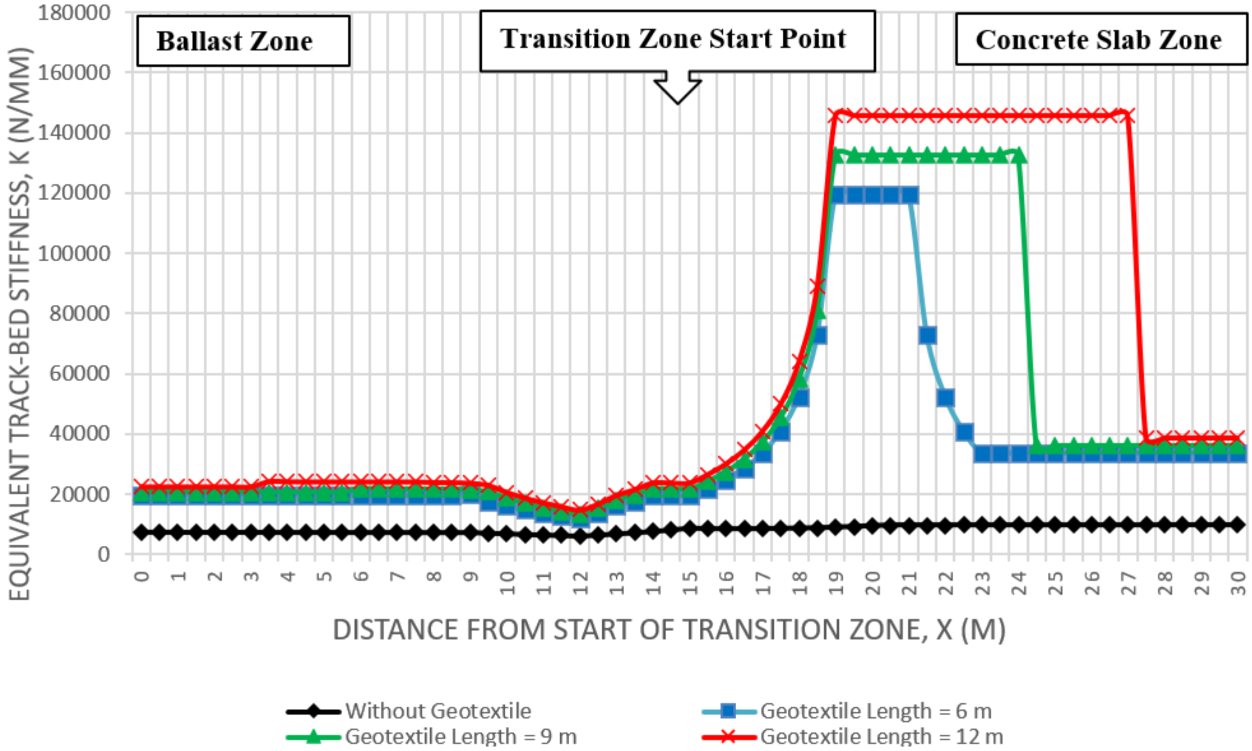
3. Validation of the numerical model

To ensure the accuracy and reliability of the finite element simulations, the developed ABAQUS model was validated against the full-scale laboratory experimental results reported by Esen et al. [12]. Their study provides a comprehensive assessment of the immediate and long-term settlement behavior of GRS railway structures under cyclic loading for both slab track and ballasted track configurations. In the present study, the numerical framework was calibrated to replicate the experimental setup, material properties, and loading conditions of the reference study. The mean vertical displacement of the rail under static loading was extracted and compared with the experimental data from (Fig. 6). As summarized in Table 4, the maximum discrepancy between the numerical predictions and the experimental measurements remained below 10%. This high degree of correlation demonstrates the capability of the developed model to accurately capture the geomechanical response of the system, thereby justifying its use for further parametric investigations.

Although the agreement between the numerical predictions and the experimental measurements demonstrates the capability of the developed model to reproduce the observed rail displacement behavior, it should be noted that the validation was performed using vertical rail displacement as the primary response parameter. This selection was dictated by the availability of experimental data reported by Esen et al. [12], which mainly focused on displacement measurements. Additional validation using ballast stress distributions, permanent settlement accumulation, and field-monitoring data from railway transition zones would provide a more comprehensive assessment of model reliability and is recommended for future studies.



(a)



(b)

Fig. 5. Spatial variation of equivalent track-bed stiffness along the slab-ballast transition zone for different geotextile reinforcement lengths (6, 9, and 12 m) under: (a) 160 km/h; and (b) 500 km/h.

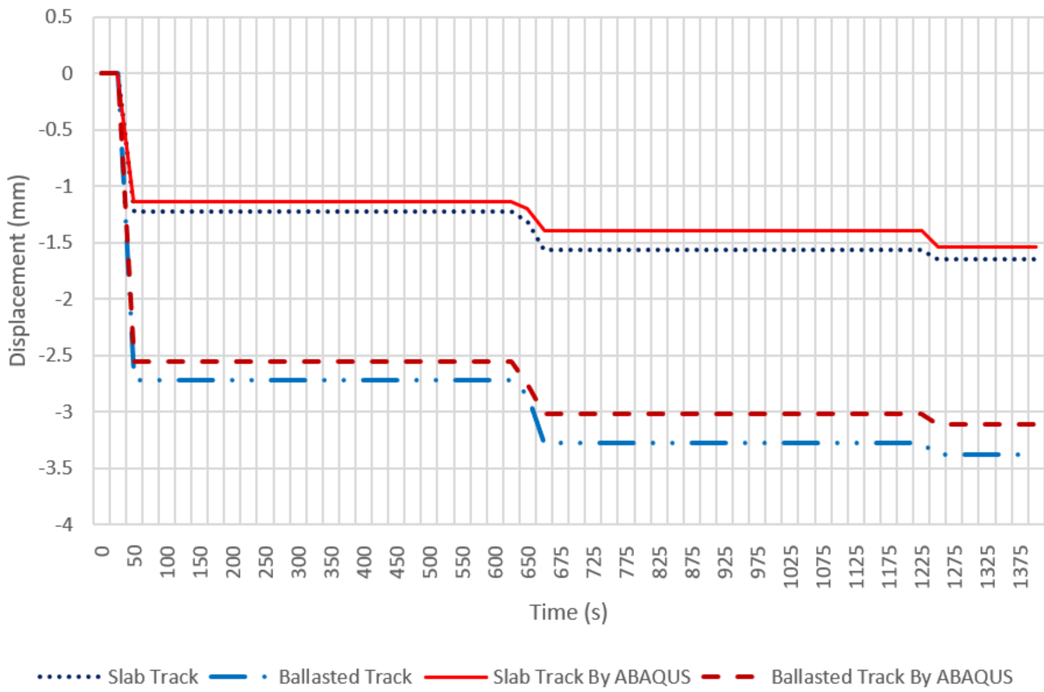


Fig. 6. Validation of the FE model: comparison of average vertical rail displacement between ABAQUS numerical results and experimental data for ballasted and slab track systems.

Table 4. Comparison of mean vertical rail displacements for different track types under static load.

Time (s)	Displacement for slab track (mm)	Displacement for slab track by ABAQUS (mm)	Percentage difference (%)	Displacement for ballasted track (mm)	Displacement for ballasted track by ABAQUS (mm)	Percentage difference (%)
50	-1.223	-1.133	7.36	-2.722	-2.552	6.24
725	-1.315	-1.2	8.74	-2.858	-2.758	3.5
750	-1.567	-1.398	10.78	-3.279	-3.018	7.96
1325	-1.651	-1.537	6.9	-3.376	-3.108	7.94

4. Results and discussion

4.1. Numerical analysis of the transition zone in the absence of a concrete approach slab

The simulation results for the transition zone without a concrete approach slab are presented in Fig. 7. The maximum vertical rail deflection is recorded at approximately 153.18 mm and 180.75 mm for operational speeds of 160 km/h and 500 km/h, respectively. This rapid manifestation of displacement within a short longitudinal distance (less than 6 m) is attributed to the stiffness discontinuity at the ballast-to-slab interface, which induces significant dynamic impact loads on the superstructure. The analysis reveals that the peak vertical displacement occurs 3 meters from the transition onset – a phenomenon governed by the rail's flexural rigidity and the spring-mass modeling of the fastening system – marking the most critical loading state. Based on the analytical framework established in Section 2.6, the calculated track bed stiffness values for the specified speeds are 3,649 N/mm and 6,202 N/mm. These results serve as a baseline for the subsequent parametric study on geotextile integration to optimize track stiffness and minimize vertical oscillations in the transition region.

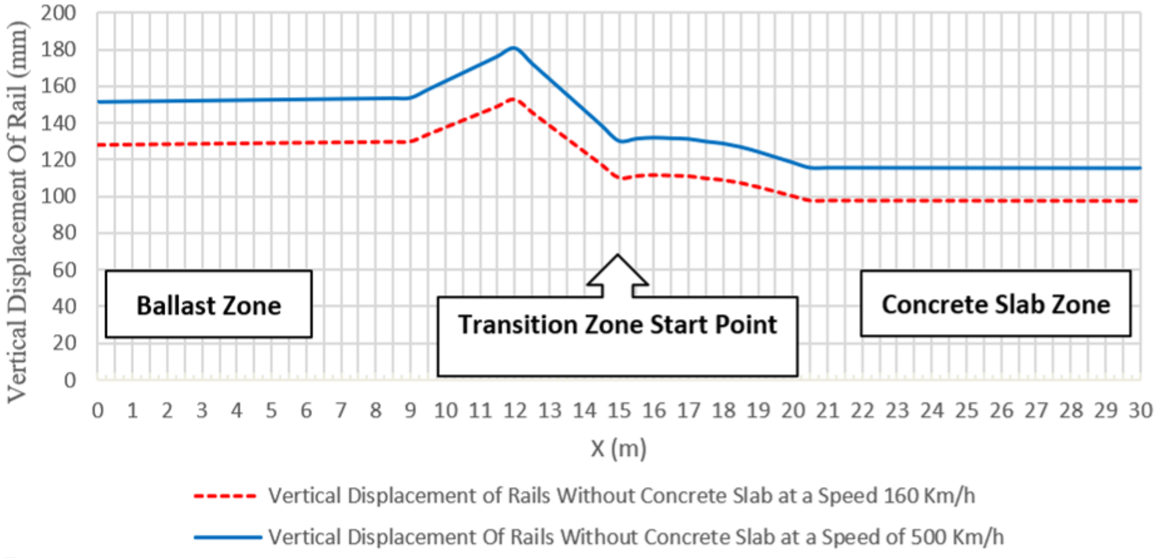


Fig. 7. Vertical rail displacement without a concrete approach slab.

4.2. Effects of train velocity and geotextile reinforcement on transition zone performance

Figs. 8 and 9 present a comparative analysis of vertical rail deflections within the transition zone, specifically investigating the efficacy of a single geotextile layer (as per Section 2.1 specifications) integrated with cement-stabilized ballast and concrete slab sections. At an operational speed of 160 km/h, the implementation of a 6-meter cement-stabilized ballast layer with geotextile results in a vertical rail deflection of approximately 85.15 mm. This deflection is reduced to 76.59 mm and 70.08 mm for 9-meter and 12-meter reinforced zones, respectively. At a higher velocity of 500 km/h, the corresponding deflections are 93.74 mm, 84.17 mm, and 75.55 mm for the 6-, 9-, and 12-meter reinforced sections. These rapid, localized deflections (< 6 m track length) are indicative of significant dynamic loading stemming from the impulsive and transient nature of high-speed train wheel-rail interaction.

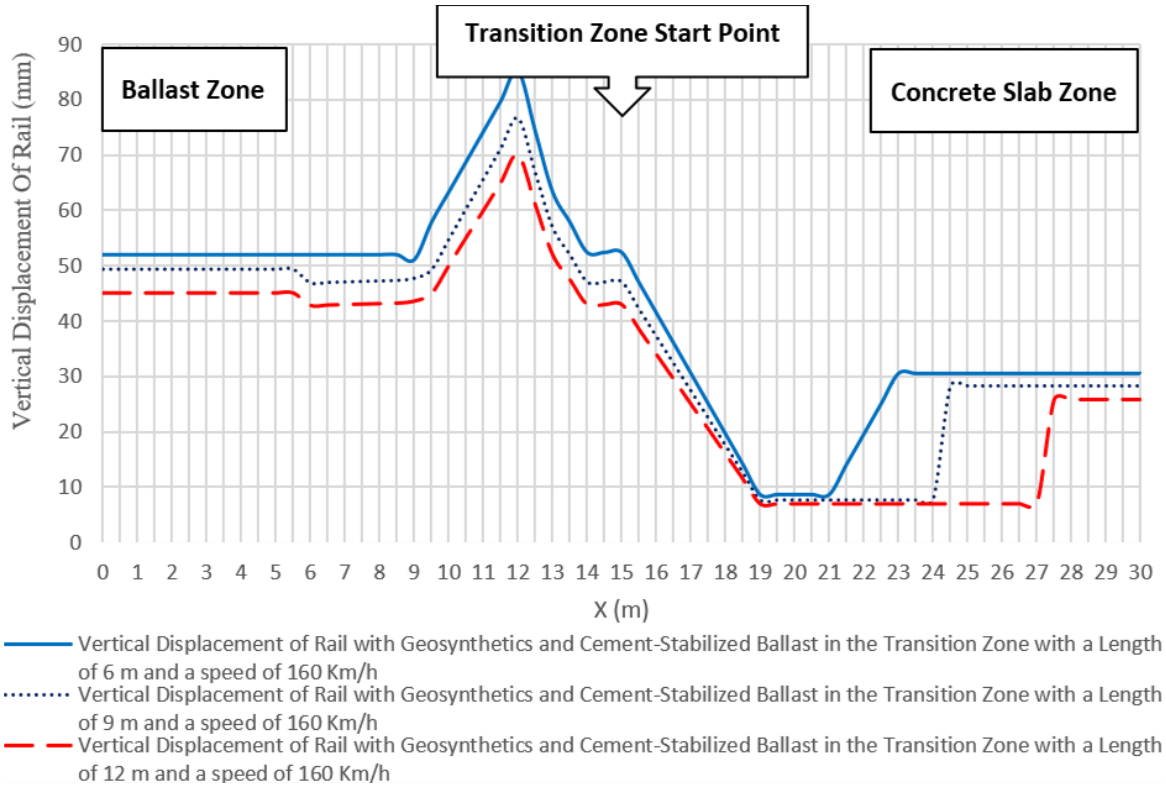


Fig. 8. Vertical rail displacement with cement-stabilized ballast implementation at 160 km/h.

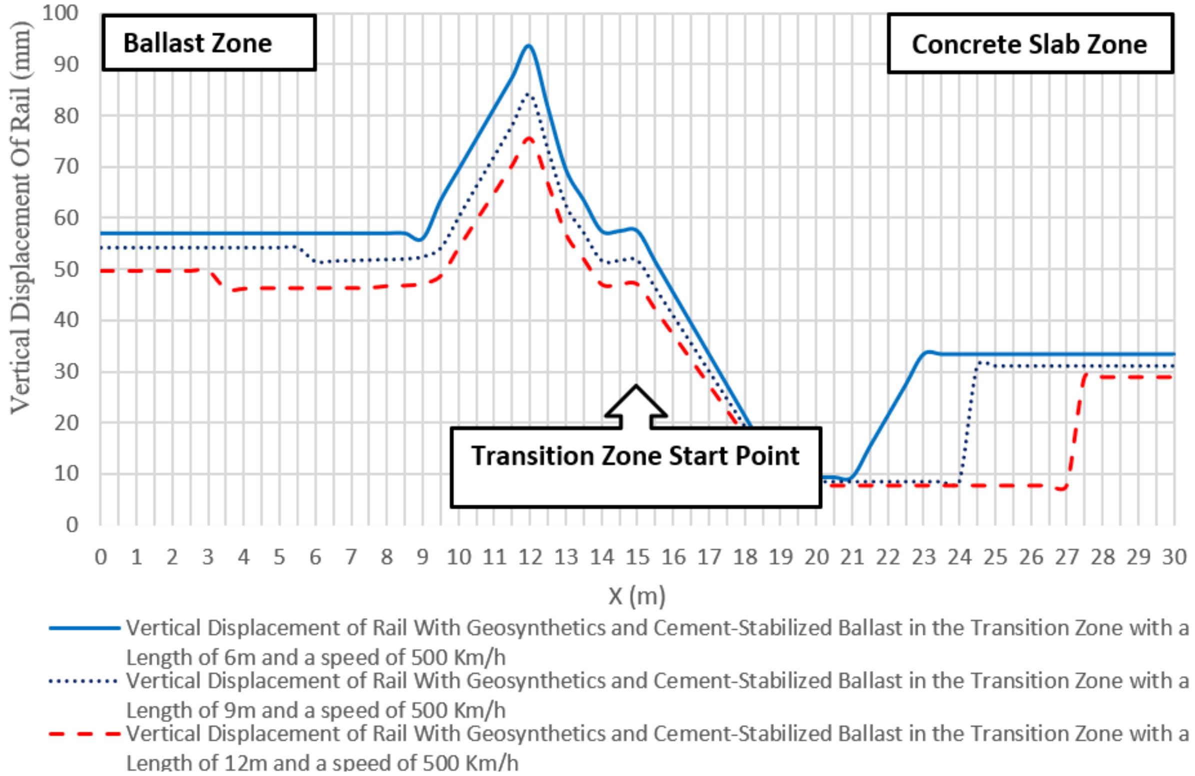


Fig. 9. Vertical rail displacement with cement-stabilized ballast implementation at 500 km/h.

The railbed stiffness for these transition zones, augmented by geotextile, demonstrates a marked increase. For the 160 km/h scenario, stiffness values ascend from 6,565 N/mm (6m) to 7,299 N/mm (9m) and 7,977 N/mm (12m). At 500 km/h, these values escalate substantially to 11,993 N/mm, 13,318 N/mm, and 14,838 N/mm, respectively. The graphical data reveal that the extension of the transition zone by incorporating geotextile – promoting confinement effects within the cement-stabilized ballast and concrete slab, and enhancing their tensile capacity – leads to a notable reduction in vertical rail deflections. At 160 km/h, the percentage reduction in deflection is 44.41%, 50%, and 54.25% for the 6-, 9-, and 12-meter transition lengths, respectively. Correspondingly, at 500 km/h, these reductions are 48.29%, 53.43%, and 58.2%. Notably, the percentage reduction in deflection does not exhibit a commensurate increase with velocity. This observation is ascribed to the inherent transience of high-speed loading; the very short wheel-rail contact duration limits the extent to which higher speeds can amplify the effectiveness of the deflection reduction strategy, even though the absolute deflection values may change.

4.3. Comparative analysis of vertical rail displacement mitigation strategies in transition zones

Table 5 provides a comparative assessment of various strategies for mitigating vertical rail displacement, quantifying the percentage reduction relative to a baseline scenario lacking a concrete approach slab.

Table 5. Comparative analysis of vertical rail displacement mitigation strategies in transition zones.

Transition zone modeling method	Design speed (km/h)	Approach slab length (transition zone, m)	Maximum vertical rail displacement in the transition zone (mm)	Percentage reduction in displacement (%)
Without concrete approach slab	160	—	153.18	—
	500	—	180.75	—
Cement-stabilized ballast	160	6	119.91	21.72
	160	9	110.93	27.59
	160	12	103.67	32.32
	500	6	126.18	30.19
	500	9	116.91	35.32
	500	12	107.79	40.36
Geosynthetic layer on cement-stabilized ballast	160	6	85.15	44.41
	160	9	76.59	50
	160	12	70.08	54.25
	500	6	93.47	48.29
	500	9	84.17	53.43
	500	12	75.55	58.2

In contrast, the implementation of a 12-meter cement-stabilized ballast layer demonstrates significant efficacy, achieving approximately a 40% reduction in vertical rail displacement within the transition zone compared to the unreinforced condition.

However, the optimal strategy identified for minimizing vertical rail deflection in the transition zone is the combined application of a geosynthetic layer atop a 12-meter cement-stabilized ballast section. This integrated approach yields a substantial approximate reduction of 58%, representing the most effective solution investigated.

4.4. Influence of geotextile reinforcement and transition length on vertical rail displacement reduction

Fig. 10 shows that the investigated track stabilization techniques, including cement-stabilized ballast and the combined use of geosynthetics with cement-stabilized ballast, can reduce vertical rail displacement by up to 54% at a constant speed of 160 km/h. This reduction is mainly associated with the improved confinement provided by the geosynthetic-reinforced cement-stabilized ballast layer.

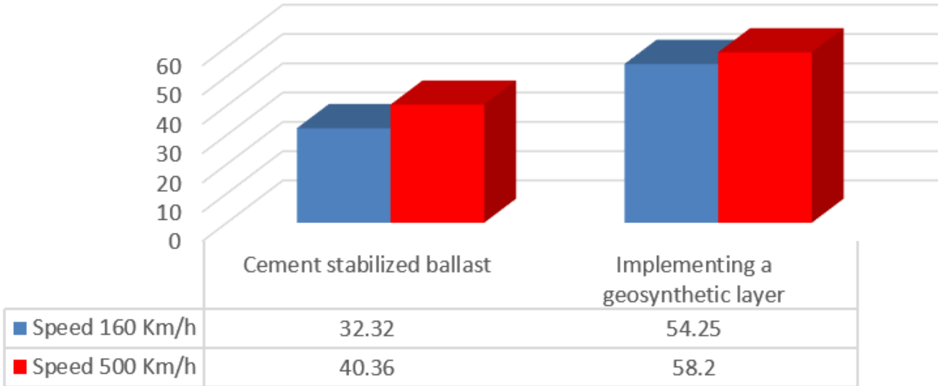


Fig. 10. Maximum percentage reduction in vertical rail displacement at speeds of 160 and 500 km/h.

In addition, extending the pre-slab zone length further decreases vertical rail displacement by approximately 10% in the cement-stabilized ballast case and by about 10% when geosynthetics are incorporated. This beneficial effect is attributed to a wider stress distribution zone and the increased effective anchorage length of the geosynthetic reinforcement.

5. Limitations and future research

The present study employs a dynamic-equivalent loading approach in which train-induced dynamic effects are represented through a Dynamic Impact Factor (DIF). Although this methodology captures the influence of speed-dependent load amplification, several aspects of railway dynamics are not explicitly considered, including moving wheel loads, axle-spacing effects, transient load redistribution, and coupled vehicle-track interaction mechanisms. Consequently, the obtained results should be interpreted within the scope of a simplified engineering assessment of transition-zone performance. Future research should extend the developed numerical framework by incorporating moving-load simulations, vehicle-track-subgrade interaction models, and time-dependent dynamic analyses. Such developments would enable a more comprehensive evaluation of vibration propagation, resonance phenomena, and the long-term behavior of geotextile-reinforced transition zones under realistic operational conditions.

Future validation efforts should extend beyond displacement-based comparisons and include additional response indicators such as stress distribution within ballast layers, permanent settlement evolution, and field-monitoring measurements from operational railway transition zones. Such multi-parameter validation would further enhance confidence in the predictive capability of the proposed numerical framework. Another limitation of the present numerical framework is the assumption of linear elastic behavior for the ballast layer. While this approach enables efficient comparison between different reinforcement configurations, it does not fully represent the nonlinear and path-dependent characteristics of ballast under repeated train loading. More advanced constitutive formulations, such as the Hardening Soil model, Drucker-Prager plasticity model, or Discrete Element Method (DEM)-based approaches, may provide more realistic predictions of ballast degradation, permanent settlement accumulation, and long-term track performance. Future studies should investigate the influence of these advanced material models on the effectiveness of geotextile-reinforced transition zones.

The present study focuses primarily on evaluating the influence of reinforcement length on the performance of railway transition zones, while the mechanical properties of the geotextile were kept constant throughout the analyses. Consequently, the effects of geotextile tensile stiffness, which is recognized as one of the most influential parameters governing reinforcement efficiency, were not explicitly investigated. Variations in tensile stiffness may significantly affect load transfer mechanisms, lateral confinement, membrane action, and the overall deformation response of the track structure. Therefore, the results presented herein should be interpreted within the context of the specific geotextile properties adopted in the numerical model. Future studies should perform comprehensive parametric analyses considering different levels of tensile stiffness (EA), soil-geotextile interface characteristics, ballast stiffness, reinforcement placement depth, and the use of multiple reinforcement layers to establish optimized design recommendations for high-speed railway transition zones.

6. Practical implications

Geotextile reinforcement is expected to enhance track performance and may contribute to reducing maintenance demands by improving the overall structural response of the railway transition zone. However, no explicit economic assessment is carried out in this study. Therefore, any statement regarding cost-effectiveness should be interpreted qualitatively rather than quantitatively. A comprehensive life-cycle cost analysis and maintenance-benefit evaluation are recommended for future research to better quantify the economic implications of the proposed reinforcement approach.

7. Conclusions

The present study investigated the performance of geotextile-reinforced slab-ballast transition zones in high-speed railways using a three-dimensional dynamic-equivalent finite element framework. Based on the obtained results, the following conclusions can be drawn:

- The developed numerical framework successfully quantified the influence of geotextile reinforcement on the mechanical response of railway transition zones and provided a practical methodology for evaluating stiffness-transition improvement strategies.
- Geotextile inclusion within the cement-stabilized ballast layer enhanced the confinement of ballast particles and improved stress redistribution across the slab-ballast interface, resulting in a smoother transition of structural stiffness.
- The results confirmed that train speed remains a dominant factor affecting transition-zone performance. Higher operating speeds increase dynamic load amplification and consequently intensify track deformation, highlighting the need for effective reinforcement measures in future high-speed railway systems.
- Reinforcement length was identified as a key design parameter governing the effectiveness of geotextile stabilization. Extending the reinforcement zone improved load-transfer efficiency and increased the equivalent track-bed stiffness, leading to superior structural performance.
- The combined application of geotextile reinforcement and cement-stabilized ballast demonstrated the most favorable response among the investigated mitigation techniques, indicating its potential as a practical reinforcement strategy for reducing vertical rail displacement and stiffness discontinuities in railway transition zones.
- From an engineering perspective, the proposed reinforcement system may contribute to improved track stability, enhanced serviceability, and reduced maintenance requirements by mitigating excessive deformation at critical slab-ballast interfaces.
- The study contributes to the existing literature by providing one of the first systematic assessments of geotextile reinforcement length in cement-stabilized railway transition zones and by clarifying the underlying confinement mechanisms responsible for performance improvement.
- Despite these findings, the adopted methodology represents a dynamic-equivalent approach and does not explicitly account for moving-load effects, coupled vehicle-track interaction, or ballast material nonlinearity. Future research should therefore incorporate advanced dynamic simulations, field-scale validation, and parametric investigations of geotextile stiffness, interface behavior, and long-term cyclic performance to establish comprehensive design recommendations for next-generation high-speed railway infrastructure.

Statements & declarations

Author contributions

Nima Moslemi Patrudi: Writing - Review & Editing, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Hossein Ghasemzadeh Tehrani: Writing - Review & Editing, Supervision, Methodology, Investigation.

Iman Aghayan: Writing - Review & Editing, Supervision, Methodology, Investigation.

Mohsen Bagheri: Writing - Original Draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Conceptualization.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Shan, Y., Albers, B., Savidis, S. A. Influence of different transition zones on the dynamic response of track–subgrade systems. Computers and Geotechnics, 2013; 48: 21–28. doi:10.1016/j.compgeo.2012.09.006.

[2] Shih, J. Y., Thompson, D. J., Zervos, A. The effect of boundary conditions, model size and damping models in the finite element modelling of a moving load on a track/ground system. Soil Dynamics and Earthquake Engineering, 2016; 89: 12–27. doi:10.1016/j.soildyn.2016.07.004.

[3] Chen, J., Zhou, Y. Dynamic vertical displacement for ballastless track-subgrade system under high-speed train moving loads. Soil Dynamics and Earthquake Engineering, 2020; 129: 105911. doi:10.1016/j.soildyn.2019.105911.

[4] Connolly, D. P., Costa, P. A. Geodynamics of very high speed transport systems. Soil Dynamics and Earthquake Engineering, 2020; 130: 105982. doi:10.1016/j.soildyn.2019.105982.

[5] Zhou, S., Wang, B., Shan, Y. Review of research on high-speed railway subgrade settlement in soft soil area. Railway Engineering Science, 2020; 28: 129–145. doi:10.1007/s40534-020-00214-x.

[6] Charoenwong, C., Connolly, D. P., Colaço, A., Alves Costa, P., Woodward, P. K., Romero, A., Galvín, P. Railway slab vs ballasted track: A comparison of track geometry degradation. Construction and Building Materials, 2023; 378: 131121. doi:10.1016/j.conbuildmat.2023.131121.

[7] Wang, R., Cheng, J.-j., Gao, L., Li, Z.-g., Qi, Y.-l., Wang, M.-t., Ding, B.-s. Research on the swelling mechanism of high-speed railway subgrade and the induced railway heave of ballastless tracks. Transportation Geotechnics, 2021; 27: 100470. doi:10.1016/j.trgeo.2020.100470.

[8] Charoenwong, C., Connolly, D. P., Woodward, P. K., Galvín, P., Alves Costa, P. Analytical forecasting of long-term railway track settlement. Computers and Geotechnics, 2022; 143: 104601. doi:10.1016/j.compgeo.2021.104601.

[9] Coelho, B., Priest, J., Powrie, W., Holscher, P. Monitoring of transition zones in railways. In: 10th international conference and exhibition; 2009 Jun 23–24; London, United Kingdom. p. 1–9.

[10] Sañudo, R., Miranda, M., Markine, V., delÍOlio, L. The Influence of Train Running Direction and Track Supports Position on the Behaviour of Transition Zones. Transportation Research Procedia, 2016; 18: 281–288. doi:10.1016/j.trpro.2016.12.037.

[11] Sadeghi, V., Bagheri, M., Abasi Hamidi, J. Earthquake-Induced Deformation of Road Embankments: A Finite Difference Method. Civil Engineering and Applied Solutions, 2026; 2: 28–41. doi:10.22080/ceas.2026.30800.1066.

[12] Esen, A. F., Woodward, P. K., Laghrouche, O., Čebašek, T. M., Brennan, A. J., Robinson, S., Connolly, D. P. Full-scale laboratory testing of a geosynthetically reinforced soil railway structure. Transportation Geotechnics, 2021; 28: 100526. doi:10.1016/j.trgeo.2021.100526.

[13] Jain, A., Metrikine, A. V., Steenbergen, M. J. M. M., van Dalen, K. N. Railway transition zones: evaluation of existing transition structures and a newly proposed transition structure. International Journal of Rail Transportation, 2024; 12: 979–999. doi:10.1080/23248378.2023.2272668.

[14] Jain, A., Metrikine, A. V., Steenbergen, M. J. M. M., van Dalen, K. N. Dynamic amplifications in railway transition zones: performance evaluation of sleeper configurations using energy criterion. Frontiers in Built Environment, 2024; 10: 1–9. doi:10.3389/fbuil.2024.1285131.

[15] Li, G., Wang, X., Wu, S., li, D., Ma, L., Ding, W. Dynamic response analysis of vehicle-track coupled system at subgrade-bridge transition zone in seasonal frozen region under multi-source excitation. Soil Dynamics and Earthquake Engineering, 2025; 190: 109216. doi:10.1016/j.soildyn.2025.109216.

[16] Cui, W., Xiao, H., Zhong, Y., Ma, C., Gao, L., Chen, H., Zhao, S. Vehicle-induced vibration and damage characteristics of ballastless track in subgrade settlement zone at 400 km/h. Transportation Geotechnics, 2026; 61: 102069. doi:10.1016/j.trgeo.2026.102069.

[17] Tutumluer, E., Kang, M., Qamhia, I. I. A. Geosynthetic stabilization of road pavements, railroads, and airfields. Transportation Geotechnics, 2025; 50: 101321. doi:10.1016/j.trgeo.2024.101321.

[18] Hasheminezhad, A., Ceylan, H., Kim, S., Tutumluer, E. Geosynthetics for resilient geotechnics: A review of applications and innovations. Transportation Geotechnics, 2025; 55: 101676. doi:10.1016/j.trgeo.2025.101676.

[19] Mohan, V., Nandan, A. A comprehensive review of the performance of pine needle geotextiles in reinforced subgrade pavement for sustainable road construction and maintenance. Environmental Science and Pollution Research, 2025; 32: 11838–11878. doi:10.1007/s11356-025-36361-z.

[20] Mohan, V., Nandan, A. Green geotechnical resilient infrastructure: a multi-scale evaluation of pine needle bio-geotextiles to improve land and airfield pavement performance. Innovative Infrastructure Solutions, 2026; 11: 235. doi:10.1007/s41062-026-02651-9.

[21] Petriaev, A., Fedorenko, E. Geosynthetics and Soil Interaction Numerical Modeling for High-speed Rail Lines Substructure. Transportation Research Procedia, 2023; 68: 694–701. doi:10.1016/j.trpro.2023.02.096.

[22] Ahmadi, A., Nasrollahi, K., Nielsen, J. C. O., Dijkstra, J. Dynamic vehicle–track interaction and differential settlement in a transition zone on railway ballast—An integrated 3D discrete–continuum model. Computers and Geotechnics, 2026; 190: 107737. doi:10.1016/j.compgeo.2025.107737.

[23] Punetha, P., Nimbalkar, S. Numerical investigation on dynamic behaviour of critical zones in railway tracks under moving train loads. Transportation Geotechnics, 2023; 41: 101009. doi:10.1016/j.trgeo.2023.101009.

[24] Asghari, K., Sotoudeh, S., Zakeri, J.-A. Numerical evaluation of approach slab influence on transition zone behavior in high-speed railway track. Transportation Geotechnics, 2021; 28: 100519. doi:10.1016/j.trgeo.2021.100519.

[25] Hassan, A. S. M., Khalil, A. A., El-Sonbaty, A. A., Ahmed, H. Y. Comparative numerical study on the performance of different mitigation techniques in high-speed railway transition zones. Soil Dynamics and Earthquake Engineering, 2026; 206: 110256. doi:10.1016/j.soildyn.2026.110256.

[26] Hou, R.-Y., Zheng, J.-J., Fang, H., You, L. An analytical model for dynamic response of geosynthetic reinforced embankment system under traffic load. Computers and Geotechnics, 2022; 142: 104555. doi:10.1016/j.compgeo.2021.104555.

[27] Alam, M. J. I., Gnanendran, C. T., Lo, S. R. Experimental and numerical investigations of the behaviour of footing on geosynthetic reinforced fill slope under cyclic loading. Geotextiles and Geomembranes, 2018; 46: 848–859. doi:10.1016/j.geotexmem.2018.08.001.

[28] Wu, D., Luo, C., Gao, Z., Li, D., Xu, C. Effect of Different Reinforced Load Transfer Platforms on Geosynthetic-Reinforced Pile-Supported Embankment: Centrifuge Model Test. KSCE Journal of Civil Engineering, 2022; 26: 630–649. doi:10.1007/s12205-021-0623-7.

[29] Ding, X., Zhao, J., Ou, Q., Liu, J. Numerical analysis of geosynthetic-reinforced embankment performance under moving loads. Journal of Rock Mechanics and Geotechnical Engineering, 2024; 16: 682–696. doi:10.1016/j.jrmge.2023.04.019.

[30] Liu, K.-F., Feng, W.-Q., Cai, Y.-H., Xu, H., Wu, P.-C. Physical model study of pile type effect on long-term settlement of geosynthetic-reinforced pile-supported embankment under traffic loading. Transportation Geotechnics, 2023; 38: 100923. doi:10.1016/j.trgeo.2022.100923.

[31] Ma, S., Lu, L., Song, A., Wang, Z., Xiao, L., Arai, K. Experimental and Numerical Investigation of Prestressed Geosynthetic-Reinforced Embankment Subjected to Traffic Loading. International Journal of Geomechanics, 2024; 24: 04023250. doi:10.1061/IJGNAI.GMENG-7815.

[32] Hou, R.-Y., Zheng, J.-J., Fang, H., Yang, W.-Y. A Mechanical Model and Solution for Dynamic Response of Geosynthetic-Reinforced Pile-Supported Embankment under Traveling Loads. International Journal of Geomechanics, 2024; 24: 04023260. doi:10.1061/IJGNAI.GMENG-8755.

[33] Alsirawan, R., Alnmr, A., Koch, E. Experimental and Numerical Investigation of Geosynthetic-Reinforced Pile-Supported Embankments for Loose Sandy Soils. Buildings, 2023; 13: 2179.

[34] Zhang, M., Zhu, H., Yang, J., Qiu, C., Javadi, A. A. Experimental study of a 3D printed geogrid embedded with FBG sensor for reinforcement of subgrade with underlying cave. Geotextiles and Geomembranes, 2023; 51: 81–92. doi:10.1016/j.geotexmem.2023.05.001.

[35] Shu, S., Zhang, F., Wang, D., Ge, B., Jiang, Y. Dynamic behavior and characteristics of geogrid-reinforced sand under cyclic loading. Soil Dynamics and Earthquake Engineering, 2024; 180: 108630. doi:10.1016/j.soildyn.2024.108630.

[36] Ghorbanzadeh, A., Nematzadeh, M., Tarkhan, M., Bagheri, M. Experimental and predictive study on shear and electrical behavior of concrete containing ceramic waste and EAF slag after high temperature exposure. Construction and Building Materials, 2026; 536: 146900. doi:10.1016/j.conbuildmat.2026.146900.

[37] Ranjbar, F., Asgari, A., Bagheri, M., Bengar, H. A. Insights into the mechanisms and seismic performance of reinforced concrete buildings on a pile foundation in liquefiable deposits. Structures, 2026; 87: 111557. doi:10.1016/j.istruc.2026.111557.

[38] Forcellini, D., Ranjbar, F., Asgari, A., Bagheri, M., Yazdani, H. Seismic resilience assessment of pile groups against liquefaction-induced lateral spreading: Influence of ground inclination and pile spacing. Engineering Structures, 2026; 360: 122786. doi:10.1016/j.engstruct.2026.122786.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Influence of Localized Beam Cracking Representation on Nonlinear Seismic Performance of RC Moment Frames

Ehsan Abbasalipour [a], Mehdi Dehestani [b*], Aref Hasanzadeh [b]

[a] *Department of Civil Engineering, Se.C., Islamic Azad University, Semnan, Iran*
[b] *Faculty of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran*

A R T I C L E I N F O

Article history:

Received 06 June 2026

Revised 28 June 2026

Accepted 01 August 2026

Available online 01 April 2027

Keywords:

Reinforced concrete cracking

Localized stiffness reduction

Seismic performance

Nonlinear static analysis

Nonlinear dynamic analysis

A B S T R A C T

Cracking in reinforced concrete (RC) members is a key mechanism governing stiffness degradation and seismic response of structural systems. This study examines the influence of different beam-cracking modeling approaches on the nonlinear seismic behavior of RC moment-resisting frames. Three RC frames with 3, 6, and 9 stories are analyzed under three modeling schemes: (1) uncracked elastic model, (2) code-based effective stiffness reduction model, and (3) localized cracking model distributed along beam elements. In the localized formulation, stiffness degradation is defined in a position-dependent manner, allowing non-uniform reduction along the member length to better represent crack concentration in plastic hinge regions and mid-span zones. This overcomes a major limitation of conventional effective stiffness methods, which assume uniform stiffness reduction and cannot capture spatial damage distribution. Nonlinear static (pushover) and nonlinear time-history analyses are performed using Imperial Valley, Duzce, and Northridge ground motions within a nonlinear analysis framework. Results show that cracking representation significantly affects seismic response. Considering cracking reduces initial stiffness, increases fundamental period, and decreases base shear capacity. In the code-based model, base shear capacity decreases by up to 32%, while roof displacement increases by 21%, 26%, and 8% for the 3-, 6-, and 9-story frames, respectively. The localized cracking model yields responses between the uncracked and code-based models and consistently predicts lower displacements than the latter. Crack distribution also significantly influences bending moments, plastic hinge formation, and energy dissipation. Overall, the localized cracking model provides an improved representation of stiffness degradation effects and a practical simplified framework for nonlinear seismic assessment of RC frames.

How to cite this article: Abbasalipour, E., Dehestani, M., Hasanzadeh, A. Influence of Localized Beam Cracking Representation on Nonlinear Seismic Performance of RC Moment Frames. Civil Engineering and Applied Solutions. 2027; 3(2): 102–121. doi:10.22080/ceas.2026.32032.1109.

1. Introduction

The bond between concrete and reinforcing steel constitutes one of the primary mechanisms responsible for force transfer and the satisfactory performance of reinforced concrete (RC) members. Perfect bond conditions do not exist in practice, and relative slip develops between concrete and reinforcement due to strain incompatibility between the two materials. This phenomenon leads to stiffness degradation, reduced energy dissipation capacity, and alterations in the ductility characteristics of structural members. Bond behavior is generally characterized through three successive stages: initial adhesion, stress transfer through reinforcement ribs following crack formation, and eventual bond failure resulting from insufficient concrete confinement. Crack distribution and strain

* Corresponding author.
E-mail addresses: dehestani@nit.ac.ir (M. Dehestani).

https://doi.org/10.22080/ceas.2026.32032.1109
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

localization in RC members are governed by the interaction between concrete softening and reinforcement-induced tension stiffening, while crack spacing is influenced by energy dissipation mechanisms within the member [1].

Cracking is recognized as one of the most influential factors affecting both the stiffness and durability of RC members. Cracks may develop due to flexure, shear, torsion, concrete shrinkage, reinforcement corrosion, thermal variations, and differential settlement, thereby influencing structural performance at both serviceability and ultimate limit states [1]. Although reinforcement contributes to controlling crack propagation and improving crack distribution, the post-cracking response remains highly dependent on concrete-steel bond characteristics, confinement conditions, and the nonlinear behavior of the constituent materials.

Extensive research has been conducted on bond-slip behavior and the modeling of concrete-steel interaction. The concrete-steel interface has been represented as a bond zone in which force transfer is simulated through bond stresses and relative slip. Despite providing accurate predictions, the increased number of degrees of freedom and the associated computational complexity have limited the applicability of such approaches in large-scale structural models [2]. Experimental bond stress-slip relationships under monotonic and cyclic loading conditions have been widely employed in the analysis of RC members and connections [3], and numerical investigations have demonstrated that accurate prediction of structural response requires the simultaneous consideration of reinforcement nonlinearity and bond-slip effects [4, 5].

To reduce computational demands, several simplified modeling approaches have been proposed. Equivalent reinforcement stiffness formulations allow bond-slip effects to be incorporated without significantly increasing model degrees of freedom [6], while fiber-based formulations represent bond-slip effects through equivalent strain components within member behavior [7-10]. Bond interface elements developed within the finite element framework have demonstrated considerable capability in reproducing the hysteretic response of RC members, although at the expense of higher computational cost [11-14]. Neglecting bond-slip effects leads to an overestimation of member stiffness and an underestimation of displacement demands in RC structures [15-17]. The representation of bond-slip behavior in regions experiencing inelastic deformations, particularly within plastic hinge zones, remains a challenge for conventional models and often requires additional modifications [18, 19]. Consequently, advanced analytical models have been proposed to predict bond-slip behavior and associated failure mechanisms with greater accuracy [20].

Alongside bond modeling, considerable attention has been devoted to crack propagation and its influence on the structural behavior of RC members. Crack location and depth can substantially affect the dynamic response and natural frequencies of structural elements [21, 22], and crack propagation in heterogeneous materials may significantly alter failure paths and damage evolution patterns [23-25]. Reinforcement corrosion and associated cracking phenomena have been identified as major contributors to stiffness degradation and durability reduction [26, 27]. Explicit formulations describing strain and slip distributions along reinforcing bars, while accounting for bond degradation mechanisms, have been developed to address localized cracking behavior [28-32], and studies on nonlinear RC frame modeling have further highlighted the significant role of connection deformability in determining seismic response characteristics [33]. Cracking has been shown to contribute to reductions in dynamic stiffness, modifications of modal characteristics, and deterioration of serviceability performance in RC structures [34]. Long-term crack development accelerates stiffness degradation and increases seismic vulnerability over time [35], while pre-existing cracks may adversely affect the residual strength of structural members subjected to fire exposure [36]. The reinforcement ratio and steel fiber reinforcement have been identified as significant parameters governing member ductility and post-cracking behavior. Fracture mechanics-based approaches and fictitious crack models have been developed to predict the nonlinear behavior of cracked RC beams, offering improved representation of these phenomena [37-40].

Despite substantial progress in this field, most existing seismic assessment approaches still represent cracking effects using uniform effective stiffness reduction factors applied along the entire member length. This simplifying assumption neglects the spatially non-uniform nature of cracking in RC members, particularly under seismic loading conditions. In reality, crack formation is highly localized and tends to concentrate in regions of maximum curvature demand, such as beam-column joints and potential plastic hinge zones.

However, conventional effective stiffness methods do not explicitly capture: (I) spatial variability of cracking along the member length, (II) interaction between localized cracking and plastic hinge formation, and (III) the resulting non-uniform stiffness degradation governing nonlinear seismic response behavior. As a result, these approaches may introduce bias in the prediction of stiffness distribution, displacement demand deformation distribution, and energy dissipation capacity of RC structural systems. Although fiber-based and detailed finite element models can represent such localized effects with higher accuracy, their implementation generally requires more detailed modeling assumptions and higher computational effort, which may limit their application in large-scale parametric seismic studies.

To address this limitation, the present study proposes a localized, position-dependent beam cracking representation, in which stiffness reduction is applied non-uniformly along beam elements based on expected damage concentration zones, including plastic hinge regions and mid-span areas. Unlike conventional code-based approaches, the proposed model explicitly accounts for the spatial distribution of cracking rather than assuming uniform stiffness loss.

The novelty of this study lies in systematically investigating the influence of spatially varying cracking representation on the nonlinear seismic response of RC moment-resisting frames with different heights. The proposed formulation provides an intermediate modeling framework between simplified code-based stiffness assumptions and more detailed nonlinear modeling approaches by incorporating spatially varying stiffness degradation within conventional frame analysis procedures. From a practical standpoint, the proposed method enhances the understanding of stiffness redistribution, plastic hinge development, and energy dissipation mechanisms, thereby offering improved insight for performance-based seismic assessment of RC structures.

Accordingly, the proposed framework introduces a position-dependent stiffness degradation strategy along RC beams that captures localized cracking effects in critical regions, providing a more detailed representation of stiffness variation compared with conventional uniform stiffness reduction approaches within nonlinear seismic analysis.

2. Numerical modeling and methodology

2.1. Design and modeling of structural frames

In this study, three-, six-, and nine-story reinforced concrete (RC) moment-resisting frames with residential occupancy and regular configurations in both plan and elevation were investigated. The structures were assumed to be located in a region of high seismic hazard and founded on Type II soil, in accordance with the Fourth Edition of the Iranian standard No. 2800. Structural modeling and analyses were performed using the SAP2000 software package.

To evaluate the influence of stiffness degradation caused by beam cracking on the seismic response of RC structures, nonlinear static (pushover) and nonlinear dynamic analyses were conducted. The nonlinear dynamic analyses were performed using three near-fault ground motion records, namely Imperial Valley, Duzce, and Northridge earthquakes. The dead loads of floors and roof were taken as 550 kg/m², while the live loads of floors and roof were assumed to be 200 kg/m^2 and 150 kg/m², respectively. A story height of 3.2 m was considered for all structures, and the lateral load-resisting system was selected as a reinforced concrete intermediate moment-resisting frame. The analytical models of the studied frames are presented in Figs. 1 to 3.

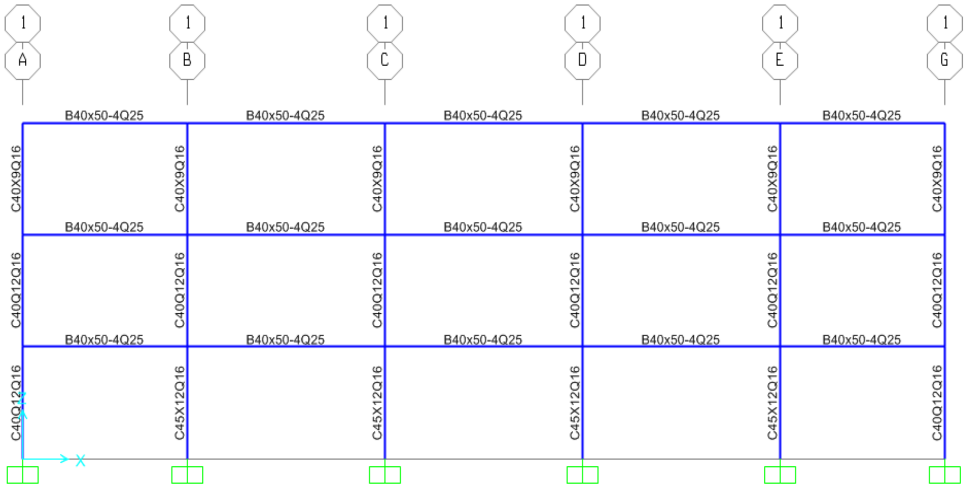


Fig. 1. Three-story intermediate reinforced concrete moment-resisting frame.

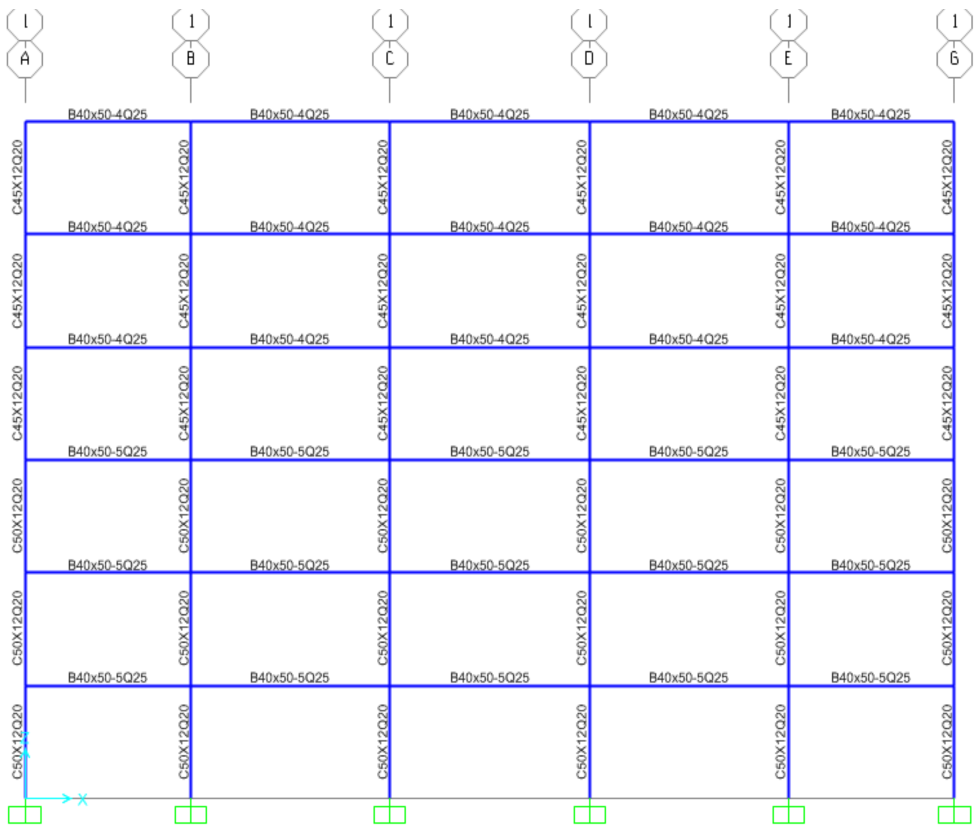


Fig. 2. Six-story intermediate reinforced concrete moment-resisting frame.

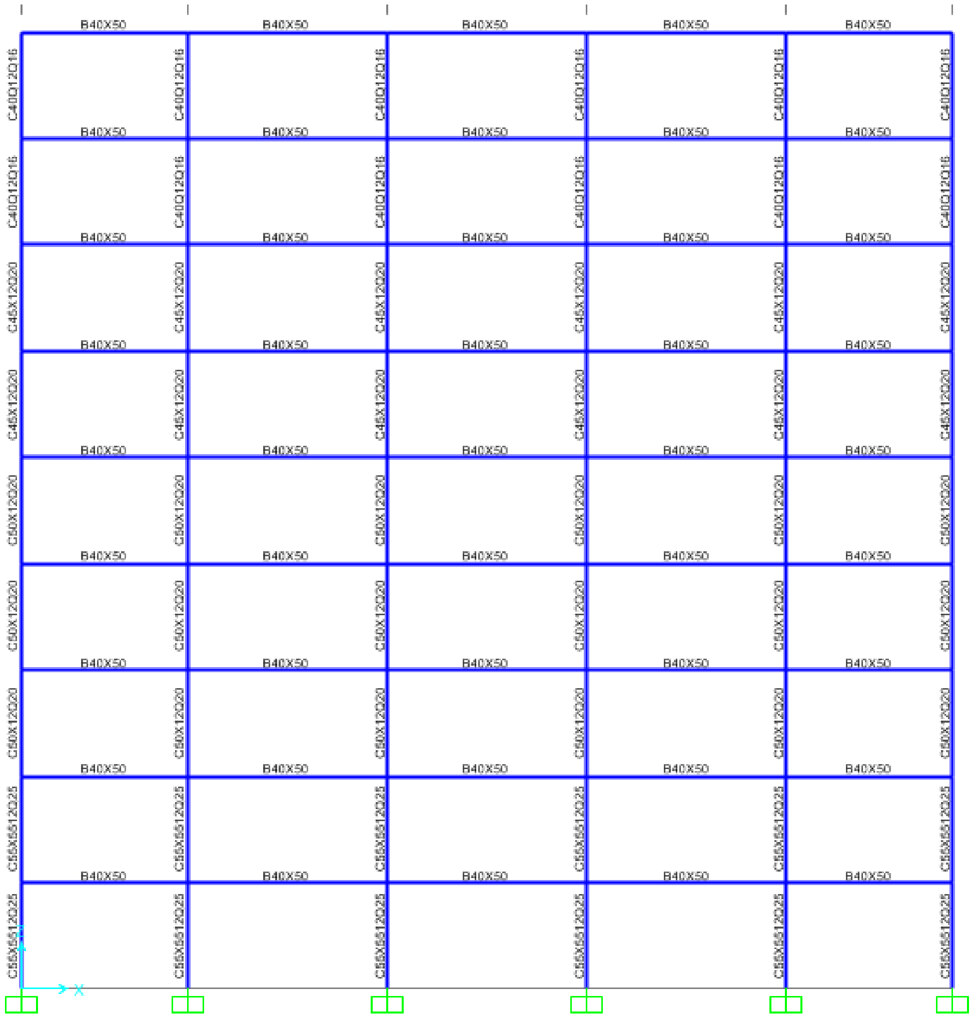


Fig. 3. Nine-story intermediate reinforced concrete moment-resisting frame.

To investigate the influence of cracking modeling approaches on the seismic behavior of the structures, three different modeling scenarios were considered: (I) an uncracked model, (II) a code-based stiffness reduction model, and (III) a localized cracking model. In the first scenario, the beam stiffness was defined without considering the effects of cracking. In the second scenario, beam stiffness was reduced according to the provisions of the Iranian standard No. 2800 by applying a stiffness modification factor of 0.35 to the flexural stiffness of beam members. In the third scenario, stiffness reduction was implemented locally in regions susceptible to crack concentration and plastic hinge formation.

Within the proposed localized cracking approach, the effective stiffness values of beam members were determined using the analytical relationships developed in the present study and subsequently assigned to critical regions of the beams, including both end zones and the mid-span region. Based on the findings reported by Forouzan [41], the stiffness degradation resulting from localized cracking and bond-related effects can be represented through an equivalent modification of the effective beam stiffness. The formulation proposed by Forouzan [41] was developed based on experimental investigations on reinforced concrete beam specimens, and its applicability was examined through comparisons with available experimental results. In this approach, the adjusted steel elastic modulus is introduced only as a numerical parameter to reproduce the experimentally supported reduction in member-level stiffness within the cracked regions of the beam model. Accordingly, the modified values were implemented within the beam elements to simulate the influence of localized deterioration effects on the global seismic response. The stiffness reduction factors employed in the analyses were calculated based on the analytical formulations proposed in this study and the reference methodology presented in Forouzan [41].

In this study, in order to account for the effect of slip in the beam-column connection region, the total beam displacement is decomposed into two components: flexural displacement, $\Delta_{bending}$, and displacement induced by connection rotational slip, Δ_{θ}. Accordingly, the total member displacement is obtained as the sum of these components, expressed as $\Delta_{total} = \Delta_{bending} + \Delta_{\theta}$.

The slip in the fixed-end region is further divided into two parts: axial slip resulting from the pull-out of embedded reinforcing bars, Δ_{axial}, and slip induced by flexural cracking in the beam end region, $\Delta_{bending}$. These two components are combined based on the principle of superposition to represent the total connection slip. The same slip-based stiffness modification framework was applied consistently across all three modeling scenarios to ensure uniform consideration of bond-slip effects.

The axial slip behavior is modeled based on the bond stress between concrete and steel along the development length of the reinforcing bars. For this purpose, analytical relationships and stress-slip models calibrated against experimental results are employed, and calibration parameters are adopted from previous studies [42]. In addition, the Gergely-Lutz empirical relationship is used to control crack width and to establish a link between cracking and slip behavior [43]. The second component of slip is associated with flexural cracking in the beam end region, where, after cracking, the tensile contribution of concrete is neglected, and force transfer is primarily carried by steel reinforcement. In this context, compatibility of strains and section equilibrium relations are used to incorporate the effect of this type of slip on the displacement response of the member. Finally, the total slip and beam displacement are implemented in the numerical model as an equivalent effect on the torsional and flexural stiffness of the member and are used in the nonlinear structural analysis.

2.2. Modification of beam stiffness due to bond slip

To incorporate the effect of end rotation in the plastic hinge region, the beam flexural stiffness (EI) is reduced within the plastic hinge length, L_P, at the member end, while the supports are modeled as fixed. In the present study, the localized cracking region is assumed to coincide with the plastic hinge zone at the beam ends, where the stiffness degradation associated with cracking and bond-slip effects is introduced. Accordingly, the length of the affected region is defined by the plastic hinge length L_P. For the determination of the plastic hinge length L_P, several empirical relationships have been proposed, among which the Sheikh and Bayrak formulation is commonly used. This expression includes parameters such as the section depth and an experimentally calibrated coefficient (ranging between 0.9 and 1.0) [44]. In the present study, an appropriate value of this coefficient is selected based on the modeling conditions. Subsequently, by equating the displacement obtained from the modified stiffness model with the displacement resulting from the rotational stiffness k_θ at the support, and employing the moment-area method, the equivalent stiffness in the plastic hinge region, EI_{eq}, is determined.

2.2.1. Evaluation of the modified stiffness in an internal continuous beam

Based on the study by Forouzan [41], in an internal continuous beam with an end rotational stiffness k_θ subjected to a uniformly distributed lateral load W, the lateral displacement consists of two components: the flexural deformation and the deformation induced by end rotation. Accordingly, the total displacement is expressed as shown in Eq. 1.

$$\Delta_1 = \frac{WL^4}{384EI} + \frac{WL^3}{48k_\theta} \quad (1)$$

The first term represents the flexural deformation of the beam, while the second term corresponds to the additional deformation due to the rotational stiffness k_θ at the beam ends. In this stage, the beam behavior is idealized as a cantilever member with an effective length of $L/2$ (Fig. 4).

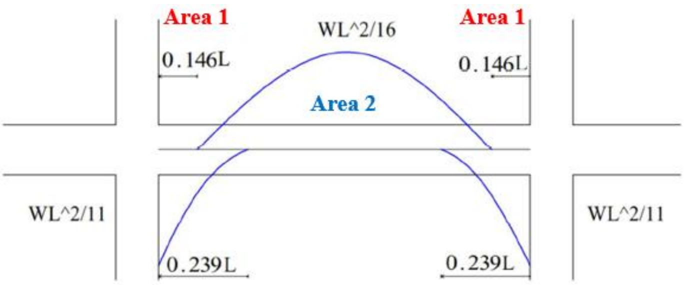


Fig. 4. Bending moment diagram of the internal continuous beam, showing the plastic hinge (localized cracking) zones at the beam ends (Area 1) and the remaining elastic region (Area 2) [41].

By equating the actual beam deflection with that of an equivalent beam having a reduced flexural rigidity EI_{eq}, and employing the moment-area method, the modified stiffness is determined. Accordingly, the bending moment distribution is evaluated considering symmetry in the two half-spans, and the moment functions are derived for two regions. Hence, the bending moment expression for Region 1 of the internal continuous beam is given by Eq. 2.

$$M_1 = -\frac{Wx^2}{2} + \frac{WLx}{2} - \frac{WL^2}{11} \quad (2)$$

The bending moment expression for Region 2 is given by Eq. 3.

$$M_2 = 0.4987Wx^2 - 0.353WLx \quad (3)$$

Boundary conditions and symmetry at the supports and midspan are then applied to determine the unknown coefficients. The deflection is evaluated using the moment-area method for half of the span only, as in Eq. 4.

$$\Delta = \int \frac{M(x)\,x}{EI}\,dx \quad (4)$$

Finally, the actual deflection is equated to the deflection of the equivalent model, as Eq. 5.

$$\Delta(EI_{eq}) = \Delta_{real} \quad (5)$$

From this equality, the equivalent flexural rigidity EI_{eq} is obtained as Eq. 6.

$$\frac{1}{EI_{eq}} = \frac{1}{EI} + \frac{1}{24\beta k_\theta L} \tag{6}$$

where $\beta = (-0.2495\alpha^3 + 0.201\alpha^2 - 0.004495\alpha)$ and $\alpha = L_P/L$. The coefficient β represents the position-dependent stiffness modification function used to describe the stiffness variation within the localized cracking region. Therefore, the stiffness reduction is not uniformly distributed along the beam length but is governed by the normalized cracked length parameter.

2.2.2. Evaluation of the modified stiffness in an exterior continuous beam

The moment distribution in an exterior continuous beam is not symmetric (Fig. 5); therefore, the rotations at the two supports are not identical. The maximum bending moment and maximum deflection occur at a distance of approximately $0.518L$ along the span. By dividing the beam into two segments, the exterior continuous beam can be idealized as two cantilever-type members with fixed-fixed-free boundary conditions, having lengths of $0.518L$ and $0.482L$, respectively. Each fixed-fixed-free beam segment with length L is equivalent to a fixed-fixed beam with an effective length of $2L$ [41].

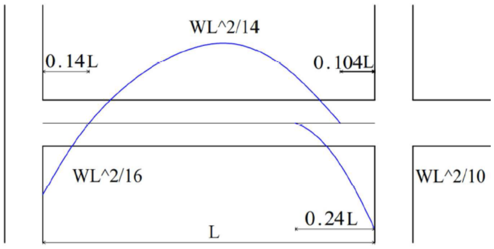


Fig. 5. Bending moment diagram of the exterior continuous beam [41].

2.2.3. Modified stiffness at 0.518L

The bending moment diagram of a fixed-fixed beam with a total length of $1.036L$ is shown in Fig. 6.

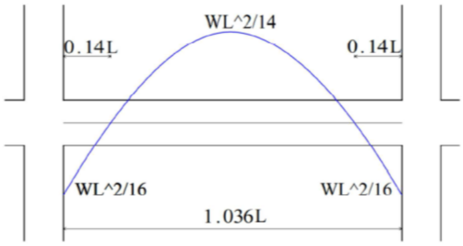


Fig. 6. Bending moment diagram of a fixed-fixed beam with a length of $1.036L$ in the exterior continuous beam [41].

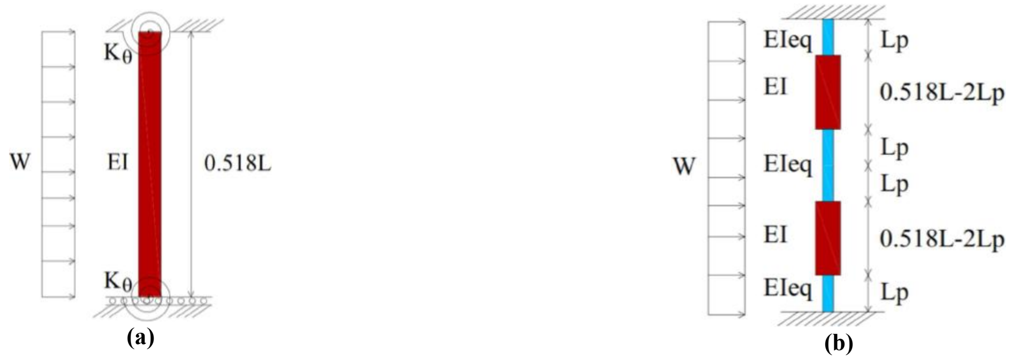


Fig. 7. (a) Simplified model of a fixed-free beam segment with length $0.518L$ subjected to a uniformly distributed load; and (b) equivalent fixed-fixed beam model with length $1.036L$ using EI_{eq} [41].

If a beam with rotational stiffness $k_{\theta 1}$ at both ends is subjected to a uniformly distributed horizontal load W, the horizontal displacement of the model shown in Fig. 7a is given as Eq. 7.

$$\Delta_1 = 0.002999\frac{WL^4}{EI_1} + 0.0347\frac{WL^3}{k_{\theta 1}} \tag{7}$$

where the first term represents the flexural deformation of the beam and the second term corresponds to the deformation induced by the end rotational stiffness $k_{\theta 1}$.

When the same load is applied to an equivalent beam with reduced flexural rigidity EI_{eq} (Fig. 7b), the horizontal displacement is obtained using the moment-area method. The bending moment function is assumed in the form $M = ax^2 + bx + c$, and the coefficients a, b, and c are determined by applying the boundary conditions. Consequently, the bending moment equation for the fixed-fixed beam under uniform loading is expressed as Eq. 8.

$$M = -0.4913Wx^2 + 0.517WLx - \frac{WL^2}{16} \tag{8}$$

After applying the boundary conditions, the coefficients are obtained. The deflection is then computed using the moment-area integration and equated to the response of the reduced-stiffness model as shown in Eq. 9.

$$\Delta(EI_{eq}) = \Delta_{real} \tag{9}$$

Accordingly, the modified stiffness EI_{eq} at $0.518L$ is obtained as shown in Eq 10.

$$\frac{1}{EI_{1eq}} = \frac{1}{EI_1} + \frac{0.01506}{\beta_1 k_\theta L} \tag{10}$$

where $\beta_1 = (0.1663\alpha_1^3 - 0.1878\alpha_1^2 + 0.02913\alpha_1)$ and $\alpha_1 = \frac{L_P}{1.036L}$.

2.2.4. Modified Stiffness at 0.482L

The bending moment diagram of a fixed-fixed beam with a length of $0.964L$ is presented in Fig. 8 [41].

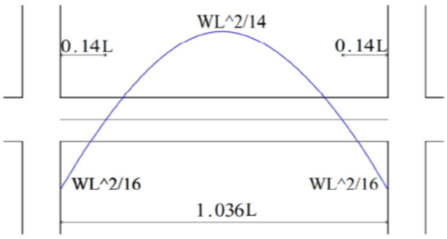


Fig. 8. Bending moment diagram of a fixed-fixed beam with a length of $0.964L$ in the side continuous beam [41].

(a) **(b)**

Fig. 9. (a) Simplified model of a fixed-roller beam with a length of $0.482L$ subjected to a uniformly distributed load; and (b) simplified model of a fixed-fixed beam with a length of $0.964L$ subjected to a uniformly distributed load [41].

When a beam with rotational stiffness at both ends is subjected to a uniformly distributed horizontal load W, as shown in Fig. 9a, the horizontal displacement is given by Eq. 11.

$$\Delta_1 = 0.00225\frac{WL^4}{EI_2} + 0.01866\frac{WL^3}{k_{\theta 2}} \tag{11}$$

The first term represents the flexural deformation of the beam, while the second term corresponds to the deformation associated with the rotational stiffness at the beam ends. When an equivalent load is applied to the beam with reduced stiffness EI_{eq} at both ends (Fig. 9b), the horizontal displacement is determined using the moment-area method. The bending moment diagram of the fixed-fixed beam with a length of $0.964L$ is divided into two segments. In each segment, the bending moment equation is expressed in the form $M = ax^2 + bx + c$, where the coefficients a, b, and $c$are obtained by applying the boundary conditions. The moment equation for the first region of the side continuous beam is expressed as Eq. 12.

$$M = -0.5722Wx^2 + 0.5516WLx - \frac{WL^2}{10} \tag{12}$$

The moment equation for the second region of the side continuous beam is given by Eq. 13.

$$M = 0.291Wx^2 - 0.2205Wx \tag{13}$$

Using symmetry, the analysis is performed for half of the beam, and the integration domain is divided into several segments. The displacement is obtained from the moment-area method and equated to that of the equivalent model as shown in Eq. 14.

$$\Delta(EI_{eq}) = \int M(x)\, x\, dx \tag{14}$$

Consequently, the modified stiffness EI_{eq} at $0.482L$ is obtained, as shown in Eq. 15.

$$\frac{1}{EI_{2eq}} = \frac{1+0.003095/\beta_2}{EI_2} + \frac{0.01041}{\beta_2 k_{\theta 2} L} \tag{15}$$

where $\beta_2 = (0.21595\alpha_2^4 - 0.2861\alpha_2^3 + 0.1094\alpha_2^2 + 0.1242\alpha_2)$ and $\alpha_2 = \frac{L_P}{0.964L}$.

2.2.5. Relationship between the stiffness of fixed-fixed beam sections at 0.518L and 0.482L

According to the study by Forouzan [41], the formulations used in the present research have been derived and verified. In the side continuous beam, the stiffness on the two sides of the support is not identical. Assuming identical beam geometry along the span, this stiffness difference arises from variations in the amount of reinforcing steel. In addition, the displacement at the $0.518L$ region is simultaneously influenced by both flexural deformation and rotational (torsional) deformation at the supports. Therefore, by having the properties of one span, the characteristics of the opposite side (including the amount of tensile and compressive reinforcement) can be determined through an iterative trial-and-error procedure.

2.3. Calculation of $E_{s,eq}$ based on the modified stiffness

In this study, the flexural stiffness of the beam is defined as the slope of the moment-curvature diagram. Based on equilibrium considerations and assuming yielding of the tensile steel, in the sectional analysis of the beam, the difference between tensile steel forces and the compressive forces contributed by concrete and reinforcing steel is evaluated by varying the neutral axis depth. This iteration is continued until the force imbalance becomes negligible. Once the neutral axis depth is determined, the yielding of the tensile reinforcement is verified. If yielding does not occur, the trial-and-error process is repeated until tensile yielding is achieved. After determining the neutral axis depth, the bending moment capacity is calculated. The stiffness is defined based on the relationship between moment and curvature and can be expressed as Eq. 16 [41].

$$EI = A_s E_s j d^2 (1 - k) \tag{16}$$

where A_s is the area of tensile reinforcement, E_s represents the elastic modulus parameter associated with the tensile reinforcement contribution in the section stiffness formulation, $k = [(\rho + \rho')(\eta)^2 + 2(\rho + \rho'\ d'/d)\eta]^{1/2} - (\rho + \rho')\eta$ and ρ is the ratio of tensile reinforcement area to concrete area, $\rho'$is the ratio of compressive reinforcement area to concrete area, and η is the ratio of the modulus of elasticity of steel to that of concrete. According to Eq. (16), the reinforcement-related parameters have a significant influence on the calculated flexural stiffness of reinforced concrete beams, whereas changes in reinforcement area alone may not provide the required stiffness modification. Whereas changes in A_s do not lead to the desired reduction in stiffness. Therefore, instead of directly modifying the constitutive behavior of reinforcing steel, an equivalent adjustment of the effective steel modulus is adopted to incorporate the stiffness degradation effect associated with cracking and bond-related effects, which consequently affects the parameters η and k. Using Eq. 16 together with the previously derived $EI_{eq}$for each beam, the equivalent steel modulus $E_{s,eq}$ is obtained [41]. It should be noted that the obtained equivalent modulus $E_{s,eq}$ is considered as a modeling parameter to represent the reduced effective stiffness of the cracked beam section rather than the actual mechanical property of reinforcing steel. For implementation in the numerical model, the calculated equivalent steel modulus $E_{s,eq}$ is assigned to the reinforcement material properties of the corresponding beam regions. Therefore, the proposed approach modifies the effective stiffness of the cracked beam segments while preserving the original constitutive behavior of reinforcing steel.

2.4. Selection and scaling of ground motion records

Near-fault ground motions exhibit distinct dynamic characteristics compared to far-field records due to the presence of velocity pulses induced by forward directivity effects and permanent displacement components. These features lead to a concentration of energy within a short time duration and can significantly amplify the nonlinear response of structures. Accordingly, the use of near-fault records is essential for a realistic assessment of seismic structural performance. In this study, three near-fault ground motions, namely Imperial Valley (1979), Northridge (1994), and Duzce (1999), were selected from the PEER database [45] to provide a consistent basis for comparative evaluation of the investigated modeling approaches. The selection was performed considering seismic characteristics such as magnitude, fault distance, faulting mechanism, and the compatibility of site conditions. The general characteristics of the selected records are presented in Table 1.

Table 1. Characteristics of selected near-fault ground motion records [45].

Record name	Event year	Location	Station	Magnitude	Fault mechanism
DUZCE	1999	Duzce, Turkey	IRIGM 487	7.14	Strike-slip
IMPERIAL VALLEY	1979	Imperial Valley	SAHOP CASA FLORES	6.53	Strike-slip
NORTHRIDGE	1994	Northridge	Pacoima Dam	6.69	Reverse

For dynamic analysis, pairs of horizontal components of the ground motions were employed. The scaling procedure was carried out in accordance with the provisions of Iranian standard 2800 as follows. First, each record was normalized such that its peak ground acceleration was equal to gravitational acceleration, g. Subsequently, the response spectra of each component were computed for 5% damping, and the combined spectrum of each pair was obtained using the square root of the sum of squares (SRSS) method. Next, scaling factors were determined such that the average of the combined spectra over the period range $0.2T$ to $1.5T$ (where T is the fundamental period of the structure) was not less than 90% of the design spectrum of standard 2800. The final scaling factors were applied to the normalized ground motions and used in nonlinear dynamic analyses. When a sufficient number of appropriate records is not available, the use of simulated ground motions compatible with the target seismic conditions is permitted. Moreover, the selected records should, as far as possible, be representative of the project site in terms of geological, tectonic, and soil conditions. The duration of strong motion is also considered to be at least 10 seconds or three times the fundamental period of the structure, whichever is greater. Finally, three pairs of scaled ground motions, in accordance with standard 2800, were used for comparative nonlinear dynamic analyses of the studied structures (3-, 6-, and 9-story buildings). The scaling results of the selected ground motions are illustrated in Figs. 10 to 12.

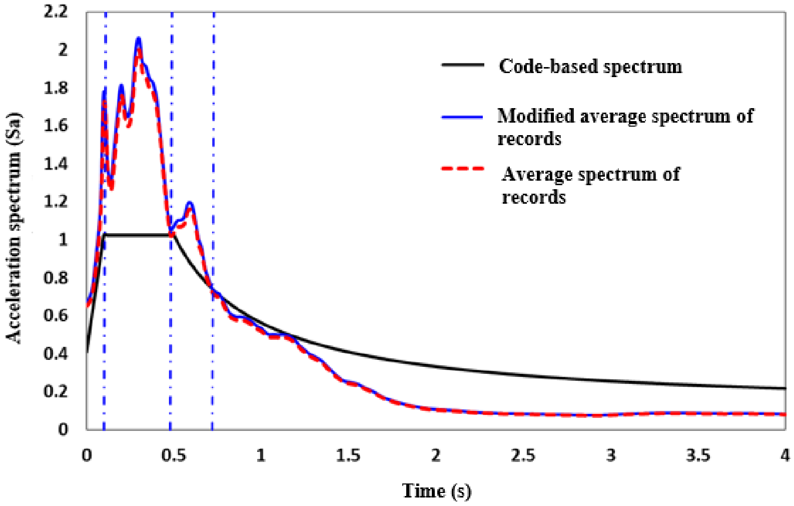


Fig. 10. Scaling of ground motions to the standard 2800 response spectrum for the 3-story structure.

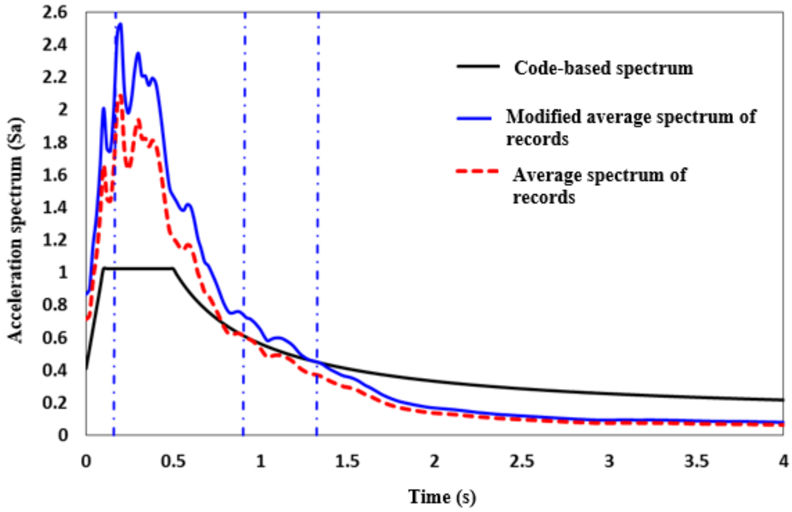


Fig. 11. Scaling of ground motions to the standard 2800 response spectrum for the 6-story structure.

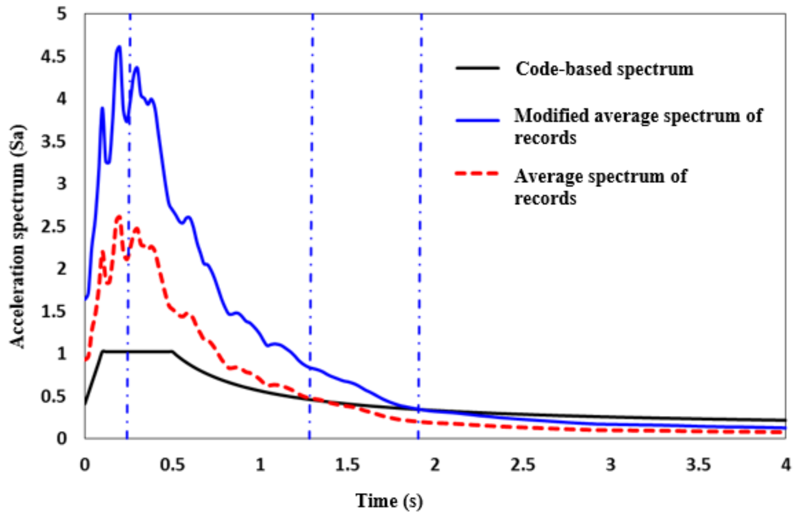


Fig. 12. Scaling of ground motions to the standard 2800 response spectrum for the 9-story structure.

The proposed localized cracking model is computationally efficient as it is implemented within the standard beam element formulation in SAP2000 without requiring mesh refinement or additional degrees of freedom. Compared with conventional detailed nonlinear modeling strategies, the proposed approach avoids the need for additional modeling complexity such as refined section discretization or increased degrees of freedom, while providing a more detailed representation of stiffness degradation than conventional uniform code-based assumptions. This makes the model suitable for practical seismic assessment of reinforced concrete moment-resisting frames.

3. Results and discussion

3.1. Static analysis results

The pushover curves for the 3-, 6-, and 9-story structures (Figs. 13 to 15) indicate that considering the cracking effect in the modeling leads to a noticeable reduction in the ultimate shear capacity and an overall degradation of the capacity curves across all investigated cases. This behavior reflects a reduction in the effective lateral stiffness of the system due to cracking. Examination of the initial slope of the pushover curves, which represents the initial stiffness of the structure, shows that the code-based cracking assumption results in a significantly more pronounced stiffness reduction, with a considerable change in the initial slope compared to the other modeling cases. This directly leads to an increase in the natural period of the structure. In contrast, in the beam-length cracking model, stiffness variations are more limited, although a reduction in the ultimate capacity is still observed.

Overall, the obtained results indicate that the cracking effect, particularly under code-based assumptions, plays a decisive role in reducing the lateral stiffness and seismic capacity of the investigated structures. Neglecting this effect may therefore lead to an unconservative estimation of the nonlinear structural response.

This behavior can be attributed to the reduction in effective flexural rigidity in cracked beam regions, which leads to a decrease in the global lateral stiffness of the structural system. In the code-based approach, the uniform stiffness reduction accelerates global softening, resulting in a more pronounced increase in the fundamental period. In contrast, the localized cracking model preserves stiffness in less damaged regions, leading to a more gradual degradation of stiffness and a smoother capacity curve response.

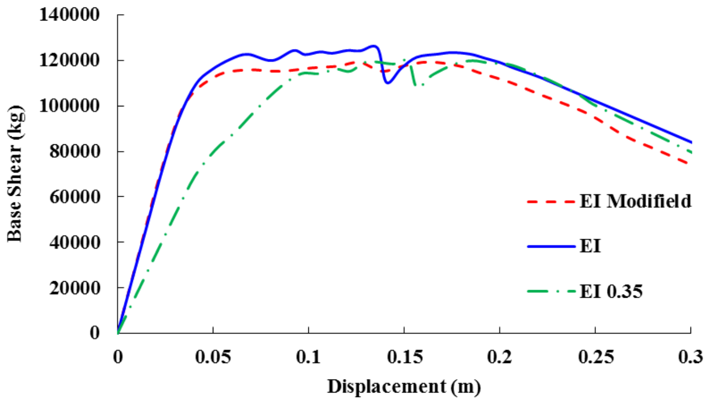


Fig. 13. Pushover curve of the 3-story structure under all three cracking assumptions.

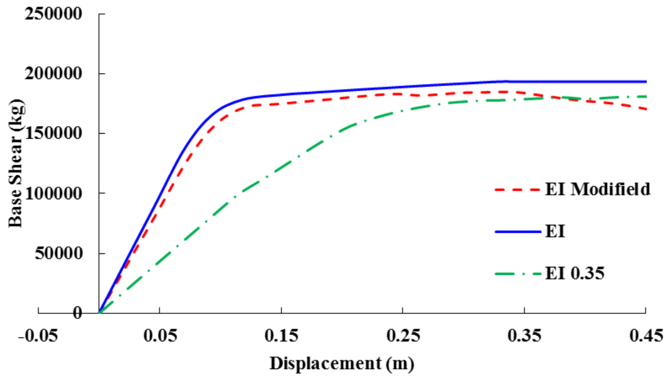


Fig. 14. Pushover curve of the 6-story structure under all three cracking assumptions.

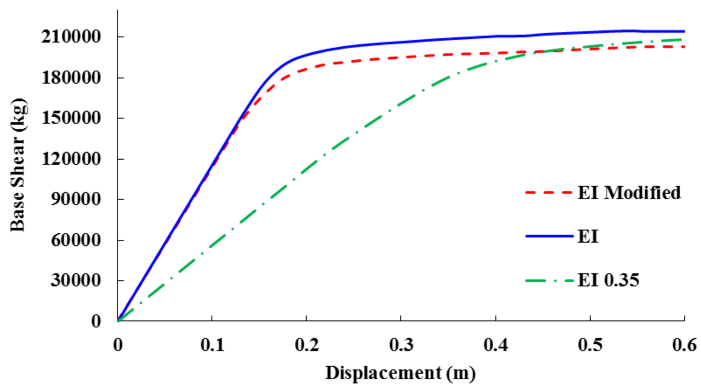


Fig. 15. Pushover curve of the 9-story structure under all three cracking assumptions.

3.2. Comparison of the effect of considering cracking in nonlinear dynamic analysis

For the 3-story structure, the inclusion of cracking had a significant influence on the seismic response under different ground motion records. In the code-based cracking model, the base shear decreased by approximately 35% based on the average response of the investigated ground motions, while the roof displacement increased by about 21% compared with the uncracked case. This reduction in base shear is primarily due to global stiffness degradation, which modifies the dynamic characteristics of the structure and reduces lateral force demand. Conversely, in the localized cracking model, stiffness degradation is distributed along beam elements rather than being uniformly applied, resulting in a more balanced redistribution of internal forces and reducing excessive global softening within the investigated cases.

The increase in roof displacement is associated with the formation of softened hinge regions at beam ends, where localized cracking significantly reduces rotational stiffness. This localized flexibility increases global deformation demand without substantially altering the overall force path of the structure. This behavior is also consistent with the formation of plastic hinge regions at beam ends, where nonlinear deformations become concentrated. The development and progression of these hinges contribute to the observed increase in global flexibility and roof displacement.

In contrast, in the beam-length cracking model, the base shear increased by approximately 10%, whereas the roof displacement remained nearly unchanged compared to the uncracked case. In terms of flexural response, the mid-span beam moment decreased by approximately 77% in the code-based cracking model, whereas an increase of about 45% was observed in the beam-length cracking model. This variation in beam moment response reflects changes in internal force redistribution mechanisms. In the code-based model, stiffness degradation reduces moment demand through global force redistribution, whereas in the localized cracking model, partial stiffness retention in beam spans allows higher moment demand to develop, particularly in mid-span regions.

Furthermore, the energy dissipation results indicate that cracking generally leads to a reduction in dissipated energy. On average, the amount of absorbed energy in the code-based cracking case was approximately 34% lower than that of the beam-length cracking model. This reduction in dissipated energy suggests a tendency toward a more concentrated inelastic response within the investigated cases, whereas the localized cracking approach promotes a more progressive development of inelastic mechanisms along beam elements, resulting in a more distributed energy dissipation pattern.

The corresponding results are illustrated in Figs. 16 to 19.

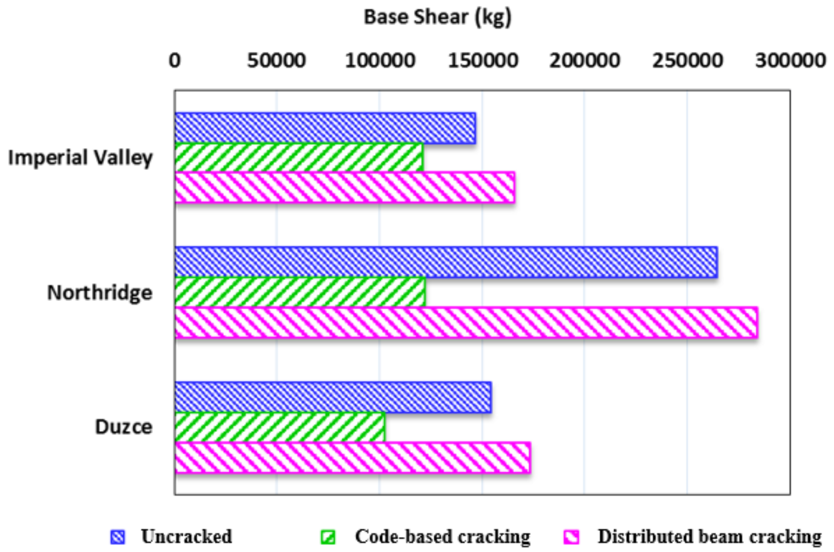


Fig. 16. Comparison of base shear responses for the 3-story structure with and without considering cracking under all ground motions.

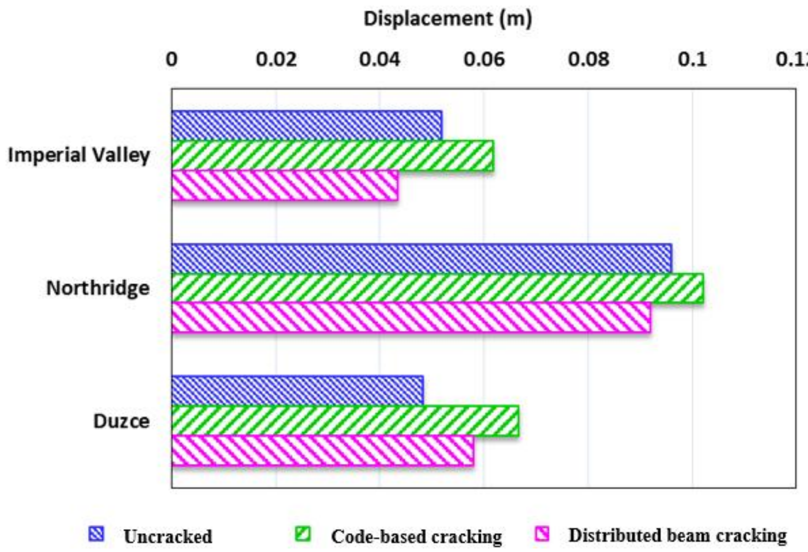


Fig. 17. Comparison of roof displacement responses for the 3-story structure with and without considering cracking under all ground motions.

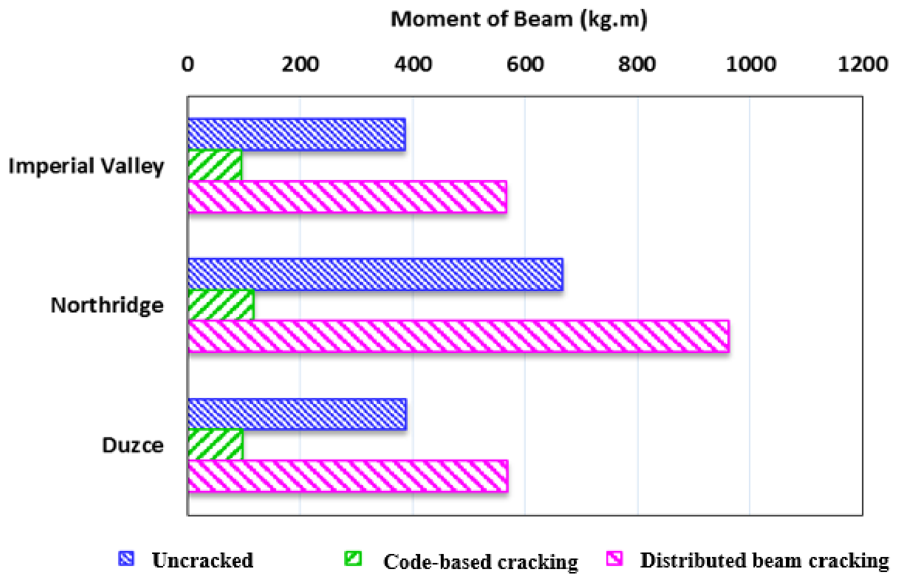


Fig. 18. Comparison of beam bending moment responses for the 3-story structure with and without considering cracking under all ground motions.

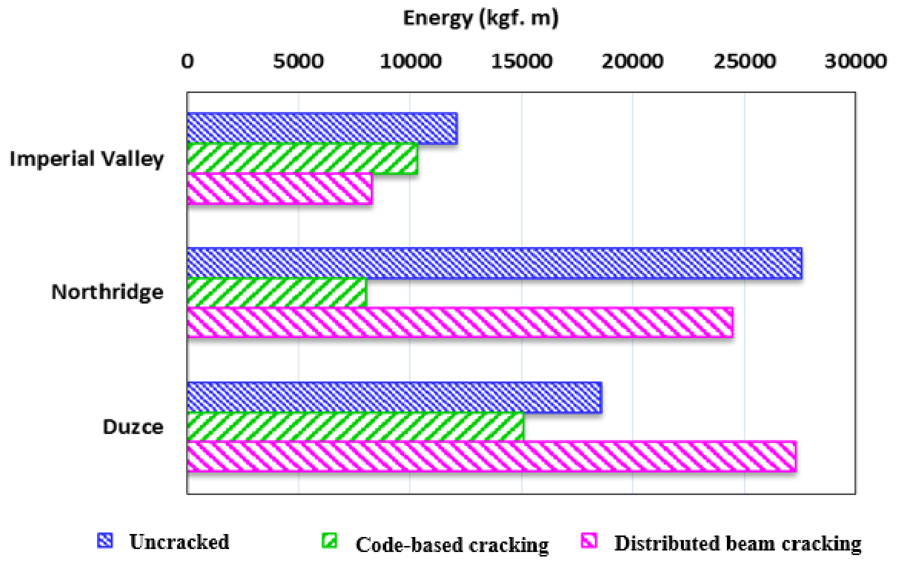


Fig. 19. Comparison of dissipated energy for the 3-story structure with and without considering cracking under all ground motions.

In the 6-story structure, consideration of cracking has led to significant variations in seismic response under different ground motion records. In the code-based cracking model, the base shear decreased by approximately 32% on average among the investigated ground motions, whereas in the beam-distributed cracking model, this reduction is about 9%. This trend is associated with increased global flexibility due to stiffness degradation, which alters the modal properties of the structure and reduces lateral force demand. In addition, the change in stiffness distribution leads to modifications in modal participation factors, where higher vibration modes become more active, resulting in a more distributed inertia force pattern along the height of the structure. However, the distributed cracking model mitigates excessive stiffness loss by preserving stiffness continuity along beam spans, leading to a more balanced global response.

Regarding roof displacement, the code-based cracking scenario resulted in an increase of approximately 26%, while in the beam-distributed cracking model, the roof displacement decreased by about 14% on average. This variation is governed by the development of localized flexibility zones in beam-end regions, which act as rotational hinges and increase the global lateral deformation response without significantly altering the global force transfer mechanism.

In terms of flexural response, the mid-span beam moment decreased by approximately 81% in the code-based cracking model and by about 16% in the beam-distributed cracking model. This significant reduction in bending moment is directly linked to the redistribution of internal forces caused by stiffness degradation, which shifts demand away from beam mid-spans toward alternative load paths within the structural system. This redistribution is governed by the progressive development of plastic hinges at beam-column connections, which serve as primary inelastic zones controlling the global response.

Furthermore, energy-related results indicate that, for most ground motion records, the dissipated energy in the code-based cracking case is, on average, approximately 19% lower than that in the beam-distributed cracking case. This reduction in dissipated energy suggests a more concentrated inelastic response in the code-based model, whereas the distributed cracking approach enables a more progressive and spatially distributed development of inelastic mechanisms along beam elements.

The corresponding results are illustrated in Figs. 20 to 23.

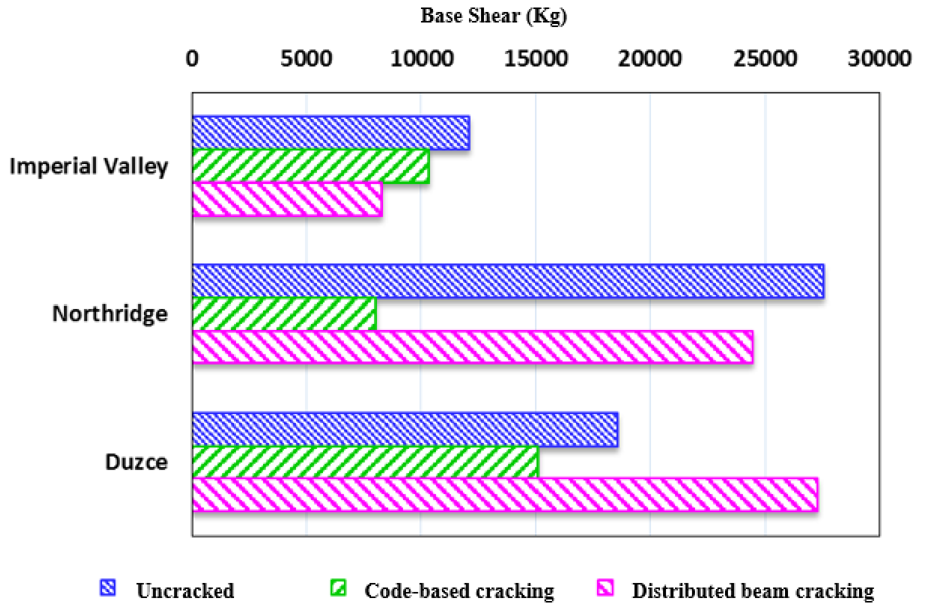


Fig. 20. Comparison of base shear results in the 6-story structure with and without considering cracking for all ground motions.

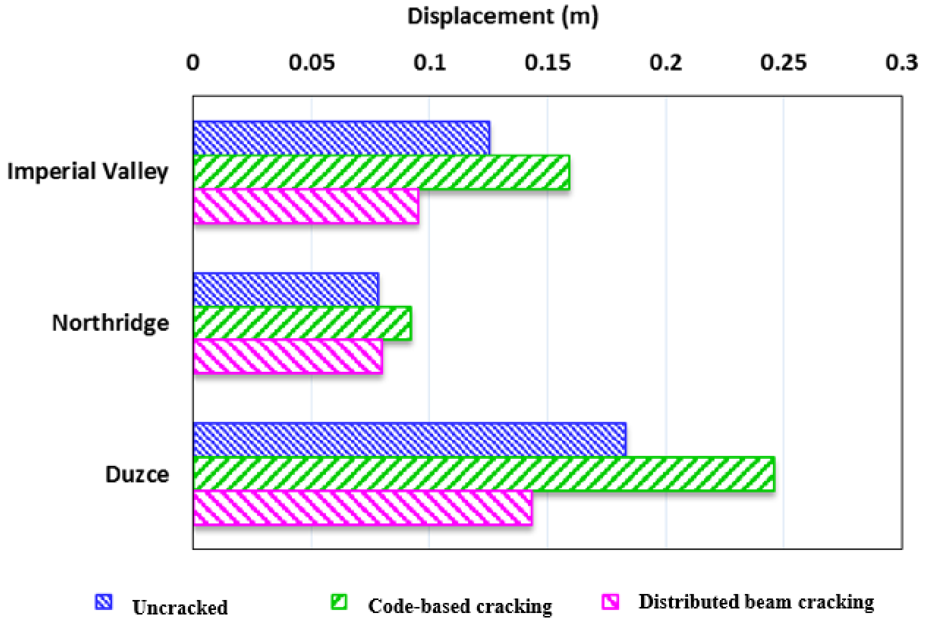


Fig. 21. Comparison of roof displacement results in the 6-story structure with and without considering cracking for all ground motions.

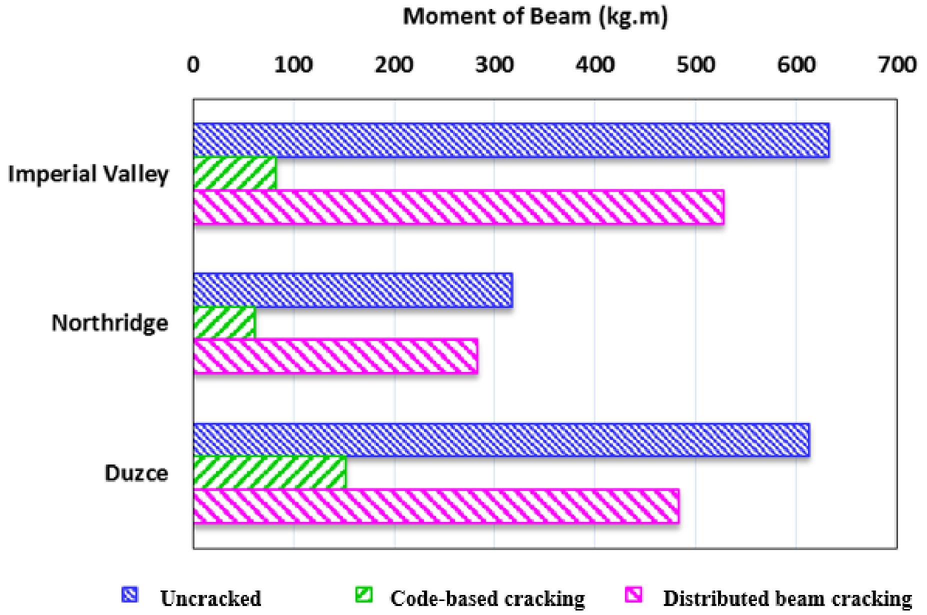


Fig. 22. Comparison of beam bending moment results in the 6-story structure with and without considering cracking for all ground motions.

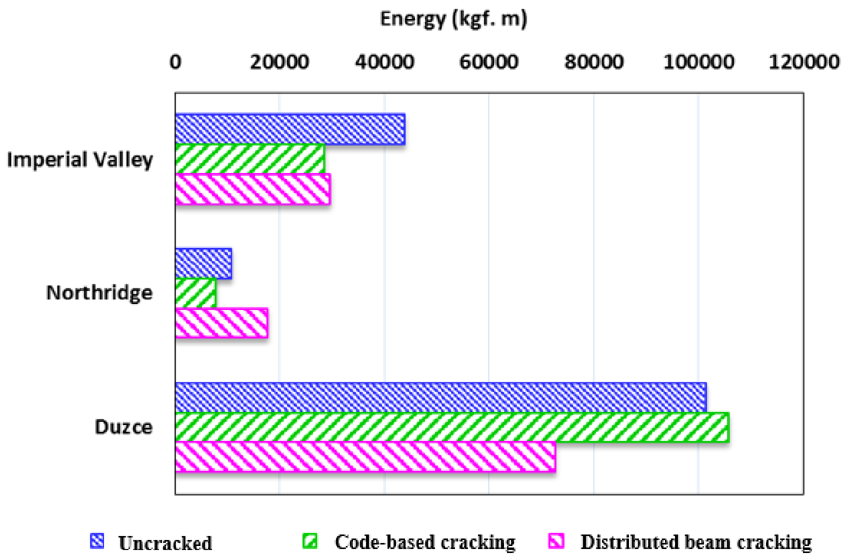


Fig. 23. Comparison of dissipated energy results in the 6-story structure with and without considering cracking for all ground motions.

In the 9-story structure, the consideration of cracking has resulted in different variations in seismic response under various ground motion records. In the code-based cracking model, the base shear decreased by approximately 31% on average among the investigated ground motions, whereas in the beam-distributed cracking model, a negligible change (approximately 1% increase) was observed compared to the uncracked case. In taller structures, the influence of higher vibration modes becomes more pronounced, which reduces the sensitivity of global response to localized stiffness changes. This increased contribution of higher modes reduces the dominance of the first mode and results in a more complex dynamic response, where stiffness changes at local levels have a reduced influence on global behavior. As a result, stiffness degradation effects are partially distributed across multiple stories rather than being concentrated in specific regions.

Regarding roof displacement, the code-based cracking model exhibited an increase of about 8%, while the beam-distributed cracking model showed an increase of approximately 5% relative to the uncracked condition. The relatively smaller increase in roof displacement compared to lower-rise structures is attributed to the higher contribution of higher vibration modes and the more distributed deformation pattern along the height of the structure, which limits the concentration of lateral deformations in specific stories.

In terms of flexural response, the mid-span beam moment decreased by approximately 81% in the code-based cracking model, whereas in the beam-distributed cracking model, an increase of about 5% was reported. This variation in bending moment response reflects a shift in internal force distribution mechanisms, where increased modal participation in taller structures leads to a more distributed demand pattern and reduces the dominance of local stiffness degradation effects.

Furthermore, energy-related results indicate that the dissipated energy in the code-based cracking case is, on average, approximately 47% lower than that in the beam-distributed cracking model. This lower sensitivity of energy dissipation suggests that in taller buildings, inelastic demands are more widely distributed across multiple stories, reducing the impact of localized cracking on global energy absorption capacity within the investigated cases.

These results are presented in Figs. 24 to 27.

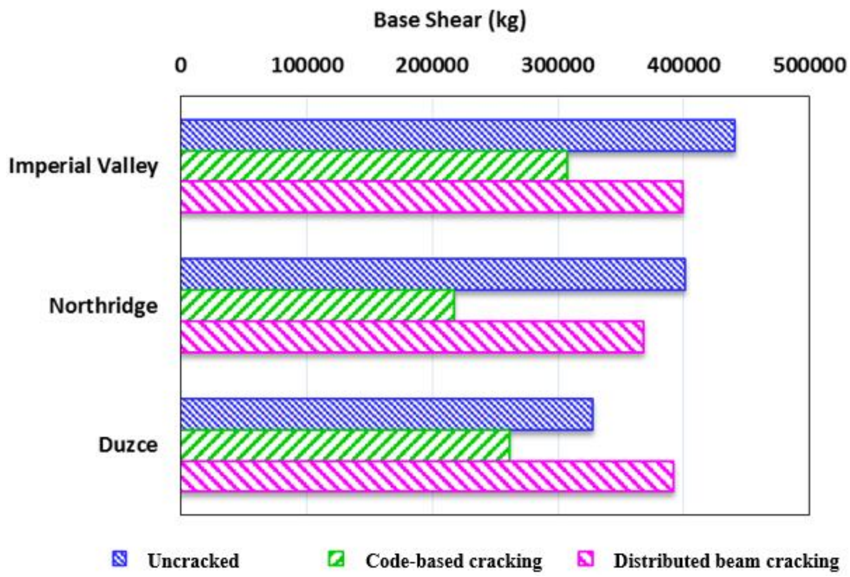


Fig. 24. Comparison of base shear results in the 9-story structure with and without considering cracking for all ground motions.

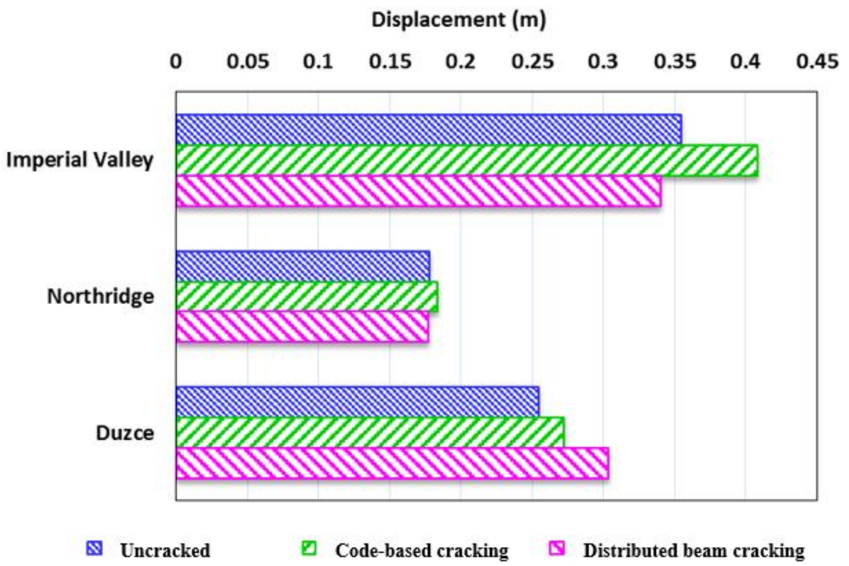


Fig. 25. Comparison of roof displacement results in the 9-story structure with and without considering cracking for all ground motions.

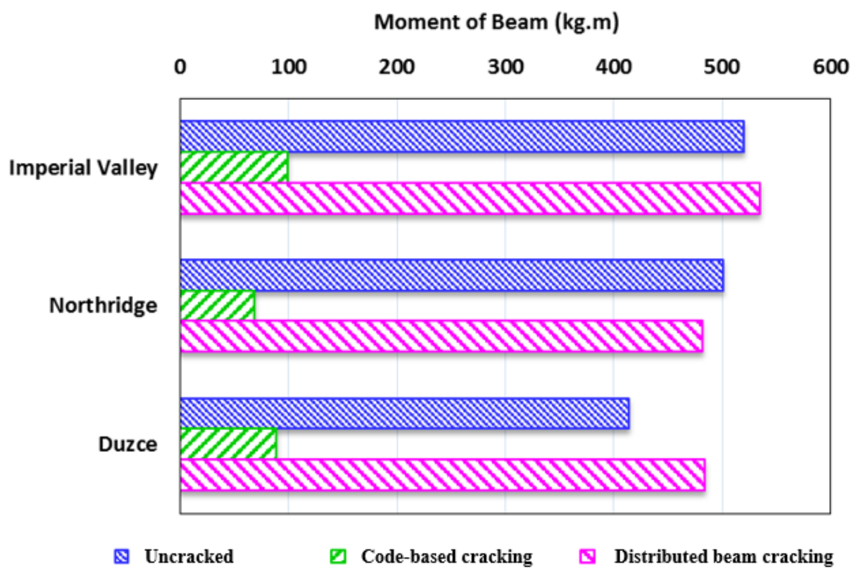


Fig. 26. Comparison of beam bending moment results in the 9-story structure with and without considering cracking for all ground motions.

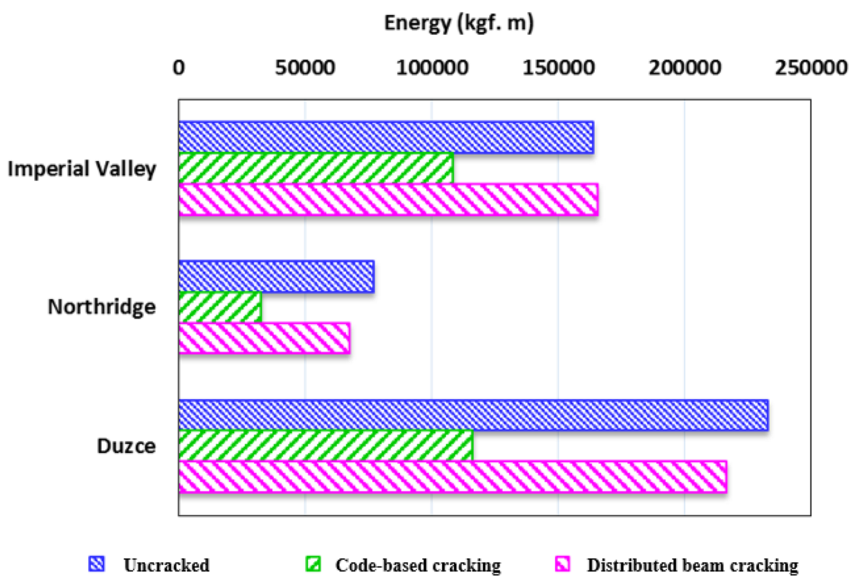


Fig. 27. Comparison of dissipated energy results in the 9-story structure with and without considering cracking for all ground motions.

To facilitate comparison among the investigated frame heights and cracking modeling approaches, Table 2 summarizes the average percentage variation of the principal response parameters relative to the uncracked model. Positive values indicate an increase, whereas negative values indicate a decrease in the corresponding response parameter.

Table 2. Summary of average percentage variations in seismic response parameters relative to the uncracked model.

Frame height	Cracking model	Base shear (%)	Roof displacement (%)	Mid-span beam moment (%)	Dissipated energy
3-Story	Code-based cracking	-35	21	-77	34% lower than beam-length cracking model
3-Story	Beam-length cracking	10	0	45	Reference
6-Story	Code-based cracking	-32	26	-81	19% lower than beam-length cracking model
6-Story	Beam-length cracking	-9	-14	-16	Reference
9-Story	Code-based cracking	-31	8	-81	47% lower than beam-length cracking model
9-Story	Beam-length cracking	1	5	5	Reference

Although interstory drift and residual deformation were not explicitly evaluated, the observed variations in roof displacement, energy dissipation, and stiffness degradation can be considered as representative indicators of global deformation demand and overall nonlinear seismic behavior. These response parameters provide a consistent basis for comparative evaluation of the investigated modeling approaches. However, these results are interpreted as comparative indicators for the considered ground motion records rather than statistically generalized seismic demand estimates. The limited number of records considered in this study did not allow a comprehensive statistical evaluation based on standard deviation, coefficient of variation, and confidence intervals.

It should also be noted that the proposed approach was not directly compared with fiber-based models in the present study. Therefore, further investigations involving experimentally validated fiber-section models are recommended to assess the accuracy and computational efficiency of the proposed modeling strategy.

4. Conclusions

In this study, the seismic response of 3-, 6-, and 9-story reinforced concrete moment-resisting frames was investigated considering different cracking representations, including uncracked conditions, code-based stiffness degradation, and beam-distributed cracking models using SAP2000. The results demonstrated that the modeling approach used to represent cracking significantly influences both global and local nonlinear structural behavior, particularly in terms of stiffness degradation patterns, plastic hinge development, modal characteristics, and energy dissipation mechanisms.

The findings indicate that the manner in which cracking is modeled plays a key role in governing the balance between global and local response characteristics. Uniform code-based stiffness reduction tends to induce more pronounced global softening and concentrated inelastic demand, whereas beam-distributed cracking provides a more gradual and spatially distributed representation of stiffness degradation, leading to a more balanced internal force redistribution and a more realistic simulation of flexural behavior in reinforced concrete beams.

From an engineering perspective, simplified uniform cracking approaches may lead to conservative estimations of seismic demand, while distributed cracking models can better capture the progressive development of nonlinear behavior in RC frames. Therefore, careful selection of cracking representation is essential for reliable seismic performance assessment of reinforced concrete structures.

The present study is subject to certain limitations that should be considered in interpreting the results. The analysis is performed on regular 3-, 6-, and 9-story reinforced concrete moment-resisting frames, and extension to irregular or taller structures requires further investigation. Only three near-fault ground motion records were used, which provide a limited dataset for capturing full record-to-record variability. The modeling primarily focuses on beam cracking effects, while other deterioration mechanisms such as column damage, shear failure, and bond-slip effects are not explicitly considered. Recent studies have shown that column shear degradation and confinement effects can significantly influence the seismic response of RC frames. It has been demonstrated that neglecting shear behavior may reduce the accuracy of nonlinear RC frame analysis [46], while ANN-based approaches have been developed for predicting the shear response of RC columns under axial loading conditions [47]. Furthermore, recent data-driven formulations have highlighted the significant role of confinement parameters and sectional characteristics in evaluating concrete column performance [48, 49]. Therefore, future studies should incorporate column shear degradation, confinement effects, and other interacting deterioration mechanisms into the proposed framework to achieve a more comprehensive seismic assessment of RC moment-resisting frames [50]. In addition, the proposed modeling approach has not been independently validated against experimental frame-level results or detailed numerical simulations. Therefore, the obtained results should be interpreted within the adopted modeling assumptions. Furthermore, the present study focuses on global response indicators, including roof displacement, base shear, bending moment, and energy dissipation. Interstory drift ratio and other local damage indicators were not explicitly evaluated and are recommended for future investigations to provide a more comprehensive seismic performance assessment.

Statements & declarations

Author contributions

Ehsan Abbasalipour: Conceptualization, Idea, Methodology, Investigation, Visualization, Writing - Original Draft, Formal analysis.

Mehdi Dehestani: Conceptualization, Methodology, Supervision, Project administration, Validation, Writing - Review & Editing.

Aref Hasanzadeh: Data curation, Methodology, Validation, Writing - Review & Editing.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The data presented in this study will be available upon request from the corresponding author.

Declarations

The authors declare no conflict of interest.

Appendix A. Supplementary relations and detailed moment-area formulations

A.1. Interior continuous beam

To compute the displacement, the moment–area method is employed, and due to symmetry, only half of the span is considered. The integration domain is divided into four segments, and the displacement relation is obtained from the first moment of area of the bending moment diagrams. The moment-area equation is expressed as Eq. A.1.

$$\bar{X}_1 = 0.5L - \frac{x_1 A_1}{A_1} = 0.5L - \frac{\int_0^{L_p} \frac{xM_1}{EI_{eq}} dx}{\int_0^{L_p} \frac{M_1}{EI_{eq}} dx} \tag{A.1a}$$

$$\bar{X}_2 = 0.5L - \frac{x_2 A_2}{A_2} = 0.5L - \frac{\int_{L_p}^{0.239L} \frac{xM_1}{EI} dx}{\int_{L_p}^{0.239L} \frac{M_1}{EI} dx} \tag{A.1b}$$

$$\bar{X}_3 = 0.354L - \frac{x_3 A_3}{A_3} = 0.354L - \frac{\int_{L_p}^{0} \frac{xM_2^4}{EI} dx}{\int_{L_p}^{0.354L-L_p} \frac{M_2}{EI} dx} \tag{A.1c}$$

$$\bar{X}_4 = 0.354L - \frac{x_4 A_4}{A_4} = 0.354L - \frac{\int_{0.354L-L_p}^{0.354L} \frac{xM_2}{EI_{eq}} dx}{\int_{0.354L-L_p}^{0.354L} \frac{M_2}{EI_{eq}} dx} \tag{A.1d}$$

The beam deflection with reduced stiffness is obtained from Eq. A.2.

$$\Delta_2 = -2\left[\frac{W}{EI_{eq}}\left(-0.2495LL_p^3 + 0.201L^2L_p^2 - 0.04495L^3L_p\right) + \frac{W}{EI}\left(0.0044L^4 + 0.2495LL_p^3 - 0.201L^2L_p^2 + 0.004495L^3L_p\right)\right] \tag{A.2}$$

By equating the two expressions, the equivalent flexural rigidity is obtained as Eq. A.3.

$$\frac{1}{EI_{eq}} = \frac{1}{EI} + \frac{1}{24\beta k_\theta L} \tag{A.3}$$

where $\beta = -0.2495\alpha^3 + 0.201\alpha^2 - 0.004495\alpha, \alpha = \frac{L_p}{L}$.

A.2. Side continuous beam

A.2.1. At 0.518L

Due to symmetry, half of the beam span is considered for moment-area calculations. At $0.14L$, the bending moment is zero; therefore, this point is included in the integration limits. Since the moment-area is evaluated with respect to the midspan while integration begins from the support, the centroid locations are defined as Eq. A.4.

$$\bar{X}_1 = 0.518L - \frac{\int_0^{L_p} \frac{M}{EI_{eq}} dx}{\int_0^{L_p} \frac{xM}{EI_{eq}} dx} \tag{A.4a}$$

$$\bar{X}_2 = 0.518L - \frac{\int_{L_p}^{0.14L} \frac{M}{EI} dx}{\int_{L_p}^{0.14L} \frac{xM}{EI} dx} \tag{A.4b}$$

$$\bar{X}_3 = 0.518L - \frac{x_3A_3}{A_3} = 0.518L - \frac{\int_{0.14L}^{0.518L-L_P}\frac{M}{EI}dx}{\int_{L_P}^{0.354L-L_P}\frac{xM}{EI}dx} \tag{A.4c}$$

$$\bar{X}_4 = 0.518L - \frac{x_4A_4}{A_4} = 0.518L - \frac{\int_{0.518L-L_P}^{0.518L}\frac{M_2^4}{(EI)_{eq}}dx}{\int_{0.518L-L_P}^{0.518L}\frac{xM_2}{(EI)_{eq}}dx} \tag{A.4d}$$

The beam deflection with reduced stiffness is given by Eq. A.5.

$$\Delta_2 = \frac{W}{EI_{eq}}\left(-0.1723LL_p^3 + 0.2016L^2L_p^2 - 0.0324L^3L_p\right) + \frac{W}{EI_1}\left(-0.003626L^4 + 0.1723LL_p^3 - 0.2016L^2L_p^2 + 0.0324L^3L_p\right) \tag{A.5}$$

By equating the expressions, the equivalent stiffness is obtained as Eq. A.6.

$$\frac{1}{EI_{1eq}} = \frac{1}{EI_1} + \frac{0.01506}{\beta_1 k_\theta L} \tag{A.6}$$

where $\beta_1 = 0.1663\alpha_1^3 - 0.1878\alpha_1^2 + 0.02913\alpha_1, \alpha_1 = \frac{L_p}{1.036L}$.

A.2.2. At 0.482L

For moment-area evaluation, half of the beam span is used due to symmetry. In addition, the integration domain is divided into four regions because of variations in bending moment expressions and stiffness along the span. The centroid locations are defined as in Eq. A.7.

$$\bar{X}_1 = 0.482L - \frac{\int_0^{L_p}\frac{M_2}{EI_2}dx}{\int_0^{L_p}\frac{xM_1}{EI_2}dx} \tag{A.7a}$$

$$\bar{X}_2 = 0.482L - \frac{\int_{L_p}^{0.242L}\frac{M_1}{EI_2}dx}{\int_{L_p}^{0.14L}\frac{xM_1}{EI_2}dx} \tag{A.7b}$$

$$\bar{X}_3 = 0.378L - \frac{\int_0^{0.378L-L_p}\frac{M_2}{EI_2}dx}{\int_0^{0.378L-L_p}\frac{xM_2}{EI_2}dx} \tag{A.7c}$$

$$\bar{X}_4 = 0.378L - \frac{\int_{0.378L-L_p}^{0.378L}\frac{M_2}{EI_2}dx}{\int_{0.378L-L_p}^{0.378L}\frac{xM_2}{EI_2}dx} \tag{A.7d}$$

The corresponding deflection expression is given by Eq. 14, and the reduced-stiffness form is expressed in Eq. A.8.

$$\Delta_2 = 2\left[\frac{W}{EI_{2eq}}\left(-0.2758LL_p^3 + 0.1016L^2L_p^2 + 0.1113L^3L_p + 0.21595L^4\right) + \frac{W}{EI_2}\left(-0.00256L^4 + 0.2758LL_p^3 - 0.10165L^2L_p^2 + 0.1113L^3L_p - 0.21595L^4\right)\right] \tag{A.8}$$

Finally, by equating the expressions, the equivalent stiffness is obtained as Eq. A.9.

$$\frac{1}{EI_{2eq}} = \frac{1+0.003095/\beta_2}{EI_2} + \frac{0.01041}{\beta_2 k_{\theta 2} L} \tag{A.9}$$

where $\beta_2 = 0.21595\alpha_2^4 - 0.2861\alpha_2^3 + 0.1094\alpha_2^2 + 0.1242\alpha_2, \alpha_2 = \frac{L_p}{0.964L}$.

References

[1] Chen, G., Baker, G., Hunt, G. W. Computational approach to localization using global energy minimization. Computers & Structures, 2000; 78: 529–536. doi:10.1016/S0045-7949(00)00043-2.

[2] Degroot, A. K., Kusters, G. M. A., Monnier, T. Concrete mechanics. Part B: Numerical modelling of bond slip behavior. Washington, D.C. (US): NASA; 1981. Report No.: 82.

[3] Eligehausen, R., Popov, E. P., Bertero, V. V. Local bond stress-slip relationships of deformed bars under generalized excitations: experimental results and analytical model. Berkeley (CA): University of California, Berkeley; 1983. Report No.: UCB/EERC-83/23.

[4] Kwak, H.-G., Filippou, F. Finite element analysis of reinforced concrete structures under monotonic loads. Berkeley (CA): University of California, Berkeley; 1990. Report No.: UCB/SEMM-90/14.

[5] Kwak, H.-G., Filippou, F. Finite element analysis of reinforced concrete structures under monotonic and cyclic loads. In: 10th World Conference in Earthquake Engineering; 1992 Jul 19–24; Madrid, Spain. p. 4227–4232.

[6] Kwak, H.-G., Filippou, F. A New Reinforcing Steel Model With Bond-Slip. Structural Engineering and Mechanics, 1995; 3: 299–312. doi:10.12989/sem.1995.3.4.299.

[7] Monti, G., Filippou, F., Spacone, E. Analysis of hysteretic behavior of anchored reinforcing bars. ACI Structural Journal, 1997; 94: 248–261. doi:10.14359/477

[8] Monti, G., Filippou Filip, C., Spacone, E. Finite Element for Anchored Bars under Cyclic Load Reversals. Journal of Structural Engineering, 1997; 123: 614–623. doi:10.1061/(ASCE)0733-9445(1997)123:5(614).

[9] Monti, G., Spacone, E. Reinforced Concrete Fiber Beam Element with Bond-Slip. Journal of Structural Engineering, 2000; 126: 654–661. doi:10.1061/(ASCE)0733-9445(2000)126:6(654).

[10] Spacone, E., Filippou, F. C., Taucer, F. F. Fibre Beam–Column Model For Non-Linear Analysis of R/C Frames: Part I. Formulation. Earthquake Engineering & Structural Dynamics, 1996; 25: 711–725. doi:10.1002/(SICI)1096-9845(199607)25:7<711::AID-EQE576>3.0.CO;2-9.

[11] Gan, Y. Bond stress and slip modeling in nonlinear finite element analysis of reinforced concrete structures [Master's Thesis]. Toronto (CA): University of Toronto; 2000.

[12] Lowes, L., Moehle, J., Govindjee, S. Concrete-steel bond model for use in finite element modeling of reinforced concrete structures. ACI Structural Journal, 2004; 101: 501–511. doi:10.14359/13336.

[13] Jendele, L., Cervenka, J. Finite element modelling of reinforcement with bond. Computers & Structures, 2006; 84: 1780–1791. doi:10.1016/j.compstruc.2006.04.010.

[14] Kwak, H.-G., Kim, J.-K. Implementation of bond-slip effect in analyses of RC frames under cyclic loads using layered section method. Engineering Structures, 2006; 28: 1715–1727. doi:10.1016/j.engstruct.2006.03.003.

[15] Beconcini, M. L., Croce, P., Formichi, P. Influence of bond-slip on the behaviour of reinforced concrete beam to column joints-Tailor Made Concrete Structures: New Solutions for our Society. In: International FIB Symposium; 2008 May 19–21; Amsterdam, Netherlands. p. 533–539.

[16] Zhou, Y. D., Cheuk, C. Y., Tham, L. G. An embedded bond-slip model for finite element modelling of soil–nail interaction. Computers and Geotechnics, 2009; 36: 1090–1097. doi:10.1016/j.compgeo.2009.03.002.

[17] Kwak, H.-G., Kim, S.-P. Simplified monotonic moment–curvature relation considering fixed-end rotation and axial force effect. Engineering Structures, 2010; 32: 69–79. doi:10.1016/j.engstruct.2009.08.017.

[18] Murcia-Delso, J., Stavridis, A., Shing, B. Modeling the bond-slip behavior of confined large-diameter reinforcing bars. In: 3rd International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering, COMPDYN; 2011 May 25–28; Corfu, Greece. p. 1–14.

[19] Liang, X. An investigation of bond-slip behavior of reinforcing steel subjected to inelastic strains [Master's Thesis]. Ames (IA): Iowa State University; 2013.

[20] Coccia, S., Di Maggio, E., Rinaldi, Z. Bond slip model in cylindrical reinforced concrete elements confined with stirrups. International Journal of Advanced Structural Engineering (IJASE), 2015; 7: 365–375. doi:10.1007/s40091-015-0104-7.

[21] Shafaie, J., Hosseini, A., Marefat, M. 3D finite element modelling of bond-slip between rebar and concrete in pull-out test. In: 3rd International Conference on Concrete & Development, Iran; 2009 April 27; Tehran, Iran. doi:10.13140/2.1.2012.1605.

[22] Alijani, A., Khomami Abadi, M., Razzaghi, J. A new technique in the modal analysis of cracked reinforced concrete (RC) beams through the finite element method. Journal of Structural and Construction Engineering, 2020; 7: 5–30. doi:10.22065/JSCE.2018.127880.1535.

[23] Khademalrasoul, A. Simulation of the effect of internal heterogeneous linear and curved boundaries on crack propagation path in mixed mode using the level set method. Journal of Structural and Construction Engineering, 2017; 4: 42–54. doi:10.22065/jsce.2017.67302.1001.

[24] Hakimi, A., Moslemi, H. Modeling of Analysis Crack using Adaptive Finite Element Method Via Stress Intensity Factor Error Estimation. Journal of Structural and Construction Engineering, 2015; 2: 57–68.

[25] Dehghan Manshadi, S. M., Nategh, S., Khojasteh, A., Rahimian, M. Fracture Modes of an Annular Crack in a Transversely Isotropic Solid. Amirkabir Journal of Civil Engineering, 2018; 50: 855–864. doi:10.22060/ceej.2017.12713.5255.

[26] Miri, M., Rakhshanimehr, M., Miri, F. The effect of Cover, Pressure Strength, and bars corrosion on Cracking of Reinforced Concrete Structures. Amirkabir Journal of Civil Engineering, 2017; 49: 251–262. doi:10.22060/ceej.2016.693.

[27] Hajighasemali, S., Ramezanianpour, A. A., Kashefizadeh, M. Analytical Modeling of Load-Deflection Curve in Reinforced Concrete Beams. Amirkabir Journal of Civil Engineering, 2016; 47: 63–70. doi:10.22060/ceej.2016.614.

[28] Broujerdian, V., Sherafati, A., Karimpour, H. Effect of Crack Cohesive Stresses on the Load-Deformation Response of Reinforced Concrete Beams. Amirkabir Journal of Civil Engineering, 2018; 50: 89–96. doi:10.22060/ceej.2017.11366.5012.

[29] Okamura, H., Maekawa, K. Nonlinear Analysis and Constitutive Models of Reinforced Concrete. 1st ed. Tokyo (JP): Giho-do Press; 1991.

[30] Shima, H., Chou, L., Okamura, H. Micro and Macro Models for Bond Behavior in Reinforced Concrete. Journal of the Faculty of Engineering, Univ. of Tokyo, 1987; 39: 133–194.

[31] Maekawa, K., Okamura, H., Pimanmas, A. Non-Linear Mechanics of Reinforced Concrete. 1st ed. London (UK): CRC Press; 2003. doi:10.1201/9781482288087.

[32] Soltani, M., Maekawa, K. Path-dependent mechanical model for deformed reinforcing bars at RC interface under coupled cyclic shear and pullout tension. Engineering Structures, 2008; 30: 1079–1091. doi:10.1016/j.engstruct.2007.06.013.

[33] Birely, A. C., Lowes, L. N., Lehman, D. E. A model for the practical nonlinear analysis of reinforced-concrete frames including joint flexibility. Engineering Structures, 2012; 34: 455–465. doi:10.1016/j.engstruct.2011.09.003.

[34] Vandecruys, E., Hendriks, M. A. N., van de Velde, M., Lombaert, G., Verstrynge, E. Dynamic stiffness parameter assessment of cracked reinforced concrete beams: A numerical and experimental study. Engineering Structures, 2024; 318: 118758. doi:10.1016/j.engstruct.2024.118758.

[35] Kordtabar, B., Dehestani, M., Yazdanpanah, O., Hasanzadeh, A. Time-dependent seismic fragility and serviceability assessment of reinforced concrete frames under chloride-induced corrosion. Scientific Reports, 2025; 15: 40575. doi:10.1038/s41598-025-24197-z.

[36] Zhang, H. Y., Li, Q. Y., Kodur, V., Lv, H. R. Effect of cracking and residual deformation on behavior of concrete beams with different scales under fire exposure. Engineering Structures, 2021; 245: 112886. doi:10.1016/j.engstruct.2021.112886.

[37] Dancygier, A. N., Karinski, Y. S. Effect of cracking localization on the structural ductility of normal strength and high strength reinforced concrete beams with steel fibers. International Journal of Protective Structures, 2019; 10: 457–469. doi:10.1177/2041419618824609.

[38] Namdar, A., Darvishi, E., Feng, X., Zakaria, I., Mat Yahaya, F. Effect of flexural crack on plain concrete beam failure mechanism: A numerical simulation. Fracture and Structural Integrity, 2016; 10: 168–181. doi:10.3221/IGF-ESIS.36.17.

[39] Karinski, Y. S., Dancygier, A. N., Gebreyesus, Y. Y. Probabilistic model for cracking localization in reinforced fibrous concrete beams. Journal of Engineering Mathematics, 2024; 144: 23. doi:10.1007/s10665-023-10330-2.

[40] Fu, C., Zhu, Y., Tong, D. Stiffness Assessment of Cracked Reinforced Concrete Beams Based on a Fictitious Crack Model. KSCE Journal of Civil Engineering, 2021; 25: 516–528. doi:10.1007/s12205-020-2056-0.

[41] Forouzan, S. Effect of bond-slip on the ductility of reinforced concrete beams [Thesis]. Babol (IR): Babol Noshirvani University of Technology; 2016.

[42] Ayoub, A., Filippou Filip, C. Mixed Formulation of Bond-Slip Problems under Cyclic Loads. Journal of Structural Engineering, 1999; 125: 661–671. doi:10.1061/(ASCE)0733-9445(1999)125:6(661).

[43] Gergely, P., Lutz, L. A. Maximum Crack Width in Reinforced Concrete Flexural Members. 318Reference, 2005; 10: 87–117. doi:10.14359/17348.

[44] Bayrak, O., Sheikh, S. High-strength concrete columns under simulated earthquake loading. 1997; 94: 708–722.

[45] PEER ground motion database. Pacific Earthquake Engineering Research Center. 2010. Available online: https://ngawest2.berkeley.edu/. (Accessed December 2010).

[46] Shayanfar, J., Akbarzadeh Bengar, H. Nonlinear analysis of RC frames considering shear behaviour of members under varying axial load. Bulletin of Earthquake Engineering, 2017; 15: 2055–2078. doi:10.1007/s10518-016-0060-z.

[47] Sadeghpour Haji, M., Niknam, R., Shayanfar, J. ANN-Based Modeling of Shear Behavior of Reinforced Concrete Columns under Constant Axial Loads. Civil Engineering and Applied Solutions, 2026; 2: 28–48. doi:10.22080/ceas.2025.30228.1050.

[48] Shayanfar, S., Ghasemi Naghibdehi, M., Shayanfar, J. Data-Driven Prediction of FRP Confinement Efficiency in Concrete Columns: Incorporating Corner Radius and Aspect Ratio Effects. Civil Engineering and Applied Solutions, 2026; 2: 1–15. doi:10.22080/ceas.2026.30750.1059.

[49] Shayanfar, S., Ghasemi Naghibdehi, M., Shayanfar, J. Data-Driven Predictive Formulation for FRP-Confined Circular Concrete Columns. Civil Engineering and Applied Solutions, 2027; 3: 25–36. doi:10.22080/ceas.2026.30751.1060.

[50] Javad, S., Meysam, O., Mahdi, N. Analysis-oriented model for seismic assessment of RC jacket retrofitted columns. Steel and Composite Structures, An International Journal, 2020; 37: 371–390. doi:10.12989/scs.2020.37.3.371.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Effects of Self-Selected Music Tempo on Young Drivers' Speed and Acceleration

Aref Ansari [a], Ehsan Ramezani-Khansari [a*], Ali Abdi Kordani [a], Hamid Mirzahossein [a]

[a] *Civil Engineering Department, Imam Khomeini International University, Qazvin, Iran*

ARTICLE INFO

Article history:

Received 27 May 2026

Revised 04 July 2026

Accepted 04 August 2026

Available online 01 April 2027

Keywords:

Music

Young

Speed

Driving simulator

ABSTRACT

This study examined the effects of participant-selected music on driving behavior using a driving simulator with 30 young male participants aged 20 to 25. Each participant drove under three conditions: no music, high-tempo music, and low-tempo music. Speed and acceleration data were recorded, and aggressive driving behavior was assessed via a questionnaire. Results showed that high-tempo music significantly increased average driving speed compared to no music, while low-tempo music had no effect. For speeds above 40 km/h (critical for pedestrian safety), both high- and low-tempo music influenced speed, with high-tempo increasing and low-tempo decreasing it. However, music type did not significantly affect maximum positive acceleration. Aggressive driving behavior, measured by questionnaire scores, showed no correlation with the impact of music on speed. The study also noted limitations in the simulator's ability to replicate braking sensations, leading to the exclusion of negative acceleration data. Statistical analyses (ANOVA and Tukey tests) confirmed significant differences in speed between high-tempo music and other conditions but no notable effects on acceleration or aggressive behavior. The findings suggest that music tempo can alter driving speed, particularly at higher velocities, but does not uniformly impact other driving metrics or aggression levels.

How to cite this article: Ansari, A., Ramezani-Khansari, E., Abdi Kordani, A., Mirzahossein, H. Effects of Self-Selected Music Tempo on Young Drivers' Speed and Acceleration. Civil Engineering and Applied Solutions. 2027; 3(2): 122–130. doi:10.22080/ceas.2026.31922.1104.

1. Introduction

Studies have shown that listening to music while driving, as a secondary activity, may cause driver distraction. This has led to research aimed at identifying the extent to which different music genres influence driving behavior [1]. Multiple studies have indicated that listening to fast-paced, high-energy music results in faster driving and increased arousal among drivers while on the road [2]. It has been observed that loud music increases driving speed and the likelihood of running red lights during driving simulations [3]. Studies showed that driving behavior is influenced by the driver's perception of the road, surrounding environment, and hazardous situations [4]. Situations are considered potentially hazardous when a driver must alter speed or direction. Safer drivers tend to anticipate such situations in advance, applying braking or steering adjustments early to avoid collisions. Mental emotions are directly linked to physical responses; both negative and positive emotions exert distinct effects on the body's physical reactions [5]. Although music and podcasts are auditory stimuli, they create mental imagery in the listener's mind [6].

Babić et al. [7] showed that young male drivers, compared to other groups, pay more attention to road signs while listening to music. In contrast, Ünal et al. [8] found that listening to music reduces drivers' concentration during driving. In this study, drivers who did not listen to music demonstrated greater focus on the road. Widén and Erlandsson [9] examined the impact of music on youth risk-taking behavior, involving 9 women and 8 men. They found that music plays a significant role in shaping personal identity. Listening to loud music may lead to norm-breaking behaviors.

In a study involving 47 participants and two driving conditions − loud music and no music − drivers' heart rates were also measured. The results indicated that music enhanced drivers' focus when following the lead vehicle. The findings suggest that music

* Corresponding author.
E-mail addresses: e.r.khansari@eng.ikiu.ac.ir (E. Ramezani-Khansari).

https://doi.org/10.22080/ceas.2026.31922.1104
ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

does not disrupt steady following behavior; in fact, it may improve driver attention [10]. A study conducted by Farrell [11] on 20 college students found that loud music can impair driver reaction times. In another study, music was found to distract drivers and reduce the following distance between vehicles. Two music rhythms − soft and loud − were played for drivers, and results showed that loud music caused driver distraction and impaired control when approaching the vehicle ahead [12].

Wang et al. [13] demonstrated that listening to music while driving led to cognitive distraction and mental overload. Increased cognitive workload caused delays in response time to changes in traffic flow or unexpected braking. Music contributed to more frequent lane departures and rear-end collisions. Distraction is a significant factor in traffic accidents. Wang et al. examined two sources of distraction: listening to music and talking on the phone. A simulated brake light was used to measure reaction time in both scenarios. Drivers who were on the phone exhibited slower response times compared to those listening to music. Thus, drivers reacted faster while listening to music than while engaged in phone conversations.

Kasprzak [14] investigated how listening to music affects driver behavior. Participants were instructed to follow a vehicle, mimicking real-life driving behavior. The speed of the leading vehicle varied throughout the simulation, and drivers were expected to adjust their speed accordingly. Results showed that drivers listening to music exhibited better responses to speed increases or decreases of the lead vehicle. Another experiment by Niu et al. [15] showed that a mentally preoccupied driver performs worse in navigating the route. However, drivers who listened to music showed better vehicle-following behavior and steering control, though they also tended to be involved in more vehicle collisions.

Babić et al. [7] aimed to examine the impact of various selected music genres on the driving behavior and visual scanning patterns of young individuals in urban environments. The genres chosen by the drivers included classical, metal, and balkan folk music. Listening to each genre of music affected the frequency of drivers' glances at roadside signs.

A study compared driving performance while listening to music selected by the driver versus music selected by the researcher. The findings showed that young drivers listening to their preferred music in urban conditions drove at higher speeds than those who did not. Those listening to calm music exhibited more relaxed driving behavior [14].

Navarro et al. [16] utilized amateur drivers to investigate the effects of self-selected music on their driving behavior. When young and inexperienced drivers began driving while listening to their favorite music, the results indicated that such music generated positive energy, leading to improved driver responsiveness. In contrast, amateur drivers who drove without music experienced greater fear and anxiety while continuing along the route.

In a study, two types of music, selected by the experimenter and selected by the participant, were played for participants. Researchers used a driving simulator with two types of roads: one with simple conditions and low traffic volume, and the other an urban road with more challenging conditions and higher traffic. Results showed that music selected by the drivers had no significant impact on driving performance, while self-selected music improved overall driver performance. Listening to preferred music led to better lane maintenance and faster driving within legal speed limits. Drivers who listened to their own music also performed better when braking for pedestrians and stopping at red lights compared to those who listened to experimenter-selected music [17].

Ghojazadeh et al. [18] conducted research in which participants aged between 18 and 60 drove using a driving simulator. The first experimental condition involved listening to music (of any genre and format) chosen by the driver, while the second condition involved driving without music. The variables examined in this study included physiological indicators such as heart rate, breathing rate, and arousal level, as well as driver performance metrics like reaction time and delay ratio. The post-experiment results showed that listening to music had a positive effect on a driver’s ability to adjust the vehicle's position. However, loud music increased heart rate, which consequently led to more driver errors. Conversely, listening to calm, self-selected music resulted in smoother breathing and a more stable heart rate.

Music is one of the most popular activities while driving. Studies have shown mixed findings; some researchers concluded that music distracts drivers, while others found that preferred music improves driver comfort and performance in response to road obstacles. Whether music is selected by the experimenter or the participant can also influence driving behavior [19].

Most studies have focused on the general effects of music on driving behavior, while fewer have examined the differences between music chosen by the experimenter and that selected by the participant. In this research, only participant-selected music is used, and its differential impact on driving behavior is studied. The principal novelty of the present study, relative to prior research, can be classified into two primary categories. First, in previous investigations, the musical stimuli were selected by the experimenter; in contrast, the present study required participants to self-select their preferred music. This approach enhances the ecological validity of the findings, rendering them more representative of real-world driving conditions. Second, beyond the variables examined in earlier work − such as travel speed − the current study additionally incorporates an analysis of other driving maneuvers, namely lane-changing and car-following behavior. These maneuvers are inherently more complex and facilitate a more thorough and precise assessment of music's influence on driving performance.

The following section addresses data collection. Section three covers the data analysis methodology. Section four presents the discussion and conclusions. Finally, the study concludes with a summary of the findings.

2. Methodology

In this study, 30 young male students (between 20 and 25 years old) with more than 3 years of driving experience participated in a driving experiment using a K.N. Toosi University of Technology driving simulator, which has been used in many studies [20-

22] (Fig. 1). A fixed-base driving simulator was used for the experiments. The simulator can record measures accounting for vehicle locations, vehicle movement (e.g., speed, acceleration/deceleration), and driver operations (e.g., steering wheel angle). Given the inherent divergence between actual driving behavior and simulated settings, the validation of driving simulators is indispensable. The driving simulator, whose validity has been demonstrated in earlier studies and which has seen extensive prior use, was utilized in this investigation [23-25].

Fig. 1. Driving simulator.

Each participant drove in three simulated scenarios: without music, with high-tempo music, and with low-tempo music. In accordance with established protocols in driver behavior research, music with a tempo exceeding 120 BPM was classified as fast or high tempo, while music below 80 BPM was designated as slow or low tempo, with moderate tempo serving as the intermediate range. Music (genre and lyrics) was selected by participants. Each scenario lasted ten minutes, in addition to a warm-up period. The environment (urban) and traffic conditions (medium density) were identical for all scenarios. The sequence of driving scenarios was randomized. Specifically, following a warm-up drive and no-music driving, participants were randomly assigned to begin with either the high-tempo or the low-tempo music condition; the remaining music condition was then administered as the third scenario. This procedure was adopted to prevent potential confounding effects of fatigue or simulator adaptation on the experimental results.

The route was generally straight, with no appreciable horizontal curvature. Along the corridor, the level of service of the traffic was such that drivers experienced no congestion-induced speed reduction, thereby enabling them to maintain their desired operating speeds. Several intersections were located along the route to facilitate the collection of vehicular acceleration and deceleration data. The length of the route was determined as a function of the drivers' average travel speed. The posted speed was 90 km/hr.

Speed and acceleration data were collected for each driver across the scenarios. Due to the small number of participants in other age and gender groups, their data were excluded from analysis. This exclusion helped eliminate the effects of other variables, allowing the study to focus solely on the influence of music.

In addition to the driving data, participants also completed an Aggressive Driving Behavior Questionnaire (ADBQ) to measure aggressive behavior. The ADBQ is a 20-item self-report instrument to assess drivers' propensity for engaging in aggressive driving behaviors. The questionnaire was developed using a factor-analytic approach that combined five previously established aggressive driving scales [26]. Participants rate each item on a 6-point Likert-type scale ranging from 1 (Never) to 6 (Nearly All the Time). The instrument has demonstrated robust psychometric properties in previous studies, with high internal consistency (Cronbach's $\alpha = 0.77$ to 0.86) and established predictive validity in both simulated and real-world driving contexts [26]. For the present sample, the ADBQ yielded a Cronbach's alpha of 0.85, indicating acceptable internal consistency.

After collecting the data from participants, the dataset was divided into two categories: speed data and acceleration data. These data were analyzed from several perspectives. Initially, the average driving speed of participants in different scenarios was examined, followed by the analysis of acceleration. Additionally, the relationship between participants' aggressive behavior and the effect of music type on average speed was evaluated. Descriptive and inferential statistics were used for this purpose. Given that there were three different data groups or scenarios, two statistical tests (ANOVA and Tukey) were applied. ANOVA served as a preliminary test to explore whether significant differences existed among the scenario data. Subsequently, the Tukey test was employed to determine the nature and direction of those differences across scenarios. Since the data were analyzed in an aggregated (cumulative) manner, analysis of variance (ANOVA) was employed.

3. Results and discussion

In this section, the speed data, acceleration data, and the results of the aggressive behavior questionnaire are analyzed with respect to the type of music played during driving.

3.1. Speed

Fig. 2 presents a sample speed-time graph for a participant under different driving scenarios. In this graph, speed is represented in three colors: yellow indicates driving with low-tempo music, green indicates driving without music, and red indicates driving with high-tempo music. The horizontal axis shows the 600-second driving period for each participant, while the vertical axis indicates the driver's speed in kilometers per hour. Speed data were recorded at intervals of 0.1 seconds.

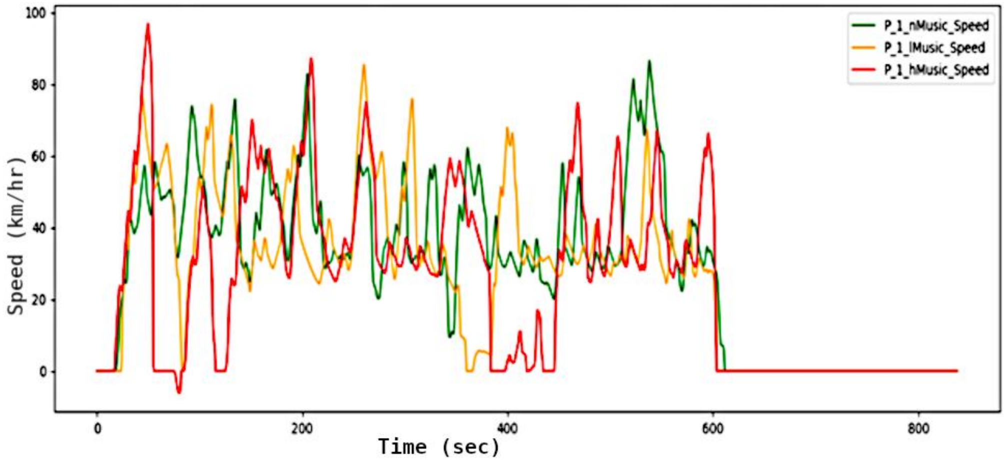


Fig. 2. Speed-time chart of a sample driver in three scenarios.

Fig. 3 illustrates the average speed of each driver across the different driving scenarios.

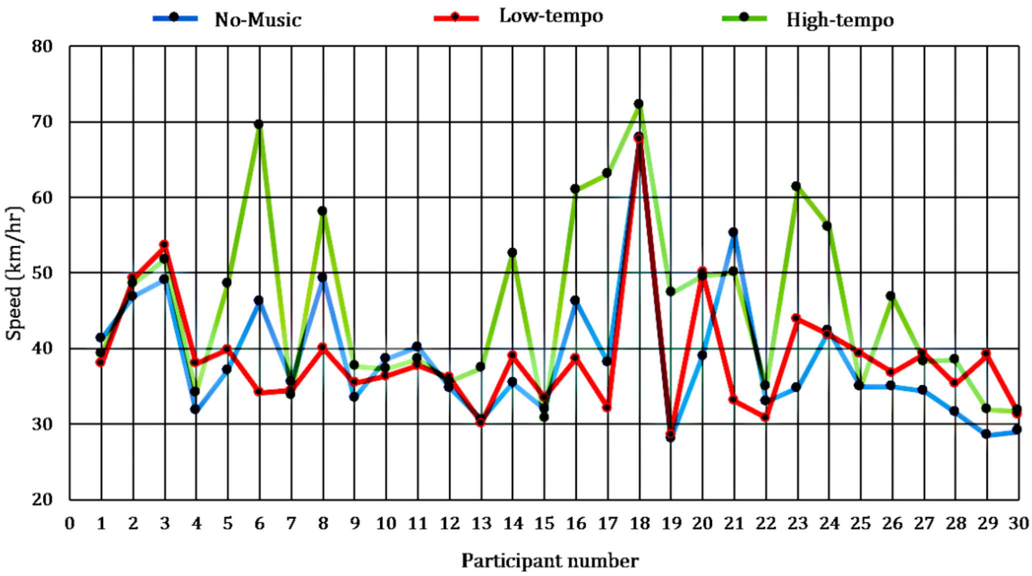


Fig. 3. Average driving speed of drivers.

Table 1 presents the results of the ANOVA test conducted on all average speed data.

Table 1. ANOVA test of average speeds in different scenarios.

Statistics	Significant level of variance analysis test (one-way ANOVA)
5.23	0.007

The ANOVA p-value was less than 0.05, indicating a statistically significant difference among the three groups of data. It is worth noting that, for deeper analysis, ANOVA was also performed individually for each participant's data. Approximately 90% of cases revealed significant differences among their scenario groups, confirming the overall ANOVA findings. To rank the average speeds across scenarios, the Tukey post hoc test was used, and its results are presented in Table 2. Pairwise comparisons using Tukey's test showed that the average speed under the high-tempo music condition differed significantly from both the no music and low-tempo music conditions. However, there was no significant difference between the no music and low-tempo music conditions. Similar results have been reported in previous studies. Brodsky demonstrated that drivers exposed to fast-paced music not only drove faster but also committed more traffic violations [27, 28].

Table 2. Comparing the average speed in different scenarios.

Tukey's HSD pairwise group comparisons (95.0% confidence interval)					
Comparison	P-value	Comparison	P-value	Comparison	P-value
(No music - low tempo music)	0.999	(Low tempo music - high tempo music)	0.018	(No music - high tempo music)	0.016

The relationship between speed and crash risk is well-established as exponential: even small increases in mean speed lead to disproportionately large rises in severe and fatal crash involvement [29, 30].

3.2. Speeds above 40 km/h

Given the role of speed in the severity of accidents, the effect of music type on speeds above 40 km/h was examined. According to a study [31], in vehicle-pedestrian collisions, the risk of fatality is less than 10% when the vehicle speed is below 40 km/h. On the other hand, given that the traffic volume was low, vehicles are not expected to have selected speeds below 40 km/h. Speeds lower than 40 km/h were predominantly associated with the periods immediately before and after a stop. Therefore, one-way ANOVA analysis was conducted again using only speed data above 40 km/h. Speeds below 40 km/h were excluded, and averages were calculated based on the remaining data. The null hypothesis was that the average speeds across the three conditions (no music, low-tempo music, and high-tempo music) were equal. Based on the findings, the null hypothesis was rejected in 97% of cases, clearly indicating that music playback significantly influenced driver behavior. In other words, high-tempo music increased average speed, while low-tempo music reduced it, compared to the no music condition (Table 3).

Table 3. Comparing the average speed in different scenarios.

Tukey's HSD pairwise group comparisons (95.0% confidence interval)					
Comparison	**P-value**	**Comparison**	**P-value**	**Comparison**	**P-value**
(No music - low tempo music)	0.046	(Low tempo music - high tempo music)	0.038	(No music - high tempo music)	0.009

3.3. Acceleration

Acceleration reflects the rate of change in speed over time. Therefore, the selected time interval directly influences the calculated acceleration value. If the time interval is too long, the speed changes within that window may be smoothed out, with only the initial and final speeds considered, which potentially causes error in the analysis. On the other hand, selecting intervals that are too short may impose unnecessary computational load on the model. Thus, acceleration was calculated using various time intervals and the results analyzed. It was observed that when the time interval was less than one second, despite a significant increase in computational complexity, no meaningful additional detail was captured. Hence, it was deemed reasonable not to use intervals shorter than one second. Conversely, when the interval exceeded three seconds, important fluctuations in acceleration were smoothed out and excluded from the analysis. Therefore, a one-second interval was identified as the optimal time frame for calculating acceleration.

Maximum and minimum accelerations were calculated. Negative acceleration indicates deceleration or braking. It was observed that the values of negative acceleration differed significantly from those of positive acceleration. This discrepancy may be attributed to the driving simulator's limited ability to convey deceleration, particularly in modeling braking force. This is a limitation of the simulator, which cannot effectively replicate the physical sensation of braking. Therefore, negative acceleration data were excluded from further analysis. Some studies demonstrated that drivers brake more abruptly and more strongly than in real driving because finding safety margins is more difficult in a driving simulator than in the real world [28, 32]. Fig. 4 displays the maximum positive acceleration for each driver. It should be noted that acceleration is typically zero and only occurs when the driver actively changes speed; hence, averaging acceleration values would be meaningless and would result in values close to zero. To determine maximum acceleration, the five highest acceleration values for each participant were averaged. Fig. 4 displays the maximum positive acceleration for each driver in different scenarios.

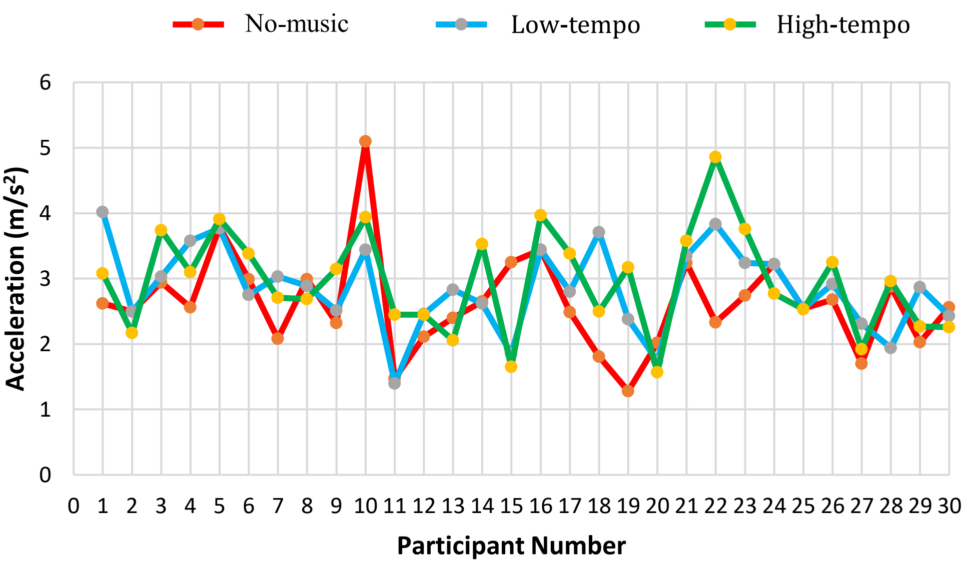


Fig. 4. Maximum acceleration during driving.

Fig. 5 and Tables 4 and 5 show the results of ANOVA and Tukey tests on the maximum positive acceleration data. The significance level criterion value was greater than 0.05, and the null hypothesis (that the means of the three groups, namely maximum acceleration under no music, low-tempo music, and high-tempo music conditions, are equal) was not rejected. Therefore, playing music does not have a significant effect on the drivers' maximum positive acceleration. The Tukey pairwise comparison test also indicates that the null hypothesis was not rejected in any of the pairwise comparisons, meaning there were no significant differences between the group means in any two-group comparisons. In other words, music did not affect the drivers' maximum positive acceleration. Niu et al. [33] reported that high-tempo music significantly increased the standard deviation of longitudinal acceleration, reflecting more volatile and aggressive control of the vehicle.

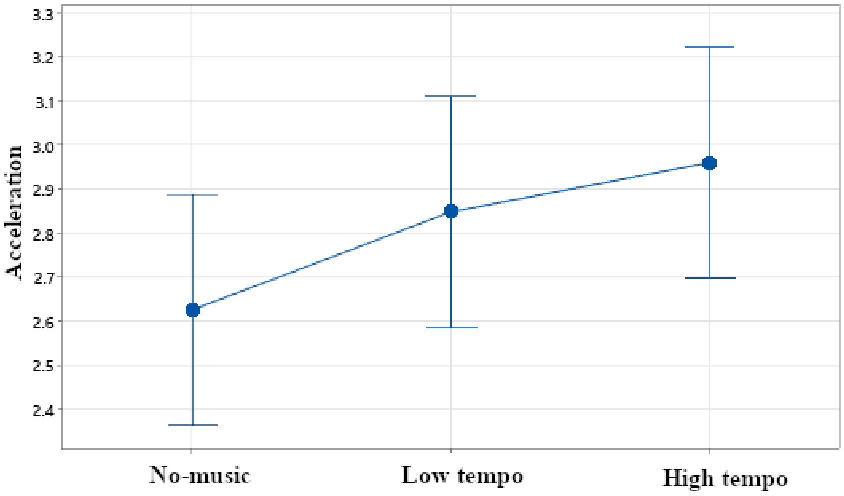


Fig. 5. Interval plot of accelerated maximization data for different scenarios.

Table 4. ANOVA of accelerated maximization data for different scenarios.

Source	DF	Adj SS	Adj MS	F-value	P-value
Factor	2	1.738	0.8688	1.67	0.195
Error	87	45.321	0.5209		
Total	89	47.059			

Table 5. Tukey simultaneous tests for differences of means.

Difference of levels	Difference of means	SE of difference	95% CI	T-value	Adjusted p-value
No music - low tempo music	0.222	0.186	(−0.222, 0.666)	1.19	0.461
No music - high tempo music	0.334	0.186	(−0.110, 0.778)	1.79	0.178
Low tempo music - high tempo music	0.112	0.186	(−0.332, 0.556)	0.60	0.820

3.4. Aggressive behavior and speed variations in scenarios

In the data collection section, it was stated that the participants completed the aggressive behavior assessment questionnaire before starting the test. Participants expressed their opinions on a scale of very low, low, moderate, high, and very high, where the "very high" option indicated the highest level of aggressive behavior and the "very low" option indicated the lowest level of aggressive behavior in response to the given question. The qualitative responses of the participants were converted into numerical scores ranging from 1 (very low) to 5 (very high). As the questionnaire score increased, the participant’s behavior was estimated to be more aggressive. The maximum and minimum scores of the aggressive behavior assessment questionnaire were 115 and 23, respectively. Fig. 6 displays the distribution of participants' aggressive driving behavior scores.

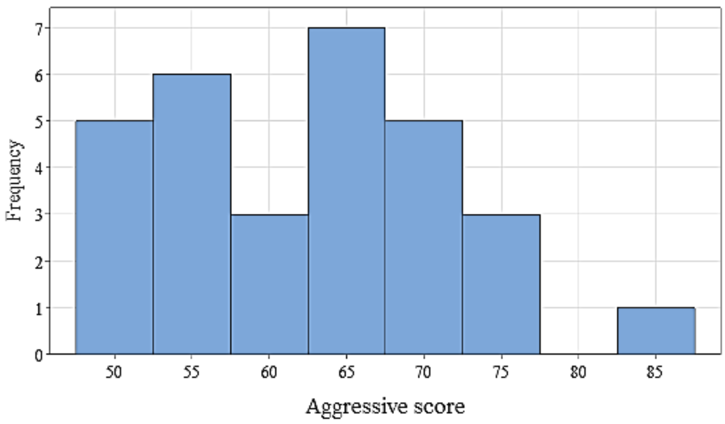


Fig. 6. Participants' aggressive driving behavior scores.

The relationship between drivers' aggressive behavior and the effect of low-tempo and high-tempo music was examined using the Pearson correlation test. The effect of low-tempo music was defined as the difference in mean speed between the no music condition and the low-tempo music condition, while the effect of high-tempo music was defined as the difference in mean speed between the no music condition and the fast music condition. It was observed that there was no significant relationship between aggressive behavior and the effect of music type on the participants’ mean speed (Table 6). Similar results to those of Brodsky’s study have been obtained [3, 27]. The aggressive driving questionnaire is designed to assess stable aggressive driving traits, whereas music can temporarily alter arousal and mood. This distinction between trait and state responses should be studied in future studies.

Table 6. Pairwise Pearson Correlations between the difference in average speed of different scenarios and the aggressive scale.

Sample 1	Sample 2	Correlation	95% CI for ρ	P-value
Difference between no music and low tempo	Aggressiveness	−0.360	(−0.638, 0.000)	0.051
Difference between high tempo and no music	Aggressiveness	0.321	(−0.045, 0.610)	0.084
Difference between high tempo and no music	Difference between no music and low tempo	−0.026	(−0.382, 0.338)	0.893

3.5. Limitations and future work

It is worth noting that the majority of university students are young adults; consequently, the study sample consisted primarily of young university students. Furthermore, when drivers are required to operate a driving simulator while simultaneously adhering to specific conditions (e.g., listening to music), recruiting a large number of participants and volunteers becomes particularly challenging. In this study, a minimum of 30 participants were recruited to enable the use of parametric tests and to enhance the statistical significance of the results. This number of participants is consistent with prior studies in which 30 or fewer participants were recruited for driving tests [34-37]. This demographic homogeneity and number of participants represent a limitation of the current research, which could be addressed in future studies.

As previously mentioned, driving simulators inherently differ from real-world driving conditions in unavoidable respects. Consequently, the current investigation could be further validated or replicated through on-road driving tests. One of the key limitations of driving simulators is the overestimation of deceleration, which can be corrected or calibrated through real-world experiments in future work. This research has not considered driver-level variables such as annual mileage, prior simulator experience, fatigue, mood, and risk-taking tendency, which may influence driving behavior and could influence the experimental outcomes.

Although tempo and loudness were controlled in this study, given that the aim was to examine the effect of drivers' preferred music, factors such as emotional intensity, arousal level, genre, and lyrics were not controlled, which could be investigated in future research.

This paper has focused on the average speed, speed above 40 km/h, and maximum positive acceleration. Additional measures such as speed variability, lane keeping, headway, time-to-collision, braking frequency, and lateral deviation can be studied in future work.

4. Conclusions

Music can influence human moods and emotions. While driving, listening to different types of music may have varying effects on the driver and influence their driving behavior. This study investigated the effect of music type on drivers' speed and acceleration. Participants took part in three driving scenarios: without music, with high-tempo music, and with low-tempo music in a driving simulator, and their driving data were collected. The music was selected by the drivers themselves to ensure familiarity and better simulation of driving conditions. Furthermore, to study the relationship between music and aggressive driving behavior, participants completed an aggressive behavior assessment questionnaire.

The main findings of this study are summarized as follows:

- It was observed that fast-tempo music increased the average driving speed compared to the no music condition, whereas slow-tempo music had no effect on the participants' average speed.
- Considering that speeds over 40 km/h can cause harm to pedestrians, average speeds above 40 km/h were also examined. In this context, it was found that both slow-tempo and fast-tempo music significantly influenced participants' average speeds, causing a decrease and increase in speed, respectively.
- No significant relationship was found between drivers' maximum positive acceleration and the type of music. Moreover, no significant relationship was observed between drivers' aggressive behavior and music type. In other words, increased aggressive behavior did not amplify the effect of music.

Statements & declarations

Author contributions

Aref Ansari: Data curation, Investigation.

Ehsan Ramezani-Khansari: Supervision, Writing - Original Draft, Visualization, Software.

Ali Abdi Kordani: Writing - Review & Editing, Formal analysis.

Hamid Mirzahossein: Validation, Methodology.

Funding

This work was not supported by a funding agency.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

The authors declare that they have no competing interests.

References

[1] Pêcher, C., Lemercier, C., Cellier, J.-M. Emotions drive attention: Effects on driver's behaviour. Safety Science, 2009; 47: 1254–1259. doi:10.1016/j.ssci.2009.03.011.

[2] Dalton, B. H. The effects of sound types and volumes on simulated driving performance, simple vigilance and heart rate [Master's Thesis]. St. John's (CA): Memorial University of Newfoundland; 2006.

[3] Brodsky, W. The effects of music tempo on simulated driving performance and vehicular control. Transportation Research Part F: Traffic Psychology and Behaviour, 2001; 4: 219–241. doi:10.1016/S1369-8478(01)00025-0.

[4] Bucchi, A., Sangiorgi, C., Vignali, V. Traffic Psychology and Driver Behavior. Procedia - Social and Behavioral Sciences, 2012; 53: 972–979. doi:10.1016/j.sbspro.2012.09.946.

[5] Cook, T., Roy, A., Welker, K. Music as an emotion regulation strategy: An examination of genres of music and their roles in emotion regulation. Psychology of Music, 2017; 47: 144–154. doi:10.1177/0305735617734627.

[6] Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., Lonsdale, C. Video Improves Learning in Higher Education: A Systematic Review. Review of Educational Research, 2021; 91: 204–236. doi:10.3102/0034654321990713.

[7] Babić, D., Babić, D., Sucha, M., Stanić, V., Toman, M. The influence of music genres on the driving behaviour of young drivers and their visual scanning of the environment. Transportation Research Part F: Traffic Psychology and Behaviour, 2021; 81: 396–407. doi:10.1016/j.trf.2021.07.001.

[8] Ünal, A. B., Steg, L., Epstude, K. The influence of music on mental effort and driving performance. Accident Analysis & Prevention, 2012; 48: 271–278. doi:10.1016/j.aap.2012.01.022.

[9] Widén, S., Erlandsson, S. Risk perception in musical settings—a qualitative study. International Journal of Qualitative Studies on Health and Well-being, 2007; 2: 33–44. doi:10.1080/17482620601121169.

[10] Ünal, A. B., de Waard, D., Epstude, K., Steg, L. Driving with music: Effects on arousal and performance. Transportation Research Part F: Traffic Psychology and Behaviour, 2013; 21: 52–65. doi:10.1016/j.trf.2013.09.004.

[11] Farrell, J. The Effect of Increasing Music Volume on Reaction Time. The Journal of Science and Medicine, 2021; 3: 1–5. doi:10.37714/josam.v3i0.62.

[12] Rijwan, M. Effect of Music Listening on Brake Reaction Time (A Case Study of Koteshwor - Sallaghari Road Section). Patan (NP): Institute of Engineering, Central Pulchowk Campus; 2016.

[13] Wang, D., Jimison, Z., Richard, D., Chuan, C.-H. Effect of Listening to Music as a Function of Driving Complexity: A Simulator Study on the Differing Effects of Music on Different Driving Tasks. In: Proceedings of the international driving symposium on human factors in driver assessment, training and vehicle design; 2015; Salt Lake City, Utah. p. 253–259. doi:10.17077/drivingassessment.1580.

[14] Kasprzak, E. Does listening to favourite music influence emotional experiences and reaction time of amateur drivers? Polskie Forum Psychologiczne, 2023; 27: 399–418. doi:10.34767/PFP.2022.04.02.

[15] Niu, J., Ma, C., Liu, J., Li, L., Hu, T., Ran, L. Is music a mediator impacting car following when driver's personalities are considered? Accident Analysis & Prevention, 2020; 147: 105774. doi:10.1016/j.aap.2020.105774.

[16] Navarro, J., Osiurak, F., Gaujoux, V., Ouimet, M. C., Reynaud, E. Driving Under the Influence: How Music Listening Affects Driving Behaviors. Journal of Visualized Experiments, 2019; 145: e58342. doi:10.3791/58342.

[17] Jimison, Z. N. The Effect of Music Familiarity on Driving: A Simulated Study of the Impact of Music Familiarity Under Different Driving Conditions [Master's Thesis]. Jacksonville (FL): University of North Florida; 2014.

[18] Ghojazadeh, M., Farhoudi, M., Rezaei, M., Rahnemayan, S., Narimani, M., Sadeghi-Bazargani, H. Effect of music on driving performance and physiological and psychological indicators: A systematic review and meta-analysis study. Health Promot Perspect, 2023; 13: 267–279. doi:10.34172/hpp.2023.32.

[19] Karageorghis, C. I., Mouchlianitis, E., Payre, W., Kuan, G., Howard, L. W., Reed, N., Parkes, A. M. Psychological, psychophysiological and behavioural effects of participant-selected vs. researcher-selected music in simulated urban driving. Applied ergonomics, 2021; 96: 103436. doi:10.1016/j.apergo.2021.103436.

[20] Javanbakht, N., Mirbaha, B. From Traits to Speed Control: Engineering Insights from a Driving-Simulator Hazard Scenario. Contributions of Science and Technology for Engineering, 2026; 3: 10–19. doi:10.22080/cste.2025.29839.1075.

[21] Javanbakht, N., Mirbaha, B. A model for evaluating drivers' mean speed: using psychological and driving simulator data. Numerical Methods in Civil Engineering, 2023; 8: 48–55. doi:10.61186/NMCE.2303.1027.

[22] Javanbakht, N., Mirbaha, B. Evaluating Drivers' Response to Road Hazard: A Simulation Study. Advances in Civil Engineering, 2024; 2024: 6788857. doi:10.1155/2024/6788857.

[23] Ramezani-Khansari, E., Tabibi, M., Moghadas Nejad, F., Mesbah, M. Comparing the Effect of Age, Gender, and Desired Speed on Car-Following Behavior by Using Driving Simulator. Journal of Advanced Transportation, 2021; 2021: 9922321. doi:10.1155/2021/9922321.

[24] Ramezani-Khansari, E., Tabibi, M., Moghadas Nejad, F. Estimating Lane Change Duration for Overtaking in Nonlane-Based Driving Behavior by Local Linear Model Trees (LOLIMOT). Mathematical Problems in Engineering, 2021; 2021: 4388776. doi:10.1155/2021/4388776.

[25] Ramezani-Khansari, E., Tabibi, M., Moghadas Nejad, F. Validating Driving Simulator for Car-Following Distance. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2021; 45: 281–290. doi:10.1007/s40996-020-00576-6.

[26] Velazquez, E. M., Mouloua, M. Understanding Aggressive Driving: The Role of Personality and Individual Differences. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2021; 65: 972–976. doi:10.1177/1071181321651345.

[27] Brodsky, W., Olivieri, D., Chekaluk, E. Music genre induced driver aggression: A case of media delinquency and risk-promoting popular culture. Music & Science, 2018; 1: 2059204317743118. doi:10.1177/2059204317743118.

[28] Hoffman, J. D., Lee, J. D., Brown, T. L., McGehee, D. V. Comparison of Driver Braking Responses in a High-Fidelity Simulator and on a Test Track. Transportation Research Record, 2002; 1803: 59–65. doi:10.3141/1803-09.

[29] Elvik, R. A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims. Accident Analysis & Prevention, 2013; 50: 854–860. doi:10.1016/j.aap.2012.07.012.

[30] Aarts, L., van Schagen, I. Driving speed and the risk of road crashes: A review. Accident Analysis & Prevention, 2006; 38: 215–224. doi:10.1016/j.aap.2005.07.004.

[31] Rosén, E., Sander, U. Pedestrian fatality risk as a function of car impact speed. Accident Analysis & Prevention, 2009; 41: 536–542. doi:10.1016/j.aap.2009.02.002.

[32] Boer, E. R., Kuge, N., Yamamura, T. Affording realistic stopping behavior: A cardinal challenge for driving simulators. In: 1st Human-Centered Transportation Simulation Conference; 2001 Nov 04–07; Iowa City, Iowa.

[33] Niu, J., Ma, C., Liu, J., Li, L., Hu, T., Ran, L. "Is music a mediator impacting car following when driver's personalities are considered," Accid. Anal. Prev., vol. 147, p. 105774, 2020.

[34] Wang, X., Liu, S., Cai, B., Hurwitz, D., Guo, Q., Wang, X. Sample Size Study of Driving Simulator Experiment for Freeway Design Safety Evaluations. Transportation Research Record, 2023; 2677: 73–92. doi:10.1177/03611981221144296.

[35] Alonso, F., Faus, M., Riera, J. V., Fernandez-Marin, M., Useche, S. A. Effectiveness of Driving Simulators for Drivers' Training: A Systematic Review. Applied Sciences, 2023; 13: 5266. doi:10.3390/app13095266.

[36] Li, G., Lai, W., Sui, X., Li, X., Qu, X., Zhang, T., Li, Y. Influence of traffic congestion on driver behavior in post-congestion driving. Accident Analysis & Prevention, 2020; 141: 105508. doi:10.1016/j.aap.2020.105508.

[37] Ramezani-Khansari, E. Analysis of the following behavior and line changes of anxious and depressed drivers. Ain Shams Engineering Journal, 2026; 17: 104294. doi:10.1016/j.asej.2026.104294.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

CEAS Civil Engineering and Applied Solutions

Reliability Assessment of Steel Structure Subjected to Blast Loading Using Monte Carlo Sampling

Zeinab Meghdadi [a], Mohammad Shabanlou [a*]

[a] *Department of Civil Engineering, Sharif University of Technology, Tehran, Iran*

A R T I C L E I N F O

Article history:

Received 29 April 2026

Revised 04 June 2026

Accepted 07 August 2026

Available online 01 April 2027

Keywords:

Steel plate shear wall

Probabilistic modeling

Monte Carlo

Blast loading

A B S T R A C T

In recent years, the occurrence of accidental or deliberate explosions has prompted engineers to consider the effect of explosive blast load on the overall performance of structures. Normally, residential structures are designed to resist conventional design loads such as gravitational, earthquakes, wind, and snow. In this research, a steel structure with a lateral load-bearing system of steel plate shear wall has been subjected to explosive blast load. Uncertainty in structural components as well as blast loading can lead to an incorrect and erroneous estimate of the structural response to blast. This study investigates the reliability of a steel structure under blast load, taking into account the uncertainty in the distance between the blast charge and the structure, the weight of the blast charge, the wall plate thickness, the yield stress of the wall plate, the yield stress of the beam and column of the structure. The probability of failure is then obtained by modeling the structure in OpenSees software and generating 1700 samples using Monte Carlo sampling method. Results indicated that the probability of failure of the mentioned system under the blast load is 55% and the most effective parameter is web plate thickness.

How to cite this article: Meghdadi, Z., Shabanlou, M. Reliability Assessment of Steel Structure Subjected to Blast Loading Using Monte Carlo Sampling. Civil Engineering and Applied Solutions. 2027; 3(2): 131–143. doi:10.22080/ceas.2026.31632.1096.

1. Introduction

Conventional steel structures are often designed considering regular loads such as earthquake and wind. However, structures are prone to internal or external explosions during their lifetime, which brings catastrophic damage to buildings and results in causalities. Therefore, the blast response of structures and identifying their residual operational performance is of great importance. Additionally, an important characteristic for blast-subjected structures is the difference in the material properties. For instance, as the strain rate is high in blast loading, the yield stress of steel increases significantly while the elastic modulus remains the same.

Recently, the use of steel plate shear wall (SPSW) as a lateral load-bearing system for steel structures has increased. The main components of a SPSW system are beams and columns, also known as boundary members and steel web plate. High strength and stiffness characteristics, stable hysteresis loops, considerable energy dissipation capacity, high ductility, and redundancy are important advantages of SPSW systems [1-3].

Preserving human lives and the protection of equipment and assets are two important aspects of engineering. Therefore, structures are designed for gravitational and lateral loads based on the current design codes. Financial concerns, as well as the low probability of high-risk loads such as impact or explosion, usually cause these loads to be excluded from the design process. Normally, only military structures or special structures are designed and controlled for the mentioned loads. However, with the increase in the possibility of terrorist attacks and blast incidents, it seems necessary to have a good insight on the performance of normal residential structures against explosions. Many studies have focused on reliability analysis of structures when exposed to explosive blast loading. Stewart et al. performed a reliability analysis to assess the risks of damage caused by blast loading to

* Corresponding author.

E-mail addresses: mohammad.shabanlou@sharif.edu (M. Shabanlou).

https://doi.org/10.22080/ceas.2026.31632.1096

ISSN: 3092-7749/© 2027 The Author(s). Published by University of Mazandaran.

infrastructure and proposed a probabilistic risk assessment procedure [4]. Al-Habahbeh and Stewart [5] conducted a structural reliability analysis of unreinforced brick masonry walls subjected to explosive loading. Using Monte Carlo sampling method, the failure probability shows that the wall has a very high reliability when subjected to explosive blast loading. In a study conducted by Shi and Stewart, blast reliability curves for RC columns under blast loading for different threat scenarios were generated using Monte Carlo sampling method [6]. The results showed that by considering spatial variability of concrete compressive strength and concrete cover, the probability of failure increases. Stewart and Netherton compared field test results with reliability-based computer analysis and proposed design load factors for different reliability levels. For instance, the load factor for a 0.99 reliability level will be as high as 1.30 [7, 8]. Stochino conducted a parametric and reliability study on the response of reinforced concrete beams under explosive loads [9]. Results showed that slenderness and peak load value were the most effective parameters affecting the beam response. Thomas et al. investigated the effect of vehicle collision and explosive loads on the reliability of bridge columns using a 1st-order 2nd-moment reliability study [10, 11]. The most affective parameter affecting the reliability of structures subjected to explosive loading was demonstrated to be the diameter of the column and the standoff distance of the blast.

The aim of the current study is to further focus on the reliability of steel structures with an SPSW system under blast loading. The parameters that are investigated here are the scaled distance, the infill plate thickness, the plate yield stress, and the yield stress of the structure's beam and column.

2. Steel plate shear wall system

The steel plate shear wall (SPSW) works as a lateral bearing system, comprising an infill web plate in a rigid frame, with the frame members serving as the boundary members of these plates. In a frame with multiple stories, the SPSW system can be likened to cantilever girders, where the columns act as flanges, floor beams like stiffeners, and at the end, steel web plates as the girders web [12-14]. However, as the stiffness of the flanges in plate girders is insufficient, SPSWs are fundamentally different from plate girders in terms of structural performance. In other words, as the stiffness of columns is very high in SPSW, the tension field expands in the whole area of the plate, but this phenomenon is not possible in a conventional plate girder [15].

The SPSW system demonstrates a stable hysteresis loop, rendering it a dependable option as a lateral structural system [16]. The shear actions are resisted by a field of tension created in the SPSW. The web plate does not provide remarkable strength in the linear elastic region; however, it is known for its high post-buckling strength [17-20]. According to the concept of capacity design, the tension field which is formed in the steel plate, imposes forces on the boundary elements. So columns and beam shall be designed in such a way to resist this force [21-23]. Also, Seismic assessment of the SPSWs reveals sufficient ductility, stiffness, and energy absorption [24-26].

3. Structural design

The structural design is performed in accordance with AISC 360 [27] and AISC 341 [28], while the seismic and gravity loads are determined based on ASCE 7-16 [29]. The considered building represents a conventional residential steel structure located in a region with high seismicity. Therefore, the structural configuration and member sizes are selected to satisfy the requirements of standard seismic design practice rather than special blast-resistant design provisions.

The perimeter beams of the SPSW system are connected to the columns using rigid connections, while the remaining beams (gravity beams) are modeled with hinged connections [30]. The structure is designed to resist the demands resulting from gravitational and seismic loading. According to AISC 341 [28], the aspect ratio of steel plate shear walls should be between 0.8 and 2.5, and this requirement is satisfied in the present model.

The prototype structure is a three-story building with a story height of 3.3 m and a bay span of 5 m. Four SPSW bays are located on the perimeter of the building, and the structure was initially analyzed and designed using a three-dimensional model. After completing the design process and finalizing the structural sections, including columns, beams, and steel plate thickness, one representative perimeter frame was extracted and modeled as a two-dimensional frame in OpenSees [31] for the nonlinear dynamic analyses. The schematic view of the structural model is shown in Fig. 1.

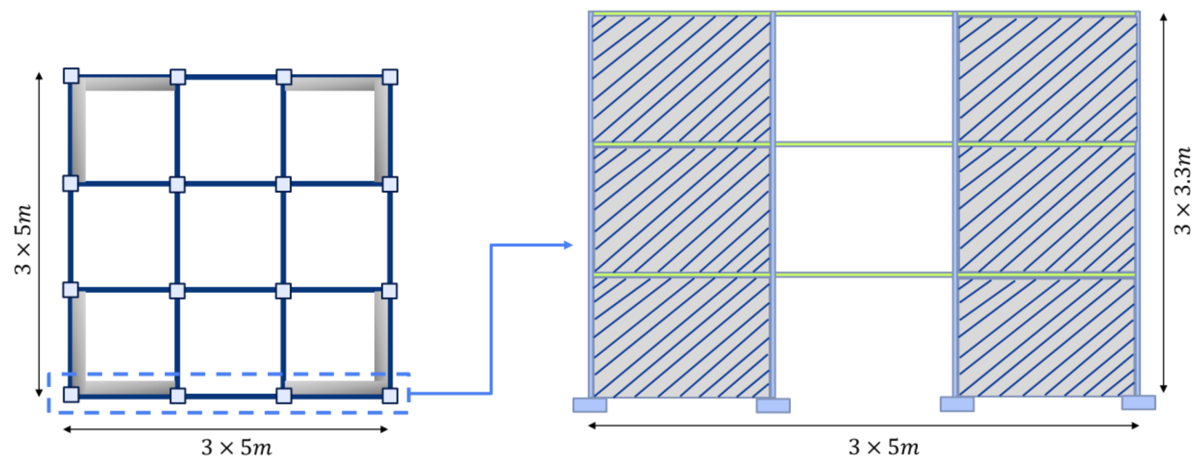


Fig. 1. Schematic view of the studied structure.

In the considered building, two SPSW systems are present in each principal direction. Therefore, the blast load applied to the structure is assumed to be equally shared between the two frames, and half of the calculated blast pressure is applied to the analyzed frame.

In the selected frame, the columns consist of box sections (BOX 200×20). The boundary beams at the first and second stories are plate girders with web dimensions of 350×8 mm and flange dimensions of 200×20 mm. The boundary beams at the third story are plate girders with web dimensions of 400×10 mm and flange dimensions of 300×25 mm.

4. OpenSees modelling and verification

4.1. Numerical model

The structure was modeled and analyzed using the nonlinear analysis software OpenSees [31]. The strip model introduced by Thorburn et al. [32] was utilized. Also, the newest material model developed by Jalali et al. [33-35]. The area of cross-section of each truss element is correlated with the steel plate thickness by Eq. 1 [30].The variables α, t_w, h, L, and n represent the web plate inclination angle, thickness of the wall, the height and length of the wall, and the number of strip truss elements, respectively.

$$A_s = \frac{[L\cos(\alpha) + h\sin(\alpha)]}{n} t_w \tag{1}$$

To capture the behavior of beams and columns, a plasticity distributed model with nonlinear elements was implemented. The modeling approach assumes that entering the nonlinear region is possible for every point of the elements, allowing for use of materials with nonlinear behavior along the length of the element for the analyze.

To simulate the steel plate shear wall, 10 truss elements were employed in each direction, meeting the minimum number of recommendations outlined in AISC 341-16 [28]. Consequently, the wall was composed of 20 truss strip elements in each infill panel.

In OpenSees, the Steel02 material, a strain-hardening isotropic steel with yield stress F_y = 235 MPa, modulus of elasticity E_s = 206 GPa, and a strain-hardening ratio of 0.01, is utilized to model beams and columns. Additionally, the SPSW02 material proposed by Jalali and Banazadeh [33, 34] is employed to model the strips equivalent to the infill plate. Also, a material with the same properties is used for the columns and beams. The cyclic behavior of this material is illustrated in Fig. 2, with $\sigma_{max,1}$ representing the ultimate stress of the main loading cycle.

The use of equivalent strip elements for the infill plate allows the numerical model to represent the post-buckling behavior of the SPSW system. After buckling, the steel plate contributes to the lateral resistance mainly through tension-field action, which is captured by the nonlinear response of the SPSW02 strip elements. Therefore, the global nonlinear behavior of the SPSW system is simulated through the interaction between the boundary beams, boundary columns, and the yielded plate strips. In the reliability analysis, the yield stress of the steel infill plate and the yield stress of the boundary beams and columns were considered as uncertain parameters. The elastic modulus was assumed to remain unchanged, which is consistent with the adopted treatment of high-strain-rate effects, where the steel yield stress is modified while the elastic stiffness is maintained.

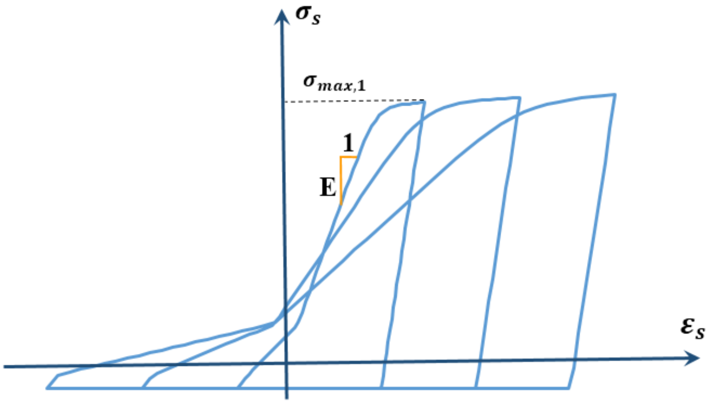


Fig. 2. Cyclic performance of SPSW02 model in OpenSees.

In the nonlinear dynamic analyses, viscous damping was modeled using Rayleigh damping implemented in OpenSees. In this approach, the damping matrix is constructed as a combination of mass-proportional and stiffness-proportional components. In the present model, the damping formulation includes a mass-proportional term and a stiffness-proportional term based on the last committed stiffness matrix of the structure.

The damping coefficients were determined to achieve a target damping ratio of 5%, which is commonly adopted for steel building structures in the literature. The coefficients were calculated using the first and fifth natural vibration modes of the structure obtained from eigenvalue analysis. The circular frequencies of these modes were determined from the corresponding eigenvalues, and the Rayleigh damping parameters were then calculated based on the selected modal frequencies and the target damping ratio.

This damping formulation was applied consistently to all nodes and elements of the numerical model, and the same damping parameters were used in all nonlinear time-history analyses performed for the blast loading simulations.

4.2. Verification

The numerical simulation in OpenSees was verified with an experimental specimen from a test performed by the study of Sabouri-Ghomi and Sajjadi [36]. The specimen was a single-story SPSW, with an aspect ratio of 1.46. The wall infill plate was modeled using the strip model corresponding to 20 trusses with material of SPSW02 as in studies by Jalali and Banazadeh [33, 34]. In addition, the boundary members were modeled using steel02 material with non-linear elements of steel. The results obtained from the cyclic experiment showed good agreement with numerical results. It is important to note that there are no experiments on the behavior of SPSW under blast loading to be used as a reference for verification purposes. Based on FEMA 427 [37], it is possible to model the explosive blast load as a quasi-static load (pressure load). So, to get the most similar reliable results, verification was carried out by implementing a cyclic loading model that is comparable to blast numerical modeling implemented herein. The comparison of test specimen and numerical results can be observed in Fig. 3.

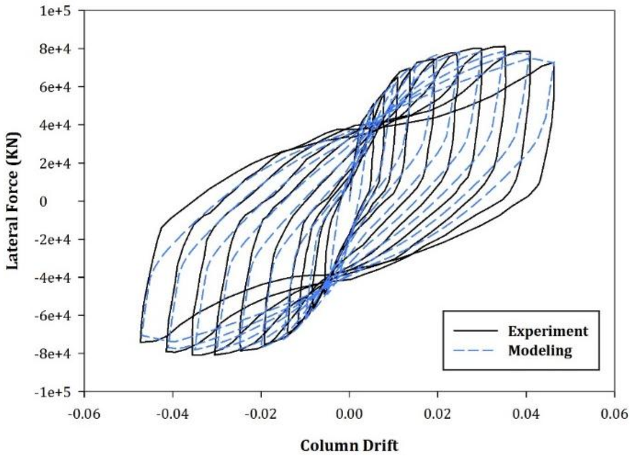


Fig. 3. Verification of experimental specimen with numerical model.

5. Blast loading

Blast loading significantly exceeds the structural design forces and is applied over a brief period of time. The structural performance under blast and ballistic loading is highly intricate, influenced by factors such as high strain rate, elevated temperature, and propagation of blast waves within the structure. This study examines the behavior of a steel structure with an SPSW system under external blast loading, which occurs when the explosive charges are situated outside the structure.

To accurately simulate these conditions, the blast loading was determined using empirical relationships based on the charge weight and standoff distance, consistent with the procedures described in UFC 3-340-02. The blast event was characterized by estimating the incident overpressure and the peak reflected pressure. For the 2D OpenSees model, the complex time-varying blast wave was simplified into an equivalent triangular pressure-time pulse, which is a standard engineering approximation for structural response analysis.

To account for the 3D nature of the building in a 2D framework, the total blast pressure acting on the building front was distributed across the primary structural frames. Since the structure contains two identical SPSW systems in each principal direction, the total blast load was halved and applied to the 2D analytical frame as a uniformly distributed time-varying load acting on the front-facing columns and beams. The spatial distribution of this pressure was calculated based on the tributary area of the structural members. In this analysis, only the positive phase of the blast pressure was considered; negative phase effects and pressures on the back face of the structure were neglected to maintain a conservative estimate of the structural response.

5.1. Dynamic properties of materials

Blast loading results in very high strain rates ranging from 10^2 to 10^4 s^{-1}, whereas the rate of strain in static condition falls between 10^{-6} to 10^{-5} s^{-1} [38]. These high strain rates change the dynamic mechanical properties of the structure, leading to different types of failure in various members of the structure. The impact of large values of strain rates on the properties of steel has been the subject of numerous studies [39].

It has been observed that high strain rates significantly influence material properties, resulting in changes in the flow of the plastic stresses and material ductility during impulsive loadings. With increasing strain rates, both the yield and ultimate tensile stress also rise significantly [40]. Yield stress (f_{dy}) and also ultimate stress of design in dynamic state (f_{du}), can be obtained by modifying the pure mentioned stresses through strength and dynamic increasing parameters, respectively (Eqs. 2 and 3) [41].

$$f_{dy} = (SIF) \times (DIF) \times f_y \tag{2}$$

$$f_{du} = (SIF) \times (DIF) \times f_u \tag{3}$$

SIF is a strength rise parameter as defined in Table 1. Additionally, DIF is the dynamic increasing parameter and is provided for steel members, as tabulated in Table 2.

Table 1. Resistance coefficients for various materials.

Material	SIF
S500 rebar	1.15
Rolled steel ST37 and ST52	1.15
Plate girders and other plate-made sections	1.15

Table 2. Dynamic increasing parameter (DIF) for common steel materials.

	Yield stress		Ultimate stress
Loading type material	Bending-shear $\frac{f_{dy,y}}{f_y}$	Compression-tension $\frac{f_{dy,y}}{f_y}$	All $\frac{f_{du,u}}{f_u}$
ST37	1.3	1.2	1.1
ST52	1.2	1.15	1.05

The dynamic increasing parameter is utilized to adjust static strength values in order to account for the impact of high strain rates on material strength. This factor is influenced by the type of stress, such as bending or direct shear. Flexural stresses develop rapidly, while shear strain rates are relatively low, resulting in delayed maximum shears. For axial members subjected to compressive or tensile forces, the strain rate is lower compared to flexural members. Experimental evidence indicates that the modulus of elasticity in dynamic condition remains unchanged in steel materials. Therefore, the yielding strength increasing coefficient for beams was 1.495, and for columns and the steel plate shear wall, this value has been considered equal to 1.38.

5.2. Blast load on front face

The front face is exposed to a two-line pressure pattern, resulting in an equivalent triangular pressure as shown in Fig 4. The maximum over-pressure on the front face, also referred to as the front wave of the explosive blast, is equal to the pressure of reflection (p_r). This pressure reaches the decay state (p_s) at the leveling time (t_c) [41-43].

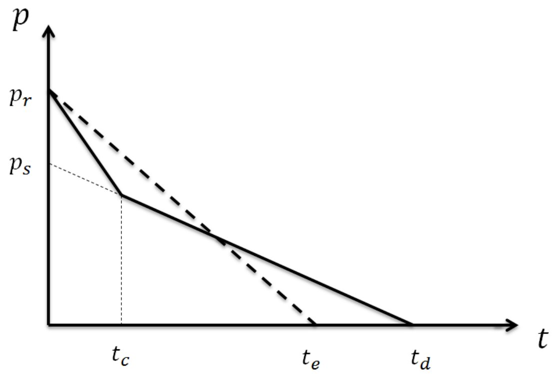


Fig. 4. Blast load on front face.

The leveling time (t_c) is defined by S, G, R, and C_r parameters (Eq. 3). S and G are levelling distances, defined by Eqs. 4 and 5, and R is the ratio of S/G. In addition, C_r is the velocity of sound, which can be obtained from the graph of Fig. 5.

$$t_c = \frac{4S}{(1+R)C_r} \tag{4}$$

$$S = min\left\{B \; ; \frac{H}{2}\right\} \tag{5}$$

$$G = max\left\{B \; ; \frac{H}{2}\right\} \tag{6}$$

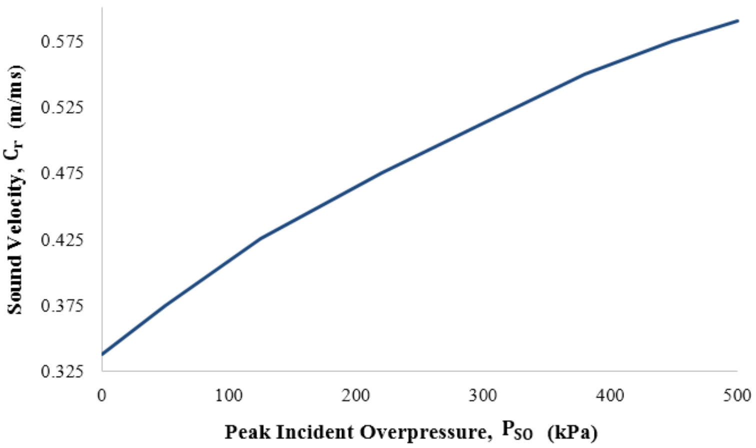


Fig. 5. C_r values as a function of peak incident over-pressure.

Furthermore, the parameters $P_{SO,}$ which is the peak over-pressure of the explosion, and t_e are obtained based on Eqs. 7 to 9.

$$P_S = P_{SO} + C_d q_s \quad (7)$$

$$q_s = \frac{5P_{SO}^2}{2(P_{SO}+7P_0)} \quad (8)$$

$$t_e = (t_d - t_c)\frac{P_S}{P_R} + t_c \quad (9)$$

5.3. Blast load on side face and roof

As there is no reflection wave, the blast load subjecting to the side faces are less than the blast loads on the front face. The triangular blast load with the maximum value of P_a is presented in Fig. 6. P_a is determined using C_e, q_s, C_d, and λ_{rw} parameters by Eq. 10. $C_{e,}$ which is the blast reduction coefficient, is obtained using Fig. 7 [41, 43]. q_s is the maximum dynamic pressure, and C_d is the drag coefficient calculated using Table 3. Also, λ_{rw} is the blast wave length calculated using Eq. 9.

$$\lambda_{rw} = U_s \times t_d \quad (10)$$

$$U_s = 340\sqrt{1 + 0.83P_{SO}} \quad (11)$$

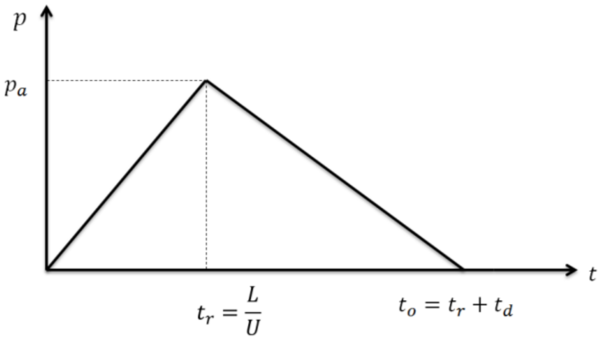


Fig. 6. Blast load on side face and roof.

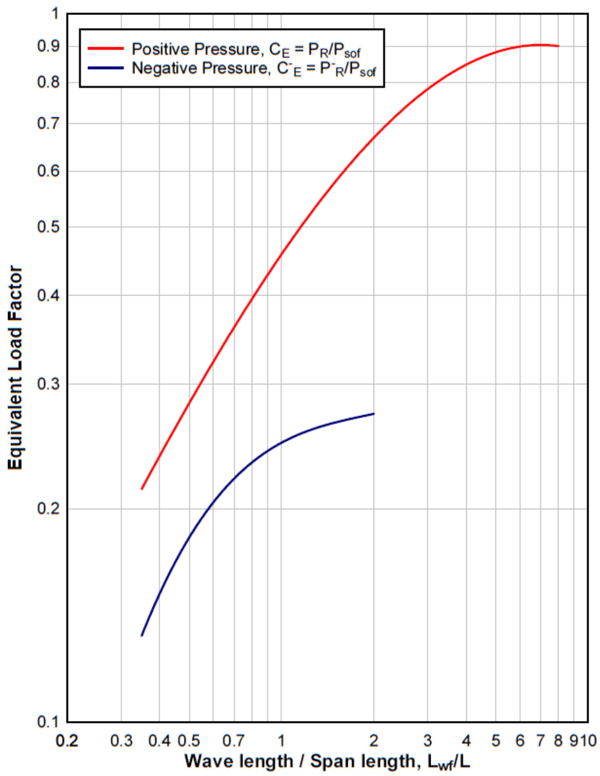


Fig. 7. Blast reduction coefficient.

Table 3. Drag coefficient (C_d).

Maximum dynamic 0ressure (q_s) (kg/cm^2)	Drag coefficient (C_d)
0-1.75	-0.4
1.75-3.5	-0.3
3.5-9	-0.2

5.4. Blast load on back face

The equation used for calculation of P_a for the back face is the same as for the side faces. However, for calculation of the C_e value, the height (H) and the length of the building (L) are summed up as shown in Fig. 8. Blast loads on the front face and back face act in the opposite direction, resulting in a reduction in the final value of blast load. Therefore, the blast load on the back face is neglected conservatively [41].

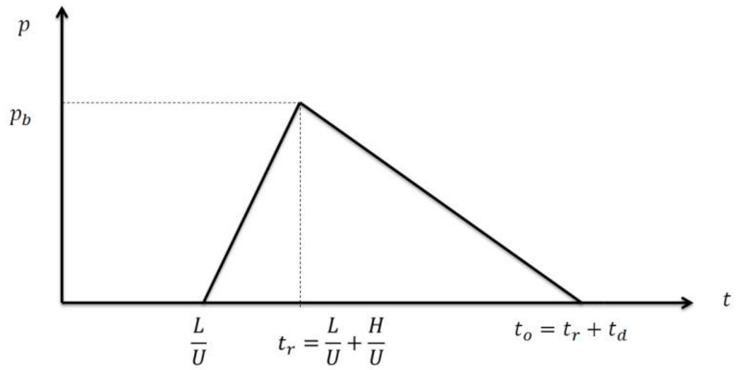


Fig. 8. Blast load on back face.

5.5. Applying blast load

The method of explosive charge and its standoff distance was used to determine the blast type and, therefore, the calculation of incident overpressure, peak reflected pressure, and blast duration. According to the US homeland security document [44], explosive charges and their improvised explosive devices can be categorized as depicted in Fig. 9. In this study, a charge weight (W) equal to 1814 kg was selected and situated at a standoff distance (R) equal to 20 m from the structure, as illustrated in Fig. 10.

Improvised Explosive Device	Explosive Capacity
	227 kg
	454 kg
	1814 kg
	4536 kg
	27215 kg

Fig. 9. Types of explosive charges.

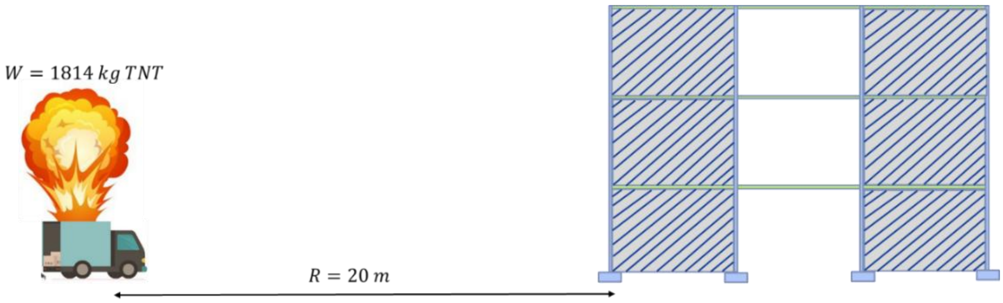


Fig. 10. Standoff distance.

The blast parameters, including the incident overpressure (P_{SO}) and blast duration (t_d), were determined using empirical relationships based on the scaled distance $Z = R/W^{1/3}$. Following the methodology of Moghaddam et al. [45], these parameters were estimated utilizing the empirical approaches proposed by Brode [46] and Newmark and Hansen [47], which are standard for surface blast analysis. These parameters can be calculated using Eqs. 12 to 14.

$$Z = \frac{R}{W^{\frac{1}{3}}} \tag{12}$$

$$t_d = W^{\frac{1}{3}} \frac{980\left[1+\left(\frac{Z}{0.54}\right)^{10}\right]}{\left(\left[1+\left(\frac{Z}{0.02}\right)^{3}\right]\times\left[1+\left(\frac{Z}{0.74}\right)^{6}\right]\times\left[1+\left(\frac{Z}{6.9}\right)^{2}\right]\right)} \tag{13}$$

$$P_{SO} = P_0 \frac{808\left[1+\left(\frac{Z}{4.5}\right)^{2}\right]}{\left(\left[1+\left(\frac{Z}{0.048}\right)^{2}\right]\times\left[1+\left(\frac{Z}{0.32}\right)^{2}\right]\times\left[1+\left(\frac{Z}{1.35}\right)^{2}\right]\right)} \tag{14}$$

6. Structural reliability analysis

As the failure modes of the structural systems are not independent statistically, nor are they entirely dependent, the analysis of reliability is intricate. The demand and capacity are not statistically independent, and the structural performance will be determined from a non-linear finite element analysis.

Optimization is the main pillar for the first-order reliability method (FORM); therefore, this method is susceptible to nonconvergence problems, particularly in dynamic problems. A crucial element for achieving convergence is the requirement for the limit-state function to possess continuous differentiability. Consequently, models containing discontinuities and piecewise functional forms are deemed unfeasible within the First Order Reliability Method (FORM) unless appropriate smoothing techniques are applied [48]. This makes FORM less appealing for engineering purposes in the construction industry. Moreover, in blast loading applications where nonlinear dynamic problems are prevalent, the calculation of the index of reliability using FORM necessitates the linearization of the function of limit-state, potentially resulting in unreliable reliability indices.

Therefore, the damage probability can be derived from a stochastic finite element analysis that will employ event-based Monte Carlo Simulation (MCS) analysis. By using MCS, it is possible to track the failure modes for each generated sample. Despite being computationally intensive, MCS is considered a robust method due to the requirement of a significant number of sample points, particularly when using crude sampling techniques. Nonetheless, this approach is favored as a benchmark tool when precision is valued more than computational efficiency.

A reliability assessment comprises two primary components: stochastic variables and limit-state functions. Stochastic variables, which are gathered in vector x, represent the uncertainties in the problem. The limit-state function, denoted as g, defines the occurrence for which the probability is being determined. For instance, in the fundamental reliability scenario, the main objective is to calculate the probability of system failure, also known as exceedance probability (EP). Failure takes place when the system's capacity, $R(x)$, surpasses the demand, $S(x)$, resulting in the limit-state function, as shown in Eq. 15.

$$g(x) = R(x) - S(x) \tag{15}$$

The exceedance probability, which is shown by G, is related to the reliability index denoted by β based on Eq. 16.

$$G = \Phi(-\beta) \tag{16}$$

where Φ is the univariate standard normal cumulative distribution function (CDF).

Finally, the exceedance probability in the sampling method is calculated using a step function, $I(x)$, by Eq. 17.

$$G = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} I(x) f_x(x) dx \tag{17}$$

where $f_x(x)$ is the joint probability density function of random variables x, and the number of integrals equals the number of variables in x, denoted here by n. Also, $I(x)$ is defined as Eq. 18.

$$I(x) = \begin{cases} 1, & g(x) \le 0 \\ 0, & g(x) > 0 \end{cases} \tag{18}$$

Reliability analysis are carried out to assess the structural performance in real-world conditions characterized by uncertainties in loadings and structural material and details. While a basic structural analysis can determine whether a structure will fail in specific conditions, in situations where the structural demand or resistance is uncertain, a reliability assessment is employed to assess the probability of failure. The sampling-based reliability assessment is carried out to investigate the vulnerability of a steel building with an SPSW system designed for seismic forces when exposed to blast explosive loading. In this study, the limit-state function has been chosen for checking the maximum drift ratio with 1.5%. In order to perform the stochastic nonlinear finite element analysis, the Rtx software has been linked to OpenSees software [49]. An overall view of the procedure has been presented in Fig. 11.

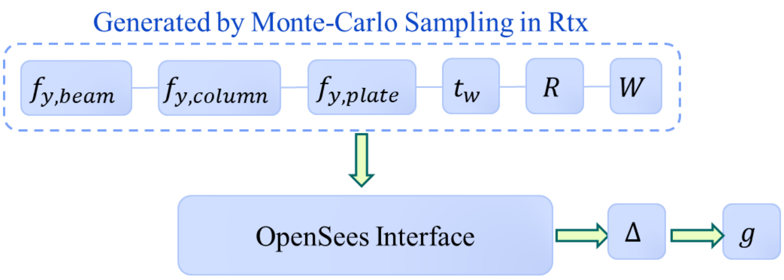


Fig. 11. Overall view of stochastic finite element analysis.

In this study, the maximum inter-story drift ratio was selected as the engineering demand parameter for defining structural performance under blast loading. The failure limit state was defined as exceedance of a 1.5% inter-story drift ratio. This value was selected based on the allowable story drift limits specified in ASCE 7 for building structures and was adopted as a practical performance-based deformation threshold. Although blast loading is different from seismic loading in terms of duration and loading mechanism, the ASCE 7 drift limit provides a recognized engineering basis for defining excessive lateral deformation.

In addition, because the present study is a probabilistic and comparative reliability assessment, a consistent drift threshold is required to classify the structural response of all Monte Carlo simulations into safe and failed cases. Therefore, the 1.5% drift ratio was used as a uniform limit state to evaluate and compare the vulnerability of the SPSW structure under different blast-loading scenarios and uncertainty realizations (Eq. 19). Exceedance of this threshold was interpreted as reaching an unacceptable deformation/damage state, not necessarily global collapse of the structure.

$$g(x) = 0.015 - \Delta(x) \tag{19}$$

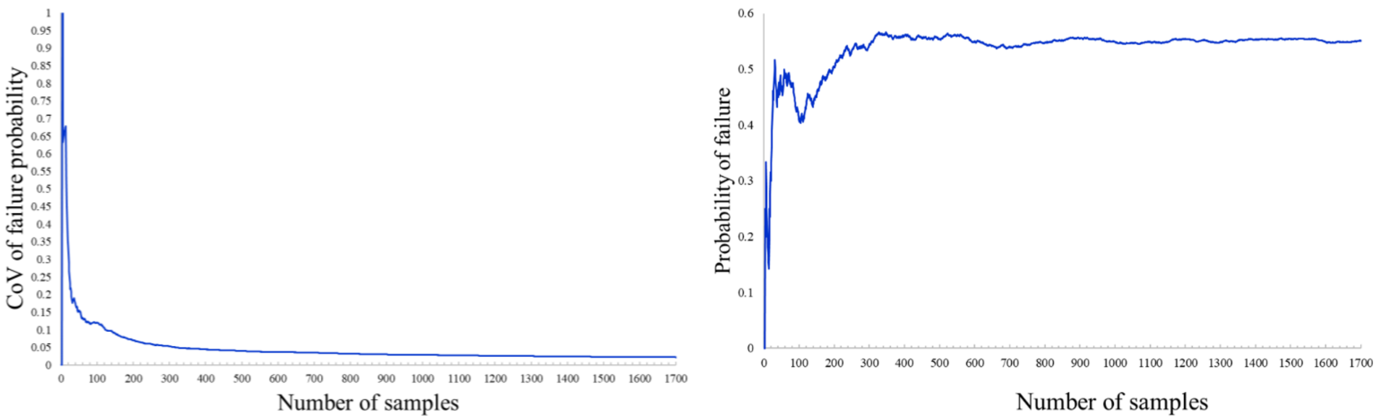
Uncertainty in material properties, geometric characteristics and blast parameters is unavoidable. Even with the advancements in manufacturing technologies and precision measuring devices, the precise values of these parameters remain challenging to ascertain. As a result, tolerance ranges for the parameters have been employed instead. Every parameter is allocated a fixed value and a distribution derived from pertinent literature or building standards, and the reliability assessment is carried out in the manner outlined in the preceding section. Table 4 provides the variables, related distributions, and values employed for the fixed analysis.

Table 4. Variable, distribution, and mean values [50].

Uncertainties			
Variable	**Distribution**	**Mean**	**CoV**
t_w	Lognormal	1 mm	0.03
$f_{y,\,plate}$	Lognormal	1.38 × 2400 kg/cm^2	0.07
$f_{y,\,column}$	Lognormal	1.38 × 2400 kg/cm^2	0.07
$f_{y,\,beam}$	Lognormal	1.495 × 2400 kg/cm^2	0.07
R	Lognormal	20 m	0.03
W	Normal	1814 kg	0.001

7. Results and discussions

Fig. 12 illustrates the convergence behavior of the Monte Carlo simulation results as the number of generated samples increases. In this analysis, each Monte Carlo sample represents a realization of the uncertain parameters listed in Table 4. For each sample, a nonlinear dynamic analysis was performed, and the maximum inter-story drift ratio was obtained. If the calculated drift ratio exceeded the selected performance limit of 1.5%, the sample was classified as failure. The cumulative failure probability was then calculated as the ratio of failed samples to the total number of generated samples. By progressively increasing the number of samples, the convergence of the estimated failure probability was evaluated, which resulted in the curve shown in Fig. 12.



(a) **(b)**

Fig. 12. (a) CoV for probability of failure; and (b) probability of failure.

The stability of the Monte Carlo estimator was verified by examining the convergence of the failure probability as the number of samples approached 1700. The observed stabilization of the estimator, characterized by a Coefficient of Variation of approximately 0.02, confirms that the adopted sample size is sufficient to provide a stable estimate for the investigated performance limit state. In addition, the probabilistic characteristics of the random variables used in the Monte Carlo simulation are summarized in Table 4. As shown in this table, the uncertainties considered in the reliability analysis include the web plate thickness, yield stress of the web plate, yield stress of columns and beams, standoff distance, and charge weight. The corresponding mean values and coefficients of variation are reported in Table 4. This information clarifies the probabilistic input parameters used for generating the Monte Carlo samples and improves the reproducibility of the reliability analysis.

1700 samples were generated using Monte Carlo sampling method. Considering the uncertainties in material properties of steel for columns, beams, and web plate, web plate thickness, standoff distance, and charge weight, the fragility curve can be obtained as shown in Fig. 13. As is evident in Fig. 13, the probability that the maximum drift ratio surpasses the 1.5% limit is 55%, and the reliability index is equal to -0.129. This shows that a typical steel structure with an SPSW system designed for a high seismic area is highly vulnerable when exposed to an explosive blast induced by a truck vehicle. However, in order to avoid non-economical design sections, it is possible to decrease the hazard by defining safety circles to stop suspicious vehicles from entering the residential areas. Fig. 12a shows that after generation of 1700 samples, the failure probability has converged to an almost constant value with a Coefficient of Variation (CoV) equal to 0.02.

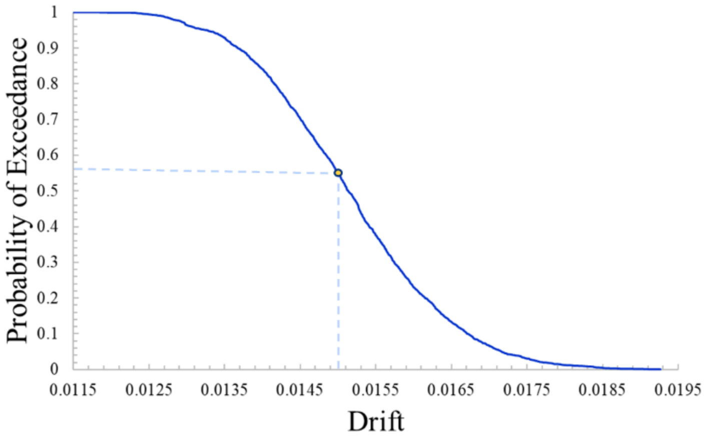


Fig. 13. Structural reliability under explosive blast considering the defined variable.

From an engineering performance perspective, the selected 1.5% inter-story drift ratio represents a deformation-based limit state for evaluating the serviceability and damage level of the SPSW system under extreme blast loading. Exceedance of this threshold does not necessarily indicate global collapse of the structure; rather, it implies that the structure has reached an unacceptable level of lateral deformation for the adopted performance criterion. Therefore, the calculated failure probability of 55% and the corresponding reliability index of $\beta = -0.129$ indicate a high likelihood of exceeding the selected drift-based performance limit under the considered truck-borne blast scenario. This result suggests that, although the SPSW building was designed for a high seismic region, its deformation demand under extreme blast loading can exceed the acceptable performance range, emphasizing the need for appropriate protective measures such as controlling the minimum standoff distance.

Also, a sensitivity analysis was performed. As depicted in Fig. 14, the results showed that the reliability index is mostly affected by the web plate thickness. To identify the relative importance of the uncertain input parameters, a sensitivity analysis was performed using the Monte Carlo simulation results. In this study, the sensitivity index was expressed as $dP_f/dStdv$, where P_f is the probability of failure, and *Stdv* denotes the standard deviation of each random variable. This index represents the rate of change in the estimated failure probability with respect to the standard deviation of a given uncertain parameter. Therefore, a larger absolute value of $dP_f/dStdv$ indicates that the probability of failure is more sensitive to the uncertainty associated with that parameter.

The parameter influence was evaluated by perturbing the standard deviation of each random variable individually while keeping the mean values and the statistical properties of the other variables unchanged. For each perturbation, the Monte Carlo simulations were repeated, and the corresponding change in failure probability was calculated. The sensitivity index was then obtained as the ratio between the change in probability of failure and the change in the standard deviation of the considered parameter.

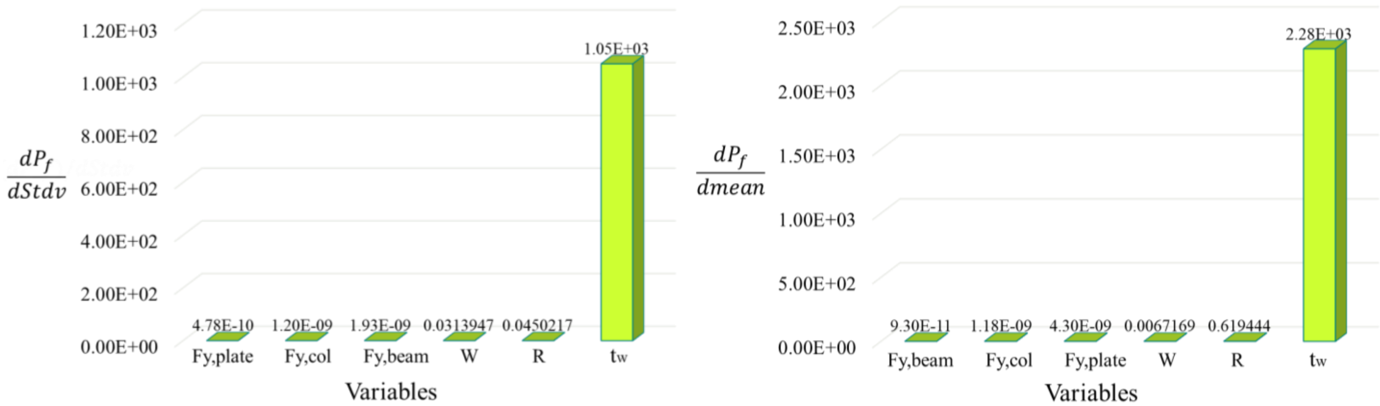


Fig. 14. Sensitivity analysis based on: (a) $\frac{dP_f}{dStdv}$; and (b) $\frac{dP_f}{dmean}$.

8. Conclusions

This paper performs a reliability analysis on a steel structure with a steel plate shear wall system that is basically designed for gravitational and seismic forces in a high seismic area based on the ASCE7-16 after being subjected to blast loading. The stochastic nonlinear dynamic finite element analysis has been carried out considering the inherent uncertainties in the variables such as web plate thickness of SPSW, yield stress of steel, standoff distance, and charge weight. A finite element model has been built and verified using the so-called strip model to capture the behavior of the steel structure after being subjected to explosive blast. 1700 analysis were done using samples that were generated based on the Monte Carlo sampling method with Rtx software. The following conclusions can be drawn:

- By considering W, R, web plate thickness, web plate yield stress, beam and column yield stress as random variables, the failure probability of a steel frame subjected to blast loading induced by a delivery truck is 55 percent. Also, the corresponding reliability index is equal to -0.129.
- Using sampling-based sensitivity analysis, web plate thickness, standoff distance, and explosion charge weight turn out to be the most important parameters in the failure probability of a steel frame.
- By performing histogram sampling, the exceedance probability curve of the drift ratio is obtained.
- Monte Carlo sampling can be used as an appropriate method to investigate the reliability level of steel structures with SPSW system.
- Rtx software can be used as an engineer-friendly reliability assessment tool to evaluate the reliability level of the designed structures exposed to blast loadings.

Statements & declarations

Author contributions

Zeinab Meghdadi: Investigation, Methodology, Resources, Writing - Original Draft, Writing – Review & Editing.

Mohammad Shabanlou: Investigation, Methodology, Resources, Writing - Original Draft, Writing – Review & Editing.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Data availability

The data presented in this study will be available on interested request from the corresponding author.

Declarations

The authors declare no conflict of interest.

References

[1] Shabanlou, M., Moghaddam, H., Saedi Daryan, A. The Effect of Geometry on Structural Behavior of Buildings with Steel Plate Shear Wall System Subjected to Blast Loading. International Journal of Steel Structures, 2021; 21: 650–665. doi:10.1007/s13296-021-00463-4.

[2] Sarcheshmehpour, M., Shabanlou, M., Meghdadi, Z., Estekanchi, H. E., Mofid, M. Seismic evaluation of steel plate shear wall systems considering soil-structure interaction. Soil Dynamics and Earthquake Engineering, 2021; 145: 106738. doi:10.1016/j.soildyn.2021.106738.

[3] Alvanchi, A., Jafari, M. A., Shabanlou, M., Meghdadi, Z. A novel public-private-people partnership framework in regeneration of old urban neighborhoods in Iran. Land Use Policy, 2021; 109: 105728. doi:10.1016/j.landusepol.2021.105728.

[4] Stewart, M. G., Netherton, M. D., Rosowsky, D. V. Terrorism Risks and Blast Damage to Built Infrastructure. Natural Hazards Review, 2006; 7: 114–122. doi:10.1061/(ASCE)1527-6988(2006)7:3(114).

[5] Al-Hababheh, O., Stewart, M. Structural Reliability Analysis of Unreinforced Brick Masonry Walls Subject to Explosive Blast Loading. In: Proceedings of the First International Conference of Protective Structures (ICPS); 2010 Sep–Oct; Manchester, United Kingdom.

[6] Shi, Y., Stewart, M. G. Spatial reliability analysis of explosive blast load damage to reinforced concrete columns. Structural Safety, 2015; 53: 13–25. doi:10.1016/j.strusafe.2014.07.003.

[7] Stewart, M. G., Netherton, M. D. Reliability-Based Design Load Factors for Explosive Blast Loading. Journal of Performance of Constructed Facilities, 2015; 29: 1–8. doi:10.1061/(ASCE)CF.1943-5509.0000709.

[8] Netherton, M. D., Stewart, M. G. Blast Load Variability and Accuracy of Blast Load Prediction Models. International Journal of Protective Structures, 2010; 1: 543–570. doi:10.1260/2041-4196.1.4.543.

[9] Stochino, F. RC beams under blast load: Reliability and sensitivity analysis. Engineering Failure Analysis, 2016; 66: 544–565. doi:10.1016/j.engfailanal.2016.05.003.

[10] Shabanlou, M., Mofid, M., Eslami, M., Kodur, V., Meghdadi, Z., Tavakoli, A. Experimental, analytical and numerical investigation on the effects of different shear reinforcement on residual post-fire behavior of reinforced concrete deep beams. Structure and Infrastructure Engineering, 2025; 1–18. doi:10.1080/15732479.2025.2456466.

[11] Thomas, R. J., Steel, K., Sorensen, A. D. Reliability analysis of circular reinforced concrete columns subject to sequential vehicular impact and blast loading. Engineering Structures, 2018; 168: 838–851. doi:10.1016/j.engstruct.2018.04.099.

[12] Astaneh-Asl, A., Zhao, Q. Cyclic tests of steel shear walls. Research project. Berkeley: Department of Civil and Environmental Engineering, University of California, 2001.

[13] Astaneh-Asl, A. Seismic Behavior and Design of Composite Steel Plate Shear Walls. Berkeley (CA): Steel Technical Information & Product Service (Steel TIPS); 2002. Report No.: 01.

[14] Astaneh-Asl, A. Steel Plate Shear Walls. In: U.S.-Japan Workshop on Seismic Fracture Issues in Steel Structures; 2000 Feb; San Francisco, California. p. 1–12.

[15] Berman, J. W., Bruneau, M. Capacity Design of Vertical Boundary Elements in Steel Plate Shear Walls. Engineering Journal, 2008; 45: 57–72. doi:10.62913/engj.v45i1.925.

[16] Elgaaly, M., Caccese, V., Du, C. C. Postbuckling Behavior of Steel-Plate Shear Walls under Cyclic Loads. Journal of Structural Engineering-asce, 1993; 119: 588–605. doi:10.1061/(ASCE)0733-9445(1993)119:2(588).

[17] Bahrebar, M., Lim, J. B. P., Clifton, G. C., Zirakian, T., Shahmohammadi, A., Hajsadeghi, M. Perforated steel plate shear walls with curved corrugated webs under cyclic loading. Structures, 2020; 24: 600–609. doi:10.1016/j.istruc.2020.01.047.

[18] Zirakian, T., Zhang, J. Seismic design and behavior of low yield point steel plate shear walls. International Journal of Steel Structures, 2015; 15: 135–151. doi:10.1007/s13296-015-3010-8.

[19] Emami, F., Mofid, M. On the hysteretic behavior of trapezoidally corrugated steel shear walls. The Structural Design of Tall and Special Buildings, 2014; 23: 94–104. doi:10.1002/tal.1025.

[20] Choi, I.-R., Park, H.-G. Steel Plate Shear Walls with Various Infill Plate Designs. Journal of Structural Engineering, 2009; 135: 785–796. doi:10.1061/(ASCE)0733-9445(2009)135:7(785).

[21] Meghdadi, Z., Khaloo, A. Experimental, theoretical and numerical study on flexural behavior of hybrid steel-GFRP reinforced concrete slabs. Structural Concrete, 2025; 26: 739–758. doi:10.1002/suco.202301085.

[22] Meghdadi, Z., Khaloo, A., Ashour, A. Residual structural performance of fire-exposed concrete slabs reinforced with hybrid GFRP-steel bars: Experimental and numerical studies. Structures, 2025; 80: 109757. doi:10.1016/j.istruc.2025.109757.

[23] Shabanlou, M., Meghdadi, Z., Ghaffar, S. H. Experimental and analytical study of the residual performance of reinforced concrete deep beams with circular web openings. Results in Engineering, 2025; 25: 104229. doi:10.1016/j.rineng.2025.104229.

[24] Shabanlou, M., Mofid, M., Tavakoli, A. Post-fire performance of GFRP reinforced concrete deep beams: experimental and numerical study. Structure and Infrastructure Engineering, 2024; 22: 1280–1298. doi:10.1080/15732479.2024.2382131.

[25] Shabanlou, M., Mofid, M., Tavakoli, A. Experimental and Numerical Study on the Behavior of Reinforced Concrete Deep Beams with Normal-Strength and High-Strength Concrete After Being Exposed to Fire. Arabian Journal for Science and Engineering, 2024; 49: 13457–13476. doi:10.1007/s13369-023-08676-x.

[26] Alinia, M. M. A study into optimization of stiffeners in plates subjected to shear loading. Thin-Walled Structures, 2005; 43: 845–860. doi:10.1016/j.tws.2004.10.008.

[27] American Institute of Steel Construction (AISC). ANSI/AISC 360-16: Specification for Structural Steel Buildings. Farmington Hills (MI): AISC; 2016.

[28] American Institute of Steel Construction (AISC). AISC 341-16: Seismic Provisions for Structural Steel Buildings. Farmington Hills (MI): AISC; 2016.

[29] American Society of Civil Engineers (ASCE). ASCE 7-10: Minimum Design Loads for Buildings and Other Structures. Reston (VA): ASCE; 2010.

[30] Sabelli, R., Bruneau, M. Steel Design Guide 20: Steel Plate Shear Wall. First ed. Farmington Hills (MI): AISC; 2007.

[31] Mazzoni, S., McKenna, F., H. Scott, M., L. Fenves, G. OpenSees Command Language Manual. Oakland (CA): University of California; 2006.

[32] Thorburn, L. J., Kulak, G. L., Montgomery, C. J. Analysis of Steel Plate Shear Walls. Calgary (CA): University of Alberta; 1983. Report No.: 107.

[33] Jalali, S. A., Banazadeh, M. Computer-based evaluation of design methods used for a steel plate shear wall system: Computer-Based Evaluation of SPSW Design Methods. The Structural Design of Tall and Special Buildings, 2016; 25: doi:10.1002/tal.1290.

[34] Jalali, S. A., Banazadeh, M. Development of a new deteriorating hysteresis model for seismic collapse assessment of thin steel plate shear walls. Thin-Walled Structures, 2016; 106: 244–257. doi:10.1016/j.tws.2016.05.008.

[35] Jalali, S. A., Darvishan, E. Seismic demand assessment of self-centering steel plate shear walls. Journal of Constructional Steel Research, 2019; 162: 105738. doi:10.1016/j.jcsr.2019.105738.

[36] Sabouri-Ghomi, S., Sajjadi, S. R. A. Experimental and theoretical studies of steel shear walls with and without stiffeners. Journal of Constructional Steel Research, 2012; 75: 152–159. doi:10.1016/j.jcsr.2012.03.018.

[37] Federal Emergency Management Agency (FEMA). FEMA 427: Risk Management Series FEMA 427 Primer for design of commercial buildings to mitigate terrorist attacks. Washington, D.C. (US): ASCE; 2003.

[38] Ngo, T., Mendis, P., Gupta, A., Ramsay, J. Blast Loading and Blast Effects on Structures – An Overview. Electronic Journal of Structural Engineering, 2007; 7: 76–91. doi:10.56748/ejse.671.

[39] Moghimi, H., Driver, R. G. Performance assessment of steel plate shear walls under accidental blast loads. Journal of Constructional Steel Research, 2015; 106: 44–56. doi:10.1016/j.jcsr.2014.11.010.

[40] Needham, C. E. Blast Waves (Shock Wave and High Pressure Phenomena). First ed. Berlin (DE): Springer, Heidelberg; 2010. doi:10.1007/978-3-642-05288-0.

[41] Iranian National Building Code, Part 21. 2nd ed. Tehran (IR): Building and Housing Research Center (BHRC); 2016.

[42] U.S. Department of Defence. UFC 3-340-02: Unified Facilities Criteria (UFC) Structures to Resist the Effects of Accidental Explosions. Arlington County (VA): UFC; 2008.

[43] Dusenberry, D. O. Handbook for Blast Resistant Design of Buildings. First ed. Hoboken (NJ): John Wiley & Sons, Inc; 2010. doi:10.1002/9780470549070.

[44] U.S. Department of Homeland Security. DHS Bomb Threat Stand-Off Chart. 2010. https://publicintelligence.net/dhs-bomb-threat-stand-off-chart/. (Accessed February 2010).

[45] Moghaddam, H., Shabanlou, M., Saedi Daryan, A. Study of the Blast Resistance of Seismically Designed Steel Plate Shear Wall Frames. Sharif Journal of Civil Engineering, 36.2: 117–129. doi:10.24200/j30.2019.53613.2562.

[46] Brode, H. L. Numerical Solutions of Spherical Blast Waves. Journal of Applied Physics, 1955; 26: 766–775. doi:10.1063/1.1722085.

[47] Newmark, N. M., Hansen, R. J. Design of blast-resistant structures. First ed. Columbus (OH): McGraw-Hill; 1961.

[48] Shoaei, P., Mahsuli, M. Reliability-based design of steel moment frame structures isolated by lead-rubber bearing systems. Structures, 2019; 20: 765–778. doi:10.1016/j.istruc.2019.06.020.

[49] Talebiyan, H., Mahsuli, M. Sampling-Based Reliability Sensitivity Analysis Using Direct Differentiation. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 2020; 6: 04020015. doi:10.1061/AJRUA6.0001050.

[50] Jordaan, I. Decisions under uncertainty: probabilistic analysis for engineering decisions. First ed. Cambridge (UK): Cambridge University Press; 2005.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

A Hybrid GMDH Neural Network-Genetic Algorithm Approach for Predicting Shear Wave Velocity Profiles

Omolbanin Ataee [a]

[a] *Department of Geography and Urban Planning, University of Mazandaran, Babolsar, Iran*

ARTICLE INFO

Article history:

Received 12 May 2026

Revised 16 June 2026

Accepted 07 August 2026

Available online 01 April 2027

Keywords:

Genetic algorithm

Group method of data handling

Mashhad city

Shear wave velocity

Standard penetration test

ABSTRACT

Shear wave velocity (V_s) is a fundamental parameter in geotechnical earthquake engineering and seismic site characterization; however, direct in-situ measurement remains expensive and spatially limited. This study proposes an interpretable hybrid group method of data handling-genetic algorithm (GMDH-GA) framework for predicting shear wave velocity profiles in heterogeneous alluvial deposits. A comprehensive geotechnical-geophysical database comprising 1991 records from 226 boreholes in Mashhad, Iran, was compiled using downhole measurements and standard penetration test (SPT) data. Depth (Z), SPT blow count (N), and upper-layer shear wave velocity (V_{su}) were selected as predictor variables. Separate soil-specific models were developed for clay, silt, sand, gravel, and combined soil datasets using a 70/30 training-testing strategy. The proposed models generated explicit polynomial equations while maintaining high predictive capability. Statistical evaluation demonstrated strong agreement between measured and predicted V_s values, with testing-stage coefficients of determination reaching 0.98-0.99 and low prediction errors across all soil groups. Comparative analyses showed that the proposed GMDH-GA models substantially outperform widely used empirical correlations in terms of accuracy and generalization capability. Sensitivity analysis indicated that V_{su} is the dominant predictor, followed by SPT resistance and depth. The results demonstrate that the proposed framework provides a reliable and interpretable alternative for estimating shear wave velocity profiles in regions where direct geophysical measurements are unavailable.

How to cite this article: Ataee, O. A Hybrid GMDH Neural Network-Genetic Algorithm Approach for Predicting Shear Wave Velocity Profiles. Civil Engineering and Applied Solutions. 2027; 3(2): 144–160. doi:10.22080/ceas.2026.31713.1099.

1. Introduction

Shear wave velocity (V_s) is one of the most important parameters used to characterize the dynamic behavior and stiffness of soil deposits. In particular, the average shear wave velocity of the upper 30 m of the ground profile (V_{s30}) is widely used for seismic site classification, site-response analysis, and seismic microzonation studies. Although V_{s30} is an important parameter for evaluating local site conditions, it should not be considered a direct measure of seismic hazard, which is controlled by multiple geological and seismological factors.

Direct measurement of V_s can be obtained using geophysical techniques such as downhole, crosshole, and surface-wave methods. However, these methods generally require specialized equipment, experienced personnel, and data interpretation or inversion procedures, which may increase project cost and complexity. Consequently, empirical correlations between V_s and routinely measured geotechnical parameters, particularly standard penetration test (SPT) blow counts, have been widely developed and applied in engineering practice.

* Corresponding author.

E-mail addresses: o.ataee@umz.ac.ir (O. Ataee).

https://doi.org/10.22080/ceas.2026.31713.1099

ISSN: 3092-7749/

Despite their practical value, conventional empirical V_s-SPT correlations often exhibit considerable uncertainty because soil behavior is influenced by multiple factors, including soil type, stress conditions, depositional environment, and regional geological characteristics. Therefore, advanced data-driven approaches capable of capturing complex nonlinear relationships may provide more reliable and site-specific predictions of shear wave velocity.

Extensive research has been conducted to establish empirical relationships between SPT blow count (N) and shear wave velocity. Early studies by Ohta and Goto [1] and Imai and Tonouchi [2] introduced power-law and stress-dependent formulations that have formed the basis of most subsequent models. These relationships were later refined by incorporating soil density, confining stress, and depositional effects [3, 4]. However, despite their simplicity and practical utility, such empirical correlations are generally site-specific and exhibit limited transferability due to variations in soil fabric, stress history, and geological conditions [5]. A number of correlations between the V_s and the SPT-N value are collected from the literature and shown in Table 1.

Table 1. Inventory of proposed empirical correlations between V_s and SPT-N value.

Authors	All soils	Gravel	Sand	Silt	Clay
Ohsaki and Iwasaki [6]	$V_s = 81.38N^{0.39}$		$V_s = 59.4N^{0.47}$		
Imai [7]	$V_s = 91N^{0.337}$		$V_s = 80.6N^{0.331}$		$V_s = 80.2N^{0.292}$
Ohta and Goto [1]	$V_s = 85.35N^{0.348}$				
Seed and Idriss [8]	$V_s = 61N^{0.5}$				
Lee Shannon [9]	$V_s = 76.2N^{0.24}$			$V_s = 104(N+1)^{0.334}$	$V_s = 138.4(N+1)^{0.242}$
Iyisan [10]	$V_s = 51.5N^{0.516}$				
Jafari et al. [5]	$V_s = 22N^{0.85}$				
Hasancebi and Ulusay [11]	$V_s = 90N^{0.309}$		$V_s = 90.82N^{0.319}$		$V_s = 97.89N^{0.269}$
Dikmen [12]	$V_s = 58N^{0.39}$		$V_s = 73N^{0.33}$	$V_s = 60N^{0.36}$	$V_s = 44N^{0.48}$
Maheswari et al. [13]	$V_s = 95.64N^{0.301}$		$V_s = 100.53N^{0.265}$		$V_s = 89.31N^{0.358}$
Hafezi et al. [14]	$V_s = 99N^{0.53}$	$V_s = 80N^{0.58}$	$V_s = 80N^{0.58}$	$V_s = 45N^{0.72}$	$V_s = 45N^{0.72}$
Chatterjee and Choudhury [15]	$V_s = 78.21N^{0.38}$		$V_s = 54.82N^{0.53}$		$V_s = 77.11N^{0.39}$
Fatehnia et al. [16]	$V_s = 77.1N^{0.355}$				
Kirar et al. [17]	$V_s = 99.5N^{0.345}$		$V_s = 100.3N^{0.338}$		$V_s = 94.4N^{0.379}$

To overcome these limitations, recent research has increasingly focused on data-driven machine learning approaches. Algorithms such as artificial neural networks (ANN), support vector machines (SVM), random forest (RF), and gradient boosting have demonstrated superior capability in capturing nonlinear relationships between geotechnical parameters and shear wave velocity. Numerous studies report that these models outperform conventional regression-based methods in terms of predictive accuracy and robustness [18-20]. Furthermore, hybrid and ensemble learning frameworks have further improved performance by enhancing generalization capability and reducing model uncertainty [21, 22].

Recent advances in machine learning have substantially expanded the application of intelligent predictive models in civil and geotechnical engineering. Beyond conventional neural networks, researchers have increasingly employed hybrid and ensemble frameworks to improve prediction accuracy, robustness, and generalization capability under complex nonlinear conditions. For example, Tabari et al. [23] developed a supervised committee neural network for estimating aquifer parameters and demonstrated that combining multiple intelligent predictors can significantly improve predictive performance compared with individual models. Their findings highlighted the potential of ensemble learning strategies for modeling highly nonlinear geotechnical and hydrogeological systems.

More recently, machine-learning-driven methodologies have been extended to structural and seismic engineering applications. Farahani [24] proposed a data-driven seismic design framework integrating wavelet transforms and machine learning algorithms for reinforced concrete frame assessment. The study demonstrated the effectiveness of hybrid intelligent systems in extracting complex patterns from engineering datasets and improving prediction reliability. Similarly, Sadeghpour Haji et al. (2026) employed artificial neural networks to model the shear behavior of reinforced concrete columns under constant axial loading and reported high predictive accuracy for nonlinear structural responses. These studies collectively indicate the growing trend toward integrating advanced machine learning techniques with engineering domain knowledge to address complex prediction problems [25].

Despite these developments, several challenges remain unresolved in shear wave velocity prediction. Existing studies predominantly focus on black-box machine learning algorithms that provide limited interpretability for engineering decision-making. Moreover, many published models have been developed using relatively small or geographically restricted datasets, which may not adequately represent the spatial variability and heterogeneity of natural alluvial deposits. In addition, only limited research has explored the integration of interpretable polynomial neural networks with evolutionary optimization algorithms for regional-scale V_s profile prediction. Consequently, there remains a need for predictive frameworks that simultaneously achieve high accuracy, transparency, and robustness while being developed from large-scale geotechnical databases. The present study addresses this research gap through the development of an interpretable hybrid GMDH-GA model calibrated using an extensive geotechnical-geophysical dataset from Mashhad, Iran.

Despite these advances, a key limitation of most machine learning models is their lack of interpretability. Many high-performance algorithms operate as black-box systems, providing accurate predictions without yielding explicit mathematical expressions, which limits their applicability in engineering design. In this context, the group method of data handling (GMDH), originally introduced by Ivakhnenko [26], offers a powerful alternative. GMDH is a self-organizing polynomial neural network that automatically constructs model structure through layer-wise selection of selected neurons. Unlike black-box models, it produces explicit mathematical equations, making it particularly suitable for engineering applications requiring both accuracy and transparency. Previous studies have confirmed its effectiveness in modeling complex geotechnical phenomena, including settlement, shear strength, liquefaction potential, and shear wave velocity [27-29]. To further improve performance, GMDH has been coupled with evolutionary optimization techniques, particularly genetic algorithms (GA). The hybrid GMDH-GA approach enhances model structure selection and parameter optimization, thereby improving generalization and avoiding local minima. Such a combination has demonstrated efficient performance in various geotechnical challenges; for instance, in soil property estimation and site characterization, such as predicting the correlation between unconfined compressive strength and standard penetration test results [30, 31] and in the assessment of soil liquefaction potential [32, 33].

Despite considerable progress in machine learning-based prediction of shear wave velocity, several important limitations remain unresolved. First, many high-performance machine learning models operate as black-box systems and therefore lack transparency and interpretability for practical geotechnical design. Second, most existing studies rely on relatively limited regional datasets that insufficiently capture the heterogeneity of complex alluvial deposits. Third, limited attention has been given to integrating interpretable polynomial neural networks with evolutionary optimization algorithms for regional V_s profile prediction.

To address these limitations, this study develops a hybrid GMDH-genetic algorithm framework using a comprehensive geotechnical-geophysical database consisting of 1991 records from 226 boreholes in Mashhad, Iran. Unlike conventional black-box approaches, the proposed method generates explicit polynomial equations while maintaining high predictive capability. In addition, separate soil-specific models are developed for clay, silt, sand, and gravel deposits to better represent soil-dependent nonlinear behavior. The main contributions of this study are:

- development of interpretable hybrid GMDH-GA models for V_s prediction;
- incorporation of sequential upper-layer velocity (V_{su}) information to capture stratigraphic continuity;
- establishment of soil-specific predictive equations using a large regional database; and
- comprehensive comparison with widely used empirical correlations.

2. General setting and geological characteristics of the site

Mashhad, the capital of Razavi Khorasan province and the second most populous city in Iran, covers an area exceeding 300 km² in northeastern Iran. The site is geographically located between latitudes 36.10°-36.24° N and longitudes 59.25°-59.43° E (Fig. 1). From a tectonic perspective, the region is situated at the boundary between the Central Iran and Koppe Dagh seismotectonic provinces. The Koppe Dagh structural zone, including the Hezar Masjed mountains, extends from the Caspian sea toward Afghanistan along a northwest-southeast trend. The Mashhad-Ghuchan depression represents a narrow compressional basin separating the folded-thrusted structures of the Koppe Dagh from the Binalood mountains. The northeastern boundary of the basin is controlled by major thrust faults, including the Kashafrood Fault, whereas the southwestern boundary is associated with the South Mashhad and Chenaran faults. The city is predominantly underlain by young alluvial deposits of the Mashhad Plain.

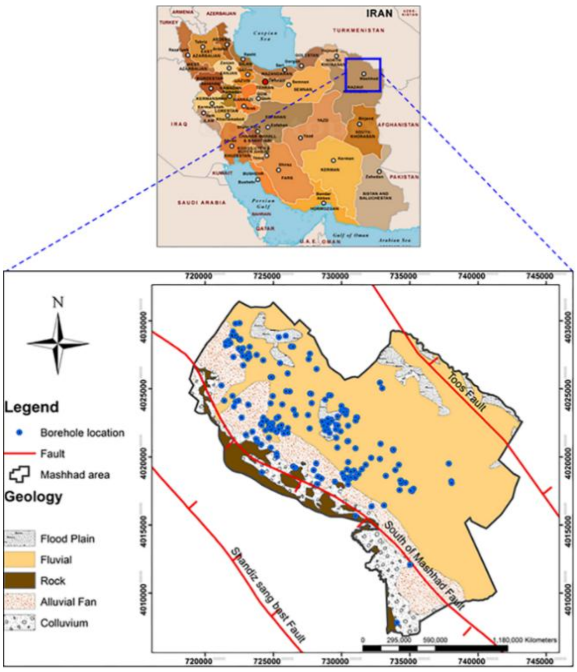


Fig. 1. Geological map of study area with location of boreholes and faults.

Based on regional seismotectonic studies and historical earthquake records, the study area is characterized by high seismic activity. Although the region is currently experiencing a relatively quiescent seismic period, historical evidence indicates the occurrence of numerous destructive earthquakes. According to the Iranian code of practice for seismic resistant design of buildings (Standard No. 2800) [34], the area is classified as a high seismic hazard zone, with a peak ground acceleration (PGA) of approximately 0.30-0.35 g for a return period of 475 years.

3. Methodology

3.1. Data acquisition and preprocessing

A comprehensive geotechnical-geophysical database was developed using site investigation reports provided by several consulting engineering companies in Mashhad, Iran. The final database comprised 1,991 data records obtained from 226 boreholes distributed across 186 investigation sites within the Mashhad Plain. The subsurface deposits primarily consist of clay, silt, sand, and gravelly alluvial materials, representing a wide range of geotechnical conditions. Standard penetration tests (SPT) were conducted at regular depth intervals in accordance with ASTM standards. In addition, downhole geophysical surveys were performed at selected locations to directly measure shear wave velocity (V_s) profiles.

The compiled database included depth (Z), soil type, SPT blow count (N), moisture content, density, Atterberg limits, and measured shear wave velocity (V_s) values. Prior to model development, a rigorous quality-control procedure was implemented to ensure data reliability. Incomplete records, physically inconsistent measurements, and erroneous entries were removed, while potential outliers were identified through statistical analysis and subsequently evaluated using geotechnical engineering judgment. Such preprocessing procedures are essential for improving the robustness and reliability of machine-learning-based geotechnical prediction models [35, 36].

Based on data availability, engineering relevance, and findings reported in previous studies, depth (Z), SPT blow count (N), and shear wave velocity of the upper soil layer (V_{su}) were selected as the input variables, whereas shear wave velocity (V_s) was considered the target output parameter. These variables have been widely recognized as key factors influencing subsurface stiffness characteristics and have been extensively utilized in both empirical and data-driven V_s prediction models [12, 13].

During the preprocessing stage, all input variables were normalized to eliminate scale dependency and improve the numerical stability and convergence behavior of the hybrid GMDH-GA framework. Data normalization is particularly important in evolutionary and neural-network-based algorithms because it facilitates efficient learning and minimizes the influence of variable magnitude disparities [37, 38]. Subsequently, the dataset was partitioned into training and testing subsets to enable an unbiased assessment of model performance.

3.2. Input variable selection

The selection of input variables was guided by geotechnical engineering principles, data availability, and evidence from previous empirical and machine-learning studies on shear wave velocity (V_s) prediction. The objective was to identify predictors that are both physically meaningful and readily obtainable from routine site investigation programs while maintaining an efficient model structure [12, 39, 40].

Accordingly, depth (Z), standard penetration test blow count (N), and shear wave velocity of the upper soil layer (V_{su}) were selected as the input variables, whereas V_s was considered the target output parameter. Depth was incorporated to account for variations in effective stress and soil stiffness with increasing overburden pressure. The SPT-N value was included due to its widespread use as an indicator of soil density, strength, and stiffness, while V_{su} was considered to capture near-surface stratigraphic conditions and wave propagation characteristics that influence subsurface shear wave velocity [11, 38].

To assess the suitability of the selected predictors, preliminary statistical analyses, including correlation and multicollinearity assessments, were performed. The results confirmed that the selected variables exhibited significant relationships with V_s while maintaining acceptable levels of interdependence, thereby minimizing redundancy within the input space [41, 42].

3.3. Model development framework (GMDH)

The group method of data handling (GMDH) is a self-organizing neural network approach capable of modeling complex nonlinear relationships between input and output variables [26]. In a classical GMDH network, neurons are generated by combining different pairs of input variables through polynomial transfer functions at each layer, thereby producing new neurons for the subsequent layer. Through this iterative process, the network gradually identifies the best-performing model structure and establishes functional relationships between the input parameters and the target output [43].

The primary objective of the GMDH algorithm is to determine an approximating function $\hat{f}$ that best represents the actual functional relationship f between the input vector $X = (x_1, x_2, x_3, \dots, x_n)$ and the output variable. Accordingly, the predicted output $\hat{y}$ is estimated such that the error between the predicted and observed values is minimized [26]. For a dataset containing M observations, the relationship between the input and output variables can be expressed as Eq. 1.

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad \text{where} \quad (i = 1, 2, \dots, M) \tag{1}$$

where y_i represents the observed output corresponding to the *i-th* observation, and x_{ij} denotes the *j-th* input variable associated with that observation.

In this step, a GMDH-type NN is trained to predict the output value ($\hat{y}_i$) for the given input vector, $X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$ as shown in Eq. 2.

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad \text{where} \quad (i = 1, 2, \dots, M) \tag{2}$$

A polynomial function is defined to minimize the squared error between observed and predicted outputs, as shown in Eq. 3.

$$\sum_{i=1}^{M}\left[\hat{f}(x_{i1}, x_{i2}, \dots, x_{in}) - y_i\right]^2 \to min \tag{3}$$

In the GMDH algorithm, a discrete form of the Volterra functional series formulates the general connection between input and output variables. This function is the discrete equivalent of the Kolmogorov-Gabor polynomial [26, 44, 45]. Hence:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \cdots \tag{4}$$

A system of partial quadratic polynomials including only two input variables can be used for representing this general form of mathematical description as shown in Eq. 5.

$$\hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \tag{5}$$

The coefficients a_i in Eq. 5 are calculated using regression techniques so that the difference between the real output, y and the predicted one, $\hat{y}$ for each pair of x_i and x_j as input variables is minimized [26, 45]. In this way, for every quadratic function, $\hat{y}_i$, the coefficients a_i are calculated to optimally fit the output in the whole set of input-output data pairs, as shown in Eq. 6.

$$E = \frac{1}{M}\sum_{i=1}^{M}(y_i - \hat{y}_i)^2 \to min \tag{6}$$

Only the neurons exhibiting the lowest prediction error according to an external validation criterion are retained for the subsequent layer, while the remaining neurons are discarded [42]. This self-organizing mechanism enables automatic determination of the best-performing network architecture and prevents excessive model complexity. In the fundamental form of the GMDH algorithm, all possible combinations of two input variables are considered to construct a regression polynomial that best fits the observed output data ($y_i, i = 1, 2, \dots, M$) based on the least-squares criterion [26]. The multiple regression analysis together with the least-squares technique is then employed to determine the optimal vector of coefficients for the polynomial model [26, 43, 45]. This can be expressed in the form of Eq. 5 as shown in Eq. 7.

$$a = (A^T A)^{-1} A^T Y \tag{7}$$

where A is the design matrix, and Y is the vector of observed outputs.

In the present study, depth (Z), standard penetration test blow count (N), and shear wave velocity of the upper soil layer (V_{su}) were used as input variables, whereas shear wave velocity (V_s) was considered the target output parameter. Compared with conventional ANN and ensemble learning methods such as random forest and XGBoost, GMDH provides explicit polynomial representations while automatically determining the best-performing network architecture. This capability enhances model interpretability, reduces the risk of overfitting, and improves predictive reliability, making GMDH particularly suitable for heterogeneous geotechnical datasets and nonlinear soil behavior modeling [42, 46].

3.4. Genetic algorithm (GA) optimization procedure

To enhance the predictive performance of the GMDH model and improve its generalization capability, a genetic algorithm (GA) was employed as a global optimization tool. GA is an evolutionary search technique that has been widely applied for optimizing machine-learning models in complex nonlinear engineering problems due to its ability to efficiently explore large solution spaces and avoid local optima [47]. Within the proposed hybrid framework, each chromosome represents a candidate configuration of the GMDH model, including parameters related to network architecture and neuron selection. An initial population of candidate solutions was randomly generated to ensure sufficient diversity and facilitate effective exploration of the search space. The fitness of each chromosome was evaluated based on the root mean square error (RMSE) between observed and predicted shear wave velocity (V_s) values, with lower RMSE values indicating superior model performance [42].

The optimization process was performed using the basic genetic operators of selection, crossover, and mutation. In each generation, candidate solutions with better fitness values were more likely to be selected for reproduction, while crossover and mutation were applied to generate new solutions and maintain diversity within the population. The process was repeated until the stopping criteria were satisfied, such as reaching the maximum number of generations or observing no significant improvement in the fitness function. The best solution obtained by the genetic algorithm (GA) was then incorporated into the GMDH framework to develop the final hybrid GMDH-GA model. By automatically optimizing the network structure and model parameters, the GA improved the prediction capability of the model and reduced the possibility of overfitting. Similar applications of evolutionary optimization techniques have demonstrated their effectiveness in enhancing the performance of GMDH-based and neural-network-based predictive models in engineering problems [38, 46]. Similar studies have shown that evolutionary optimization algorithms can substantially improve the predictive capability of GMDH-based and neural-network models by optimizing model structure, enhancing convergence behavior, and reducing the likelihood of premature convergence and overfitting [38, 46].

3.5. Training and testing strategy (70/30 Split)

To evaluate the predictive capability of the proposed GMDH-GA model, the available dataset was partitioned into two mutually exclusive subsets using a holdout validation strategy. Specifically, 70% of the data were randomly assigned to the training set for model development, while the remaining 30% were reserved as an independent testing set for performance evaluation. This partitioning ratio is commonly adopted in machine-learning applications because it provides an effective balance between model calibration and unbiased performance assessment [38, 48]. The training subset was used to establish the GMDH network structure and optimize its parameters through the GA procedure. To prevent data leakage, no information from the testing dataset was utilized during model development or parameter tuning. The testing subset was subsequently employed to evaluate the predictive performance of the developed model on previously unseen observations, thereby providing an objective assessment of its practical applicability and generalization capability [39, 40, 47].

Prior to data partitioning, the dataset was randomly shuffled to preserve the statistical characteristics of the original database and minimize sampling bias. Furthermore, a 5-fold cross-validation procedure was incorporated within the training phase to enhance model robustness and reduce the risk of overfitting [49]. This additional validation step improves the stability and reliability of the optimization process, particularly when dealing with heterogeneous geotechnical datasets [35].

3.6. Model validation and performance evaluation metrics

The performance of the proposed GMDH-GA model was evaluated using a comprehensive validation framework designed to assess both predictive accuracy and generalization capability. To ensure an unbiased assessment, model performance was examined using an independent testing dataset that was not involved in either model training or parameter optimization. Such an approach is widely recognized as an effective means of evaluating the robustness and practical applicability of machine-learning models on previously unseen data [38, 48]. To quantitatively assess prediction performance, four widely used statistical indicators were employed: the coefficient of determination (R^2), root mean square error (RMSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), and mean squared error (MSE). These metrics provide complementary information regarding prediction accuracy, error magnitude, and model reliability.

The performance indicators were calculated as follows:

$$\mathrm{R}^2 = 1 - \left[\frac{\sum_{i=0}^{n}\left(Y_{i(Predicted)} - Y_{i(Measured)}\right)^2}{\sum_{i=1}^{n}\left(Y_{i(Actual)}\right)^2}\right] \tag{8}$$

$$\mathrm{RMSE} = \left[\frac{\sum_{i=0}^{n}\left(Y_{i(Predicted)} - Y_{i(Measured)}\right)^2}{n}\right]^{1/2} \tag{9}$$

$$\mathrm{MAD} = \frac{\sum_{i=1}^{n}\left|Y_{i(Predicted)} - Y_{i(Measured)}\right|}{n} \tag{10}$$

$$\mathrm{MAPE} = \frac{1}{n}\left[\frac{\sum_{i=1}^{n}\left|Y_{i(Predicted)} - Y_{i(Measured)}\right|}{\sum_{i=1}^{n} Y_{i(Measured)}}\right] \times 100 \tag{11}$$

$$\mathrm{MSE} = \frac{\sum_{i=0}^{n}\left(Y_{i(Predicted)} - Y_{i(Measured)}\right)^2}{n} \tag{12}$$

where n is the number of observations. The coefficient of determination (R^2) measures the proportion of variance in the observed data explained by the model, with values approaching unity indicating superior predictive capability. RMSE quantifies the overall magnitude of prediction errors and is particularly sensitive to large deviations. MAE provides a direct measure of the average absolute prediction error, while MAPE expresses prediction errors in percentage terms, enabling relative performance comparisons across different datasets and modeling approaches [50, 51].

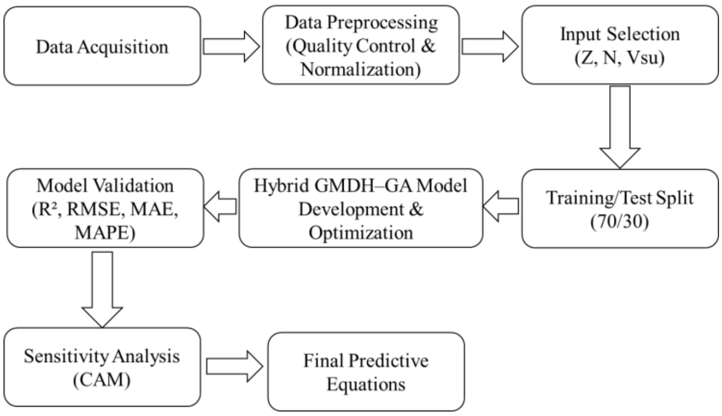


Fig. 2. Overall methodological framework of the proposed hybrid GMDH-GA model for shear wave velocity (V_s) prediction.

The overall methodological framework of the proposed GMDH-GA hybrid model for shear wave velocity (V_s) prediction, including data collection, preprocessing, input variable selection, data normalization, training/testing partitioning, GMDH model development, GA-based optimization, model validation, comparison with existing empirical correlations, and sensitivity analysis, is shown in Fig. 2.

4. Modelling shear wave velocity using GMDH neural networks

The feed-forward GMDH-type neural network for predicting shear wave velocity (V_s) was developed using the experimental dataset described in the previous sections. The input variables included depth (Z), SPT blow count (N), and shear wave velocity of the upper soil layer (V_{su}), while the target output was the corresponding shear wave velocity of the soil (V_s). To evaluate the predictive performance of the model, the dataset was randomly divided into training and testing subsets with similar statistical characteristics. For each soil group (all soils, clay, silt, sand, and gravel), 70% of the data were allocated for model training and 30% for testing and validation (Table 2) [48, 49].

Table 2. Descriptive statistics of input and output parameters used in the testing set.

Variable	Minimum	Maximum	Mean	Groups
Z	1	49	16.65	(1) All soils
N	8	86	45.19	
V_{su}	110	975	493.19	
V_s	141	990	527.68	
Z	1.50	56	20.75	(2) Clay
N	12.00	84	41.76	
V_{su}	111.30	858	513.16	
V_s	166.00	886.8	548.76	
Z	1.5	45	20.29	(3) Silt
N	6	76	34.44	
V_{su}	161	810	495.39	
V_s	187	875	521.50	
Z	1.5	41.5	15.47	(4) Sand
N	7	86	47.40	
V_{su}	110	859.9	508.47	
V_s	141	945	538.08	
Z	1	38	15.07	(5) Gravel
N	10	86	48.83	
V_{su}	134.57	947	504.32	
V_s	236.5	975	541.99	

The GMDH-type neural networks were employed to construct explicit polynomial relationships between input variables and V_s. In this study, a genetic algorithm (GA) was integrated into the GMDH framework to optimize network architecture and model parameters. The GA controls key evolutionary parameters, including population size, number of generations, crossover probability, and mutation probability. After preliminary sensitivity analysis, a population size of 100, crossover probability of 0.9, and mutation probability of 0.1 were adopted over 300 generations, as no significant improvement was observed beyond this configuration, which is consistent with commonly reported GA settings in engineering optimization studies [47].

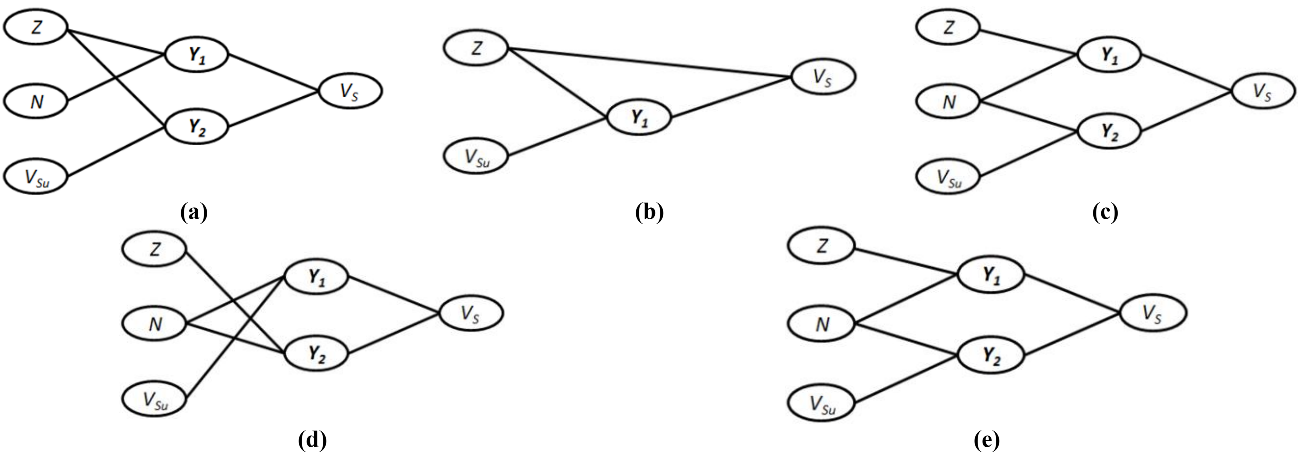


Fig. 3. Optimized single-hidden-layer GMDH network structures for the V_s: (a) all soils; (b) clay; (c) silt; (d) sand; and (e) gravel.

The hybrid GMDH-GA framework was applied separately to five soil groups in order to capture soil-specific behavior. Although multiple network architectures with different numbers of hidden layers were tested, single-hidden-layer structures were selected as they provided an optimal balance between predictive accuracy and model simplicity. The final optimized network structures for each soil group are illustrated in Fig. 3. Although most models retained all input variables during construction, the GMDH algorithm inherently performs automated feature selection by retaining only the most influential terms during polynomial formation. This capability allows the generation of explicit mathematical expressions while maintaining high predictive performance [26, 46].

The resulting polynomial equations for shear wave velocity prediction are presented in Eqs. 13 to 17. Similar explicit polynomial formulations have been successfully reported in previous GMDH-based geotechnical prediction studies. For the case of all soil types, intermediate variables Y_1 and Y_2 are defined as Eqs. 13a and 13b.

for all soils:

$$Y_1 = 203.0504 + 20.8002Z + 3.6121N - 0.2484Z^2 - 0.02666N^2 - 0.02974ZN \quad (13a)$$

$$Y_2 = 124.1691 + 1.6931Z + 0.8119V_{su} - 0.05386Z^2 - 0.000089{V_{su}}^2 + 0.007775ZV_{su} \quad (13b)$$

$$V_s = -88.1546 + 0.5184Y_1 + 0.8366Y_2 - 0.000642{Y_1}^2 - 0.0000214{Y_2}^2 + 0.000322Y_1Y_2 \quad (13c)$$

Similar formulations were obtained for clayey, silty, sandy, and gravelly soils, as shown in Eqs. 14 to 17. In all cases, depth (Z) and shear wave velocity of the upper layer (V_{su}) are expressed in meters and m/s, respectively. The coefficients of the polynomial equations were automatically determined through the combined GMDH-GA optimization process.

For clayey soils:

$$Y_1 = 101.2230 - 3.1997Z + 0.9595V_{su} - 0.077789Z^2 - 0.0003378{V_{su}}^2 + 0.0130567ZV_{su} \quad (14a)$$

$$V_s = 8.1441 + 0.8703Z + 0.9329Y_1 + 0.01513Z^2 + 0.000119{Y_1}^2 - 0.002897ZY_1 \quad (14b)$$

For silty soils:

$$Y_1 = 219.3582 + 19.3295Z + 1.4077N - 0.223459Z^2 - 0.021702N^2 + 0.033419ZN \quad (15a)$$

$$Y_2 = 101.36088 - 0.64436N + 0.824197V_{su} + 0.0065467N^2 + 0.0000761{V_{su}}^2 + 0.0000631NV_{su} \quad (15b)$$

$$V_s = 48.47287 - 0.27349Y_1 + 1.071459Y_2 - 0.0000827{Y_1}^2 - 0.000437{Y_2}^2 + 0.000735Y_1Y_2 \quad (15c)$$

For sandy soils:

$$Y_1 = 127.4764 + 0.400969N + 0.68966V_{su} - 0.0019389N^2 + 0.0001691{V_{su}}^2 - 0.0001319NV_{su} \quad (16a)$$

$$Y_2 = 195.4441 + 23.76196Z + 3.32919N - 0.379395Z^2 - 0.0270525N^2 - 0.007853967ZN \quad (16b)$$

$$V_s = -127.93006 + 0.55972Y_1 + 0.961276Y_2 - 0.000461927{Y_1}^2 - 0.00176247{Y_2}^2 + 0.00172848Y_1Y_2 \quad (16c)$$

For gravely soils:

$$Y_1 = 244.4176 + 24.5336Z + 1.93860N - 0.312301Z^2 - 0.0136057N^2 - 0.038851ZN \quad (17a)$$

$$Y_2 = 226.2265 - 1.516388N + 0.51278V_{su} + 0.00534585N^2 + 0.000228235{V_{su}}^2 + 0.00197189NV_{su} \quad (17b)$$

$$V_s = -160.70616 + 0.877126Y_1 + 0.688160Y_2 - 0.000411841{Y_1}^2 + 0.000452851{Y_2}^2 - 0.000531079Y_1Y_2 \quad (17c)$$

To evaluate model performance, predicted and measured values of V_s were compared for both training and testing datasets. Scatter plots (Figs. 4 and 5) indicate that the predicted values are closely distributed around the 1:1 reference line, confirming strong agreement between measured and modeled data.

The predictive performance of the developed models was quantitatively evaluated using standard statistical metrics, including the coefficient of determination (R^2), root mean square error (RMSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), and mean squared error (MSE), defined in Eqs. 8 to 12. The corresponding performance results for all soil groups are summarized in Table 3.

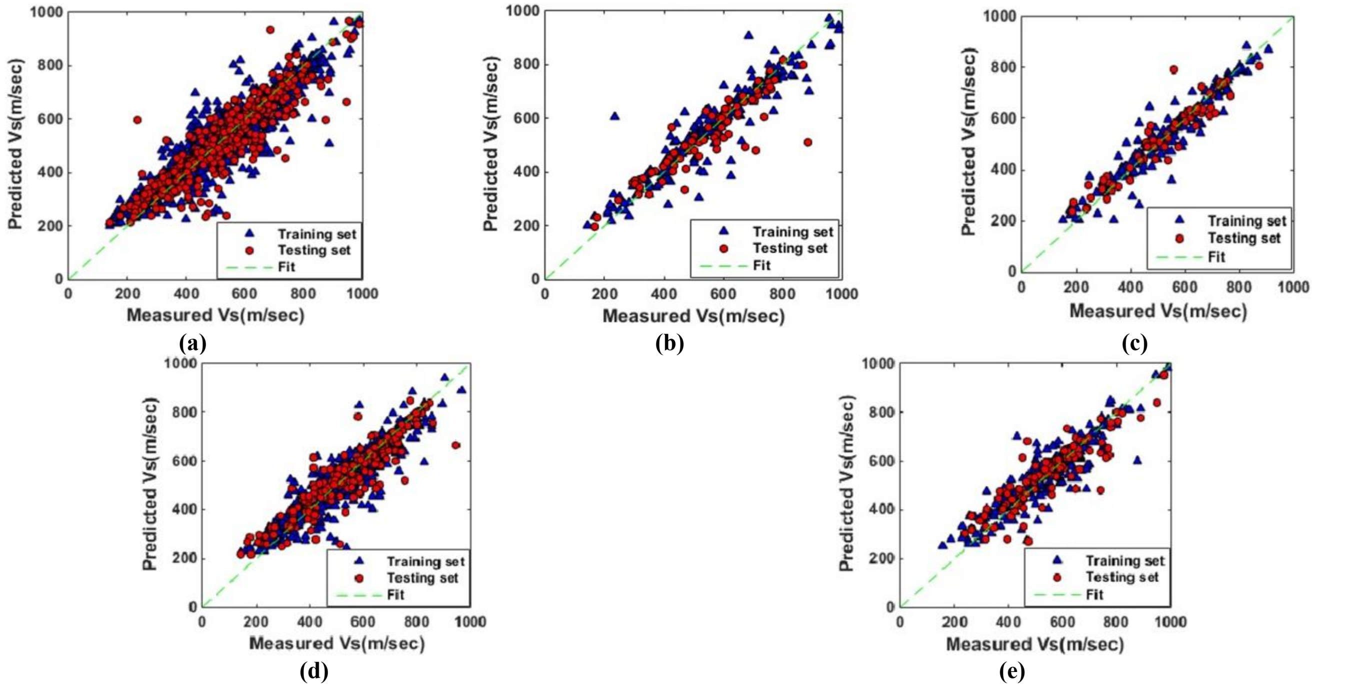


Fig. 4. The relationships between measured and predicted V_s by GMDH through the training and testing process: (a) all soils; (b) clay; (c) silt; (d) sand; and (e) gravel.

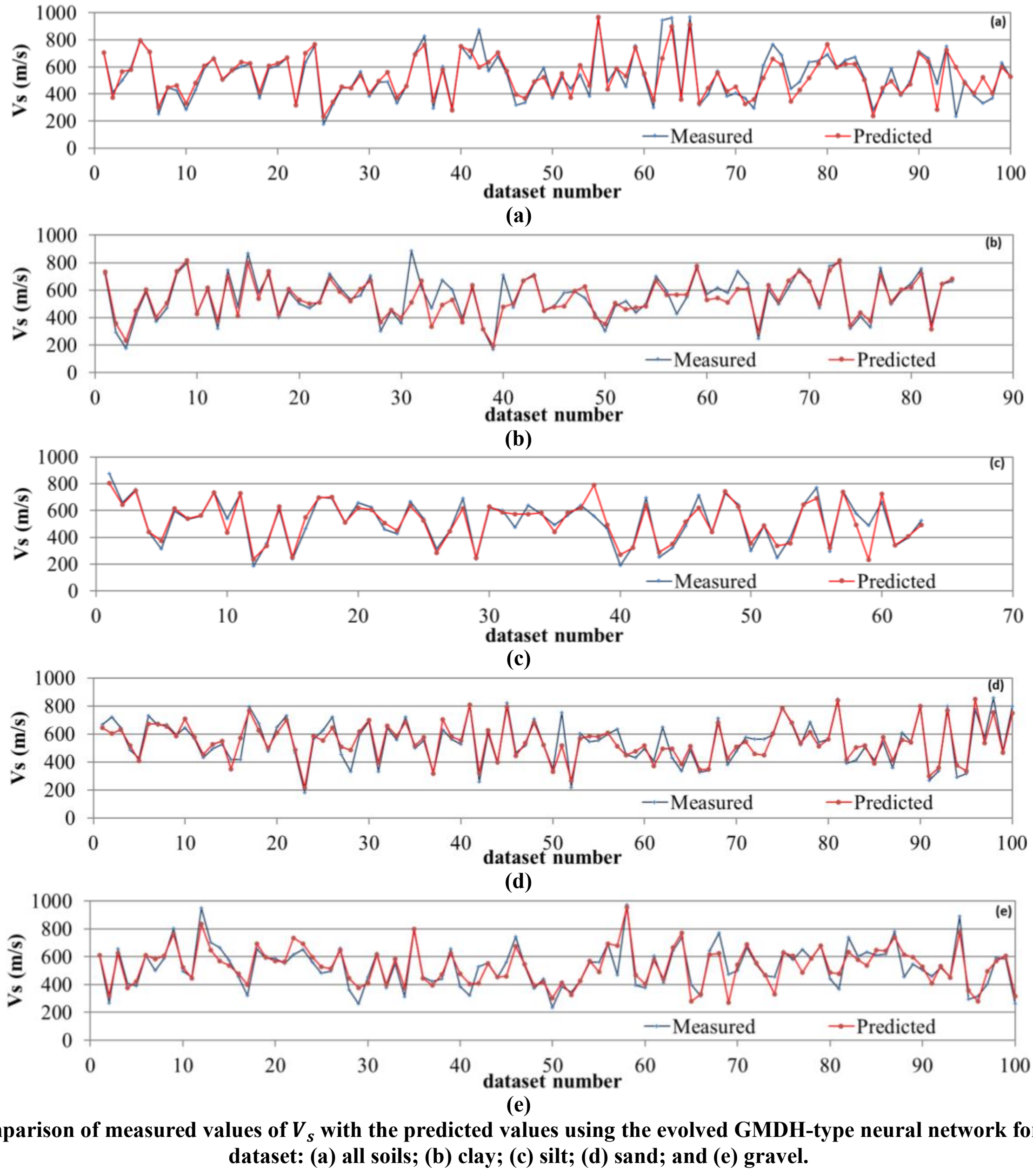


Fig. 5. Comparison of measured values of V_s with the predicted values using the evolved GMDH-type neural network for the testing dataset: (a) all soils; (b) clay; (c) silt; (d) sand; and (e) gravel.

Table 3. The amount of performance indexes for the GMDH models.

Groups	Set	Number of data	MAD (m/s)	MAPE (%)	MSE (m/s)^2	RMSE (m/s)	R^2	Correlation coefficient
All soils	Testing set	597	44.5	8.43	4275.91	65.39	0.99	0.91
	Training set	1394	44.22	8.3	3862.85	62.15	0.99	0.92
Clay	Testing set	125	44.15	8.05	4748.68	68.91	0.99	0.90
	Training set	293	42.95	7.69	4276.98	65.4	0.99	0.93
Silt	Testing set	51	39.22	7.52	3781.47	61.49	0.99	0.93
	Training set	118	39.79	7.69	2909.26	53.94	0.99	0.95
Sand	Testing set	281	44.34	8.24	4003.52	63.27	0.99	0.91
	Training set	657	43.15	8.39	3643.2	60.36	0.99	0.91
Gravel	Testing set	140	49.89	9.21	4794.01	69.24	0.98	0.89
	Training set	326	44.95	8.32	3947.26	62.83	0.99	0.90

These performance indicators are widely employed for evaluating machine-learning models in geotechnical engineering and provide complementary information regarding prediction accuracy, error magnitude, and model robustness.

The obtained results demonstrate high predictive accuracy and low error levels for both training and testing datasets across all soil categories. The consistency between training and testing performance indicates that the proposed GMDH-GA models exhibit strong generalization capability and indicate no significant evidence of overfitting. Moreover, the ability of the model to generate explicit polynomial equations enhances its practical applicability in geotechnical engineering design. Overall, the results confirm that the hybrid GMDH-GA framework provides a reliable, accurate, and interpretable tool for predicting shear wave velocity profiles in heterogeneous alluvial deposits.

A comparison between measured and predicted V_S values revealed that the developed GMDH-GA models successfully captured the nonlinear relationships among depth, SPT-*N* value, and upper-layer shear wave velocity across different soil categories. The resulting polynomial expressions provide transparent predictive formulations that can be directly implemented in engineering practice without requiring retraining procedures or specialized computational platforms.

The accuracy and predictive performance of the proposed GMDH model for group 1 (all soils) were statistically evaluated and compared with previously published empirical correlations listed in Table 1. The analysis was carried out using the full dataset of 1,991 samples, and the results presented in Table 4 clearly indicate that the GMDH-based approach outperforms all existing empirical correlations. Similar improvements achieved by hybrid machine-learning models over conventional empirical relationships have been reported in recent geotechnical studies. Accordingly, the developed model demonstrates higher predictive accuracy and enhanced reliability in estimating shear wave velocity (V_S).

Table 4. Performance indexes for GMDH model compared with other empirical correlations.

Equation of this author for all soils	No. of data	MAD (m/s)	MAPE (%)	MSE (m/s)^2	RMSE (m/s)	R^2	Correlation coefficient
Ohsaki and Iwasaki [6]	1991	194.25	36.57	54930.87	234.37	0.82	0.28
Imai [7]	1991	218.24	41.08	66361.26	257.61	0.78	0.28
Ohta and Goto [1]	1991	224.21	42.21	69410.08	263.46	0.77	0.28
Seed and Idriss [8]	1991	161.96	30.49	41250.88	203.10	0.87	0.28
Lee Shannon [9]	1991	344.42	64.84	141689.82	376.42	0.54	0.28
Iyisan [10]	1991	188.93	35.57	53192.95	230.64	0.83	0.28
Jafari et al. [5]	1991	171.82	32.35	44857.26	211.80	0.85	0.27
Hasancebi and Ulusay [11]	1991	249.30	46.93	82512.25	287.25	0.73	0.28
Dikmen [12]	1991	282.22	53.13	101196.46	318.11	0.67	0.28
Maheswari et al. [13]	1991	241.33	45.43	78236.13	279.71	0.74	0.28
Hafezi et al. [14]	1991	231.26	43.54	75992.59	275.67	0.75	0.28
Chatterjee and Choudhury [15]	1991	215.07	40.49	64911.56	254.78	0.79	0.28
Fatehnia et al. [16]	1991	244.19	45.97	79786.08	282.46	0.74	0.28
This study (GMDH model)	1991	186.21	35.06	51013.35	225.86	0.83	0.28
This study (GMDH model)	1991	44.30	8.34	3986.74	63.14	0.99	0.92

The scatter plots presented in Fig. 6 further illustrate the comparison between measured and predicted V_S values. It is evident that most of the existing empirical correlations systematically underestimate V_S and exhibit considerable data dispersion around the 1:1 reference line. Among the evaluated relationships, the correlation proposed by Jafari et al. [5] exhibits comparatively better performance; however, noticeable scatter remains, indicating limitations in its predictive capability. This observation is consistent with previous findings suggesting that empirical V_S-SPT relationships often struggle to capture the inherent heterogeneity and nonlinear behavior of soil deposits.

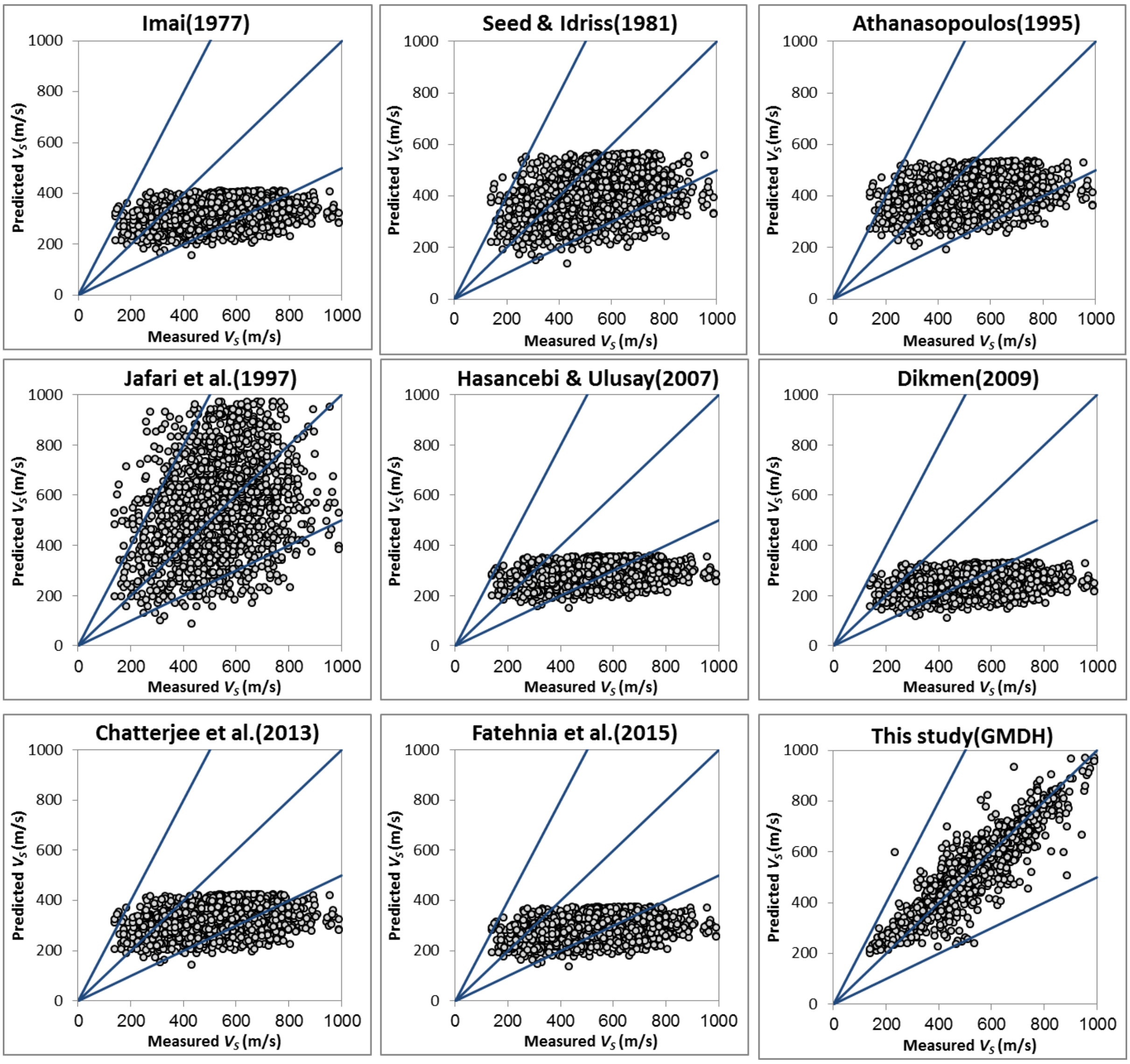


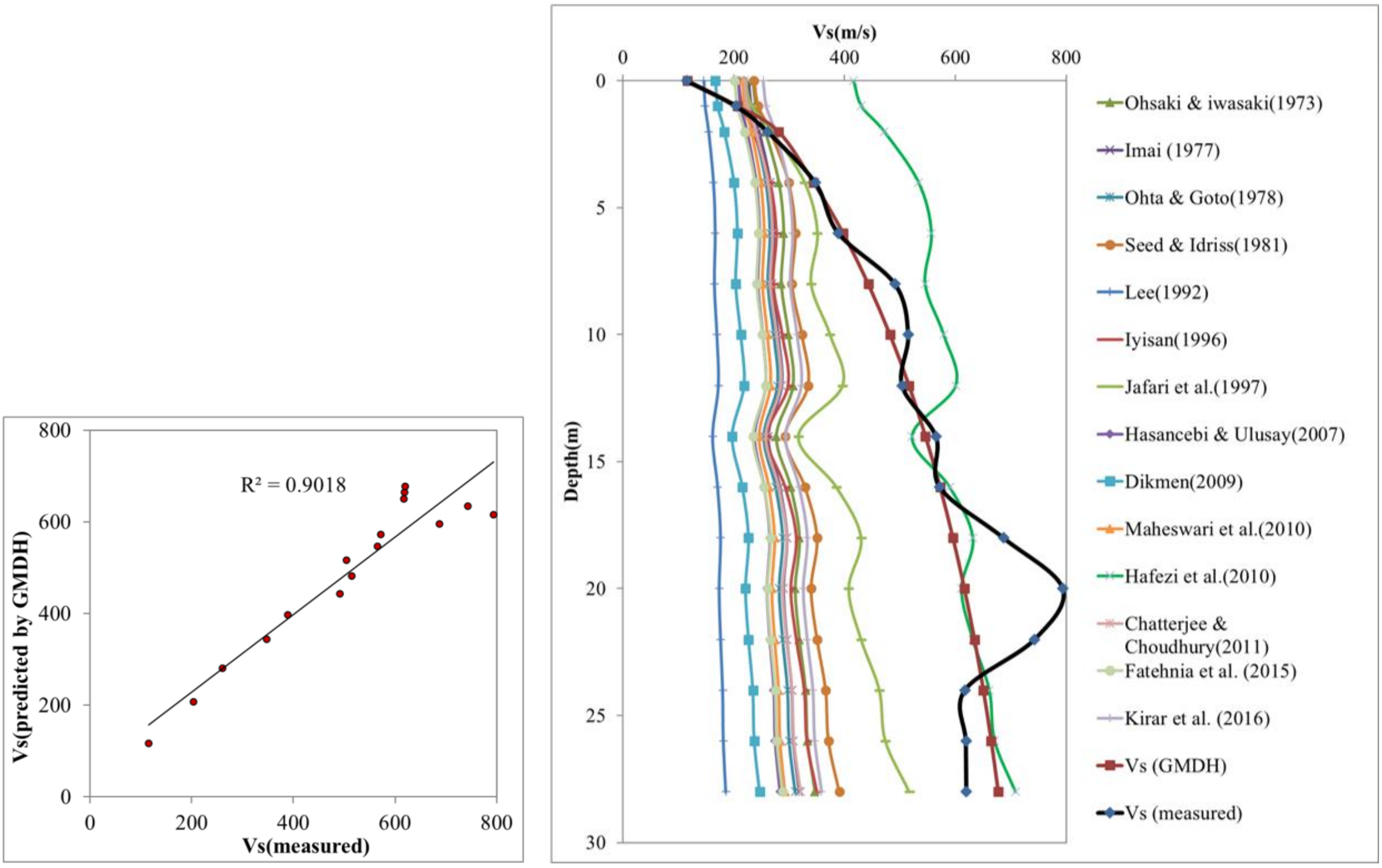
Fig. 6. Comparison of measured values of V_s with predicted values using different methods.

To further evaluate the capability of the proposed model in reproducing shear wave velocity variations with depth, a representative 30-m-deep borehole located in central Mashhad was analyzed using Eqs. 13a to 13c (Table 5). The coefficient of determination shown in Fig. 7a confirms that the GMDH model provides a close approximation to the measured V_s profile. In the stepwise implementation procedure, the assumption of $V_s = 0$ at the ground surface was adopted to initiate the calculations, and the predicted upper-layer velocity (V_{su}) was subsequently utilized as an input for estimating deeper soil layers.

Table 5. A sample of bore log profile and calculating procedure using Eqs. 13a to 13c.

Borehole ID	X	Y	Soil type	Depth	SPT	V_{su}	V_s (measured)	Y_1	Y_2	V_s (GMDH)
I6-3-1	726055	4026141	SM	0.5	15	0	115	261.4	123.3	116.69
I6-3-1	726055	4026141	SP-SM	1.5	16	116.7	204	284	216.4	207.10
I6-3-1	726055	4026141	SW-SM	3	19	207.1	261	320.5	287.8	280.72
I6-3-1	726055	4026141	SP	5	24	280.7	347	368.6	346.2	343.83
I6-3-1	726055	4026141	GP	7	26	343.8	389	407	397	397.29
I6-3-1	726055	4026141	GW	9	25	397.3	491	437.1	440.9	442.51
I6-3-1	726055	4026141	SP-SM	11	28	442.5	515	472.9	478.7	481.91
I6-3-1	726055	4026141	SC	13	30	481.9	504	504.2	512.4	516.22
I6-3-1	726055	4026141	SW-SM	15	23	516.2	565	517.9	542.3	545.91
I6-3-1	726055	4026141	SW-SM	17	29	545.9	572	552.5	568.7	572.29
I6-3-1	726055	4026141	SP-SM	19	33	572.3	687	580.1	592.6	595.46
I6-3-1	726055	4026141	SP-SM	21	31	595.5	794	597.3	614	616.11
I6-3-1	726055	4026141	GP-GM	23	33	616.1	743	617.6	633.3	634.35
I6-3-1	726055	4026141	GP	25	36	634.3	617	636.5	650.7	650.37
I6-3-1	726055	4026141	SM	27	37	650.4	619	651	666.1	664.64

Figure 7-b compares the measured downhole V_s profile with predictions obtained from the proposed GMDH model and previously developed empirical correlations. The results demonstrate that none of the existing relationships accurately reproduces the complete velocity profile throughout the investigated depth range. Although the model proposed by Hafezi et al. [14] exhibits partial agreement at greater depths, it tends to overestimate V_s values in the upper layers. In contrast, the proposed GMDH-based model consistently follows the measured trend and provides a more reliable representation of the full shear wave velocity profile. The improved performance can be attributed to the capability of GMDH to capture nonlinear interactions among depth, SPT-N values, and upper-layer shear wave velocity while simultaneously controlling model complexity through its self-organizing architecture.



(a) **(b)**

Fig. 7. (a) The relationship between the measured and calculated V_s using Eq. 8 for borehole I6-3-1; and (b) the measured and predicted V_s profile for borehole I6-3-1.

5. Sensitivity analysis

Since the polynomial equations generated by the evolved GMDH-type neural network are highly nonlinear and do not explicitly quantify the relative contribution of each input variable to the predicted shear wave velocity (V_s), a sensitivity analysis was performed to evaluate the influence of predictor variables on model output. Sensitivity analysis is widely used in machine learning-based geotechnical and geophysical studies to improve model interpretability and quantify the relative importance of input parameters [52, 53]. In this study, the cosine amplitude method (CAM) was employed to determine the relative strength of association between input variables and the output variable. The CAM is a simple yet effective statistical technique commonly used in artificial intelligence and data-driven modeling because of its ability to quantify variable dependency in multidimensional datasets [27, 54]. The relationship coefficient is defined as Eq. 18.

$$R_{ij} = \frac{\sum_{k=1}^{n}(x_{ik} \times x_{jk})}{\sqrt{\sum_{k=1}^{n} x_{ik}^2 \sum_{k=1}^{n} x_{jk}^2}} \tag{18}$$

where x_i and x_j represent the input and output variables, respectively, and n is the number of data samples. The coefficient R_{ij} ranges between 0 and 1, where values closer to zero indicate weak influence, while values approaching unity indicate strong correlation with the predicted shear wave velocity (V_s).

The sensitivity analysis results for the developed one-hidden-layer GMDH models are presented in Fig. 8. The results indicate that the shear wave velocity of the upper soil layer (V_{su}) is the most influential parameter in predicting V_s. Furthermore, the SPT blow count (N) exhibits higher sensitivity in coarse-grained soils, whereas its influence decreases in fine-grained deposits. In contrast, depth (Z) shows relatively lower sensitivity compared to the other input variables. Similar findings regarding the dominant role of near-surface shear wave velocity and SPT-N in V_s prediction have been reported in previous studies, confirming the physical consistency of the proposed model [12].

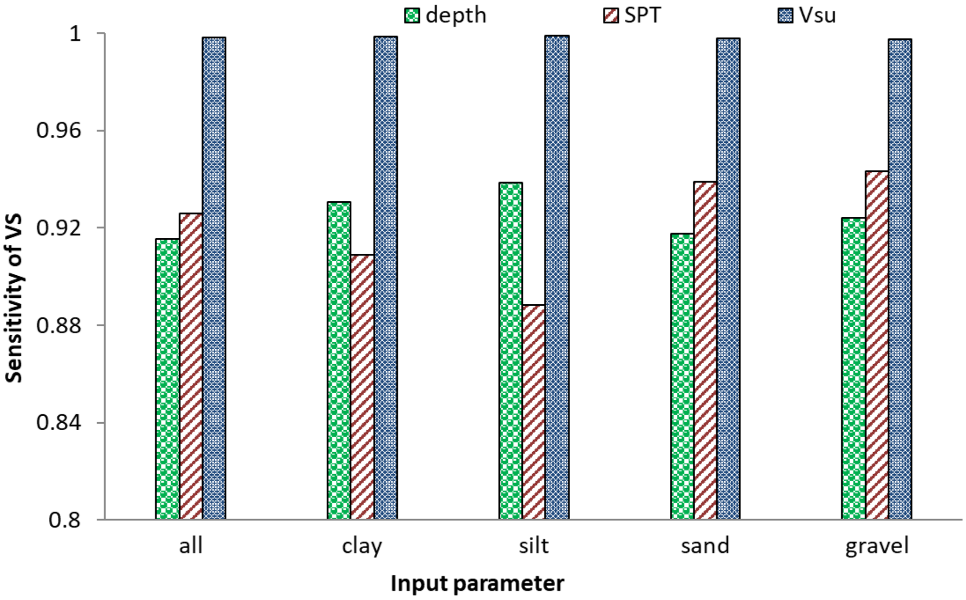


Fig. 8. Contributions of the predictor variables in the best-selected GMDH models for five soil groups under study.

6. Results and discussion

High-performance predictive models were developed for estimating shear wave velocity (V_s) using a hybrid GMDH-genetic algorithm (GMDH-GA) framework. The main objective was to derive explicit and reliable formulations for estimating the shear wave velocity of the upper 30 m soil profile (V_{s30}) in the absence of in-situ geophysical measurements. The V_{s30} parameter is widely recognized as a key index in seismic site classification and ground response analysis. Accordingly, V_s was modeled as a function of depth (Z), SPT blow count (N), and shear wave velocity of the upper soil layer (V_{su}), which were identified as dominant predictors in the sensitivity analysis.

The dataset was compiled from geotechnical and geophysical investigations conducted in Mashhad, comprising 226 boreholes and 1,991 data records representing clay, silt, sand, and gravel deposits. Separate GMDH-based polynomial models were developed for each soil group as well as for the combined dataset. For each case, 70% of the data were used for training and 30% for testing. Model performance was evaluated against previously published empirical correlations, demonstrating that the proposed GMDH-GA approach consistently provides higher accuracy and improved predictive capability across all soil types. Although deeper network structures with additional hidden layers showed marginal improvements in accuracy, single-hidden-layer models were selected as the final formulations due to their superior generalization capability and higher interpretability, which is a key requirement in engineering design applications.

Sensitivity analysis confirmed that V_{su} is the most influential predictor, and its exclusion significantly reduces model accuracy. This finding is consistent with previous studies highlighting the dominant role of near-surface shear wave velocity in soil stiffness characterization and wave propagation behavior. Additionally, SPT blow count exhibits a stronger influence in coarse-grained soils, whereas its effect decreases in fine-grained materials due to differences in fabric and compressibility characteristics. Depth (Z) shows a relatively lower but consistent contribution across all soil groups.

The inclusion of V_{su} is physically justified by the layered continuity of alluvial deposits and the progressive increase in soil stiffness with depth. In sedimentary environments, mechanical properties of adjacent layers are strongly interdependent due to stress history and depositional processes. Therefore, incorporating V_{su} captures the sequential nature of shear wave propagation in stratified media rather than introducing redundant information. However, this variable may limit direct applicability in sites lacking preliminary V_s measurements; hence, the proposed framework is particularly suitable for profile reconstruction, interpolation, and data-enhanced geotechnical site characterization.

Overall, the results demonstrate that the proposed GMDH-GA models provide reliable and robust predictions of shear wave velocity profiles and can serve as a complementary or alternative tool to conventional in-situ testing methods such as downhole and crosshole measurements. Since V_{su} profiles play a critical role in seismic microzonation, site response analysis, liquefaction potential assessment, and seismic site classification, the proposed models offer an efficient and practical tool for geotechnical earthquake engineering applications. Comparison with existing empirical correlations further confirms the superiority of region-specific modeling approaches, as global empirical relationships show significant deviations from measured data, whereas the proposed hybrid V_{su} framework exhibits strong agreement with downhole observations.

7. Conclusions

This study developed a hybrid GMDH-genetic algorithm (GMDH-GA) framework for predicting shear wave velocity (V_s) profiles in heterogeneous alluvial deposits using a comprehensive geotechnical database from Mashhad, Iran. The proposed models were constructed using depth (Z), standard penetration test blow count (SPT-N), and upper-layer shear wave velocity (V_{su}) as input parameters and were evaluated across different soil groups, including clay, silt, sand, and gravel.

The results demonstrated that the proposed GMDH-GA models achieve high predictive accuracy and consistently outperform conventional empirical correlations in terms of error reduction and generalization capability. These findings are consistent with recent studies highlighting the superiority of hybrid machine learning approaches over traditional empirical relationships in geotechnical parameter estimation. Sensitivity analysis indicated that V_{su} is the most influential predictor in V_s estimation, followed by SPT blow count, particularly in coarse-grained soils. In contrast, depth exhibited a comparatively lower but stable contribution across all soil types, reflecting the gradual variation of stress-dependent soil stiffness with depth.

Unlike conventional black-box machine learning models, the proposed GMDH-GA framework provides explicit polynomial expressions, thereby enhancing model transparency, interpretability, and practical usability in geotechnical engineering design. Overall, the developed models offer a reliable and efficient tool for seismic site characterization, seismic microzonation, ground response analysis, and preliminary site classification in regions where direct geophysical measurements are limited. The proposed approach bridges the gap between empirical correlations and advanced artificial intelligence techniques by combining predictive accuracy with physical interpretability.

8. Limitations and future work

Although the proposed GMDH-GA framework demonstrates strong predictive performance, several limitations should be acknowledged. First, the developed models are based on data from a specific geological setting (Mashhad alluvial deposits), which may limit their direct applicability to regions with different geological and stratigraphic conditions. Therefore, external validation in other geological environments is necessary before broader application. Second, the inclusion of V_{su} improves predictive accuracy but may limit the applicability of the model in sites where prior shear wave velocity information is not available. This reflects a trade-off between model accuracy and field data dependency. Third, as with most data-driven approaches, model performance is influenced by the quality and uncertainty of input data, particularly SPT measurements and in-situ geophysical observations.

Furthermore, future research should focus on validating the proposed framework across diverse geological conditions and incorporating additional geotechnical and geophysical parameters such as CPT data and soil index properties. A direct comparative analysis with modern machine-learning algorithms such as ANN, random forest, and XGBoost was beyond the scope of the present study and should be addressed in future investigations.

Statements & declarations

Author contributions

The author contributed to all stages of the research, including conceptualization, methodology, analysis, interpretation of results, writing the original draft, and revising and editing the manuscript. The author has read and approved the final version of the manuscript.

Acknowledgments

The author expresses sincere appreciation to all individuals who provided scientific, technical, or logistical assistance during the course of this research.

Funding

This research received no financial support from any governmental, commercial, or non-profit funding agency.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

The author declares no conflict of interest.

References

[1] Ohta, Y., Goto, N. Empirical shear wave velocity equations in terms of characteristic soil indexes. Earthquake Engineering & Structural Dynamics, 1978; 6: 167–187. doi:10.1002/eqe.4290060205.

[2] Imai, T., Tonouchi, K. Correlation of N-value with S-wave velocity and shear modulus. In: Proceedings of the 2nd European Symposium on Penetration Testing (ESOPT II); 1982 May 24–27; Amsterdam, Netherlands. p. 67–72.

[3] Seed, H. B., Idriss, I. M., Arango, I. Evaluation of Liquefaction Potential Using Field Performance Data. Journal of Geotechnical Engineering, 1983; 109: 458–482. doi:10.1061/(ASCE)0733-9410(1983)109:3(458).

[4] Sykora, D. E., H., S. K. Correlations of in-situ measurements in sands of shear wave velocity with standard penetration test resistance. Geotechnical Testing Journal, 1983; 6: 107–114. doi:10.1520/GTJ10700J.

[5] Jafari, M. K., Asghari, A., Rahmani, I. Empirical correlation between shear wave velocity (Vs) and SPT-N value for south of Tehran soils. In: Proceedings of 4th International Conference on Civil Engineering; 1997 May 04–06; Tehran, Iran.

[6] Ohsaki, Y., Iwasaki, R. On Dynamic Shear Moduli and Poisson's Ratio of Soil Deposits. Soils and Foundations. Soils and Foundations, 1973; 13: 61–73. doi:10.3208/sandf1972.13.4_61.

[7] Imai, T. P- and S-Wave Velocities of the Ground in Japan. In: 9th International Conference on Soil Mechanics and Foundation Engineering; 1977 Jul 10–15; Tokyo, Japan. p. 257–260.

[8] Seed, H. B., Idriss, I. M. Evaluation of Liquefaction Potential Sand Deposits Based on Observation of Performance in Previous Earthquakes. ASCE National Convention (MO), 481-544.

[9] Lee Shannon, H.-H. Analysis of the multicollinearity of regression equations of shear wave velocities. Soils and Foundations, 1992; 32: 205–214. doi:10.3208/sandf1972.32.205.

[10] Iyisan, R. Correlations between shear wave velocity and in-situ penetration test results. Teknik Dergi-tmmob Insaat Muhendisleri Odas, 1996; 7: 371–374.

[11] Hasancebi, N., Ulusay, R. Empirical correlations between shear wave velocity and penetration resistance for ground shaking assessments. Bulletin of Engineering Geology and the Environment, 2007; 66: 203–213. doi:10.1007/s10064-006-0063-0.

[12] Dikmen, Ü. Statistical correlations of shear wave velocity and penetration resistance for soils. Journal of Geophysics and Engineering, 2009; 6: 61. doi:10.1088/1742-2132/6/1/007.

[13] Maheswari, R. U., Boominathan, A., Dodagoudar, G. R. Use of Surface Waves in Statistical Correlations of Shear Wave Velocity and Penetration Resistance of Chennai Soils. Geotechnical and Geological Engineering, 2010; 28: 119–137. doi:10.1007/s10706-009-9285-9.

[14] Hafezi, M. N., Azadi, A., Amanian, M. Assessment of shear wave velocity based on standard penetration and accuracy of the results in the range of Mashhad City. In: Fourteenth Conference of Iran Geophysics, Geophysical Institute (in Persian); 2010 May 13; Tehran, Iran.

[15] Chatterjee, K., Choudhury, D. Variations in shear wave velocity and soil site class in Kolkata city using regression and sensitivity analysis. Natural Hazards, 2013; 69: 2057–2082. doi:10.1007/s11069-013-0795-7.

[16] Fatehnia, M., Hayden, M., Landschoot, M. Correlation between Shear Wave Velocity and SPT-N Values for North Florida Soils. Electronic Journal of Geotechnical Engineering, 2015; 20: 12421–12430.

[17] Kirar, B., Maheshwari, B. K., Muley, P. Correlation Between Shear Wave Velocity (Vs) and SPT Resistance (N) for Roorkee Region. International Journal of Geosynthetics and Ground Engineering, 2016; 2: 9. doi:10.1007/s40891-016-0047-5.

[18] Sarkar, A., Mazumdar, A., Bhowmik, D. Application of Soft Computing Techniques for Prediction of Seismic Wave Velocity. In: 9th International Conference on Control, Decision and Information Technologies (CoDIT); 2023 Jul 03–06; Rome, Italy. p. 2532–2537. doi:10.1109/CoDIT58514.2023.10284120.

[19] Mousavi, Z., Bayat, M., Yang, J., Feng, W.-Q. Evaluation of empirical and machine learning models for predicting shear wave velocity of granular soils based on laboratory element tests. Soil Dynamics and Earthquake Engineering, 2024; 183: 108805. doi:10.1016/j.soildyn.2024.108805.

[20] Jin, L., Fuggle, A., Roberts, H., Armstrong, C., Pua, L.-M. Estimating In Situ Shear Wave Velocity Using Machine Learning Techniques. In: Geo-Congress 2024; 2024 Feb 25–28; Vancouver, Canada. p. 436–444. doi:10.1061/9780784485347.044.

[21] Joshi, A., Raman, B., Mohan, C. K., Cenkeramaddi, L. R. A new machine learning approach for estimating shear wave velocity profile using borelog data. Soil Dynamics and Earthquake Engineering, 2024; 177: 108424. doi:10.1016/j.soildyn.2023.108424.

[22] Zhong, K., Tan, X., Liu, S., Lu, Z., Hou, X., Wang, Q. Prediction of slope failure probability based on machine learning with genetic-ADASYN algorithm. Engineering Geology, 2025; 346: 107885. doi:10.1016/j.enggeo.2024.107885.

[23] Tabari, M. M. R., Azari, T., Dehghan, V. A supervised committee neural network for the determination of aquifer parameters: a case study of Katasbes aquifer in Shiraz plain, Iran. Soft Computing, 2021; 25: 4785–4798. doi:10.1007/s00500-020-05487-2.

[24] Farahani, S. A Data-Driven Seismic Design Framework for RC Frames Using Wavelet Transforms and Machine Learning. Civil Engineering and Applied Solutions, 2027; 3: 1–24. doi:10.22080/ceas.2026.30950.1073.

[25] Sadeghpour Haji, M., Niknam, R., Shayanfar, J. ANN-Based Modeling of Shear Behavior of Reinforced Concrete Columns under Constant Axial Loads. Civil Engineering and Applied Solutions, 2026; 2: 28–48. doi:10.22080/ceas.2025.30228.1050.

[26] Ivakhnenko, A. G. Polynomial Theory of Complex Systems. IEEE Transactions on Systems, Man, and Cybernetics, 1971; SMC-1: 364–378. doi:10.1109/TSMC.1971.4308320.

[27] Gandomi, A. H., Roke, D. A. Assessment of artificial neural network and genetic programming as predictive tools. Advances in Engineering Software, 2015; 88: 63–72. doi:10.1016/j.advengsoft.2015.05.007.

[28] MolaAbasi, H., Khajeh, A., Jamshidi Chenari, R. Use of GMDH-type neural network to model the mechanical behavior of a cement-treated sand. Neural Computing and Applications, 2021; 33: 15305–15318. doi:10.1007/s00521-021-06157-6.

[29] Kordnaeij, A., Kalantary, F., Kordtabar, B., Mola-Abasi, H. Prediction of recompression index using GMDH-type neural network based on geotechnical soil properties. Soils and Foundations, 2015; 55: 1335–1345. doi:10.1016/j.sandf.2015.10.001.

[30] Kalantary, F., Ardalan, H., Nariman-Zadeh, N. An investigation on the Su–NSPT correlation using GMDH type neural networks and genetic algorithms. Engineering Geology, 2009; 104: 144–155. doi:10.1016/j.enggeo.2008.09.006.

[31] MolaAbasi, H., Saberian, M., Kordnaeij, A., Omer, J., Li, J., Kharazmi, P. Predicting the stress-strain behaviour of zeolite-cemented sand based on the unconfined compression test using GMDH type neural network. Journal of Adhesion Science and Technology, 2019; 33: 945–962. doi:10.1080/01694243.2019.1571659.

[32] Farrokhi, F., Firoozfar, A., Maghsoudi, M. S. Evaluation of liquefaction-induced lateral displacement using a GMDH-type neural network optimized by genetic algorithm. Arabian Journal of Geosciences, 2019; 13: 4. doi:10.1007/s12517-019-4980-1.

[33] Javdanian, H., Heidari, A., Kamgar, R. Energy-Based Estimation of Soil Liquefaction Potential Using GMDH Algorithm. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2017; 41: 283–295. doi:10.1007/s40996-017-0061-4.

[34] Building & housing research center (BHRC). standard No. 2800: Iranian code of practice for seismic resistant design of buildings. Tehran (IR): BHRC; 2007 (in Persian).

[35] Koopialipoor, M., Armaghani, D. J., Hedayat, A., Marto, A., Gordan, B. Applying various hybrid intelligent systems to evaluate and predict slope stability under static and dynamic conditions. Soft Computing, 2019; 23: 5913–5929. doi:10.1007/s00500-018-3253-3.

[36] Nguyen, M. D., Costache, R., Sy, A. H., Ahmadzadeh, H., Van Le, H., Prakash, I., Pham, B. T. Novel approach for soil classification using machine learning methods. Bulletin of Engineering Geology and the Environment, 2022; 81: 468. doi:10.1007/s10064-022-02967-7.

[37] Jiawei, H., Micheline, K., Jian, P. Data mining: Concepts and techniques. Waltham: Morgan Kaufmann; 2006. p. 559–569.

[38] Lee, S., Kim, J., Kang, H., Kang, D.-Y., Park, J. Genetic algorithm based deep learning neural network structure and hyperparameter optimization. Applied Sciences, 2021; 11: 744. doi:10.3390/app11020744.

[39] Kumar, A., Satyannarayana, R., Rajesh, B. G. Correlation between SPT-N and shear wave velocity (VS) and seismic site classification for Amaravati city, India. Journal of Applied Geophysics, 2022; 205: 104757. doi:10.1016/j.jappgeo.2022.104757.

[40] Daag, S., Halasan, O. P., Magnaye, A., Grutas, R., Solidum, R. Empirical Correlation between Standard Penetration Resistance (SPT-N) and Shear Wave Velocity (Vs) for Soils in Metro Manila, Philippines. Applied Sciences, 2022; 12: 8067. doi:10.3390/app12168067.

[41] Cheng, J., Sun, J., Yao, K., Xu, M., Cao, Y. A variable selection method based on mutual information and variance inflation factor. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2022; 268: 120652. doi:10.1016/j.saa.2021.120652.

[42] Sugiyama, M. Introduction to Statistical Machine Learning. 1st ed. Waltham (MA): Morgan Kaufmann; 2016. doi:10.1016/B978-0-12-802121-7.00012-1.

[43] Ivakhnenko, A., Müller, J. A. Self-organization of nets of active neurons. Systems Analysis Modelling Simulation, 1995; 20: 93–106.

[44] Madala, H. R. Inductive Learning Algorithms for Complex Systems Modeling. 1st ed. Boca Raton (FL): CRC Press; 1993. doi:10.1201/9781351073493.

[45] Lemke, F., Müller, J. A. Self-Organising Data Mining. Systems Analysis Modelling Simulation, 2003; 43: 231–240. doi:10.1080/0232929031000136135.

[46] Taheri, A., Makarian, E., Manaman, N. S., Ju, H., Kim, T.-H., Geem, Z. W., RahimiZadeh, K. A Fully-Self-Adaptive Harmony Search GMDH-Type Neural Network Algorithm to Estimate Shear-Wave Velocity in Porous Media. Applied Sciences, 2022; 12: 6339. doi:10.3390/app12136339.

[47] Katoch, S., Chauhan, S. S., Kumar, V. A review on genetic algorithm: past, present, and future. Multimedia Tools and Applications, 2021; 80: 8091–8126. doi:10.1007/s11042-020-10139-6.

[48] Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning. 2nd ed. New York (NY): Springer; 2009. doi:10.1007/978-0-387-84858-7.

[49] Arlot, S., Celisse, A. A survey of cross-validation procedures for model selection. Statistics Surveys, 2010; 4: 40–79. doi:10.1214/09-SS054.

[50] Chai, T., Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 2014; 7: 1247–1250. doi:10.5194/gmd-7-1247-2014.

[51] Botchkarev, A. Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. Interdisciplinary Journal of Information, Knowledge, and Management, 2019; 14: 45–76. doi:10.28945/4184.

[52] Asteris, P. G., Rizal, F. I., Koopialipoor, M., Roussis, P. C., Ferentinou, M., Armaghani, D. J., Gordan, B. Slope Stability Classification under Seismic Conditions Using Several Tree-Based Intelligent Techniques. Applied Sciences, 2022; 12: 1753. doi:10.3390/app12031753.

[53] Onyelowe, K. C., Mojtahedi, F. F., Ebid, A. M., Rezaei, A., Osinubi, K. J., Eberemu, A. O., Salahudeen, B., Gadzama, E. W., Rezazadeh, D., Jahangir, H., Yohanna, P., Onyia, M. E., Jalal, F. E., Iqbal, M., Ikpa, C., Obianyo, I. I., Rehman, Z. U. Selected AI optimization techniques and applications in geotechnical engineering. Cogent Engineering, 2023; 10: 2153419. doi:10.1080/23311916.2022.2153419.

[54] Jolfaei, S., Lakirouhani, A. Sensitivity analysis of effective parameters in borehole failure, using neural network. Advances in Civil Engineering, 2022; 2022: 4958004. doi:10.1155/2022/4958004.

Civil Engineering and Applied Solutions

Research Article

journal homepage: ceas.journals.umz.ac.ir

GMDH-Derived Correlation for Investigating Soil Internal Friction Angle from SPT and Physical Indices

Hassan Esmaeilzadeh [a], Issa Shooshpasha [b*], Hossein MolaAbasi [c]

[a] *Department of Civil Engineering, Na. C., Islamic Azad University, Najafabad, Iran*
[b] *Department of Civil Engineering, Babol Noshirvani University of Technology, Babol, Iran*
[c] *Department of Civil Engineering, Gonbad Kavous University, Gonbad Kavous, Iran*

A R T I C L E I N F O

Article history:

Received 12 June 2026

Revised 03 July 2026

Accepted 28 August 2026

Available online 01 April 2027

Keywords:

GMDH

Internal friction angle

Sensitivity analysis

SPT

Atterberg

Correlation

A B S T R A C T

This study develops and evaluates a Group Method of Data Handling (GMDH)–based polynomial model to estimate the internal friction angle (ϕ) of soils from in-situ measurements. The modeling framework uses 318 data sets compiled from multiple geographic regions, thereby capturing broad variability in depositional history and soil fabric. Predictor variables comprise the Standard Penetration Test (SPT) blow count (N) together with seven influential in-situ/geotechnical parameters: effective overburden stress, fines content, Atterberg limits, water content, dry unit weight, and particle-size descriptors, including the percentage passing the No. 4 sieve. The GMDH algorithm automatically selects structure and coefficients, yielding an interpretable polynomial network that balances accuracy and parsimony. Model performance was assessed using standard statistical indicators and external verification procedures, and input significance was examined through sensitivity analysis to quantify the relative contribution of each parameter. Results demonstrate that the proposed model attains high predictive accuracy for ϕ across the heterogeneous, globally sourced database, underscoring its robustness to site-specific biases. Given its reliance on commonly available field and laboratory indices, the approach offers a practical and powerful tool for preliminary design and data-driven characterization of shear strength parameters where direct shear or triaxial testing is limited. Overall, the findings support GMDH as an effective methodology for identifying reliable empirical relationships between ϕ and routinely measured soil properties.

How to cite this article: Esmaeilzadeh, H., Shooshpasha, I., MolaAbasi, H. GMDH-Derived Correlation for Investigating Soil Internal Friction Angle from SPT and Physical Indices. Civil Engineering and Applied Solutions. 2027; 3(2): 161–173. doi:10.22080/ceas.2026.32094.1112.

1. Introduction

Standard Penetration Test (SPT) is extensively employed in geotechnical research and exploration projects to quickly and cost-effectively assess the mechanical characteristics of soil layers. In this test, a 76 mm borehole is drilled, and at specified intervals, a split spoon sampler with a diameter of 50.5 mm is lowered into the borehole and impacted. The number of blows to penetrate 450 mm is registered. The number of blows in the first 150 mm is neglected because of the material disturbance during the drilling process. The SPT N value is therefore the number of blows required to penetrate 300 mm [1]. It should be noted that the results of SPT in granular soils have a limited application. For example, coarse grains of sand and rubble do not enter the SPT sampler, and the obtained blows will not be acceptable. The other limitation of SPT is that it does not provide an index of cohesion due to cementation of granular soils. SPT provides a suitable sample for performing soil-related laboratory tests. SPT N value and soil engineering behavior can be related through local correlations, as well as extensive empirical relationships.

* Corresponding author.
E-mail addresses: shooshpasha@nit.ac.ir (I. Shooshpasha).

https://doi.org/10.22080/ceas.2026.32094.1112
ISSN: 3092-7749/

Identification of soil properties plays an important role in geotechnical engineering, such as soil classification, quantitative and qualitative expression of soil, and estimation of soil behavior during swelling or settlement. The laboratory and field techniques reveal the true value of soil features such as the internal friction angle (ϕ) and cohesion (C). Standard Penetration Test (SPT) and Cone Penetration Test (CPT) are two examples of in-situ tests that are favored for direct determination of the internal friction angle of soil due to the challenges associated with soil collection and the high prices of representative undisturbed samples [2-5].

SPT has three advantages over CPT:

1. SPT supplies a sample of the soil in which the experiment was conducted and, hence, determines the general categorization of the soil, while CPT only has a cone tip and is not capable of doing this. In these situations, soil classification can be determined only based on the correlation between experiment results or by using the findings from adjacent pits and boreholes.
2. In geotechnical investigation projects, SPT is often carried out during drilling boreholes. As a result, it does not add any more costs to the design.
3. The SPT's patented equipment isn't very extensive, and almost all of the test tools are also utilized for drilling. CPT, however, frequently needs expensive proprietary equipment [6-8].

Because determining soil shear strength parameters − particularly in depth − is both costly and time-consuming, it is essential to develop models that can easily forecast these characteristics [9]. Therefore, a reliable correlation based on commonly available SPT and soil index data can provide a practical preliminary estimate of the internal friction angle when direct laboratory shear strength tests are limited. Such estimates may also be useful in seismic geotechnical applications, where preliminary evaluation of bearing capacity and soil-micropile interaction in liquefiable sandy deposits is required. In the literature, various modeling techniques based on multivariate regression and pattern recognition have been reported for system identification. Some of these include genetic algorithms, artificial neural networks, neuro-fuzzy systems, and polynomial classifiers. These techniques employ learning paradigms to assess the system parameters.

When simulating a highly nonlinear system, the application of polynomial classifiers becomes problematic in terms of processing and storage perspectives. Besides, the requirement for high-order polynomials may lead to numerical instability. The Group Method of Data Handling (GMDH) method was then presented. A multilayered second-order polynomial network called GMDH is organized according to a training scenario.

Without prior knowledge of the process, this technique can be applied to model and analyze complex and specific systems where there is an unknown relationship between variables. Recently, GMDH has been successfully used in geotechnical engineering thanks to the utilization of such self-organizing networks [10-16].

GMDH has been used by researchers in geotechnical engineering for different purposes. For example, Naeini et al. [17] predicted the subgrade reaction modulus (K_s) of clayey soil through GMDH. They found that K_s obtained by Plate Load Test (PLT) and GMDH-based equations were in excellent agreement. Moreover, the suggested GMDH model's soil liquid limit (LL) was the strongest factor for forecasting K_s.

Jahed Armaghani et al. [18] used GMDH for the prediction of rock deformation. They concluded that GMDH is a strong and reliable method for predicting rock deformation. Ardakani and Kordnaeij [19] applied GMDH for the prediction of soil compaction parameters. The results demonstrated that GMDH is very good at forecasting soil compaction parameters. Additionally, they discovered that the parameters that have the most influence on soil compaction are LL and PL (soil plastic limit). Kim et al. [5] utilized GMDH for the prediction of undrained shear strength of soil. They found that GMDH is an efficient approach in achieving a new empirical mathematical model to provide an appropriate prediction of the undrained shear strength of soils. MolaAbasi et al. [20] used GMDH to forecast the behavior of cement-treated sands. They strongly recommended GMDH as a potent approach to assess the mechanical properties of soil such as stiffness, failure strain, and brittle index. Related studies on treated/reinforced soils and data-driven prediction approaches in civil engineering have also been reported by Ghadakpour et al. [21], Ghadakpour et al. [22], Kutanaei Saman et al. [23], Nematzadeh et al. [24], and Memarzadeh et al. [25], which provide useful background for the present research.

The purpose for the creation and development of an GMDH artificial neural network in this study is to estimate the internal friction angle, taking into account different soil parameters such as the number of modified blows for the hammer energy ratio of 70% (N'_{70}), effective overburden stress (σ'_v), fine content (FC or F_{200}), LL, PI (plasticity index), water content (w), dry unit weight (γ_d), the soil particle diameter below which 10, 30, 50, and 60% of the particles are smaller than this size (D_{10}, D_{30}, D_{50}, D_{60}) and the percentage passing sieve No. 4 (F_4). The reason for choosing these parameters is based on their effect on the strength and stability of the soil. Some input variables, particularly LL, PI, fines content, and water content, may be interrelated. However, in the GMDH framework, the final polynomial structure is selected based on predictive performance, and the sensitivity analysis provides additional insight into the relative contribution of each parameter. The possible influence of multicollinearity has therefore been acknowledged as a factor that should be considered when interpreting the model. Some of the selected geotechnical input variables may show statistical dependency because soil index properties are often interrelated. For example, fines content, Atterberg limits, water content, and particle-size distribution parameters can be partially correlated depending on soil type and depositional conditions. Therefore, the final GMDH model and the sensitivity analysis results were interpreted with consideration of possible multicollinearity among the input variables. The selected parameters were retained because they are commonly available in geotechnical investigations and represent different aspects of soil behavior, including penetration resistance, stress condition, plasticity, density, water content, and gradation. These parameters were selected because they represent penetration resistance, stress condition, plasticity, density, water content, and particle-size distribution, which are directly or indirectly related to the shear strength behavior of soils. In this regard, at first, the previous studies about the correlation between ϕ and other soil parameters were presented. Then, the modeling method with GMDH was described. Finally, the accuracy of the GMDH model was evaluated using sensitivity analysis. It should be mentioned that several methods have been proposed for the determination of the modified standard penetration number. In this research, the Bowles method was used for this purpose because in the Bowles method, 70% of the energy is transmitted to the sampler considering the advanced equipment of SPT [26].

2. Review of previously proposed correlations and the novelty of this research

A series of studies about the application of the standard penetration blow counts for assessing the internal friction angle of soil are presented in this section. So far, numerous empirical relationships between the internal friction angle of the soil, the results of SPT, and other soil parameters have been presented in various scientific investigations. Table 1 presents a summary of previous investigations and empirical correlations between internal friction angle, standard penetration number, and other parameters.

Table 1. Correlation between SPT number, soil parameters, and internal friction angle.

Row	Researchers	Proposed correlation	Remark
1	Dunham [27]	$\phi_d = (12N)^{0.5} + 25$	Angular and well-graded soil particles
		$\phi_d = (12N)^{0.5} + 15$	Round and uniform-graded soil particles
		$\phi_d = (12N)^{0.5} + 20$	Round and well-graded or angular and uniform-graded soil particles
2	Peck et al. [28]	$\phi_d = 53.881 - 27.6034\exp(0.0147N_{60})$	—
3	Shioi and Fukui [29]	$\phi = \sqrt{18N'_{70} + 15}$	Roads
		$\phi = 0.36N'_{70} + 27$	Bridges
		$\phi = 0.45N'_{70} + 20$	Buildings
4	Wolff [30]	$\phi = 27.1 + 0.3N_{60} - 0.00054{N_{60}}^2$	—
5	Kulhawy and Mayne [31]	$\phi = \tan^{-1}\left(N_{60}\left(12.2 + 20.3(\sigma'/p_a)\right)\right)^{0.34}$	Sandy soils
6	Japan Road Association [32]	$\phi_d = (15N)^{0.5} + 15$	$N > 5$ & $\phi_d < 45°$
7	Hatanaka and Uchida [33]	$\phi_d = (20N_1)^{0.5} + 20$	—
8	Zekkos et al. [34]	$\phi' = 3.5\sqrt{N_{1,60} + 15} + 22.3 \pm \varepsilon$	—
9	Hettiarachchi and Brown [35]	$\phi' = \beta' \, tan^{-1}\left(\frac{0.2N_{60}}{K(\sigma'/p_a)} - 0.68B\right)$	Sandy soils
10	Esmaeilzadeh et al. [36]	$\phi = 0.66N'_{70} + 8.52$	Sandy soils
		$\phi = 13.56\ln N'_{70} - 18.20$	
11	Mahmoud [3]	$\phi = 0.209N'' + 19.68$	Silty clay with sand
12	Shooshpasha et al. [37]	$\phi' = -3.574 - 0.183FC + 1.346Y_1 + 0.00111FC^2 - 0.0056{Y_1}^2 + 0.00127Y_1FC$	Granular soils
13	Salari et al. [38]	$\phi = 0.474N + 16.188$	Granular soil (GW)
		$\phi = 0.3556N + 20.703$	Granular soil (GC)
14	Mujtaba et al. [39]	$\phi = 0.7N_{60} + 18$	—
15	Kasim and Raheem [40]	$\phi = 0.8531N + 0.1581$	—
16	Dalai and Patra [41]	$\phi = 8.103N^{0.458}$	Cohesionless soil

Compared with previous studies, the novelties of the current research are its simultaneous treatment of both fine- and coarse-grained soils, ensuring applicability across common engineering materials; the compilation of a globally sourced database (318 data sets) that captures wide variability in depositional and stress histories; the implementation of a two-hidden-layer GMDH architecture that improves representational capacity while retaining interpretability of the polynomial network; and the targeted use of seven routinely measured geotechnical predictors, which collectively elevate prediction accuracy for the internal friction angle. Together, these advances deliver a robust, data-driven correlation that is practical for preliminary design where direct shear or triaxial testing is constrained, yet rigorous enough to generalize across diverse sites and soil types. Although modern machine-learning models such as random forest, XGBoost, support vector regression, and artificial neural networks may also be used for this prediction task, the present study focused on GMDH because it provides an explicit and interpretable polynomial equation that can be readily used in preliminary geotechnical design. A detailed comparison with other machine-learning algorithms is recommended for future studies.

3. The data and case histories

The data of this study were collected from the results of geotechnical tests in different parts of Iran and other countries of the world, including 42 sites and 318 geotechnical data sets. These tests include in-situ and laboratory tests on samples taken from the boreholes. The data collected from the databases are derived from the usual reports of geotechnical studies by prominent and reputable consulting engineers. Fig. 1 shows the geographical range of data on the world map. The database was compiled from different geographical regions and therefore includes variability in depositional history and geological conditions. Nevertheless, regional data imbalance may influence the fitted correlation; therefore, the model should be interpreted as a global preliminary correlation rather than a region-specific design equation. The collected database includes soil records from different geological conditions and covers a broad range of SPT resistance, effective stress levels, soil index properties, and particle-size distributions. These variations help improve the representativeness of the data and support the generalization capability of the proposed correlation within the range of the database.

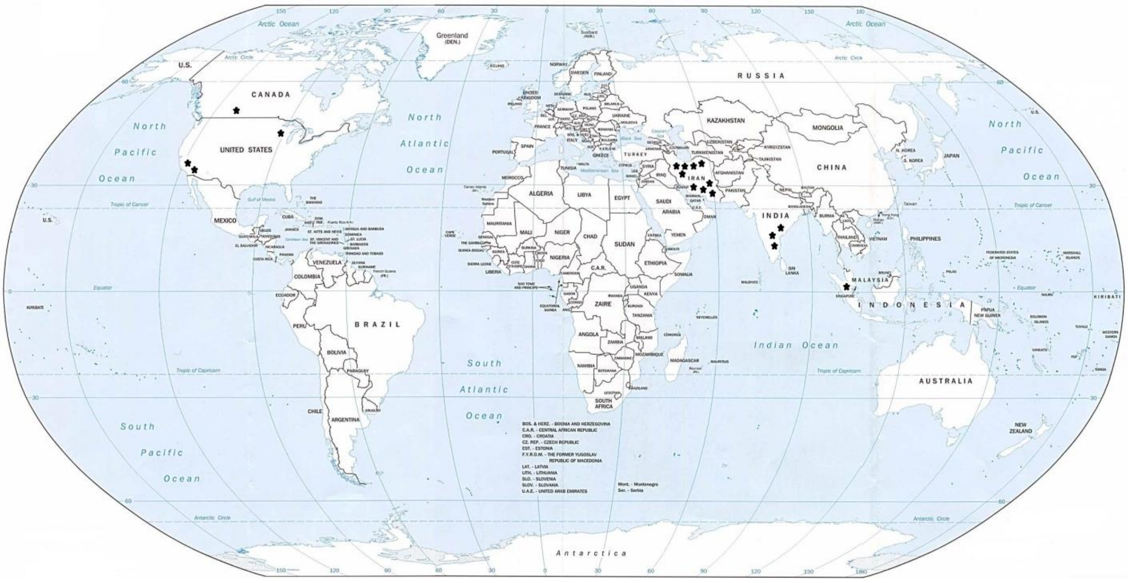


Fig. 1. Geographical distribution of the data on the world map.

It should also be emphasized that many consulting engineers in Iran have used safety hammers in the SPT, and the standard penetration sampler has no inner tube [12]. The data collected from the SPT have relatively similar drilling method and the hammer type. The collected database covers a wide range of SPT values, fines contents, unit weights, water contents, Atterberg limits, particle-size indices, and measured internal friction angles, which supports the development of a preliminary predictive correlation. Furthermore, direct shear test was carried out in the laboratory for determination of the shear strength parameters. Table 2 presents some of the data collected in this research, which are the results of reports by trusted geotechnical consultants.

To provide a clearer overview of the database, a statistical summary of the sample data presented in Table 2 is provided in Table 3. As shown in Table 3, the sample data cover a relatively wide range of corrected SPT values, fines contents, unit weights, water contents, Atterberg limits, particle-size indices, and measured internal friction angles. Therefore, the proposed correlation should be applied with consideration of the range of variables represented in the database. The proposed GMDH correlation should be applied with consideration of the range of input variables represented in the database summarized in Table 3. Predictions made outside these ranges should be considered extrapolations and may not provide reliable estimates for practical geotechnical design. The proposed model was developed using a database containing different soil categories; however, its application should remain within the range of soil types represented in the database. For practical use, the model should be applied with caution to soil categories or geological conditions that are not well represented in the collected data.

4. Description of modeling using GMDH

GMDH is a supervised machine learning tool that was suggested by [42, 43]. This tool is able to automatically synthesize the models. Moreover, it includes several processes such as planting, breeding, duplicating, and selecting or rejecting input variables, setting the model structure and its parameters, and the selection of the model based on the termination principle [42, 44].

With the help of the self-organizing GMDH algorithm, sophisticated models can be created by evaluating how well they perform on a set of multi-input, single-output data pairs (x_i, y_i) $(i = 1, 2, \ldots, M)$. The fundamental concept of GMDH is the creation of an analytical function in a progressive network according to a second-order transmission function, the coefficients of which are acquired by regression analysis [45].

Flexible structure is one of the characteristics of the GMDH network, which can be combined with repetitive and evolutionary algorithms such as back-propagation, genetic programming, particle group optimization, and genetic algorithm [13, 46-52].

Using GMDH, a model can be described as a group of neurons connected in different pairs in each layer by a second-order polynomial. As a result, in the following layer, new neurons are produced. For plotting inputs to outputs, this representation can be utilized. The important point is the challenge of finding a function (f), which can be used roughly instead of observations of the f.

A successful prediction should have an output (y) that is as close as possible to the observed output (y') for a given input vector $X = (x_1, x_2, x_3, \ldots, x_n)$.

Hence, according to M observations, a multi-input and single-output data set is obtained as Eq. 1.

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})\ (i = 1, 2, 3, \dots, M) \tag{1}$$

Table 2. A sample of the database extracted from 318 data sets.

N'_{70}	σ'_v (kg/m²)	F_{200} (%)	F_4 (%)	γ_d (kg/m³)	w (%)	PI (%)	LL (%)	D_{10} (mm)	D_{30} (mm)	D_{50} (mm)	D_{60} (mm)	ϕ (°)
5	3381	17.4	99.8	1550	26.30	0	0	0	0.18	0.26	0.35	32.0
3	4415	55	0	1440	32.00	10.00	29.00	0	0.01	0.052	0.08	31.0
12	5925	1	100	1490	23.31	0	0	0.111	0.172	0.212	0.235	27.8
49	5730	95	100	1620	18.00	30.00	31.50	0.003	0.02	0.05	0.12	24.0
58	6300	89	100	1890	16.90	15.00	35.00	0	0.006	0.018	0.025	20.1
11	2250	38.1	97	1370	24.50	0	0	0	0	0.14	0.18	33.0
18	14000	91	96	1700	14.00	15.00	32.00	0	0.005	0.018	0.033	5.0
19	7350	5.8	87	1700	2.00	0	0	0.12	0.6	1.6	2.1	35.0
14	1920	95	0	1510	25.70	26.00	51.00	0	0.0063	0.015	0.026	11.0
7	4936	29	97	1670	19.00	23.00	49.00	0	0.14	0.25	0.8	0.0
40	15600	92	100	1730	14.00	21.00	42.70	0	0.06	0.08	0.09	6.9
6	5730	95	0	1470	30.40	18.00	39.00	0	0.0015	0.005	0.008	3.0
6	30240	85	100	1830	23.91	33.00	57.00	0	0	0.002	0.004	13.0
7	7920	35	64	1600	14.00	9.00	28.00	0.008	0.06	0.9	2.833	21.0
6	19700	100	100	1630	26.00	11.00	34.00	0	0.005	0.0095	0.013	4.0
21	10192	89	99.5	1710	11.00	14.00	23.50	0	0.003	0.005	0.02	28.0
4	3800	95.5	100	1110	49.34	20.00	56.00	0	0	0.00	0.00	22.9
3	3534	4.6	99.9	1000	51.06	0	0	0.09	0.16	0.22	0.25	36.5
22	6152	7	87.5	1800	3.50	0	0	0.13	0.7	1.5	1.8	33.0
21	17820	24	49	1720	10.50	12.00	29.00	0.05	0.271	5	8.158	25.0
9	6000	95	98	1660	25.00	15.00	33.00	0	0.004	0.009	0.014	5.0
4	8700	63	88	1490	27.00	10.00	23.00	0.006	0.046	0.06	0.072	16.0
13	3652	0.3	100	1600	19.85	0	0	0.12	0.177	0.217	0.237	30.6
12	13790	95	100	1660	20.00	15.50	35.00	0	0.005	0.015	0.022	19.0
7	2995	92	100	1250	39.25	0	0	0	0	0.00	0.00	15.8
7	3900	86	0	1580	23.00	11.00	33.00	0	0.003	0.0085	0.015	7.0
16	5610	62	86	1590	17.00	6.00	26.00	0.006	0.037	0.07	1.496	7.0
13	10781	50	100	1460	30.00	29.00	59.00	0	0	0.16	0.3	0.0
12	10440	99	100	1660	25.00	13.00	32.00	0	0.004	0.0085	0.014	14.0
20	3748	0.1	100	1630	20.89	0	0	0.126	0.18	0.219	0.238	29.8
5	13890	58	79	1590	26.00	5.00	23.00	0.005	0.047	0.065	0.14	15.0
7	8510	86	0	1560	19.00	30.00	50.00	0	0	0.0013	0.0023	29.0

Table 3. Statistical summary of the sample database values presented in Table 2.

Variable	Symbol	Unit	Minimum	Maximum	Mean	Standard deviation
Corrected SPT blow count	N'_{70}	-	3.00	58.00	14.28	12.91
Effective overburden stress	σ'_v	kg/m²	1920.00	30240.00	8403.47	6130.41
Fines content	F_{200}/FC	%	0.10	100.00	60.31	36.78
Passing sieve No. 4	F_4	%	0.00	100.00	78.99	36.33
Dry unit weight	γ_d	kg/m³	1000.00	1890.00	1570.94	188.87
Water content	w	%	2.00	51.06	22.73	10.64
Plasticity index	PI	%	0.00	33.00	12.23	10.35
Liquid limit	LL	%	0.00	59.00	26.58	19.41
Particle size at 10% finer	D_{10}	mm	0.00	0.13	0.02	0.05
Particle size at 30% finer	D_{30}	mm	0.00	0.70	0.09	0.17
Particle size at 50% finer	D_{50}	mm	0.00	5.00	0.35	0.94
Particle size at 60% finer	D_{60}	mm	0.00	8.16	0.62	1.54
Internal friction angle	ϕ	Degree	0.00	36.50	18.76	11.43

Then, GMDH should be trained to forecast output values ($\hat{y}_i$) for each given input vector $X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$ as shown in Eq. 2.

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})\ (i = 1, 2, 3, \dots, M) \tag{2}$$

A reference, a general relationship between the input and output variables, is created by GMDH in the form of a mathematical description. Finding the GMDH network that minimizes the square of the differences between the actual and anticipated outputs is now the key challenge, that is shown in Eq. 3.

$$\sum_{i=1}^{M} \left[\hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) - y_i\right]^2 \rightarrow min \tag{3}$$

A sophisticated discrete form of the Volterra function can be used to express the general relationship between input and output variables. Eq. 4 represents a series of this function, which is referred to as Kolmogorov-Gabor polynomials [35, 36, 42, 47].

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \cdots \tag{4}$$

In this investigation, the second-order polynomial in the GMDH network is applied as shown in Eq. 5.

$$\text{Quadratic: } \hat{y} = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \tag{5}$$

Therefore, to establish the general mathematical relationship between outputs and inputs in Eq. 4, a partial second-order polynomial is applied in a network of connected neurons during a back calculation method. Regression methods are used to derive the coefficients ai in Eq. 5 so that for each pair of x_i, y_i as input variables, the difference between the observed (y) and the calculated ($\hat{y}$) output is minimized.

The second-order form provided in Eq. 5 is typically used to create polynomial trees, and the least squares approach is applied to achieve the coefficients. In order to fit the output in the entire set of input-output data pairs, the coefficients of each G_i quadratic function are generated as shown in Eq. 6.

$$E = \frac{\sum_{i=1}^{M}(y_i - G_i(\ ))^2}{M} \rightarrow min \tag{6}$$

The regression polynomial in the form of Eq. 5 that best fits the dependent observations ($y_i = 1, 2, \dots, M$) in least squares is established using the GMDH by taking into account all potential combinations of two independent variables out of the total input variables (n).

In the first hidden layer of the predictor network of observations $\{(y_i, x_{ip}, x_{iq}); (i = 1, 2, \dots, M)\}$ for various $p, q \in \{1, 2, \dots, n\}$ will therefore $\binom{n}{2} = \frac{n(n-1)}{2}$ neurons be formed. As a result, it is currently possible to establish M data triples $\{(y_i, x_{ip}, x_{iq}); (i = 1, 2, \dots M)\}$ from observations through $p, q \in \{1, 2, \dots, n\}$, as shown in Eq. 7.

$$\begin{bmatrix} x_{1p} & x_{1q} & y_1 \\ x_{2p} & x_{2q} & y_2 \\ x_{Mp} & x_{Mq} & y_M \end{bmatrix} \tag{7}$$

The following matrix equation can be utilized with the employment of the quadratic equation (Eq. 5) for each row of M (Eq. 8).

$$Aa = Y \tag{8a}$$

$$a = \{a_0, a_1, a_2, a_3, a_4, a_5\} \tag{8b}$$

$$Y = \{y_1, y_2, y_3, \dots, y_M\}^T \tag{8c}$$

where Y denotes a vector of the output values of the observations and a is a vector of unidentified coefficients for a quadratic polynomial in Eq. 5. Eq. 9 can be discovered as follows:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^2 & x_{Mq}^2 \end{bmatrix} \tag{9}$$

When performing a multiple regression analysis, the normal equations are solved using the least squares method (Eq. 10).

$$a = (A^T A)^{-1} A^T Y \tag{11}$$

The vector of the best coefficients from Eq. 5 is determined for the entire set of M, by Eq. 11. It must be clarified that, in accordance with the connection topology of the network, this process is repeated for each neuron in the subsequent hidden layer.

However, a solution derived directly from normal equations is more prone to round-off errors and, more critically, to the singularity of these equations [53]. In a GMDH algorithm, there are six coefficients per estimated model. The coefficients are determined by regular least squares estimation [54]. The flowchart of the GMDH algorithm has been depicted in Fig. 2.

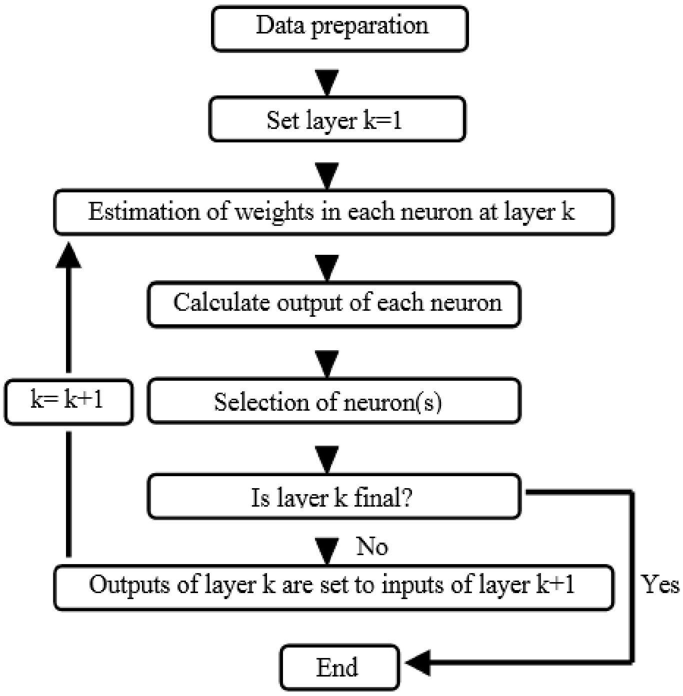


Fig. 2. Flowchart of GMDH algorithm.

5. New proposed model

Training and testing data are separated into two groups in the database. The data trained in the mentioned neural network include 200 input-output data pairs. The complete database was divided into training, testing, and validation subsets, and the validation subset was used only for the external evaluation of the final proposed model. Testing data include 86 input-output data which have not been applied in the prediction training cycle and are only trained to test GMDH models. It is necessary to mention that the selection of training and testing data sets is random, and these data sets have almost identical statistical characteristics. The training and testing subsets were selected randomly with comparable statistical characteristics. To further improve robustness assessment, future work may apply k-fold cross-validation or repeated random subsampling using larger extended databases.

Fig. 3 depicts the performance of model prediction versus observed data for the training and testing data sets. As observed, the predicted and measured values from the training and testing sets are not significantly different. The main trends obtained from the model are generally consistent with established geotechnical understanding. An increase in corrected SPT resistance usually reflects denser or stronger soil conditions and is therefore expected to be associated with a higher internal friction angle. In contrast, higher plasticity and fines-related characteristics may reduce the frictional component of shear strength. The influence of effective overburden stress should be interpreted carefully, as it may also depend on soil fabric, depositional history, and stress history.

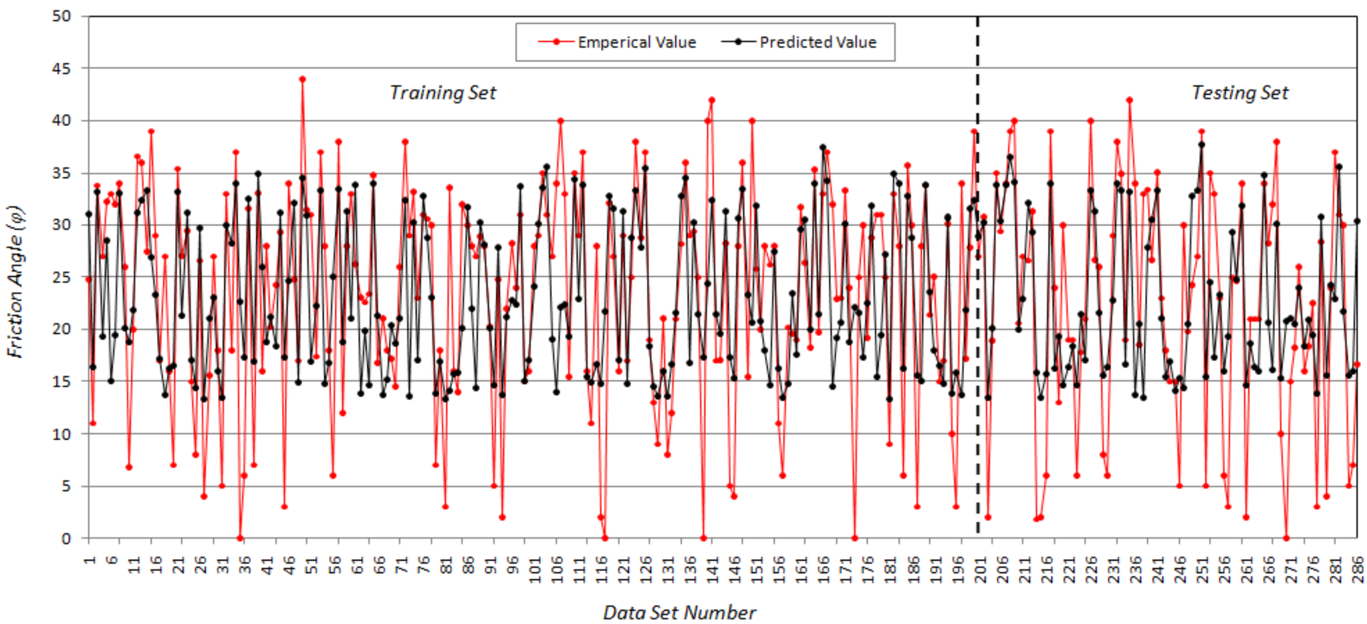


Fig. 3. Comparison between empirical and predicted values of ϕ for training and testing sets.

The developed architecture of GMDH for this research is shown in Fig. 4.

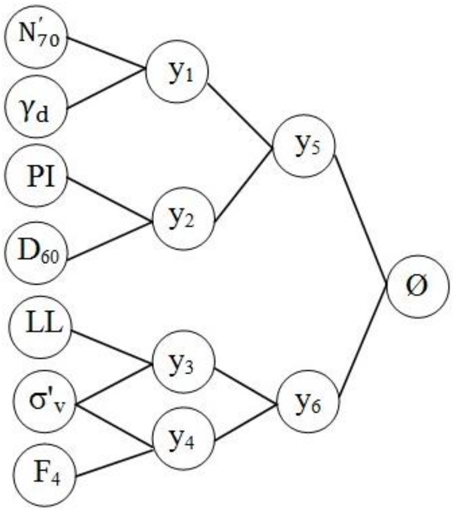


Fig. 4. The evolved architecture of a generalized GMDH for internal friction angle.

Such a model for the internal friction angle has a polynomial representation that is shown in Eq. 12.

$$\phi = -2.350050 + 0.005795 y_5 + 0.957297\, y_6 - 0.005976\, {y_5}^2 - 0.024938 {y_6}^2 + 0.037769 y_5 y_6 \quad (12a)$$

$$y_1 = 0.0828 + 0.018199 N'_{70} + 1.303645 \gamma_d - 0.000001 {N'_{70}}^2 - 0.019810 {\gamma_d}^2 - 0.000578 N'_{70} \gamma_d \quad (12b)$$

$$y_2 = 29.565296 - 0.571619 PI + 0.018102 D_{60} + 0.006224 PI^2 - 0.000011 {D_{60}}^2 + 0.000480 PI D_{60} \quad (12c)$$

$$y_3 = 62.153586 - 62.908345 LL + 58.309302 \sigma'_v + 22.280136 LL^2 - 23.030937 {\sigma'_v}^2 + 13.101200 LL \sigma'_v \quad (12d)$$

$$y_4 = 73.990484 - 73.453004 \sigma'_v + 26.393534 F_4 + 29.256098 {\sigma'_v}^2 + 24.005152 {F_4}^2 - 40.105051 \sigma'_v F_4 \quad (12e)$$

$$y_5 = 4.011870 - 0.578298 y_1 + 1.302297 y_2 + 0.008694 {y_1}^2 - 0.007880 {y_2}^2 + 0.003496 y_1 y_2 \quad (12f)$$

$$y_6 = -39.569747 + 4.094071 y_3 - 0.178640 y_4 - 0.028284 {y_3}^2 + 0.048508 {y_4}^2 - 0.070782 y_3 y_4 \quad (12g)$$

The accuracy of the statistical evaluation of this model is determined through MSE (mean squared error), RMSE (root mean square error), R^2 (coefficient of determination) and MAD (mean absolute deviation) according to the following relationships:

$$R^2 = 1 - \left[\frac{\sum_{i=0}^{M} \left(Y_{i(Model)} - Y_{i(Actual)}\right)^2}{\sum_{i=1}^{M} \left(Y_{i(Actual)}\right)^2}\right] \quad (13)$$

$$MSE = \left[\frac{\sum_{i=0}^{M} \left(Y_{i(Model)} - Y_{i(Actual)}\right)^2}{M}\right] \quad (14)$$

$$MAD = \left[\frac{\left|\sum_{i=0}^{M} \left(Y_{i(Model)} - Y_{i(Actual)}\right)\right|}{M}\right] \quad (15)$$

$$RMSE = \left[\frac{\sum_{i=0}^{M} \left(Y_{i(Model)} - Y_{i(Actual)}\right)^2}{M}\right]^{\frac{1}{2}} \quad (16)$$

In Eq. 12, N'_{70} is the corrected SPT blow count, σ'_v is the effective overburden stress, γ_d is the dry unit weight, LL is the liquid limit, PI is the plasticity index, D_{60} is the particle size corresponding to 60% finer, F_4 is the percentage passing sieve No. 4, and y_1 to y_6 are intermediate variables generated by the GMDH network. Table 4 presents the values of these parameters in the current study:

Table 4. Statistical information of the model.

Set	R^2	RMSE	MAD	MSE
Training	0.99	6.52	3.95	42.55
Testing	0.98	6.43	3.74	41.32

RMSE and MAD are reported in degrees, while MSE is reported in squared degrees. All error indices were calculated from the difference between the measured and predicted internal friction angles. The reported RMSE values were checked to ensure consistency with the corresponding MSE values. When GMDH is being designed, it should be remembered that in each layer of the network, the neurons' polynomials are all similar to each other and the network is designed based on similar processes [55]. The optimal performance of the polynomial model for the prediction of a set of data is carried out based on the statistical evaluation of the training and testing data sets (according to Table 4).

6. Validation and sensitivity analysis

The suggested model's validity and accuracy in predicting ϕ were compared with the correlations previously proposed by some researchers [3, 28-30, 33, 36, 38]. This comparison was performed for all data items considered for validation purposes in this section. This set of data is separate from the testing and training data set. Fig. 5 demonstrates the dispersion chart for the measured and predicted values of ϕ. It should be noted that the sensitivity analysis represents the relative influence of the input variables within the developed GMDH model and does not necessarily indicate direct causal relationships. Because some geotechnical variables may be correlated, the apparent importance of one parameter may partly reflect its interaction or dependency with other input variables.

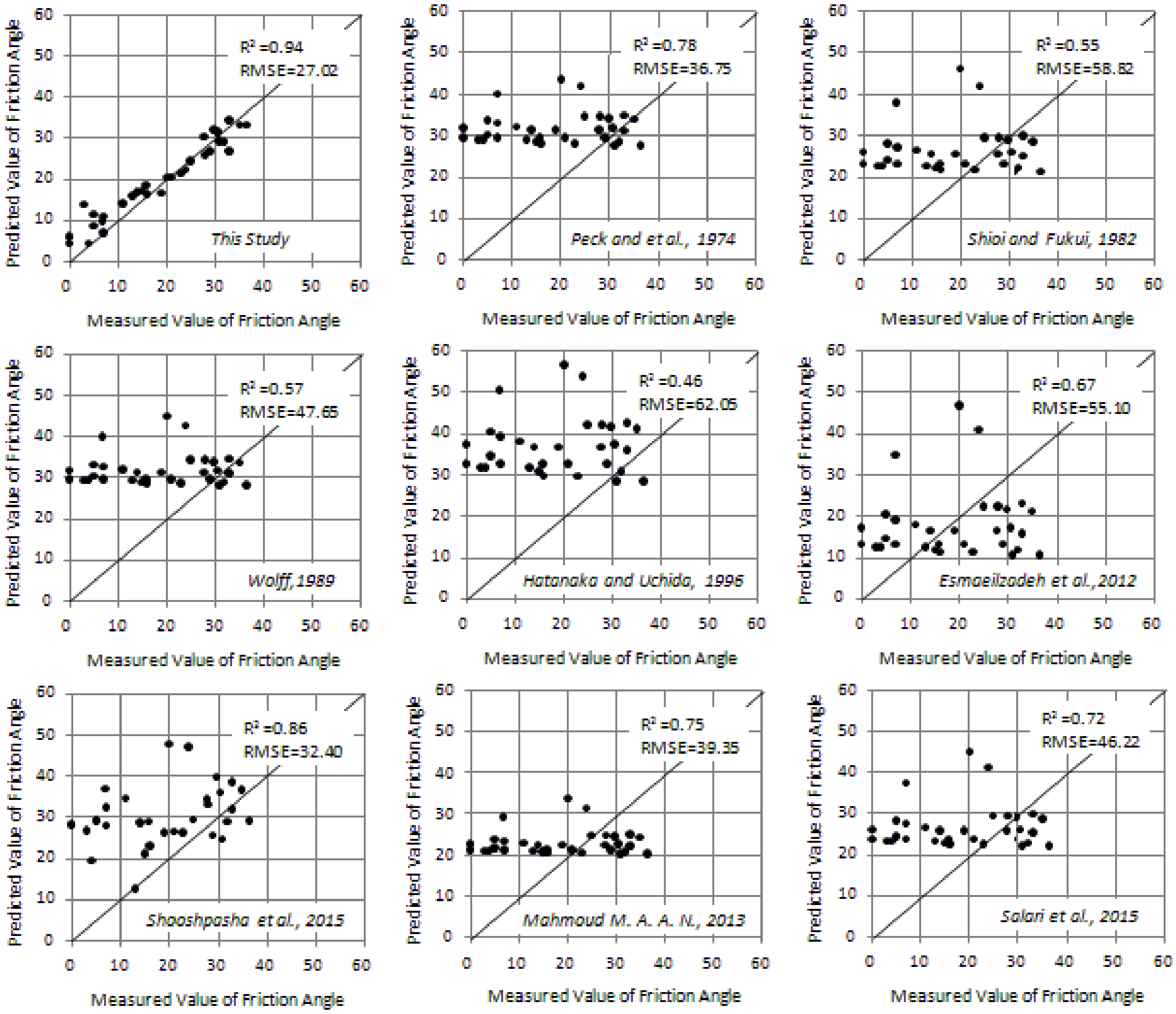


Fig. 5. Comparison of predicted and measured values of internal friction angle presented by different researchers.

The dispersion chart is a useful way to show the relationship between the data. This chart shows cases that have a positive and strong impact on the regression line. In other words, it indicates the cases in which the data are far from the range or focused. Also, a curve or any non-linear relationship can be deduced from this chart. As seen, some researchers have predicted the values of soil internal friction angle higher or lower than the measured values. The GMDH model in the present study predicts the internal friction angle with high precision. The analysis of the overall trend of points shows that there is a linear and positive relationship between variables. In addition, the focused points around the regression line are relatively good, and there is no deviation between the points. Although the present study reports deterministic prediction indices, the scatter of measured and predicted values provides a preliminary indication of prediction uncertainty. Future studies may further enhance the reliability of the proposed model by developing prediction intervals or confidence bounds for practical design applications. Sensitivity analysis in this study was performed in a similar method with partial derivatives [54]. The basis of this analysis is the partial derivative of the output of Y_j for the input x_i which represents a mathematical definition that reflects the sensitivity of Y_j to x_i [56].As a result, a function has a stronger sensitivity to a specific input variable, which causes a greater proportion of its partial derivative to be related to that variable. Also, the impact of input changes on the final output values of a network can be expressed using this method. For example, if the partial derivative value of a variable is positive at a particular point, consequently, a small increase in the input variable will result in a rise in the network's output. Additionally, the model's sensitivity analysis can be utilized to determine how input parameters affect the model's output. This mechanism involves changing each input neuron individually at a steady rate while maintaining other factors constant. Constant variables with the rates of -0.1, 0.05, 0, $+0.05$ and $+0.1$ were selected for the present research. The percentage change in the output for each input neuron is seen as a result of the input neuron's change. The sensitivity of each input neuron is found using Eq. 17 [57]:

$$\text{Sensitivity level of } X\ (\%) = \frac{1}{M}\sum_{j=1}^{M}\left(\frac{\%\ Change\ in\ Output}{\%\ Change\ in\ Input}\right)_i \times 100 \tag{17}$$

Fig. 6 depicts the results of the model's sensitivity analysis. It is clear that ϕ is significantly affected by N'_{70}. Also, LL has the least effect on the value of internal friction angle. The dominant influence of N'_{70} is reasonable from a geotechnical perspective because it directly reflects penetration resistance, relative density/consistency, and the in-situ strength condition of soil. In contrast, parameters such as LL and PI describe plasticity characteristics and may influence the friction angle more indirectly through soil fabric, fines content, and water-soil interaction.

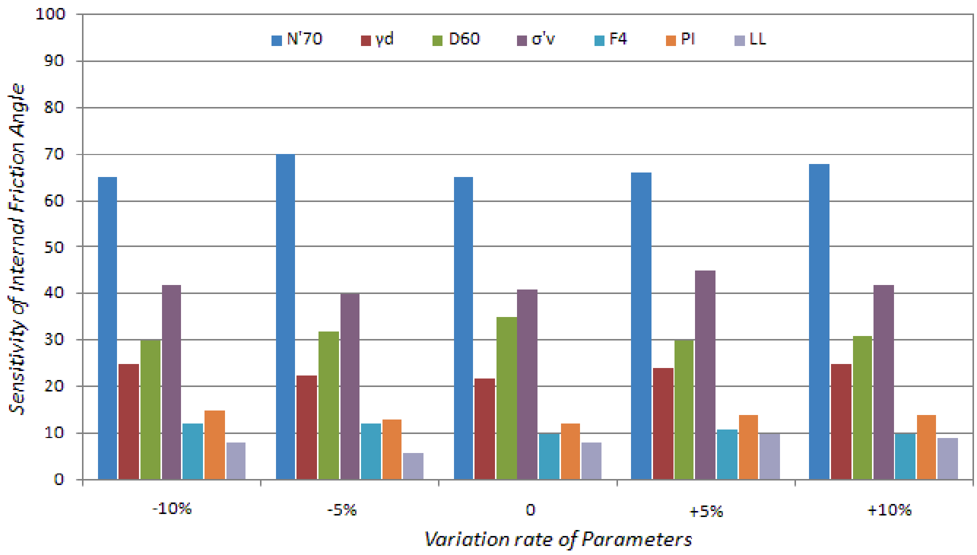


Fig. 6. Comparison of predicted and measured values of internal friction angle presented by different researchers.

7. Summary and conclusions

This research presents a new approach for predicting internal friction angle according to SPT results and geotechnical properties of soil. In order to achieve the proposed model for the internal friction angle in the current research, the developed GMDH neural network has been employed. Moreover, a polynomial mathematical model was proposed for the internal friction angle of the soil based on seven parameters including N'_{70}, LL, PI, γ_d, σ'_v, D_{60} and F_4. The new proposed model was compared with the results of other researchers, and its validation was evaluated. The main trends obtained from the model are generally consistent with established geotechnical understanding. An increase in corrected SPT resistance usually reflects denser or stronger soil conditions and is therefore expected to be associated with a higher internal friction angle. In contrast, higher plasticity and fines-related characteristics may reduce the frictional component of shear strength. The influence of effective overburden stress should be interpreted carefully, as it may also depend on soil fabric, depositional history, and stress history. The results show that the prediction of ϕ has low variance and high accuracy. The model's sensitivity analysis was also performed to examine the impact of the input parameters on the model's output. The results of the sensitivity analysis showed the effectiveness of input parameters for the initial prediction of internal friction angle and also reflected the fact that the geotechnical properties of the soil play an important role in the prediction of ϕ. It was found that ϕ was significantly influenced by N'_{70} and also it decreases with increasing PI, F_4 and LL. In other words, the internal friction angle has the highest and lowest dependency on N'_{70} and LL, respectively. Therefore, the proposed GMDH-based correlation can be considered a practical preliminary estimation tool for geotechnical applications, provided that it is used within the range of soil conditions represented in the database. Future research may extend the proposed GMDH framework to post-liquefaction soil-property estimation and to the seismic performance assessment of micropile foundations in liquefiable sandy deposits.

Statements & declarations

Author contributions

Hassan Esmaeilzadeh: Data curation, Investigation, Formal analysis, Writing - Original Draft, Writing- Reviewing & Editing.

Issa Shooshpasha: Project administration, Supervision.

Hossein MolaAbasi: Data Curation, Investigation, Formal analysis, Writing - Original Draft, Writing - Reviewing & Editing.

Funding

The authors declare that there is no funding for this research.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] MacRobert, C., Kalumba, D., Beales, P. Correlating Standard Penetration Test and Dynamic Probe Super Heavy penetration resistance values in sandy soils. Journal of the South African Institution of Civil Engineering, 2011; 53: 46–54.

[2] Holtz, R. D., Kovacs, W. D. An Introduction to Geotechnical Engineering. 1st ed. Saddle River (NJ): Prentice Hall; 1981.

[3] Mahmoud, M. A. A. N. Reliability of using standard penetration test (SPT) in predicting properties of silty clay with sand soil. International Journal of Civil and Structural Engineering, 2013; 3: 545–556.

[4] Kumar, D., Kumar, P., Wipulanusat, W., Tran, D. T., Keawsawasvong, S. Seismic Stability of Unsupported Rectangular Excavations in Cohesive-Frictional Soils: FELA Simulations and Comparative Analysis Using MARS, GP, and GMDH Models. Geotechnical and Geological Engineering, 2025; 43: 136. doi:10.1007/s10706-025-03094-2.

[5] Kim, M., Okuyucu, O., Ordu, E., Ordu, S., Arslan, Ö., Ko, J. Prediction of Undrained Shear Strength by the GMDH-Type Neural Network Using SPT-Value and Soil Physical Properties. Materials, 2022; 15: 6385. doi:10.3390/ma15186385.

[6] Shooshpasha, I., Hasanzadeh, A., Taghavi, A. Prediction of the Axial Bearing Capacity of Piles by SPT-based and Numerical Design Methods. GEOMATE Journal, 2013; 4: 560–564. doi:10.21660/2013.8.2118.

[7] Alizadeh Mansouri, M., Dabiri, R. Predicting the liquefaction potential of soil layers in Tabriz city via artificial neural network analysis. SN Applied Sciences, 2021; 3: 719. doi:10.1007/s42452-021-04704-3.

[8] Duan, W., Congress, S. S. C., Cai, G., Zhao, Z., Liu, S., Dong, X., Chen, R., Qiao, H. Prediction of in situ state parameter of sandy deposits from CPT measurements using optimized GMDH-type neural networks. Acta Geotechnica, 2022; 17: 4515–4535. doi:10.1007/s11440-022-01540-6.

[9] Mola-Abasi, H., Eslami, A. Prediction of drained soil shear strength parameters of marine deposit from CPTu data using GMDH-type neural network. Marine Georesources & Geotechnology, 2019; 37: 180–189. doi:10.1080/1064119X.2017.1415400.

[10] Ardalan, H., Eslami, A., Nariman-Zadeh, N. Piles shaft capacity from CPT and CPTu data by polynomial neural networks and genetic algorithms. Computers and Geotechnics, 2009; 36: 616–625. doi:10.1016/j.compgeo.2008.09.003.

[11] Kalantary, F., Ardalan, H., Nariman-Zadeh, N. An investigation on the Su–NSPT correlation using GMDH type neural networks and genetic algorithms. Engineering Geology, 2009; 104: 144–155. doi:10.1016/j.enggeo.2008.09.006.

[12] Shooshpasha, I., MolaAbasi, H. Prediction of Liquefaction Induced Lateral Displacements Using Polynomial Neural Networks and Genetic Algorithms. In: 15th World Conference on Earthquake Engineering; 2012 Sep 24–28; Lisbon, Portugal.

[13] Najafzadeh, M., Azamathulla, H. M. Group method of data handling to predict scour depth around bridge piers. Neural Computing and Applications, 2013; 23: 2107–2112. doi:10.1007/s00521-012-1160-6.

[14] Fattahi, H., Zandy Ilghani, N. Application of Monte Carlo Markov Chain and GMDH Neural Network for Estimating the Behavior of Suction Caissons in Clay. Geotechnical and Geological Engineering, 2023; 41: 3305–3319. doi:10.1007/s10706-023-02455-z.

[15] Mu'azu, M. A. Prediction of Soil Parameters of Waste Fibers with Hybrid Metaheuristic Algorithm. Mathematical Geosciences, 2026; doi:10.1007/s11004-025-10268-4.

[16] Zhang, S. A Hybrid Optimization Approach for Pile Capacity Estimation Using Radial Basis Functions and Time-Dependent Variables. Informatica, 2025; 49: doi:10.31449/inf.v46i19.8460.

[17] Naeini, S. A., Ziaie Moayed, R., Kordnaeij, A., Mola-Abasi, H. Prediction of subgrade reaction modulus of clayey soil using group method of data handling. Scientia Iranica, 2020; 27: 1740–1750. doi:10.24200/sci.2018.5586.1357.

[18] Jahed Armaghani, D., Momeni, E., Asteris, P. Application of group method of data handling technique in assessing deformation of rock mass. Metaheuristic Computing and Applications, 2020; 1: 1–18. doi:10.12989/mca.2020.1.1.001.

[19] Ardakani, A., Kordnaeij, A. Soil compaction parameters prediction using GMDH-type neural network and genetic algorithm. European Journal of Environmental and Civil Engineering, 2017; 23: 449–462. doi:10.1080/19648189.2017.1304269.

[20] MolaAbasi, H., Khajeh, A., Jamshidi Chenari, R. Use of GMDH-type neural network to model the mechanical behavior of a cement-treated sand. Neural Computing and Applications, 2021; 33: 15305–15318. doi:10.1007/s00521-021-06157-6.

[21] Ghadakpour, M., Janalizadeh Choobbasti, A., Soleimani Kutanaei, S. Investigation of the deformability properties of fiber-reinforced cemented sand. Journal of Adhesion Science and Technology, 2019; 33: 1913–1938. doi:10.1080/01694243.2019.1619224.

[22] Ghadakpour, M., Fakhrabadi, A., Janalizadeh Choobbasti, A., Soleimani Kutanaei, S., Vafaei, A., Taslimi Paein Afrakoti, M., Eisazadeh, N. Effect of post-construction moisture condition on mechanical behaviour of Fiber-reinforced-cemented-sand (FRCS). Geomechanics and Geoengineering, 2022; 17: 1852–1864. doi:10.1080/17486025.2021.1980230.

[23] Kutanaei Saman, S., Choobbasti Asskar, J., Fakhrabadi, A., Ghadakpour, M., Vafaei, A., Afrakoti MobinaTaslimi, P. Application of LRBF-DQ and CVBFEM Methods for Evaluating Saturated Sand Liquefaction around Buried Pipeline. Journal of Pipeline Systems Engineering and Practice, 2022; 13: 04021077. doi:10.1061/(ASCE)PS.1949-1204.0000625.

[24] Nematzadeh, M., Shahmansouri, A. A., Fakoor, M. Post-fire compressive strength of recycled PET aggregate concrete reinforced with steel fibers: Optimization and prediction via RSM and GEP. Construction and Building Materials, 2020; 252: 119057. doi:10.1016/j.conbuildmat.2020.119057.

[25] Memarzadeh, A., Nazari, A., Sabetifar, H., Nematzadeh, M. Advanced predictive modeling for masonry walls: A comparative study of six AI models and existing empirical formulas. Results in Engineering, 2025; 28: 107634. doi:10.1016/j.rineng.2025.107634.

[26] Bowles, J. E., Guo, Y. Foundation analysis and design. 5th ed. New York (NY): McGraw-Hill Companies; 1997.

[27] Dunham, J. W. Pile foundations for buildings. In: Proceedings of the American Society of Civil Engineers, Soil Mechanics & Foundations Division; 1954; Oakland, California. p. 1–21.

[28] Peck, R. B., Hanson, W. E., Thornburn, T. H. Foundation engineering. 2nd ed. Hoboken (NJ): John Wiley & Sons; 1991.

[29] Shioi, Y., Fukui, J. Application of N-value to Design of Foundations in Japan. In: European Symposium on Penetration Testing; 1982 May 24–27; Amsterdam, Netherlands. p. 159–164.

[30] Wolff, T. F. Pile capacity prediction using parameter functions. In: Predicted and observed axial behavior of piles: results of a pile prediction symposium; 1989; Oakland, California. p. 96–106.

[31] Kulhawy, F. H., Mayne, P. W. Manual on estimating soil properties for foundation design. Palo Alto (CA): Electric Power Research Institute; 1990. Report No.: EL-6800.

[32] Japan Road Association. Specifications for highway bridges part IV; 1990; Tokyo (JP).

[33] Hatanaka, M., Uchida, A. Empirical Correlation Between Penetration Resistance and Internal Friction Angle of Sandy Soils. Soils and Foundations, 1996; 36: 1–9. doi:10.3208/sandf.36.4_1.

[34] Zekkos, D., Bray, J., Der Kiureghian, A. Reliability of shallow foundation design using the standard penetration test. In: Proceedings ISC-2 on geotechnical and geophysical site characterization; 2004 Sep 19–22; Porto, Portugal. p. 1575–1582.

[35] Hettiarachchi, H., Brown, T. Use of SPT Blow Counts to Estimate Shear Strength Properties of Soils: Energy Balance Approach. Journal of Geotechnical and Geoenvironmental Engineering, 2009; 135: 830–834. doi:10.1061/(ASCE)GT.1943-5606.0000016.

[36] Esmaeilzadeh, H., Shooshpasha, I., Ghazavi, M. Estimating internal friction angle of soil using standard penetration test data (a case study in Babolsar). In: 9th International Congress on Civil Engineering; 2012 May 08; Isfahan, Iran.

[37] Shooshpasha, I., Amiri, I., MolaAbasi, H. An Investigation of Friction Angle Correlation with Geotechnical Properties for Granular Soils Using GMDH Type Neural Networks. Scientia Iranica, 2015; 22: 157–164.

[38] Salari, P., Lashkaripour, G. R., Ghafoori, M. Presentation of Empirical Equations for Estimating Internal Friction Angle of GW and GC Soils in Mashhad, Iran Using Standard Penetration and Direct Shear Tests and Comparison with Previous Equations. Open Journal of Geology, 2015; 5: 231–238. doi:10.4236/ojg.2015.55021.

[39] Mujtaba, H., Farooq, K., Sivakugan, N., Das, B. M. Evaluation of relative density and friction angle based on SPT-N values. KSCE Journal of Civil Engineering, 2018; 22: 572–581. doi:10.1007/s12205-017-1899-5.

[40] Kasim, M. N., Raheem, A. M. Evaluation of some soil characteristics from field SPT values using random number generation technique. In: IOP Conference series: earth and environmental science; 2021; Baghdad, Iraq & Istanbul, Turkey. p. 012017.

[41] Dalai, S., Patra, C. Prediction of Angle of Internal Friction Based on SPT N Values. In: Proceedings of the Indian Geotechnical Conference 2019; 2021; Singapore, Singapore. p. 471–477. doi:10.1007/978-981-33-6466-0_43.

[42] Ivakhnenko, A. G. Polynomial Theory of Complex Systems. IEEE Transactions on Systems, Man, and Cybernetics, 1971; SMC-1: 364–378. doi:10.1109/TSMC.1971.4308320.

[43] Ivakhnenko, A., Ivakhnenko, G. The review of problems solvable by algorithms of the group method of data handling (GMDH). Pattern recognition and image analysis, 1995; 5: 527–535.

[44] Madala, H. R. Inductive Learning Algorithms for Complex Systems Modeling. 1st ed. Boca Raton (FL): CRC Press; 2019.

[45] Farlow, S. J. Self-Organizing Methods in Modeling: GMDH Type Algorithms. 1st ed. Boca Raton (FL): CRC Press; 1984.

[46] Iba, H., De Garis, H. Extending genetic programming with recombinative guidance. Advances in genetic programming, 1996; 2: 69–88.

[47] Sakaguchi, A., Yamamoto, T. A GMDH network using backpropagation and its application to a controller design. In: SMC 2000 conference proceedings. 2000 IEEE International Conference on systems, man and cybernetics. Cybernetics evolving to systems, humans, organizations, and their complex interactions; 2000; Nashville, Tennessee. p. 2691–2696. doi:10.1109/ICSMC.2000.884402.

[48] Onwubolu, G. C. Design of hybrid differential evolution and group method of data handling networks for modeling and prediction. Information Sciences, 2008; 178: 3616–3634. doi:10.1016/j.ins.2008.05.013.

[49] Srinivasan, D. Energy demand prediction using GMDH networks. Neurocomputing, 2008; 72: 625–629. doi:10.1016/j.neucom.2008.08.006.

[50] Najafzadeh, M., Barani, G. A. Comparison of group method of data handling based genetic programming and back propagation systems to predict scour depth around bridge piers. Scientia Iranica, 2011; 18: 1207–1213. doi:10.1016/j.scient.2011.11.017.

[51] Najafzadeh, M., Barani, G. A., Hessami-Kermani, M. R. Group method of data handling to predict scour depth around vertical piles under regular waves. Scientia Iranica, 2013; 20: 406–413. doi:10.1016/j.scient.2013.04.005.

[52] Mehrara, M., Moeini, A., Ahrari, M., Erfanifard, A. RETRACTED: Investigating the efficiency in oil futures market based on GMDH approach. Expert Systems with Applications, 2009; 36: 7479–7483. doi:10.1016/j.eswa.2008.09.055.

[53] Sanchez, E., Shibata, T., Zadeh, L. A. Genetic Algorithms and Fuzzy Logic Systems, Soft Computing Perspectives. 1st ed. Singapore (SG): World Scientific; 1997. doi:10.1142/2896.

[54] Dag, O., Yozgatligil, C. GMDH: An R Package for Short Term Forecasting via GMDH-Type Neural Network Algorithms. The R Journal, 2016; 8: 379–386. doi:10.32614/RJ-2016-028.

[55] Dimopoulos, I., Chronopoulos, J., Chronopoulou-Sereli, A., Lek, S. Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). Ecological Modelling, 1999; 120: 157–165. doi:10.1016/S0304-3800(99)00099-X.

[56] Castillo, E., Gutierrez, J. M., Hadi, A. S. Sensitivity analysis in discrete Bayesian networks. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 1997; 27: 412–423. doi:10.1109/3468.594909.

[57] Liong, S.-Y., Lim, W.-H., Paudyal Guna, N. River Stage Forecasting in Bangladesh: Neural Network Approach. Journal of Computing in Civil Engineering, 2000; 14: 1–8. doi:10.1061/(ASCE)0887-3801(2000)14:1(1).