ARTICLE | Article in Press

Nano clay mitigation of wind induced soil loss across aggregate sizes through wind tunnel experiments and explainable machine learning

Abdirashid Ali Wehliye · Sema Kaplan

Received: 11 July 2026 / Accepted: 22 September 2026
© The Author(s) 2026

Abstract

This study aimed to investigate, under wind-tunnel conditions, the effects of different nano-clay application rates (0, 0.94, 1.88, and 3.75 g m^{-2}) on wind-induced soil loss across different aggregate sizes (0.5, 1.0, and 2.0 mm) and wind speeds (8, 10, and 12 m s^{-1}). In addition, the study aimed to model wind-induced soil loss processes using machine learning and explainable artificial intelligence methods. The wind-tunnel experiments revealed that nano-clay applications significantly reduced soil loss across all aggregate sizes and wind-speed conditions. The highest protective effect was obtained at the 3.75 g m^{-2} nano-clay rate; soil loss decreased by 80.8% in the 2 mm aggregates, by 93.2% in the 1 mm aggregates, and by 95.9% in the 0.5 mm aggregates. The protective effect of the nano-clay applications was found to be particularly pronounced at high wind speeds and in fine aggregates. Among the evaluated models, XGBoost showed the strongest development-stage performance under repeated nested group-aware cross-validation (R^2 = 0.835±0.064, RMSE = 11.896±2.229, MAE = 6.686±0.854) and was selected before independent holdout evaluation. On the group-held-out test set, the selected XGBoost model achieved R^2 = 0.882, RMSE = 13.298, and MAE = 9.443. The lower performance of the linear baseline suggested that the tree-based models captured predictive structure not adequately represented by a simple linear formulation. SHAP analysis indicated that nano-clay dose made the largest mean absolute contribution to the fitted model predictions, followed by wind speed and aggregate size. Overall, nano-clay substantially reduced wind-induced soil loss within the controlled experimental conditions examined here; broader practical applicability requires validation under field conditions.

Keywords Wind erosion · Nano-clay · Aggregate size · Wind speed · Machine learning

Introduction

Wind erosion is a critical global problem that substantially contributes to soil degradation and adversely affects agricultural productivity, soil fertility, and environmental sustainability[1,2]. By physically removing nutrient-rich topsoil, it leads to considerable soil loss and the depletion of essential nutrients, thereby reducing crop production and increasing vulnerability to drought stress[3]. This degradation increases food insecurity, particularly in

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.

https://doi.org/10.1038/s41598-026-73412-y

Published online: 27 September 2026

scientific reports

developing regions that depend on worn and degraded soils. Moreover, climate change and prolonged drought increase the risk of wind erosion and accelerate desertification by altering precipitation patterns and intensifying soil drought[4,5]. Wind-borne soil particles also impair air quality, posing risks to human health and ecosystem functions. Sustainable soil management practices, including conservation tillage and cover cropping, are of vital importance for mitigating these adverse effects and enhancing resilience under changing climatic conditions[6,7].

Factors affecting wind erosion include wind speed, soil texture, aggregate stability, moisture content, and surface cover[8]. While higher wind speeds increase the erosion potential, finer soil texture and lower surface stabilization make soils more susceptible[9,10]. Vegetation cover and microcrusts stabilize surfaces by binding soil particles and preventing saltation, thereby reducing soil loss[11–13]. Erosion control encompasses physical methods (protecting surface cover and microphytic crusts), chemical stabilizers (biopolymers and synthetic polymers), and biological approaches (supporting soil microbiota and aggregate stability)[14]. By increasing soil cohesion, surface stabilizers form protective crusts that resist wind-induced detachment and are ideally cost-effective and environmentally friendly[15,16]. Building on this classification, chemical methods such as biopolymers and synthetic polymers enhance interparticle cohesion and have shown promising erosion-reduction performance, although concerns remain regarding their durability and cost at field scale[17]. Biologically based approaches, such as microbially induced calcium carbonate precipitation and microbiota-based treatments, offer an environmentally friendly alternative but are often limited by sensitivity to environmental conditions and complex field implementation[18].

The increasing importance of nanotechnology in agriculture stems from the unique physicochemical properties of nanomaterials[19]. Distinguished by their high surface area, adsorption capacity, and layered structures, nano-clays show potential for soil conservation by stabilizing soil aggregates and modifying surface crust formation. By binding soil particles, nano-clays can enhance surface cohesion and reduce erosion by decreasing detachment caused by raindrops or wind[20]. Surface crust formation critically influences infiltration and runoff, and chemical and physical amendments also play important roles in regulating these processes[21]. The fine particles of nano-clays can function similarly to natural polymers by forming thin protective layers that increase soil strength and water-stable aggregates, reduce erosion, and improve infiltration[20]. Nano clay-based materials, either used alone or combined with industrial by-products such as blast furnace slag, have been shown to substantially reduce wind erosion of sandy soils while also improving water retention and mechanical strength[22]. Beyond nano-clay, chemically engineered polymer- and nanocellulose-based nanocomposites have likewise emerged as effective sandy-soil stabilizers. For instance, a chitin-assisted polymeric nanocomposite mulch markedly reduced wind-tunnel erosion rates while increasing the compressive and shear strength of sand-dune surfaces[23], and a related walnut-shell-derived nanocellulose–polymer binder achieved complete suppression of wind-tunnel soil loss together with pronounced gains in crust strength[24]. Together, these studies indicate that, alongside nano-clay, chemically engineered composite stabilizers constitute an actively developing alternative for erosion control in erodible sandy soils.

Data-driven approaches have increasingly complemented conventional methods for modelling and predicting wind erosion. Wind-tunnel experiments provide fundamental insights into erosion mechanics by offering controlled environments for examining the detachment and deposition patterns of soil particles at varying wind speeds[25]. However, traditional erosion models based on field measurements generally face challenges such as sparse spatio-temporal coverage and insufficient resolution to capture complex environmental interactions. Artificial intelligence (AI)-based prediction models are becoming increasingly important for the effective processing and analysis of complex, high-dimensional datasets by leveraging machine learning (ML), deep learning (DL), and hybrid soft-computing techniques[26,27]. These AI-based models can capture nonlinear relationships and may improve predictive performance when appropriately trained and validated. Data-driven approaches can complement conventional erosion-modelling frameworks by capturing nonlinear predictive relationships and integrating multiple predictor types. When adequately trained and independently validated, such models may provide useful support for erosion assessment and management[26–28].

Machine learning methods have been widely applied in soil sciences to predict soil properties and erosion susceptibility, generally providing high accuracy and robustness across diverse settings[29–31]. Machine and deep learning architectures have demonstrated superior performance in environmental modelling by capturing complex spatial and temporal patterns, particularly improving erosion risk mapping and LS-factor prediction[30,32]. Successes include high area-under-the-curve (AUC) values (> 0.90) and the accurate classification of erosion-prone areas, which facilitates better soil conservation planning[29,33]. However, gaps remain in the literature regarding model interpretability, the prevention of overfitting, and transferability across diverse soil types and climatic regimes[34]. Moreover, the integration of hybrid artificial intelligence models and comprehensive spatio-temporal datasets has not been sufficiently investigated. There is an urgent need for more explainable artificial intelligence frameworks and multi-scale validation to enhance practical adoption in soil management[30,34].

A distinctive feature of the present study is the integration of a controlled factorial wind-tunnel experiment with treatment-combination-aware machine-learning validation, allowing predictive patterns to be evaluated without separating replicates from the same experimental condition across validation partitions. Novel and sustainable surface-stabilization approaches are needed to mitigate wind erosion, and nano-clay has potential in this context because of its high surface area and binding properties. However, relatively few studies have examined nano-clay effects across contrasting wind speeds and aggregate sizes while also evaluating the predictive structure of the resulting soil-loss response. Accordingly, the present study experimentally quantified soil loss across different nano-clay doses, wind speeds, and aggregate-size classes and complemented these measurements with a group-aware machine-learning framework designed to characterize predictive relationships within the investigated experimental domain.

Materials and methods

Collection and general properties of the soil material

The soil material used in this study was collected from a wind-erosion-prone sandy agricultural area in the İncesu district of Kayseri province, at a depth of 0–10 cm, using a random sampling method to represent the surface soil. According to long-term climate data, the annual average temperature is 10.7 °C and the total annual precipitation is 389 mm[35], indicating a semi-arid climate characterized by widespread light-textured soils susceptible to wind erosion. The site therefore provides suitable material for evaluating wind erosion and surface-stabilization applications. After being brought to the laboratory, the samples were left to reach an air-dry state without exposure to direct sunlight or sudden temperature changes. Plant residues, root fragments, stones, and coarse organic remains were then removed by hand, improving the homogeneity of the soil material used in the wind-tunnel experiments.

Soil moisture content was determined by drying the samples at 105 °C until they reached constant weight (1.5%). Soil texture was determined by quantifying the sand, silt, and clay fractions using the hydrometer method[36], and the resulting particle-size composition was assigned to a USDA textural class. Soil reaction (pH) and electrical conductivity (EC) values were measured in a 1:2.5 soil: distilled-water suspension using a pH meter and an EC meter, respectively[37]. Organic matter content was determined by the Walkley–Black method[38], and lime ($CaCO_3$) content by the calcimetric method[38].

The general properties of the soil used in the study are given in Table 1. The pH of the soil was determined as 7.98 and its electrical conductivity (EC) as 0.612 dS m^{-1}. The organic matter and lime ($CaCO_3$) contents were 0.56 and 0.92%, respectively, and both parameters were found to be at low levels. The soil texture consisted largely of the sand fraction, with a total sand content of 92.2%. Within the sand fraction, the proportions of very coarse, coarse, fine, and very fine sand were 4%, 11%, 41%, and 22%, respectively. By contrast, the silt and clay contents were quite low, at 4.8% and 2%, respectively. Dry aggregate size distribution analysis was performed on a 250 g air-dried soil sample using a Retsch AS 200 Basic sieve shaker (1.0, 0.5, 0.25, 0.125, and 0.063 mm mesh sizes, 10 min duration, amplitude setting of 100), and the mass retained on each sieve was weighed. As a 0.84

Table 1 Some physical and chemical properties of the soil used in the study.

Soil properties	Value	Soil properties	Value
pH	7.98	Very coarse sand (%)	4
EC (dS m^{-1})	0.612	Coarse sand (%)	11
OM (%)	0.56	Medium sand (%)	14
$CaCO_3$ (%)	0.92	Fine sand (%)	41
Sand (%)	92.2	Very fine sand (%)	22
Silt (%)	4.80	<0.84 mm aggregates (%)	84.6
Clay (%)	2.00	Texture class	Sand

EC, electrical conductivity; OM, organic matter.

mm sieve was not included in the set, the erodible fraction (EF), defined as the percentage of aggregates < 0.84 mm[39,40], was estimated by log-linear interpolation between the 1.0 and 0.5 mm cumulative passing percentages. The estimated erodible fraction (EF) of the studied soil was 84.6%. In line with these properties, the study soil falls into the sand textural class according to the USDA texture classification and, owing to its low organic matter content, high sand content, and high amount of erodible aggregates, exhibits a structure that is highly susceptible to wind erosion.

Preparation of the aggregate fractions

Consistent with standard dry-sieving methodology in wind erosion research[40,41], the term "aggregates" is used throughout this study in an operational sense to refer to the size fractions obtained by dry sieving, regardless of their internal composition (whether predominantly composed of primary sand particles or of clay/silt-cemented structural units). Given the low clay and silt content of the studied soil, these fractions likely consist largely of loosely bound sand particles and weak sand aggregates rather than clay-cemented structural aggregates.

In order to determine the effect of aggregate size on wind-induced soil loss, three different aggregate fractions were used in the study. These fractions were separated from the bulk soil by dry sieving and selected as 0.5 mm, 1.0 mm, and 2.0 mm. These aggregate sizes were chosen to represent different levels of transportability and erosion susceptibility in sandy soils, spanning both below (0.5 mm) and above (1.0, 2.0 mm) the conventional 0.84 mm erodibility threshold used to define the erodible fraction (EF) in wind erosion studies.

Preparation of the experimental trays and placement of the soil into the trays

Experimental trays with a surface area of 0.16 m^2 (0.75 m × 0.21 m × 0.025 m) were used in the wind-tunnel experiments. For each aggregate fraction, the amount of soil required to fill the tray homogeneously was determined through preliminary experiments. Because the packing pattern and volumetric filling of the soils differed as the aggregate size changed, it was found inappropriate to use the same soil mass for each aggregate class. Accordingly, 6000 g of soil was weighed into each tray for the 2.0 mm aggregate fraction, 5500 g for the 1.0 mm aggregate fraction, and 5000 g for the 0.5 mm aggregate fraction (Fig. 1).

Nano-clay material

Commercially supplied kaolin-based nano-clay powder was used in the study. The nano-clay had an average particle size of approximately 800 nm and a purity of 99.9%. Kaolinite is a non-expanding 1:1 layered aluminosilicate mineral [$Al_2Si_2O_5(OH)_4$] characterized by a relatively low cation exchange capacity (5 cmol (+) kg^{-1}) and a BET specific surface area of approximately 30 m^2 g^{-1}. Its particle density was 2.6 g cm^{-3}, with a pH of 5.5 and a composition of 46.5% SiO_2 and 39.5% Al_2O_3. Owing to its small particle size and high surface area, it was assumed that the nano-clay could play a binding role between soil particles, form a thin coating on the surface, and reduce the detachment of aggregates under wind action.

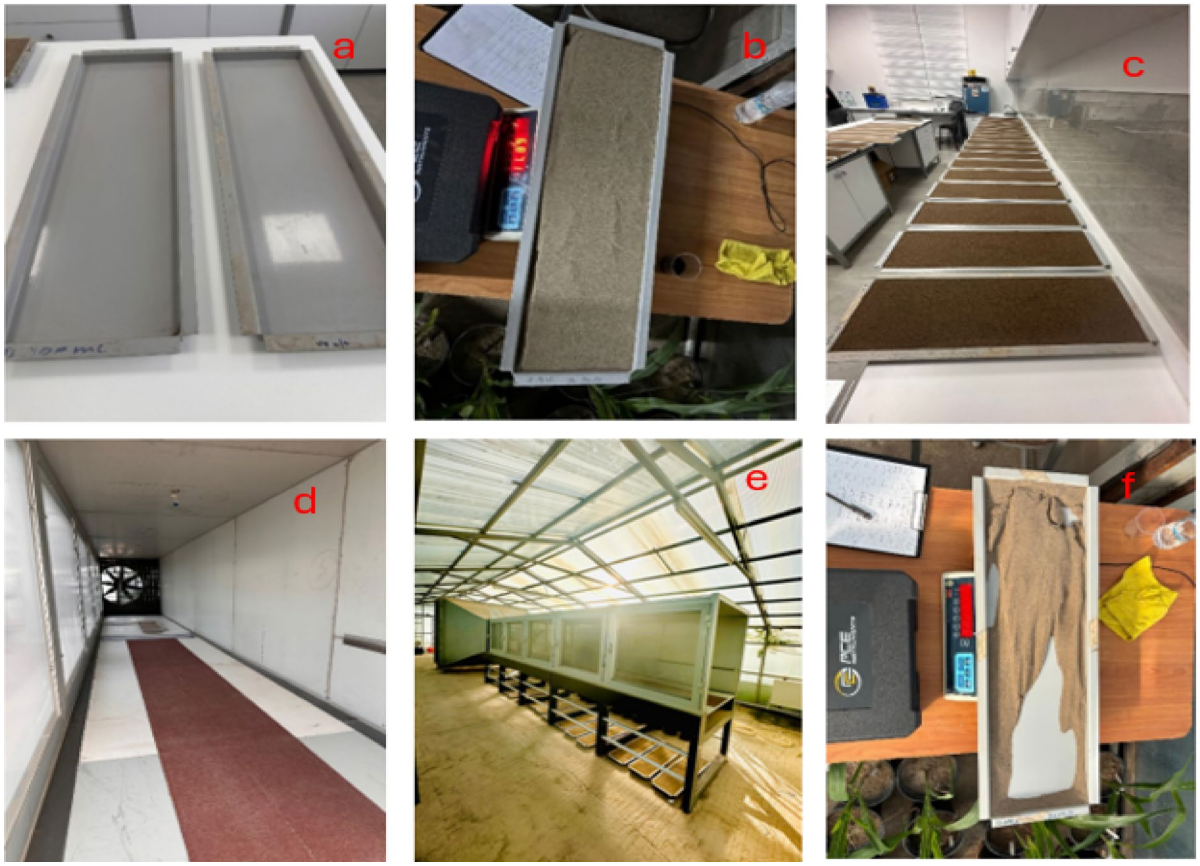


Fig. 1 Wind-tunnel experimental setup and measurement procedure used to determine the effect of nano-clay applications on soil loss under different aggregate (mm) sizes and wind speeds (m s^{-1}).

Preparation of the nano-clay suspensions

The nano-clay applications were carried out as suspensions prepared in distilled water. To support a more homogeneous dispersion of the nano-clay particles in the water and to ensure more uniform spreading over the surface during application, the suspensions were thoroughly stirred before use. A total of 300 mL of distilled water was used for each application. This volume was kept constant across all doses so that the difference between applications arose from the amount of nano-clay rather than the amount of water. In the control treatment, 300 mL of distilled water containing no nano-clay was applied.

When preparing the suspensions, the relevant amounts of nano-clay were first weighed on a precision balance. The weighed materials were added to 300 mL of distilled water and mixed until a homogeneous mixture was obtained. To prevent the suspension from settling during application, the mixture was re-shaken before each application. This step is important for ensuring that the nano-clay particles are distributed as evenly as possible over the tray surface.

Preliminary dose-determination study

Before proceeding to the main experiment, preliminary experiments were conducted to determine the appropriate nano-clay doses. In addition to the 0.94, 1.88, and 3.75 g m^{-2} doses used in the main study, the higher doses of 9.38, 18.75, and 37.5 g m^{-2} were also evaluated in the preliminary experiments.

The aim of the preliminary dose-determination study was not solely to identify the dose providing the highest surface stabilization. A surface-stabilization application suitable for use in agricultural areas should reduce soil loss without completely preventing plant emergence, should not cause excessive crusting, and should be economically and practically applicable. Therefore, in the preliminary experiments, post-application surface integrity, the appearance of the crust layer formed on the surface, surface homogeneity, and agricultural applicability were evaluated together.

It was observed that high-dose applications could provide more pronounced bonding and crust formation on the surface. However, such high doses were considered potentially limiting for germination and seedling emergence under agricultural production conditions. This concern was explored in a preliminary emergence screening (20 seeds per application), in which seedling emergence declined from 15 at 3.75 g m^{-2} to 11, 9, and 0 at 9.38, 18.75, and 37.50 g m^{-2}, respectively. This screening was used only to inform dose selection and was not treated as a replicated inferential experiment. In addition, the higher doses were considered less practical in terms of material requirement and potential field-scale applicability.

During the preliminary experiments, it was also observed that the homogeneity of the spray application had an important effect on the results. It was noted that non-uniform spray applications could lead to marked differences between replicates. Therefore, in the main experiments, a spray distance of approximately 25 cm, an application angle of 45°, and a standardized volume of 300 mL per tray were used consistently across all applications.

Wind-tunnel setup

The wind-erosion experiments were carried out in the wind-tunnel facility located at Erciyes University. The wind tunnel used consists of a fan section, a contraction cone, a working section, and an outlet section. The tunnel was designed to allow wind-erosion processes to be examined under controlled laboratory conditions (Fig. 2).

The test section of the wind tunnel is 5 m long. The inlet cross-section of the tunnel measures 0.80 × 0.90 m. The system operates with a high-capacity axial fan with a fan power of 7.5 kW. The fan can generate an air flow rate of approximately 60,000 m^3 h^{-1}. Wind speeds were set via a previously calibrated digital control panel, consistent with the growing use of digitally controlled airflow systems in aeolian erosion research to achieve more

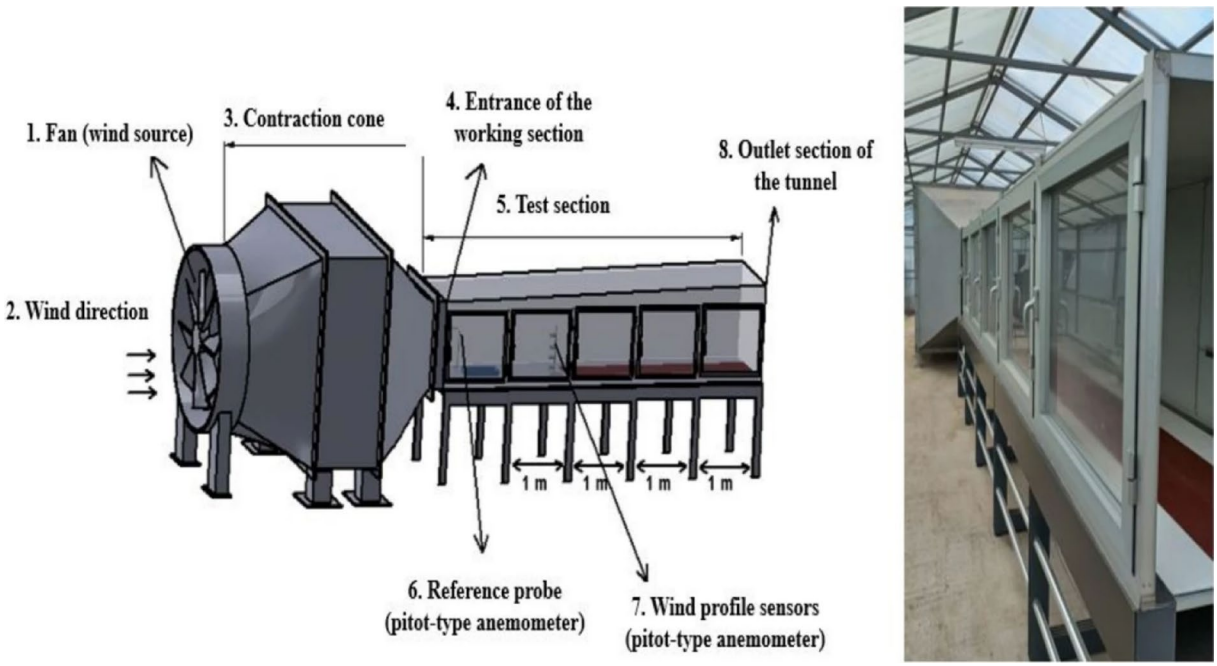


Fig. 2 Sections of the wind tunnel.

precise and reproducible experimental conditions[42–44]. The correspondence between the panel settings and the actual wind speed delivered to the sample tray was verified prior to each experimental run: airflow was measured 5 cm above the tray surface using a Testo 0635 2145 pitot-tube anemometer, and across three independent measurements the panel settings of 8, 10, and 12 m s^{-1} corresponded to actual mean wind speeds of 7.96 ± 0.04, 10.00 ± 0.05, and 11.99 ± 0.05 m s^{-1}, respectively. The thickness of the developing boundary layer was observed to vary between 0.10 and 0.24 m, corresponding to shear velocities between 0.146 and 0.504 m s^{-1}, consistent with previous calibration and validation studies conducted in the same facility[45].

Conduct of the wind-tunnel experiments

The trays, for which the nano-clay application had been completed and had been kept under laboratory conditions for 48 h, were made ready for the wind-tunnel experiments. Before the experiments, the initial weight of each tray was recorded using an analytical balance. The tray was then placed in the working section of the wind tunnel.

Each tray was exposed to wind for 5 min at the corresponding wind speed. Three wind speeds, namely 8, 10, and 12 m s^{-1}, were used in the experiments. This range closely matches the reference wind velocities (9 and 11 m s^{-1}) used in a wind-tunnel study on soils from the same region by Erci et al.,[14], as well as wind speeds reported in other regional field and wind-tunnel studies[46–48]. A 5-min exposure duration was used given the finite, shallow soil layer in the trays, which limits the erodible material available compared to an unlimited field-scale reservoir.

During the experiment, care was taken to keep the position of the tray within the tunnel similar across all treatments. In this way, differences that could arise from wind speed, surface conditions, and the position within the tunnel were minimized. After each treatment was completed, the tray was carefully removed from the tunnel and re-weighed to determine the amount of remaining soil.

The experiments were conducted in three replicates. The total number of experiments comprised 108 experimental units: 3 aggregate sizes × 4 nano-clay doses × 3 wind speeds × 3 replicates.

Determination of soil loss

In the wind-tunnel experiments, soil loss was determined based on the principle of mass loss. For each tray, the initial weight was recorded before the wind-tunnel test and the remaining weight after the test. The difference between the initial and final weights was taken as the amount of soil loss by wind.

Soil loss was calculated using the following equation:

$$\text{Soil loss}\,(\%) = [(\text{Initial weight} - \text{Final weight})/\text{Initial weight}] \times 100$$

Because directly comparing losses in grams was considered misleading, losses were expressed as percentages rather than absolute mass. The main reason for this is that the initial amounts of soil used to fill the tray differed across the aggregate sizes. The percentage loss values allowed a comparable evaluation among the treatments.

Field emission scanning electron microscopy (FE-SEM)

The surface morphology of untreated (control) and nano-clay-treated soil aggregates was examined using a field-emission scanning electron microscope (GeminiSEM 500, Zeiss, Germany) at the Technology Research and Application Center (TAUM), Erciyes University, Kayseri, Türkiye. Samples were imaged across three aggregate size fractions (0.5, 1.0, and 2.0 mm) and four nano-clay dose levels (0, 0.94, 1.88, and 3.75 g m^{-2}), with each sample examined at multiple magnifications ranging from 2000× to 30,000× to characterize both overall surface morphology and fine-scale particle interactions. Imaging was performed using an InLens secondary electron detector at an accelerating voltage of 3.00 kV and a working distance of approximately 3.4–3.6 mm. Before imaging, samples were sputter-coated with a thin conductive layer to prevent charging. Representative micrographs were selected from each treatment group for comparative analysis. To specifically investigate the microstructural mechanism underlying the observed erosion reduction, control and highest-dose (3.75 g m^{-2}) samples were

Table 2 Three-way ANOVA for the effects of aggregate size, nano-clay application rate, and wind speed, and their interactions, on soil loss (%).

Source	df	SS	MS	F	*p*
Aggregate size (A)	2	104.30	52.15	18.58	<0.001
Wind speed (W)	2	24685.00	12342.50	4398.13	<0.001
Nano-clay rate (N)	3	77739.13	25913.04	9233.86	<0.001
A × W	4	979.19	244.80	87.23	<0.001
A × N	6	4232.39	705.40	251.36	<0.001
W × N	6	7621.92	1270.32	452.67	<0.001
A × W × N	12	1075.13	89.59	31.93	<0.001
Error	72	202.05	2.81		

A, aggregate size; N, nano-clay application rate; W, wind speed; df, degrees of freedom; SS, sum of squares; MS, mean square. The error term had 72 degrees of freedom. All reported main and interaction effects had $p<0.001$.

Table 3 Effects of nano-clay application rates and wind speeds on soil loss in 2 mm aggregates (%).

Nano Clay (g m^{-2})	Wind Speed (m s^{-1})			Mean
	8	Mean	12	
0	18.31 ± 1.09^{ef}	72.37 ± 0.63^{b}	92.10 ± 0.22^{a}	60.93 ± 33.09^{A}
0.94	7.16 ± 1.76^{gh}	20.91 ± 2.18^{e}	57.08 ± 6.78^{c}	28.38 ± 22.63^{B}
1.88	6.97 ± 0.70^{gh}	11.70 ± 1.30^{fg}	47.29 ± 0.39^{d}	21.99 ± 19.10^{C}
3.75	2.84 ± 0.21^{h}	8.35 ± 0.33^{gh}	23.83 ± 3.58^{e}	11.67 ± 9.59^{D}
Mean	8.82 ± 6.07^{C}	28.33 ± 27.01^{B}	55.07 ± 25.86^{A}	

Values are expressed as mean ± standard deviation ($n=3$). Means followed by different lowercase letters (a–h) are significantly different among the nano-clay rate × wind speed interaction means. In the marginal means, different uppercase letters in the "Mean" column indicate significant differences among nano-clay application rates, and different uppercase letters in the "Mean" row indicate significant differences among wind speeds, according to the Tukey multiple comparison test ($p\leq0.01$).

Table 4 Effects of nano-clay application rates and wind speeds on soil loss in 1 mm aggregates (%).

Nano Clay (g m^{-2})	Wind Speed (m s^{-1})			Mean
	8	10	12	
0	59.38 ± 1.33^{b}	92.71 ± 0.76^{a}	95.74 ± 0.60^{a}	82.61 ± 17.49^{A}
0.94	5.25 ± 0.36^{ef}	13.67 ± 2.01^{d}	61.84 ± 3.45^{b}	26.92 ± 26.52^{B}
1.88	3.65 ± 0.16^{ef}	6.61 ± 0.29^{e}	38.50 ± 0.42^{c}	16.25 ± 16.74^{C}
3.75	1.59 ± 0.12^{f}	2.15 ± 0.05^{f}	12.99 ± 0.58^{d}	5.58 ± 5.58^{D}
Mean	17.47 ± 25.32^{C}	28.78 ± 38.80^{B}	52.27 ± 31.86^{A}	

Values are expressed as mean ± standard deviation ($n=3$). Means followed by different lowercase letters (a–h) are significantly different among the nano-clay rate × wind speed interaction means. In the marginal means, different uppercase letters in the "Mean" column indicate significant differences among nano-clay application rates, and different uppercase letters in the "Mean" row indicate significant differences among wind speeds, according to the Tukey multiple comparison test ($p\leq0.01$).

compared at 10,000× magnification across all three aggregate size classes. The FE-SEM assessment was qualitative and was used to compare representative surface features; no quantitative image-based morphometric analysis was performed.

Statistical analysis, machine learning, and explainable artificial intelligence analyses

To evaluate the effects of wind speed, aggregate size, and nano-clay application dose on soil loss, a three-way analysis of variance (ANOVA) was first performed (Table 2). Wind speed, aggregate size, and nano-clay dose were treated as fixed factors. The model included the three main effects, all two-way interactions (wind speed × aggregate size, wind speed × nano-clay dose, and aggregate size × nano-clay dose), and the three-way interaction among the factors. Because the three-way interaction was significant, post-hoc interpretation was presented separately within each aggregate-size class. For Tables 3, 4 and 5, Tukey multiple-comparison tests were used

Table 5 Effects of nano-clay application rates and wind speeds on soil loss in 0.5 mm aggregates (%).

Nano Clay (g m^{-2})	Wind Speed (m s^{-1})			Mean
	8	10	12	
0	66.19 ± 2.03^{c}	94.19 ± 1.58^{b}	97.60 ± 0.59^{a}	85.99 ± 14.98^{A}
0.94	3.25 ± 1.00^{gh}	9.05 ± 0.17^{f}	50.91 ± 0.74^{d}	21.07 ± 22.53^{B}
1.88	2.15 ± 0.31^{gh}	4.32 ± 0.35^{g}	30.85 ± 1.36^{e}	12.44 ± 13.86^{C}
3.75	1.10 ± 0.25^{h}	1.47 ± 0.20^{gh}	8.09 ± 0.17^{f}	3.55 ± 3.41^{D}
Mean	18.17 ± 28.99^{C}	27.26 ± 40.47^{B}	46.86 ± 34.45^{A}	

Values are expressed as mean ± standard deviation ($n = 3$). Means followed by different lowercase letters (a–h) are significantly different among the nano-clay rate × wind speed interaction means. In the marginal means, different uppercase letters in the "Mean" column indicate significant differences among nano-clay application rates, and different uppercase letters in the "Mean" row indicate significant differences among wind speeds, according to the Tukey multiple comparison test ($p \leq 0.01$).

to compare the nano-clay dose × wind-speed cell means and the corresponding marginal means at $\alpha = 0.01$. The Tukey tests and associated letter groupings were obtained using JMP 18.0 (SAS Institute Inc., Cary, NC, USA).

The assumptions of the three-way ANOVA were evaluated using model residuals. Residual normality was assessed using the Shapiro–Wilk test together with a normal Q–Q plot, while homogeneity of variance across the 36 experimental treatment combinations was examined using the median-centered Levene test (Brown–Forsythe form). Standardized residuals were also plotted against fitted values to identify possible systematic patterns or changes in residual dispersion. Formal diagnostics indicated departures from both residual normality and variance homogeneity. Therefore, the robustness of the conventional ANOVA inference was further examined by re-evaluating the same full-factorial Type II model using HC3 heteroskedasticity-robust covariance estimation. The HC3 analysis was used as a sensitivity analysis rather than as a replacement for the prespecified classical ANOVA.

Machine-learning analyses were subsequently conducted to characterize the predictive relationships between the experimental factors and soil loss. Wind speed (m s^{-1}), aggregate size (mm), and nano-clay dose (g m^{-2}) were used as predictors, and soil loss (%) was used as the response variable. Three modelling approaches were evaluated: Linear Regression, Random Forest, and Extreme Gradient Boosting (XGBoost). Linear Regression was included as an untuned baseline model, whereas Random Forest and XGBoost were used to capture potentially nonlinear predictive relationships and higher-order structure among the predictors.

To minimize information leakage among experimental replicates, each unique combination of wind speed, aggregate size, and nano-clay dose was treated as a single grouping unit throughout machine-learning validation. The complete dataset comprised 108 observations representing 36 unique treatment combinations, with three replicates per combination. All replicates belonging to the same treatment combination were therefore kept together during data partitioning and cross-validation.

A group-aware 75:25 split was first applied, resulting in a development dataset containing 81 observations from 27 treatment combinations and an independent holdout dataset containing 27 observations from nine treatment combinations. No treatment combination was shared between the development and holdout datasets. The holdout dataset was isolated before any model fitting, hyperparameter optimization, model comparison, or model selection was performed.

All model development was conducted exclusively within the development dataset. Predictive performance was evaluated using repeated nested group-aware cross-validation, consisting of five outer folds repeated ten times, yielding 50 outer validation partitions for each model. Within each outer-training partition, Random Forest and XGBoost hyperparameters were optimized using four-fold group-aware inner cross-validation, with RMSE used as the optimization criterion. Model selection was based solely on development-stage validation performance, specifically the lowest mean outer-validation RMSE. After the model type and hyperparameter configuration had been selected and fixed, the selected model was refitted using the complete development dataset and evaluated once on the untouched independent holdout dataset.

For Random Forest, the predefined search grid included 500, 1000, and 1500 trees (n_estimators); maximum tree depths (max_depth) of None, 5, and 10; and minimum leaf sizes (min_samples_leaf) of 1 and 2. For XGBoost, the search grid included 500 and 1000 boosting iterations (n_estimators), learning rates of 0.03 and 0.05, maximum tree depths of 2 and 3, and subsampling fractions of 0.80 and 1.00. Hyperparameter selection was performed exclusively within the inner cross-validation loop. The final Random Forest configuration selected from the development-stage optimization consisted of 1500 trees, max_depth = None, and min_samples_leaf = 1. The final XGBoost configuration consisted of 1000 boosting iterations, a learning rate of 0.03, a maximum tree depth of 2, and a subsampling fraction of 0.80. Linear Regression was retained as an untuned baseline model.

For each complete cross-validation repeat, out-of-fold predictions from the five outer folds were pooled, and R^2, RMSE, and MAE were calculated. Cross-validation performance was reported as the mean ± standard deviation across the ten repeated out-of-fold evaluations. Fold-level performance statistics were additionally retained as a supplementary stability diagnostic. The independent holdout dataset was evaluated only once after model selection had been completed.

Variance Inflation Factor (VIF) values were calculated for wind speed, aggregate size, and nano-clay dose as a diagnostic of multicollinearity among the predictors.

Explainable artificial intelligence methods were used to further examine fitted-model behaviour. Shapley Additive Explanations (SHAP) were calculated to quantify the contribution of each predictor to individual model predictions. SHAP summary and dependence plots were used to visualize the magnitude and direction of these model-based contributions. SHAP values were interpreted as contributions to fitted predictions rather than as direct measures of causal or physical importance.

Partial Dependence Plots (PDPs) were generated to examine marginal patterns in model predictions across the experimental ranges of the predictors. Two-dimensional PDPs were additionally used to describe joint fitted-model response surfaces for pairs of predictors. Accumulated Local Effects (ALE) analyses were performed as a complementary interpretation method that summarizes average local changes in model predictions over the observed predictor distribution. PDP and ALE results were therefore interpreted as descriptions of fitted-model behaviour and not as estimates of causal effects.

ALE profiles were generated for wind speed, aggregate size, and nano-clay dose using the PyALE implementation. Because the experimental predictors were evaluated at a limited number of discrete treatment levels, lines connecting ALE estimates were used only as visual guides, while continuous colour transitions in the two-dimensional PDP surfaces were treated as model-based visualizations between experimentally represented levels. Neither representation was interpreted as direct experimental evidence at untested intermediate values, a continuous causal response function, or evidence of specific threshold values.

An exploratory scenario analysis was conducted using the selected explanatory model and was restricted to the 36 factorial combinations actually represented in the experiment (3 wind speeds × 3 aggregate sizes × 4 nano-clay doses). Model-predicted soil loss was calculated for each experimentally represented combination, and the combinations associated with the lowest and highest fitted predictions were identified. The analysis was used only to summarize fitted-model behaviour within the observed factorial domain; the identified combinations were not interpreted as experimentally validated optima or as recommendations beyond the tested conditions.

For visualization of model predictions at the experimentally represented treatment combinations, heatmaps were generated separately for each aggregate-size class using the fitted explanatory model. Each heatmap displayed predictions for the 12 observed wind-speed × nano-clay combinations within that aggregate-size class. No untested intermediate factor levels were introduced in these heatmaps.

For the model-interpretation analyses, an explanatory version of the selected machine-learning model was fitted to the complete dataset using the model type and hyperparameter configuration determined during development-stage model selection. This full-data model was used only to generate SHAP, PDP, ALE, exploratory scenario, and heatmap-based model interpretations; predictive performance was not estimated from this model. All reported measures of predictive performance were derived exclusively from the repeated nested group-aware cross-validation and the independent holdout evaluation described above.

The three-way ANOVA, ANOVA assumption diagnostics, HC3 robustness analysis, machine-learning modelling, and explainable artificial intelligence analyses were performed in Python. Tukey multiple-comparison tests used for the treatment-mean groupings in Tables 3, 4 and 5 were performed in JMP 18.0 (SAS Institute Inc., Cary, NC, USA).

Results

Effect of nano-clay applications and wind speeds on soil loss

The three-way ANOVA showed that the main effects of aggregate size, nano-clay application rate, and wind speed, as well as all of their two-way and three-way interactions, had statistically significant effects on soil loss (all $p<0.001$) (Table 2). Because the three-way aggregate size × nano-clay rate × wind-speed interaction was significant, the effects of nano-clay rate and wind speed on soil loss varied with aggregate size. Therefore, the treatment means are presented separately for each aggregate size in Tables 3 and 4, and 5 to allow a clear interpretation of these interaction effects.

Diagnostic assessment of the full-factorial ANOVA indicated departures from residual normality (Shapiro–Wilk $W=0.7847$, $p<0.001$) and homogeneity of variance across the 36 treatment combinations (Brown–Forsythe/median-centered Levene statistic $=2.1355$, $p=0.00335$). Accordingly, the inferential robustness of the same Type II model was examined using HC3 heteroskedasticity-robust covariance estimation. All seven main and interaction effects retained the same significance decisions under HC3 inference, yielding 7/7 (100%) concordance with the classical ANOVA.

The effects of nano-clay applications on percentage soil loss at different aggregate sizes are presented in Tables 3, 4 and 5. In soils with a 2 mm aggregate size, the effects of nano-clay application rate, wind speed, and their interaction on soil loss were all statistically significant ($p\leq0.01$) (Table 3). In general, an increase in wind speed significantly increased soil loss, whereas the nano-clay applications reduced it. The highest soil loss was obtained in the control at a wind speed of 12 m s^{-1} (92.10%), while the lowest value was obtained with the 3.75 g m^{-2} nano-clay application at a wind speed of 8 m s^{-1} (2.84%). According to the mean values, soil loss decreased as the nano-clay dose increased, and the lowest mean soil loss was obtained with the 3.75 g m^{-2} nano-clay application (11.67%). When the wind speeds were examined, the highest mean soil loss was determined at 12 m s^{-1} (55.07%) and the lowest at 8 m s^{-1} (8.82%).

Table 4 presents the effects of different nano-clay doses and wind speeds on soil loss at a 1 mm aggregate size. The effects of nano-clay dose, wind speed, and the nano-clay × wind-speed interaction on soil loss were all statistically significant ($p\leq0.01$). The findings show that soil loss increased significantly with increasing wind speed, whereas the nano-clay applications significantly reduced this loss. The highest soil loss was determined in the control treatment at a wind speed of 12 m s^{-1} (95.74%), while the lowest soil-loss value was obtained with the 3.75 g m^{-2} nano-clay application at a wind speed of 8 m s^{-1} (1.59%). According to the mean values, the control treatment had the highest soil loss at 82.61%, whereas the 3.75 g m^{-2} nano-clay application showed the lowest value at 5.58%. In terms of wind speeds, the mean soil loss was 17.47% at 8 m s^{-1} and rose to 52.27% at 12 m s^{-1}. The findings obtained indicate that the nano-clay application increased the resistance of the aggregates to wind action and significantly limited soil loss, particularly at high speeds.

The soil-loss values occurring under different nano-clay amounts and different wind-speed applications in soils with a 0.5 mm aggregate size are given in Table 5. When Table 5 is examined, the effects of nano-clay amount, wind speed, and the nano-clay × wind-speed interaction on soil loss were found to be statistically significant ($p\leq0.01$). According to the research findings, the nano-clay applications significantly reduced wind-induced soil loss despite the decrease in aggregate size. The highest soil loss was determined in the control treatment at a wind speed of 12 m s^{-1} (97.60%), while the lowest value was found with the 3.75 g m^{-2} nano-clay application at a wind speed of 8 m s^{-1} (1.10%). Considering the mean values, in the control treatment the soil loss was 85.99, whereas

with the 3.75 g m^{-2} nano-clay application this value fell to 3.55%. Similarly, the mean soil loss rose from 18.17 to 46.86% as the wind speed increased. In particular, the 1.88 and 3.75 g m^{-2} nano-clay applications were observed to reduce soil loss substantially at all wind speeds, and it was determined that nano-clay provides surface stabilization in fine-aggregate soils, thereby creating effective protection against wind erosion.

Notably, the wind tunnel results (Tables 3, 4 and 5) showed that the relative protective effect of nano-clay treatment increased with decreasing aggregate size, with the greatest reduction in soil loss observed for the 0.5 mm fraction. However, SEM-based bridging evidence (Fig. 3) was visually less distinct at this aggregate size compared with the 1.0 and 2.0 mm fractions.

Performance of the machine learning models

Repeated nested group-aware cross-validation showed that XGBoost yielded the strongest development-stage predictive performance (Table 6), with a pooled out-of-fold R^2 of 0.835 ± 0.064, RMSE of 11.896 ± 2.229, and MAE of 6.686 ± 0.854. Random Forest showed intermediate performance (R^2 = 0.683 ± 0.051, RMSE = 16.696 ± 1.354, and MAE = 9.830 ± 1.074), whereas Linear Regression produced the lowest cross-validated performance (R^2 = 0.441 ± 0.036, RMSE = 22.240 ± 0.711, and MAE = 18.912 ± 0.528).

Based exclusively on development-stage validation performance, XGBoost was selected and its configuration was fixed before the independent holdout dataset was evaluated. On the group-held-out test set, the selected XGBoost model achieved an R^2 of 0.882, RMSE of 13.298, and MAE of 9.443. Random Forest yielded a higher R^2 (0.968) and lower RMSE (6.914) and MAE (4.942) on this particular holdout set; however, these test-set results were not used retrospectively for model selection. The difference between the cross-validation and holdout rankings indicates sensitivity of model performance to the specific treatment combinations included in the evaluation set and supports a cautious interpretation of predictive generalizability beyond the experimental domain.

Variance Inflation Factor analysis showed no evidence of multicollinearity among the three predictors. Wind speed, aggregate size, and nano-clay dose each had a VIF of 1.00, consistent with the balanced factorial structure of the experiment.

Observed and predicted soil-loss values

The relationship between observed and predicted soil-loss values for the selected XGBoost model on the independent group-held-out test dataset is shown in Fig. 4. The predictions broadly followed the observed response pattern, although deviations were evident for some intermediate and high soil-loss observations. In particular, several predictions exceeded the physical 100% response range; these values were retained without truncation when calculating model-performance metrics and are indicated at the upper plotting boundary in Fig. 4. Overall, the independent-test results should therefore be interpreted together with the reported R^2, RMSE, and MAE values rather than as evidence of uniformly accurate prediction across all treatment combinations.

SHAP-based predictor contributions and model interpretation

SHAP analysis of the selected XGBoost model showed that nano-clay dose had the largest mean absolute contribution to model predictions (mean |SHAP| = 22.661), followed by wind speed (13.866) and aggregate size (2.717) (Fig. 5). Higher nano-clay doses were generally associated with negative SHAP values, indicating contributions toward lower fitted soil-loss predictions, whereas higher wind speeds were associated with positive SHAP values, indicating contributions toward higher predictions. Aggregate size showed a comparatively smaller contribution to model output. These SHAP values quantify the magnitude and direction of predictor contributions to the fitted XGBoost predictions and should not be interpreted as direct measures of causal, mechanistic, or physical importance in the erosion process.

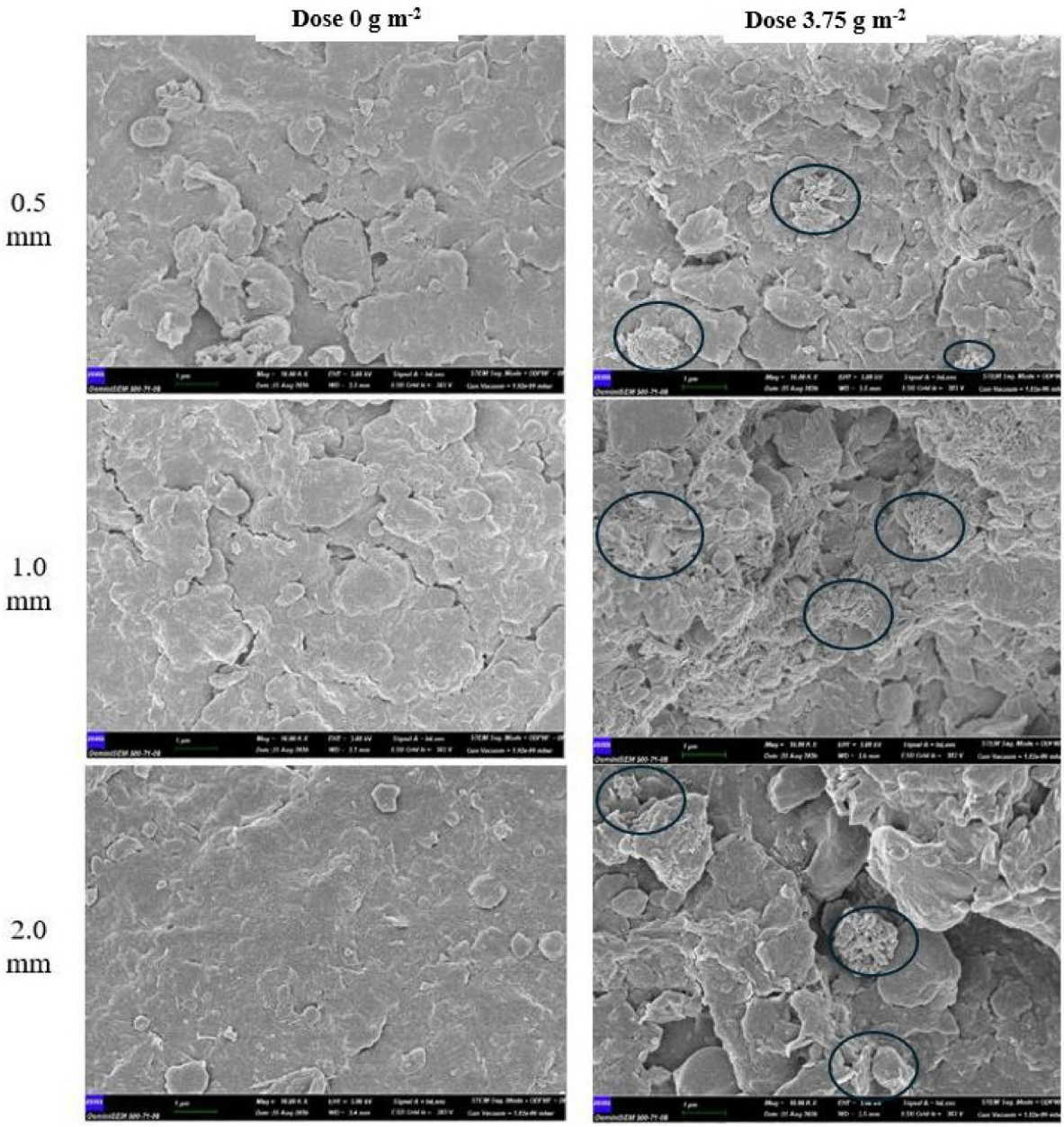


Fig. 3 Field-emission scanning electron microscopy (FE-SEM) images comparing untreated control (Dose 0) and highest-dose nano-clay-treated (3.75 g m^{-2}) sand aggregates across three aggregate size fractions: (**a**, **b**) 0.5 mm, (**c**, **d**) 1.0 mm, and (**e**, **f**) 2.0 mm. Left column: control samples; right column: 3.75 g m^{-2} samples. The circles highlight representative regions showing apparent particle bridging and clustering in the treated samples. These qualitative features are presented as visual observations and were not quantified morphometrically. Images were acquired at 10,000× magnification using an InLens secondary electron detector at an accelerating voltage of 3.00 kV.

Joint modelled relationship between wind speed and nano-clay application

The joint fitted-model relationship between wind speed and nano-clay dose was examined using a two-dimensional partial dependence plot (PDP) (Fig. 6). The horizontal axis represents the experimentally imposed wind speeds (8, 10, and 12 m s^{-1}), whereas the vertical axis represents nano-clay dose. The color scale and contour levels indicate the magnitude of soil loss predicted by the fitted XGBoost model, with higher values corresponding

Table 6 Predictive performance of the machine-learning models based on repeated nested group-aware cross-validation and independent holdout testing.

Model	Linear regression	Random forest	XGBoost†
Repeated nested CV R^2	0.441 ± 0.036	0.683 ± 0.051	**0.835 ± 0.064**
Repeated nested CV RMSE	22.240 ± 0.711	16.696 ± 1.354	**11.896 ± 2.229**
Repeated nested CV MAE	18.912 ± 0.528	9.830 ± 1.074	**6.686 ± 0.854**
Independent test R^2	0.779	**0.968**	0.882
Independent test RMSE	18.17	**6.914**	13.298
Independent test MAE	15.406	**4.942**	9.443

Values for repeated nested cross-validation are reported as mean ± standard deviation across ten complete out-of-fold evaluations. Independent-test metrics were calculated once on the untouched group-held-out dataset and are therefore reported without standard deviations. †XGBoost was selected exclusively on the basis of development-stage nested cross-validation performance before the independent test dataset was evaluated.

Significant values are in [bold].

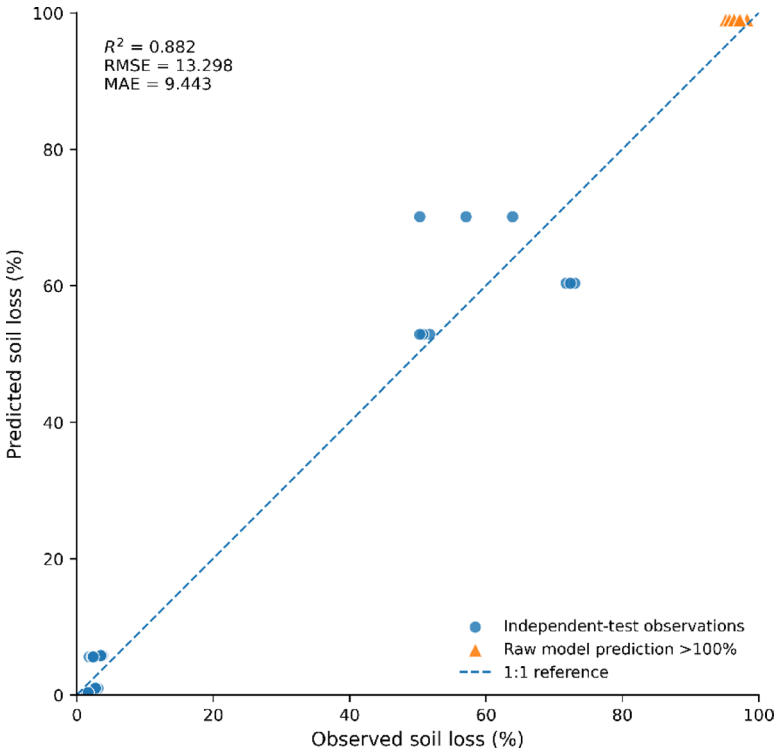


Fig. 4 Observed versus predicted soil-loss values for the selected XGBoost model on the independent group-held-out test dataset. The dashed line represents the 1:1 reference line. Triangle symbols denote raw model predictions exceeding 100%; these predictions are displayed at the upper plotting boundary for visualization only and were retained without truncation when calculating model-performance metrics.

to greater predicted soil loss and lower values corresponding to lower predicted soil loss. Across the investigated experimental range, predicted soil loss generally increased with increasing wind speed and decreased with increasing nano-clay dose. The highest predicted soil-loss values occurred at 12 m s^{-1} in the absence of nano-clay, whereas substantially lower predicted values were observed at the higher nano-clay doses. The response surface further indicated that the modelled reduction associated with nano-clay application was particularly evident under the higher wind-speed conditions. The value of 10 m s^{-1} represents one of the three experimentally imposed wind-speed levels and should not be interpreted as a model-derived threshold or critical transition point. Overall, the PDP describes the joint response surface learned by the XGBoost model and should not be interpreted as direct evidence of causal effects.

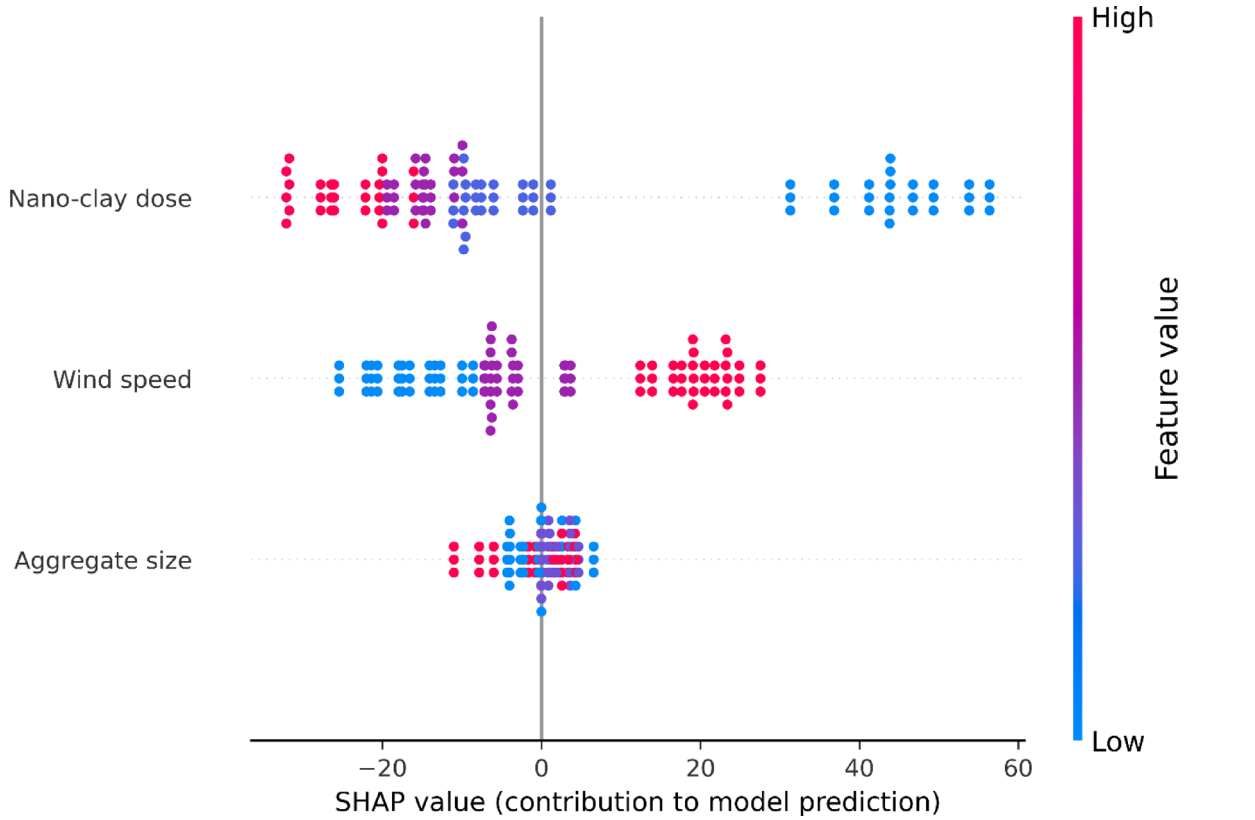


Fig. 5 SHAP summary plot for the selected XGBoost model showing the magnitude and direction of predictor contributions to fitted soil-loss predictions. Each point represents an observation, the horizontal position indicates the SHAP contribution to the model prediction, and color represents the corresponding predictor value from low to high. Positive SHAP values indicate contributions toward higher predicted soil loss, whereas negative values indicate contributions toward lower predicted soil loss. SHAP values describe fitted-model contributions and should not be interpreted as direct causal or physical effects.

Predicted soil loss at different aggregate sizes

The heatmaps showed that predicted soil loss generally increased with wind speed and decreased with increasing nano-clay dose across all aggregate sizes. The highest predicted losses occurred in the absence of nano-clay, particularly at 12 m s^{-1}, whereas lower predicted values were observed at the higher nano-clay doses. Predicted soil loss also tended to be lower for the larger aggregate class. These heatmaps provide a visual summary of fitted-model predictions across the investigated experimental combinations and should not be interpreted as independent evidence of treatment significance or physical mechanisms.

Exploratory model-based scenario ranking was restricted to the 36 treatment combinations represented in the experiment. Within this domain, the lowest raw fitted XGBoost prediction was obtained at 10 m s^{-1} wind speed, 0.5 mm aggregate size, and 3.75 g m^{-2} nano-clay (predicted soil loss = −0.796%), whereas the highest raw fitted prediction occurred at 12 m s^{-1}, 0.5 mm, and 0 g m^{-2} nano-clay (predicted soil loss = 101.035%). Because the regression model was not constrained to the physical 0–100% response range, these boundary-exceeding values were retained to preserve the model output but were interpreted only as a ranking of fitted scenarios. They should not be regarded as physically realizable soil-loss percentages, experimentally observed values, or validated optimum/worst-case conditions.

The supplementary PDP and ALE analyses were broadly consistent with the patterns observed in the main model interpretation. Across the fitted model, increasing nano-clay dose was associated with lower predicted soil loss, whereas increasing wind speed was associated with higher predicted soil loss. Aggregate size showed a

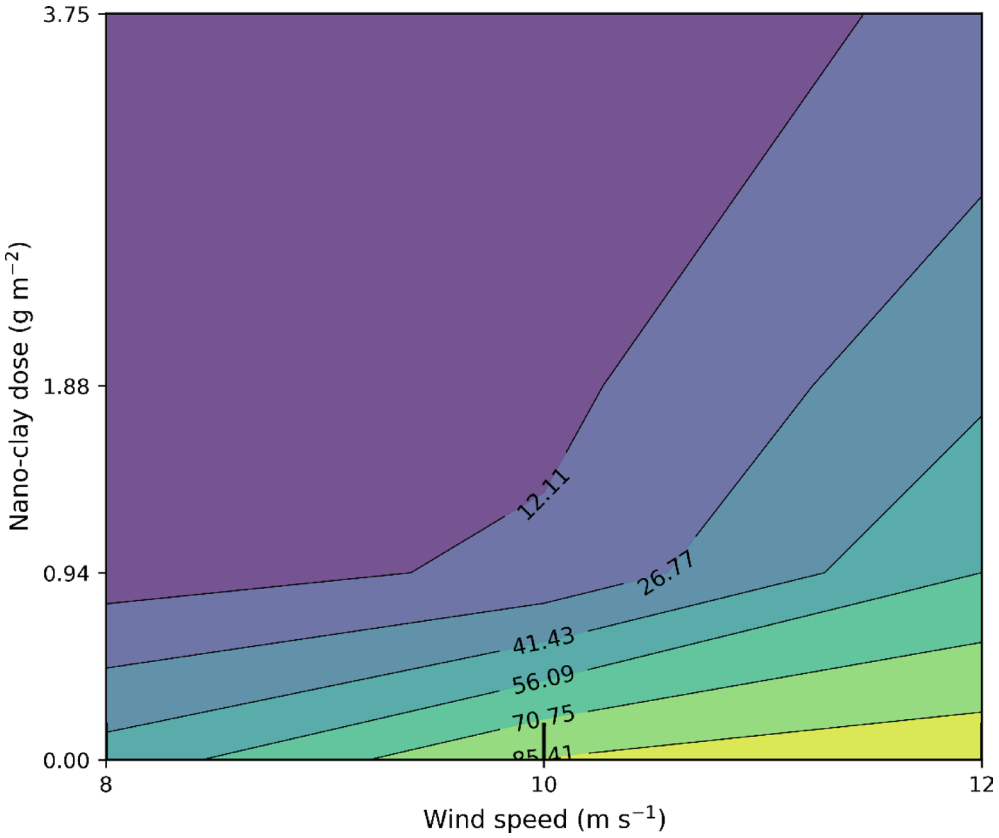


Fig. 6 Two-dimensional partial dependence plot of the fitted XGBoost model showing the joint relationship of wind speed and nano-clay dose with predicted soil loss. Wind speed is shown on the x-axis and nano-clay dose on the y-axis. Colors and contour levels represent model-predicted soil loss, with higher values indicating greater predicted soil loss. Wind speeds of 8, 10, and 12 m s^{-1} correspond to the experimentally imposed treatment levels; the value of 10 m s^{-1} does not represent an independently identified threshold. The plot describes fitted-model behaviour and should not be interpreted as direct evidence of causal effects.

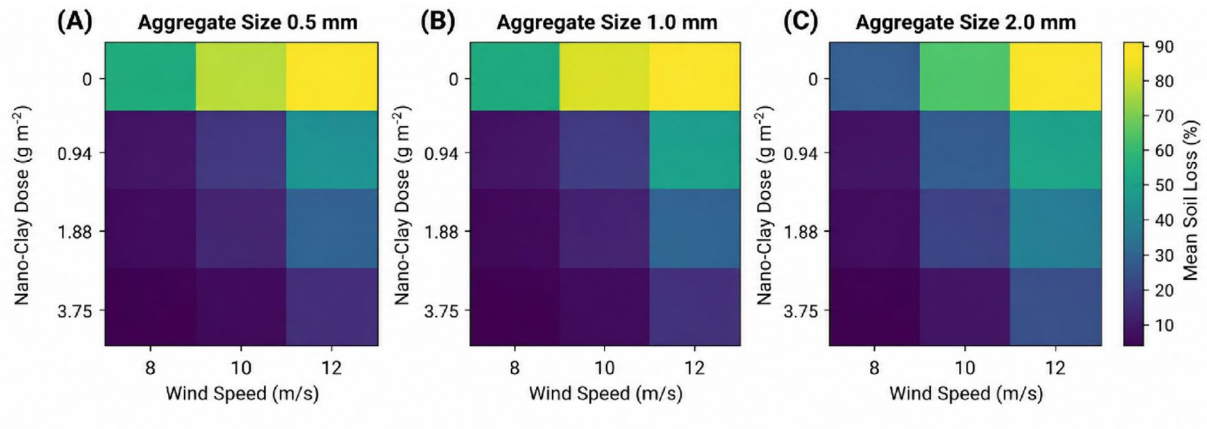


Fig. 7 Heatmap visualization of model-predicted soil loss under the experimentally represented wind-speed and nano-clay dose combinations for aggregate sizes of (**A**) 0.5 mm, (**B**) 1.0 mm, and (**C**) 2.0 mm. The heatmaps summarize fitted XGBoost predictions and should not be interpreted as independent experimental evidence.

comparatively weaker modelled effect on the predicted response. These supplementary analyses provide complementary descriptions of fitted-model behaviour and should not be interpreted as direct evidence of causal effects. The two-dimensional PDPs in the supplementary material represent the remaining pairwise predictor relationships, namely wind speed × aggregate size and nano-clay dose × aggregate size, complementing the wind speed × nano-clay dose relationship presented in Fig. 6.

Discussion

In this study, the wind-tunnel experiments conducted under different aggregate sizes (0.5, 1, and 2 mm) and wind speeds (8, 10, and 12 m s^{-1}) revealed that nano-clay applications significantly reduced soil loss. Soil loss increased markedly with increasing wind speed across all aggregate sizes; by contrast, the nano-clay applications largely limited this adverse effect.

The effect of wind speed on soil loss was clearly demonstrated by the experimental findings. In the 2 mm aggregates, soil loss in the control treatment was 18.31% at 8 m s^{-1} and rose to 92.10% at 12 m s^{-1}. Under the same conditions, soil loss increased from 59.38 to 95.74% in the 1 mm aggregates and from 66.19 to 97.60% in the 0.5 mm aggregates. Under the controlled wind-tunnel conditions used here, increasing wind speed caused substantially greater soil loss, in agreement with wind-erosion research[49]. This response is physically consistent with the greater aerodynamic shear exerted at higher wind speeds, which increases the likelihood of particle entrainment once the relevant initiation threshold is exceeded.

In the 2 mm aggregates, the mean soil loss with the 3.75 g m^{-2} nano-clay application fell to 11.67% compared with the control (60.93%), corresponding to an 80.8% reduction. This reduction was 93.2% in the 1 mm aggregates (from 82.61 to 5.58%) and 95.9% in the 0.5 mm aggregates (from 85.99 to 3.55%). These results show that soil loss decreased gradually and consistently with increasing nano-clay dose. This experimental pattern was also captured by the fitted machine-learning model: the two-dimensional PDP showed a directionally consistent trend, with predicted soil loss decreasing as nano-clay dose increased across the investigated wind-speed levels (Fig. 6).

Nano-clay may influence soil-surface behaviour through several complementary mechanisms. Owing to its small particle size and high specific surface area, nano-clay particles can occupy spaces between larger particles and increase the number of interparticle contact sites, potentially promoting adhesion and surface cohesion. After drying, these interactions may contribute to formation of a more cohesive surface layer that is less susceptible to aerodynamic detachment. Such mechanisms are consistent with the observed reductions in soil loss and with previous reports for surface-stabilizing amendments[15,50,51]. However, crust strength, interparticle bond strength, and threshold friction velocity were not directly measured in the present study; the proposed mechanisms should therefore be regarded as physically plausible interpretations supported by the qualitative FE-SEM observations and prior literature rather than as directly quantified mechanistic endpoints.

A particularly striking finding is that the protective effect of nano-clay became more pronounced at high wind speeds. This suggests that nano-clay-treated soil aggregates can better resist the mechanical forces exerted by high-speed winds, owing to enhanced surface stability or surface crust formation. It is reported that soils with increased surface stabilization exhibit lower erosion rates at high wind speeds because their larger and more cohesive structures resist fragmentation[52]. By binding soil components more tightly, the nano-clay applications reduce the availability of erodible-sized particles and thereby limit soil loss, particularly under high-wind conditions. Furthermore, this mechanism would also be expected to limit the number of particles mobilized by saltation—the principal mechanism initiating and amplifying wind erosion—thereby reducing the abrasive impact of saltating grains on the stabilized surface crust[53].

Aggregate size was also found to have a marked effect on soil loss. In the control treatments, the highest mean soil loss was determined in the 0.5 mm aggregates (85.99%) and the lowest in the 2 mm aggregates (60.93%). In this sand-dominated material, the greater resistance of the larger size fractions is most plausibly related to their greater particle or aggregate mass and the correspondingly higher aerodynamic force required for entrainment,

rather than to an assumption of strong intrinsic cohesion. Conversely, the smaller fractions are more readily mobilized under wind action. These findings are consistent with previous studies showing that aggregate or particle size strongly influences wind-erodibility[54]. Notably, the 2 mm fraction lies above the conventional 0.84 mm erodibility threshold, so its comparatively lower soil loss is consistent with expectation. In future field studies, the relative proportion of erodible and non-erodible fractions should be considered when evaluating whether the application-rate responses observed here extend to naturally mixed aggregate-size distributions.

It should be noted that, given the very low clay and silt content of the studied soil (92.2% sand), the 0.5–2.0 mm size fractions used in this study are best understood as loosely bound sand particles and weak sand aggregates rather than clay-cemented aggregates. This distinction is relevant because interparticle bonding in such fractions relies mainly on weak physical contact rather than the stronger electrostatic and cohesive forces typical of clay- or silt-rich aggregates[55]. The weak natural cohesion of this sand-dominated material may have contributed to the large relative reductions observed after nano-clay treatment. However, because only one soil was tested, the magnitude of the treatment response cannot be compared reliably with that expected for soils containing more clay, silt, or organic matter; such contrasts require direct experimental evaluation.

The apparent difference between the wind-tunnel and FE-SEM observations is not necessarily contradictory. Smaller particles and weak aggregates generally present a larger specific surface area and a greater number of potential contact sites. Accordingly, a surface-applied nano-clay suspension may interact with many exposed fine particles and particle contacts, although the completeness and thickness of surface coverage were not quantified in this study. At the 0.5 mm scale, the nano-clay may therefore have been distributed within a dense network of fine mineral particles rather than forming visually distinct clusters that were consistently resolvable by qualitative FE-SEM inspection. Previous FE-SEM studies of nano-clay-based stabilizers in sandy soils have reported bridging or cohesive structures between particles[22], which supports the plausibility of this interpretation but does not by itself establish the same mechanism quantitatively in the present samples.

The relative reduction in soil loss between the control and the highest nano-clay dose was greatest in the 0.5 mm fraction. The model-predicted heatmaps reproduced the broader dependence of predicted soil loss on aggregate size, wind speed, and nano-clay dose, but they do not provide independent confirmation of the experimental treatment effects. The larger relative response of the fine fraction is physically consistent with its greater erodibility and the possibility that additional interparticle contacts generated by the amendment preferentially stabilize the most readily mobilized material[15,50,56,57].

Tree-based algorithms such as Random Forest and XGBoost can capture complex nonlinear predictive relationships without requiring a prespecified functional form. In the present study, however, model ranking differed between repeated nested cross-validation and the independent holdout set, indicating that predictive performance was sensitive to the specific treatment combinations used for evaluation. This finding supports a cautious interpretation of apparent model superiority in relatively small experimental datasets. Similar patterns have been reported in environmental modelling studies, where the relative performance of ensemble methods varied with the evaluation setting. For example, in the Shihmen Reservoir watershed, gradient boosting achieved the lowest prediction errors, while both bagging- and boosting-based methods showed competitive performance[58,59].

SHAP analysis of the selected XGBoost model indicated that nano-clay dose had the largest mean absolute contribution to model predictions, followed by wind speed and aggregate size. Higher nano-clay doses were generally associated with negative contributions to predicted soil loss, whereas higher wind speeds were associated with positive contributions. Aggregate size showed a smaller overall contribution to model output. These patterns describe the predictive behaviour of the fitted model and should not be interpreted as direct evidence that the SHAP ranking represents causal or mechanistic importance in the erosion process. In another study on soil loss hazard mapping, SHAP analyses likewise identified key soil variables, such as sand content and elevation, that strongly influenced the risk-prediction models and enabled targeted conservation efforts[60]. Similarly, deep-active learning combined with interpretation techniques has been applied to assess global land susceptibility to wind erosion, further illustrating the growing role of explainable machine-learning approaches in erosion research[61].

The Partial Dependence Plots (PDP) and Accumulated Local Effects (ALE) analyses provided complementary descriptions of fitted-model behaviour. Across the investigated predictor ranges, higher nano-clay doses were associated with lower predicted soil loss, whereas higher wind speeds were associated with higher predicted loss. These patterns were directionally consistent with the experimental results but should not be regarded as independent confirmation of treatment efficacy or causal effects. It is reported that the integration of SHAP with gradient-boosting models also improves the understanding of the importance of environmental drivers in various contexts, such as flood-susceptibility and wind-speed prediction[62,63].

The present findings also carry important practical implications for wind erosion management. The marked reduction in soil loss observed under these controlled conditions suggests that nano-clay warrants further evaluation as a potential component of field-scale erosion-control strategies, particularly for highly erodible sandy soils[64]. Since it is applied as a surface treatment, its use could complement rather than replace existing conservation practices such as tillage management, residue retention, cover cropping, and other stabilization measures[65]. From an economic standpoint, the relatively low application rates tested here were intentionally selected to reflect agronomically realistic scenarios, thereby enhancing the feasibility of field adoption while minimizing material requirements and costs. Wider adoption will ultimately depend on factors such as application method, treatment durability, cost-effectiveness, and environmental sustainability. Taken together, nano-clay appears to hold considerable promise as a complementary stabilization technique that could be integrated into broader wind erosion management frameworks for vulnerable agricultural soils[15].

It should be emphasized that these results were obtained under fully controlled wind-tunnel conditions, and the practical application of the nano-clay suspension over large agricultural areas was not directly tested in this study. Before large-scale field implementation can be considered, the treatment should first be evaluated in small field plots under real environmental conditions, including natural rainfall, ambient wind, and freeze-thaw cycles, using varying application doses and standard agricultural spraying equipment (e.g., tractor-mounted or boom sprayers). Field-scale delivery would also require an appropriate mixing or agitation procedure to limit suspension settling during application; this aspect was not evaluated in the present study. Such field experiments would allow the assessment of application feasibility, coverage rate, operational cost, treatment persistence, and erosion-control performance under realistic conditions before any recommendation for large-scale agricultural use.

Study limitations

Despite these encouraging results, the present study has limitations that warrant acknowledgment.

First, the discrete and uniform aggregate size classes (0.5, 1.0, and 2.0 mm) used in this study represent a practical simplification for controlled wind-tunnel testing and may not fully reflect the more continuous particle-size distribution typically observed in field soils. Particle-size analysis of the studied soil, collected from an agricultural area, indicated a sand-dominated texture (4% very coarse sand, 11% coarse sand, 14% medium sand, 41% fine sand, and 22% very fine sand, with only 6.8% silt and clay combined), with 84.6% of aggregates smaller than 0.84 mm. This naturally fine and erosion-prone structure, despite the land being under agricultural use, underscores the practical relevance of developing effective stabilization strategies for such soils. Accordingly, the soil-loss responses obtained from these isolated aggregate fractions should be interpreted with caution, because field soils typically contain broader and continuously distributed aggregate-size mixtures that may behave differently under natural wind-erosion conditions. Second, all experiments were conducted in a wind tunnel, which provides a high degree of experimental control and repeatability but cannot fully reproduce the complexity of natural erosion dynamics. Third, only a sandy soil characterized by low organic matter content and high inherent erodibility was examined; therefore, the effectiveness of nano-clay may differ among soils with contrasting textures, mineralogical compositions, or organic matter levels[66,67]. In addition, several factors typical of field conditions, including natural rainfall events, wetting–drying cycles, vegetation cover, biological soil crusts, and seasonal weather variation, were deliberately excluded in order to isolate the effects of nano-clay dose, wind

speed, and aggregate size. These factors may nonetheless influence both the persistence of the treatment and its long-term effectiveness.

Limitations also apply to the machine-learning component of this study. Although the complete experiment comprised 108 observations, these observations represented 36 unique treatment combinations, each evaluated in three replicates. Replicates from the same treatment combination were therefore kept together during model validation rather than being treated as fully independent observations. The group-aware nested cross-validation framework reduces the likelihood of overly optimistic performance estimates, but it cannot eliminate the risk of overfitting or establish transferability beyond the experimental domain[62,68]. Moreover, model training was restricted to the specific ranges of wind speed, aggregate size, nano-clay application rate, and soil properties examined in this study. Predictions should therefore be interpreted with caution when extrapolated beyond these ranges[69,70]. Independent validation using field-derived datasets representing broader soil and environmental conditions will be required to assess the transferability and generalizability of the modelling framework before its wider application in wind-erosion prediction and decision support[71].

Conclusion

This study showed that nano-clay application substantially reduced wind-induced soil loss across the aggregate sizes and wind speeds investigated under controlled wind-tunnel conditions. Within the tested range, the 3.75 g m^{-2} treatment produced the largest reduction relative to the control, with reductions ranging from 80.8 to 95.9%. The protective response was particularly pronounced for finer aggregates and at higher wind speeds.

From a predictive perspective, XGBoost showed the strongest development-stage performance under repeated nested group-aware cross-validation ($R^2 = 0.835 \pm 0.064$, RMSE = 11.896 ± 2.229, MAE = 6.686 ± 0.854) and was selected before independent holdout evaluation. The selected model achieved $R^2 = 0.882$, RMSE = 13.298, and MAE = 9.443 on the group-held-out test set. SHAP analysis indicated that nano-clay dose made the largest mean absolute contribution to fitted model predictions, followed by wind speed and aggregate size; these values represent model-based predictive contributions rather than causal or mechanistic importance.

Taken together, the controlled experiment and predictive analyses indicate that nano-clay surface treatment has potential for reducing wind-induced soil loss within the conditions examined here. However, the experiment was limited to one sandy soil and a defined range of wind speeds, aggregate sizes, and nano-clay doses. Field validation across contrasting soils and environmental conditions is therefore required before treatment durability, agronomic suitability, and broader practical applicability can be established.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-73412-y.

Acknowledgements The authors thank the Proofreading & Editing Office of the Dean for Research at Erciyes University for their valuable assistance with the language editing and proofreading of this manuscript.

Author contributions Abdirashid Ali Wehliye: Formal analysis, Data curation. Sema Kaplan: Supervision, Methodology, Formal analysis, Data curation, Conceptualization, Writing—review & editing, Writing—original draft.

Funding This research received no specific grant, financial support, or funding from any public, commercial, or not-for-profit funding agency. No external funding was received for the design, implementation, or conduct of the experiments, or for the preparation and publication of this manuscript.

Data availability The experimental dataset supporting the findings of this study, together with the final Python code used for the three-way ANOVA, ANOVA assumption diagnostics, HC3 robustness analysis, group-aware machine-learning validation, SHAP, partial dependence, accumulated local effects, exploratory scenario analysis, and model-predicted heatmaps, has been deposited in Zenodo and is publicly accessible at 10.5281/zenodo.22871116. Tukey multiple-comparison tests used for the treatment-mean groupings in Tables 3, 4 and 5 were performed in JMP 18.0; the corresponding JMP output documenting

these comparisons will be included with the repository materials.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Basaran, M., Uzun, O., Kaplan, S., Görmez, F. & Erpul, G. Tillage-induced wind erosion in semi-arid fallow lands of central anatolia, Turkey. *Soil. Water Res.*, **12**(3). (2017).
2. Sun, R. et al. Global wind erosion reduction driven by changing climate and land use. *Earth's Future*. **12** (10), e2024EF004930 (2024).
3. Dubey, A. et al. *Impact of climate change and land degradation on agriculture. In Eco-Resilience: Climate Change, Land Degradation and Sustainable Solutions.* 49–95 (Springer Nature Switzerland, 2025).
4. Dahiru, A., Abdullahi, A. H., Bello, N. & Abubakar, M. L. Assessment of the influence of climate change on desertification in Katsina state. *Nigeria Discover Geosci*. **4** (1), 29 (2026).
5. Yao, F., Ding, J., Bao, A. & Li, J. Attribution and risk assessment of wind erosion in the Aral sea regions using multi-source remote sensing and RWEQ on GEE. *Remote Sens.* **17** (16), 2788 (2025).
6. Lal, R. Climate change and soil degradation mitigation by sustainable management of soils and other natural resources. *Agricultural Res.* **1** (3), 199–212. https://doi.org/10.1007/s40003-012-0031-9 (2012).
7. Wheaton, E. & Kulshreshtha, S. Environmental sustainability of agriculture stressed by changing extremes of drought and excess moisture: A conceptual review. *Sustainability* **9** (6), 970. https://doi.org/10.3390/su9060970 (2017).
8. Xu, Y. et al. *Effects of Wind Velocity and Aggregate Size on Wind Erosion Characteristics of Loose Subsoil From the Mollisols Region of China* (A Wind Tunnel Assessment. Land Degradation & Development, 2026).
9. Marticorena, B., Bergametti, G., Gillette, D. & Belnap, J. Factors controlling threshold friction velocity in semiarid and arid areas of the United States. *J. Geophys. Research: Atmos.* **102** (D19), 23277–23287 (1997).
10. Sharratt, B. S. & Vaddella, V. K. Threshold friction velocity of soils within the Columbia Plateau. *Aeolian Res.* **6**, 13–20 (2012).
11. Alharbi, H. S. Geotechnical controls on land degradation in drylands: Indicators and mitigation for infrastructure and renewable energy. *Sustainability* **18** (1), 242 (2025).
12. Avecilla, F., Panebianco, J. E. & Buschiazzo, D. E. Variable effects of saltation and soil properties on wind erosion of different textured soils. *Aeolian Res.* **18**, 145–153 (2015).
13. Patil, M., Dalal, P. H., Salifu, E., Iyer, K. K. & Dave, T. N. Biostabilization of soils as sustainable pathway for anti-desertification: Present and future perspectives. *Mater. Today: Proceed*. (2023).
14. Erci, V., Seker, C., Basaran, M. & Erpul, G. Determining the effectiveness of some soil stabilizers in wind erosion prevention using wind tunnel experiments. *Land. Degrad. Dev.* **32** (10), 2962–2977 (2021).
15. Kavazanjian Jr, E., Iglesias, E. & Karatas, I. Biopolymer soil stabilization for wind erosion control. In *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering (Volumes 1, 2, 3 and 4)* 881–884. (IOS press, 2009). https://doi.org/10.3233/978-1-60750-031-5-881
16. Ravi, S., Zobeck, T. M., Over, T. M., Okin, G. S. & D'Odorico, P. On the effect of moisture bonding forces in air-dry soils on threshold friction velocity of wind erosion. *Sedimentology* **53** (3), 597–609. https://doi.org/10.1111/j.1365-3091.2006.00775.x (2006).
17. Fatehi, H., Ong, D. E., Yu, J. & Chang, I. Biopolymers as green binders for soil improvement in geotechnical applications: A review. *Geosciences* **11** (7), 291 (2021).
18. Almajed, A., Lateef, M. A., Moghal, A. A. B. & Lemboye, K. State-of-the-art review of the applicability and challenges of microbial-induced calcite precipitation (MICP) and enzyme-induced calcite precipitation (EICP) techniques for geotechnical and geoenvironmental applications. *Crystals* **11** (4), 370 (2021).

19. Revathi, J. et al. Applications of nanotechnology in agriculture: A Review. *J. Adv. Biology Biotechnol.* **28** (2), 877–895. https://doi.org/10.9734/jabb/2025/v28i22049 (2025).
20. Liu, J., Wang, Z. & Li, Y. Efficacy of natural polymer derivatives on soil physical properties and erosion on an experimental loess hillslope. *Int. J. Environ. Res. Public Health*. **15** (1), 9. https://doi.org/10.3390/ijerph15010009 (2017).
21. Zhang, X., Miller, W., Nearing, M. & Norton, L. Effects of surface treatment on surface sealing, runoff, and interrill erosion. *Trans. ASAE*. **41** (4), 989–994. https://doi.org/10.13031/2013.17271 (1998).
22. Izadi, R., Mahinroosta, M., Allahverdi, A. & Ghadir, P. Stabilizers based on nanoclay and blast furnace slag to reduce wind erosion of sandy soil green stabilization of sandy soil. *Int. J. Environ. Sci. Technol.* **21** (14), 9075–9095 (2024).
23. Feizi, Z., Fordoei, A. R., Shakeri, A. & Sepahvand, S. Effect of polymeric nanocomposite on sandy soil Sta-Bilization. *J. Polym. Environ.* **32** (2), 842–853. https://doi.org/10.1007/s10924-023-03008-4 (2024).
24. Feizi, Z., Sepahvand, S., Ranjbar, A. & Shakeri, A. Effect of nanocellulose content in a polymer-based soil stabilizer on the wind erosion resistance of sandy soil under laboratory conditions. *Sci. Rep.* **15** (1), 42112. https://doi.org/10.1038/s41598-025-26174-y (2025).
25. Sommer, C. G., Lehning, M. & Fierz, C. Wind tunnel experiments: influence of erosion and deposition on wind-packing of new snow. *Front. Earth Sci.* **6** https://doi.org/10.3389/feart.2018.00004 (2018).
26. Myakala, P. K., Jonnalagadda, A. K. & Naayini, P. Revolutionizing Big Data with AI-Driven Hybrid Soft Computing Techniques. *Mach. Learn. Applications: Int. J.* **12** (1), 01–13. https://doi.org/10.5121/mlaij.2025.12101 (2025).
27. Rahmani, A. M. et al. Artificial intelligence approaches and mechanisms for big data analytics: a systematic study. *PeerJ Comput. Sci.* **7** (2), e488. https://doi.org/10.7717/peerj-cs.488 (2021).
28. Selmy, S. A. H., Kucher, D. E. & Moursy, A. R. A. Integrating remote sensing, gis, and ai technologies in soil erosion studies. In *Earth Sciences*. (Intechopen, 2025). https://doi.org/10.5772/intechopen.1008677
29. Khan, I., Khan, K., Bęcek, K. & Bilal, M. F. Evaluating water-induced soil erosion using machine learning: XGBoost as the most effective model. *Land. Degrad. Dev.* **37** (3), 946–960. https://doi.org/10.1002/ldr.70152 (2025).
30. Saha, S., Sarkar, R., Thapa, G. & Roy, J. Modeling gully erosion susceptibility in Phuentsholing, Bhutan using deep learning and basic machine learning algorithms. *Environ. Earth Sci.* **80** (8). https://doi.org/10.1007/s12665-021-09599-2 (2021).
31. Tarek, Z. et al. Soil erosion status prediction using a novel random forest model optimized by random search method. *Sustainability* **15** (9), 7114. https://doi.org/10.3390/su15097114 (2023).
32. Seitkazy, M. et al. Soil erosion prediction in western kazakhstan through deep learning with a neural network approach to LS-factor analysis. *J. Indian Soc. Remote Sens.* **53** (4), 1215–1226. https://doi.org/10.1007/s12524-024-02080-0 (2024).
33. Saha, A. et al. Threats of soil erosion under CMIP6 SSPs scenarios: An integrated data mining techniques and geospatial approaches. *Geocarto Int.* **37** (27), 17307–17339. https://doi.org/10.1080/10106049.2022.2127925 (2022).
34. Padarian, J., Minasny, B. & Mcbratney, A. B. Machine learning and soil sciences: a review aided by machine learning tools. *SOIL* **6** (1), 35–52. https://doi.org/10.5194/soil-6-35-2020 (2020).
35. Available at. https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceler-istatistik.aspx?k=A.&m=KAYSERI (accessed [23/08/2026]).
36. Bouyoucos, G. J. The hydrometer method for studying soils. *Soil Sci.* **25** (5), 365–370 (1928).
37. Richards, L. A. Diagnosis and Improvement of Saline and Alkali Soils. USDA Handbook No. 60. (Washington, DC, 1954).
38. Nelson, D. W. & Sommers, L. E. Total carbon, organic carbon, and organic matter. In (ed Page, A. L.) Methods of Soil Analysis, Part 2. ASA–SSSA, Madison, 539–579. (1982).
39. Chepil, W. S. Measurement of wind erosiveness by dry sieving procedure. *Sci. Agric.* **23**, 154–160 (1942).
40. Zobeck, T. M. et al. Aggregate-mean diameter and wind-erodible soil predictions using dry aggregate-size distributions. *Soil. Sci. Soc. Am. J.* **67**, 425–436 (2003).
41. Chepil, W. S. A compact rotary sieve and the importance of dry sieving in physical soil analysis. *Soil Sci. Soc. Am. J.* **26** (1), 4–6 (1962).
42. Giménez, A., Lozano, F. J., Torres, J. A. & Asensio, C. Automated system for soil wind erosion studies. *Comput. Electron. Agric.* **164**, 104889 (2019).
43. Mao, X., Cao, Y., Ding, Z., Yang, J. & Zheng, J. Combining a saltation impact sensor and a wind tunnel to explore wind erosion processes—A case study in the Zhundong mining area, Xinjiang, China. *J. Soil Water Conserv.* **77** (4), 389–402 (2022).
44. Yan, P., Wang, X., Zheng, S., Wang, Y. & Li, X. Research on wind erosion processes and controlling factors based on wind tunnel test and 3D laser scanning technology. *J. Arid Land.* **14** (9), 1009–1021 (2022).
45. Kaplan, S., Basaran, M., Uzun, O. & Erpul, G. Design and characterization of a wind tunnel for wind erosion research. *Pure Appl. Geophy.* 1–17. (2026).
46. Nouri, A. et al. The effect of fallow tillage management on aeolian soil losses in semiarid Central Anatolia, Turkey. *Agrosystems Geosci. Environ.* **1** (1), 1–13 (2018).
47. Youssef, F., Saskıa, V., Karssenberg, D. & Erpul, G. The spatial variation in wind-blown sediment transport in small scales in Karapinar-Turkey. *Anadolu Tarım Bilimleri Dergisi*. **25** (2), 80–88 (2010).

48. Youssef, F. et al. The effect of vegetation patterns on wind-blown mass transport at the regional scale: A wind tunnel experiment. *Geomorphology* **159**, 178–188 (2012).
49. Zamani, S. & Mahmoodabadi, M. Effect of particle-size distribution on wind erosion rate and soil erodibility. *Arch. Agron. Soil. Sci.* **59** (12), 1743–1753. https://doi.org/10.1080/03650340.2012.748984 (2013).
50. Wang, Z. et al. Effect of biochar and cyanobacteria crust incorporation on soil wind erosion in arid mining area under freeze-thaw action. *Sci. Rep.* **15** (1), 16363. https://doi.org/10.1038/s41598-025-96688-y (2025).
51. Yakupoglu, T., Rodrigo-Comino, J. & Cerdà, A. Potential benefits of polymers in soil erosion control for agronomical plans: A laboratory experiment. *Agronomy* **9** (6), 276. https://doi.org/10.3390/agronomy9060276 (2019).
52. Hagen, L. Soil aggregate abrasion by impacting sand and soil particles. *Trans. ASAE*. **27** (3), 805–808. https://doi.org/10.13031/2013.32875 (1984).
53. Rice, M. A., Mullins, C. E. & McEwan, I. K. An analysis of soil crust strength in relation to potential abrasion by saltating particles. *Earth Surf. Processes Landforms: J. Br. Geomorpholog. Group.* **22** (9), 869–883 (1997).
54. Skidmore, E. L. & Layton, J. B. Dry-soil aggregate stability as influenced by selected soil properties. *Soil Sci. Soc. Am. J.* **56** (2), 557–561. https://doi.org/10.2136/sssaj1992.03615995005600020034x (1992).
55. Webb, N. P. & Strong, C. L. Soil erodibility dynamics and its representation for wind erosion and dust emission models. *Aeolian Res.* **3** (2), 165–179 (2011).
56. Pi, H., Webb, N. P., Huggins, D. R. & Li, S. Effects of secondary soil aggregates on threshold friction velocity and wind erosion. *Land. Degrad. Dev.* **34** (1), 16–27 (2023).
57. Sharratt, B. S. & Vaddella, V. Threshold friction velocity of crusted windblown soils in the Columbia Plateau. *Aeolian Res.* **15**, 227–234 (2014).
58. Nguyen, K. A., Chen, W., Lin, B. S. & Seeboonruang, U. Comparison of Ensemble machine learning methods for soil erosion pin measurements. *ISPRS Int. J. Geo-Information*. **10** (1), 42. https://doi.org/10.3390/ijgi10010042 (2021).
59. Pahno, S., Yang, J. J. & Kim, S. S. Use of machine learning algorithms to predict subgrade resilient modulus. *Infrastructures* **6** (6), 78. https://doi.org/10.3390/infrastructures6060078 (2021).
60. Gholami, H. et al. An interpretable (explainable) model based on machine learning and SHAP interpretation technique for mapping wind erosion hazard. *Environ. Sci. Pollut. Res. Int.* **31** (56), 64628–64643. https://doi.org/10.1007/s11356-024-35521-x (2024a).
61. Gholami, H. et al. An assessment of global land susceptibility to wind erosion based on deep-active learning modelling and interpretation techniques. *Sci. Rep.* **14** (1), 18951. https://doi.org/10.1038/s41598-024-70125-y (2024b).
62. Durap, A. Explainable deep learning techniques for wind speed forecasting in coastal areas: Integrating model configuration, regularization, early stopping, and SHAP analysis. *Neural Comput. Appl.* **37** (25), 21219–21257. https://doi.org/10.1007/s00521-025-11433-w (2025).
63. Tella, A., Pham, Q., Zahidi, I., Fai, C. & Ibrahim, K. Advancing flood susceptibility mapping with explainable AI: A novel application of accumulated local effects (ALE). *Water Resour. Manage.* https://doi.org/10.1007/s11269-025-04430-0 (2026).
64. Abdelhak, M. Innovative techniques for soil and water conservation. *Ecosyst. Manag.: Clim. Change Sustain.* 291–326. (2024).
65. Sorum, M. G. & Sorum, N. G. Use of nanomaterials in underlying and subgrade soil stabilisation: A review. *Road. Mater. Pavement Des.* **27** (4), 919–963 (2026).
66. Vahidi, M., Bayat, H. & Ebrahimi, M. Evaluating the impact of organic amendments on soil erosion dynamics: A comprehensive examination of application methods and timing. *Can. J. Soil Sci.* **105**, 1–14 (2025).
67. Xiao, L. et al. Effects of conservation practices on global wind erosion control: Evidence from experimental data. *Land. Degrad. Dev.* **34** (14), 4386–4398 (2023).
68. Mohammadagha, M. *Hyperparameter optimization strategies for tree-based machine learning models prediction: a comparative study of AdaBoost, decision trees, and random forest* (Decision trees, and random forest, 2025).
69. Bouzekraoui, M. E. et al. Performance assessment of individual and ensemble learning models for gully erosion susceptibility mapping in a mountainous and semi-arid region. *Land* **13** (12), 1–30 (2024).
70. Gangu, K. et al. An analysis on performance scale and model validation in machine learning. In *2024 7th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 7, pp. 1107–1113). (2024)
71. Gholami, H., Mohammadifar, A., Bui, D. T. & Collins, A. L. Mapping wind erosion hazard with regression-based machine learning algorithms. *Sci. Rep.* **10** (1), 20494 (2020).
72 Khan, N. et al. Prediction of oil palm yield using machine learning in the perspective of fluctuating weather and soil moisture conditions: Evaluation of a generic workflow. *Plants (Basel Switzerland)*. **11** (13), 1697. https://doi.org/10.3390/plants11131697 (2022).
73 Li, H., Chang, L., Wei, Y. & Li, Y. Interacting effects of land use type, soil attributes, and environmental factors on aggregate stability. *Land* **12** (7), 1286. https://doi.org/10.3390/land12071286 (2023).
74 Nelson, R. E. Carbonate and gypsum. *Methods soil. analysis: Part. 2 Chem. Microbiol. Prop.* **9**, 181–197 (1982).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Abdirashid Ali Wehliye[1] · Sema Kaplan[2]

✉ Sema Kaplan
semakarabag@erciyes.edu.tr

[1] Graduate School of Natural and Applied Sciences, Soil Science and Plant Nutrition, Erciyes University, Kayseri, Türkiye

[2] Department of Soil Science and Plant Nutrition, Faculty of Agriculture, Erciyes University, Kayseri, Türkiye