CETRA2026 27-30 May 2026, Bol, Island Brač, Croatia
9th International Conference on Road and Rail Infrastructure

MACHINE-LEARNING-BASED PREDICTION OF SKID RESISTANCE FROM 3D PAVEMENT SURFACE TEXTURE: IDENTIFYING AN EFFECTIVE TEXTURE DEPTH

Matúš Kováč, Matej Brna, Peter Pisca
University of Žilina, Faculty of Civil Engineering, Slovakia

Abstract

Road surface friction is influenced by pavement texture across multiple spatial scales. This study examines how well the British Pendulum Tester (BPT) Pendulum Test Value (PTV) can be predicted from 3D pavement texture descriptors derived from a static road scanner, with a particular focus on the effect of vertical truncation depth. Forty-five asphalt surfaces were measured in situ. For each surface, BPT was performed at three locations (n = 135 point-level samples). From each measured height map, depth-truncated topographies were generated by clipping the surface below a truncation depth d to approximate the load-bearing contact zone. Micro- and macro-texture components were separated using an areal Gaussian filter (cut-off wavelength 0.50 mm), and a set of 98 micro and 137 macro texture descriptors was computed. Two regression models were evaluated: Random Forest (RF) and Gradient Boosted Trees (GBT). Dimensionality was reduced by train-only feature pre-selection (top N features) using Spearman rank correlation with PTV. Model performance was quantified using pooled out-of-sample coefficient of determination (R^2) under group-aware validation schemes: grouped k-fold cross-validation (groupKfold; groups = surfaces) and leave-one-surface-out (LOSO). Across models and validation schemes, predictive performance consistently peaked at d ≈ 1.4–1.6 mm (pooled R^2 up to ≈0.31 in groupKfold and ≈0.47 in LOSO for point-level modelling). A surface-aggregated variant (per-surface averaging of three measurements) yielded comparable performance (pooled R^2 up to ≈0.32 in groupKfold and ≈0.42 in LOSO). Overall, the results suggest that depth-limited texture descriptors contain measurable, albeit moderate, predictive information about friction, and that selecting an effective truncation depth (around 1.5 mm) is critical under BPT conditions for building more generalizable texture-friction models.

Keywords: pavement surface texture, skid resistance, British Pendulum, 3D laser scanning, Random Forest, Gradient Boosting

1 Introduction

Pavement skid resistance is a key safety-related surface characteristic, particularly under wet conditions, where insufficient friction increases the risk of loss of control and accidents. Skid resistance is strongly influenced by pavement surface texture [1], conventionally decomposed into microtexture (short wavelengths at the aggregate scale) and macrotexture (longer wavelengths associated with voids and surface relief). Microtexture primarily governs low-speed adhesion, whereas macrotexture affects drainage and hydrodynamic effects [2]. With the growing availability of high-resolution non-contact 3D measurements, areal texture descriptors have been explored for skid-resistance modelling using data-driven and machine-learning approaches [3, 4]. Recent studies also highlight that only a depth-limited part of the texture within the load-bearing contact zone may be most relevant to friction (effective texture depth) [5].

DOI: https://doi.org/10.5592/CO/CETRA.2026.1998

In this study, texture measurements were obtained using a non-contact Static Road Scanner (SRS) developed specifically for research applications. The primary motivation for using SRS is its ability to acquire high-resolution surface geometry under fully controlled stand-still conditions, avoiding boundary-condition variability inherent to friction-based devices. By eliminating operator effects and factors such as speed, water film thickness or rubber properties, SRS provides repeatable geometric data suitable for isolating the texture component of the friction mechanism. The objective of this work is to evaluate whether BPT friction (PTV) can be predicted from SRS-derived micro- and macrotexture descriptors and to determine how vertical truncation depth affects model performance. Two ensemble regressors - Random Forest [6] and Gradient Boosted Trees [7] are compared under group-aware validation to assess predictive potential and identify a practically effective texture depth for future datasets.

2 Methodology

2.1 Road sections, measurement tools, friction and texture measurements

To obtain a dataset that adequately represents the variability of pavement surface conditions, 45 asphalt test sections were selected to span a broad range of microtexture and macrotexture characteristics and, consequently, a wide spectrum of British Pendulum skid-resistance values. On each road section, measurement grids were established, within which 3 test points were defined and marked at a regular spacing of 0.5 m (figure 1). The resulting measurement campaign yielded a total of 135 spatially co-located observations.

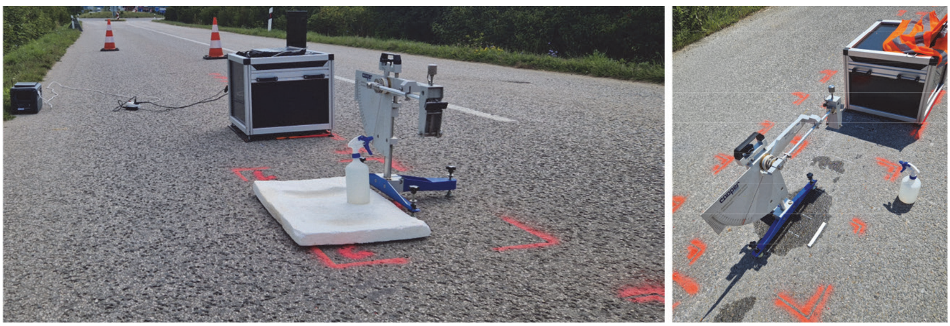

Figure 1 Static Road Scanner (SRS) and British Pendulum Tester (BPT)

For each observation, high-resolution 3D surface texture data were collected together with the corresponding Pendulum Test Value (PTV). This comprehensive dataset formed the basis for all subsequent analyses relating pavement surface texture descriptors to measured skid resistance. High-resolution characterization of pavement surface texture was performed prior to friction testing using a Static Road Scanner (SRS) developed at the University of Žilina (figure 1). The SRS captures surface topography by means of dual laser profilometry, recording the pavement morphology as a series of closely spaced scan stripes that are subsequently merged into a continuous three-dimensional surface representation. The nominal spatial resolution of the system is on the order of tens of micrometres, enabling the reliable characterization of both micro- and macro-scale surface features. Following texture acquisition, the exact measurement location was transferred to the friction-testing stage using a positioning grid applied during scanning. Based on this reference, guide markings were drawn directly on the pavement surface to delineate the slip path of the British Pendulum slider. This procedure ensured accurate relocation of the scanned area and consistent alignment between the texture measurement and the subsequent friction test.

For each measurement point, the evaluated surface was confined to a rectangular patch of approximately 126 × 76 mm, corresponding to the effective contact area of the pendulum slider. Skid resistance was assessed using a British Pendulum Tester, with all measurements carried out on a wetted pavement surface. The test yields the Pendulum Test Value (PTV).

2.2 Data pre-processing

Surface data obtained at each measurement location consisted of a set of partially overlapping scan tracks, which were first co-registered and combined to form a single, spatially coherent surface model. The pre-processing workflow was designed to eliminate artefacts related to the scanning configuration and environmental influences, while ensuring that all samples were evaluated on a uniform and comparable spatial basis. Following scan alignment and merging, auxiliary geometric elements associated with the positioning and reference system were removed. The surface model was then cropped to a standardized evaluation area of 126 × 76 mm, corresponding to the predefined region of interest. Subsequently, artefacts and spurious measurement points were identified and filtered out. Finally, large-scale form was removed by levelling the surface with respect to a best-fit reference plane. The resulting processed surface patches constituted the input data for subsequent texture filtering and computation of areal surface descriptors.

2.3 Texture filtering and derivation of micro- and macrotexture descriptors

To investigate which part of the 3D surface contributes most to the friction signal, the scanned height maps were vertically truncated at nine depth levels (d = 0.5–5.0 mm; figure 2) below the highest point. The truncated 3D surface patches were then spatially filtered. An areal Gaussian filter following the framework ISO 16610-61 [8] with a 0.50 mm cut-off wavelength was used to obtain micro-texture (high-pass component, wavelengths < 0.50 mm) and macro-texture (low-pass component, wavelengths > 0.50 mm). For each component, a set of areal descriptors (e.g. amplitude, spatial, and hybrid parameters) was extracted in accordance with ISO 25178-2 [9] and exported for machine-learning analysis. Based on the separated microtexture and macrotexture surfaces and the derived descriptors, three dataset variants were constructed: (i) MICRO, containing the 98 microtexture predictors; (ii) MACRO, containing the 137 macrotexture predictors; and (iii) COMBINED, comprising all 235 predictors. Prior to modelling, predictors with missing values and near-constant variance were removed, and a correlation filter was applied to reduce redundancy; therefore, the effective number of predictors varied slightly with truncation depth and split.

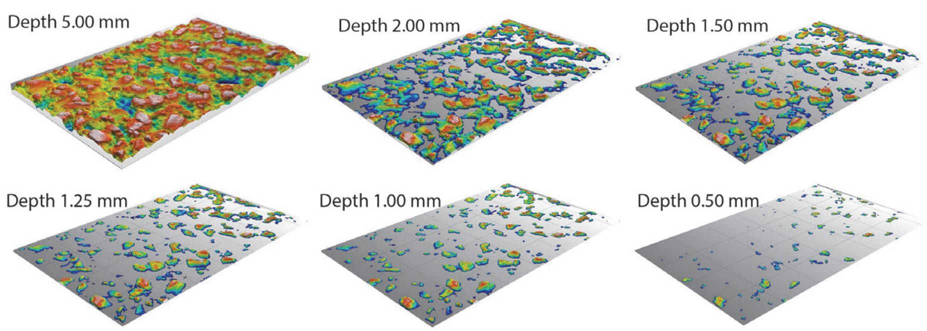

Figure 2 Illustration of six depth-limited (truncated) topographies of one surface

3 Machine-learning models and validation protocol

Two regression models were trained on the texture descriptors: Random Forest (RF) [6] and Gradient Boosted Trees (GBT) [7]. All modelling was performed in MATLAB using the Statistics and Machine Learning Toolbox (RF: TreeBagger; GBT: fitrensemble with LSBoost and regression-tree learners). To avoid leakage between measurements from the same surface, we primarily report results from group-aware splits: groupKfold and leave-one-surface-out (LOSO). In repeated groupKfold, pooled R^2 was computed by concatenating all out-of-sample predictions from all folds and repeats (each observation contributes one out-of-sample prediction per repeat) and comparing them to the corresponding measurements. Two main validation schemes were used in this paper: (i) groupKfold ($k = 5$, repeated 10 times), where folds are created by grouping the three measurements belonging to the same surface, and (ii) LOSO, where all measurements of one surface are left out for testing. For reference only, we also ran standard random holdout and row-wise holdout splits. However, these can yield optimistic estimates because measurements from the same surface may appear in both training and testing. All preprocessing and feature-selection steps (including missing/correlation filters and TopN) were performed using training data only within each split. The learned feature subset was then applied to the corresponding test fold. Figures 3 and 4 summarise point-level modelling results for the COMBINED predictor set ($n = 135$). Under groupKfold (figure 3), pooled R^2 peaked at about 0.31 (RF: TopN = 10 at $d = 1.4$ mm; GBT: TopN = ∞ at $d = 1.6$ mm). Under LOSO (figure 4), the maximum pooled R^2 reached approximately 0.47 (GBT: TopN = ∞ at $d = 1.6$ mm), while RF achieved up to about 0.33 (TopN = 10 at $d = 1.4$ mm). In most cases RF benefited from stronger feature reduction (TopN≈10–30), whereas GBT often preferred larger feature sets.

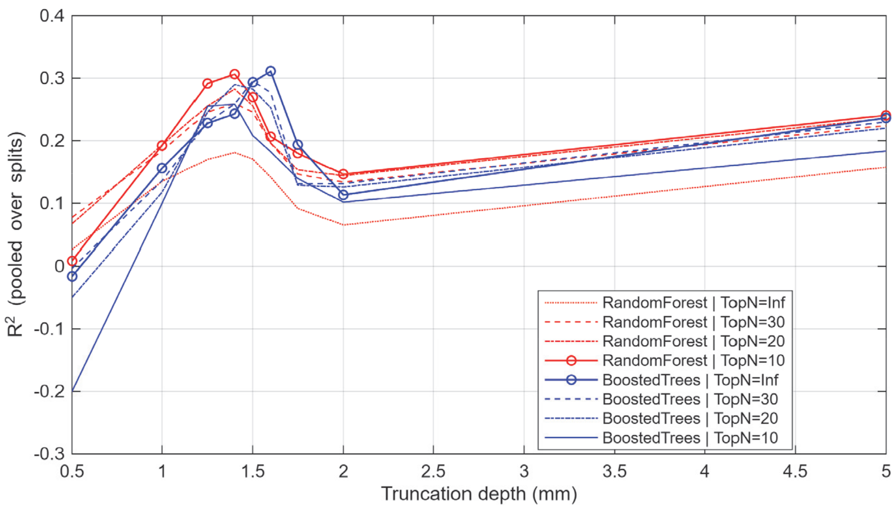

Figure 3 Point-level COMBINED dataset ($n = 135$): pooled R^2 vs truncation depth d for groupKfold (groups = surfaces)

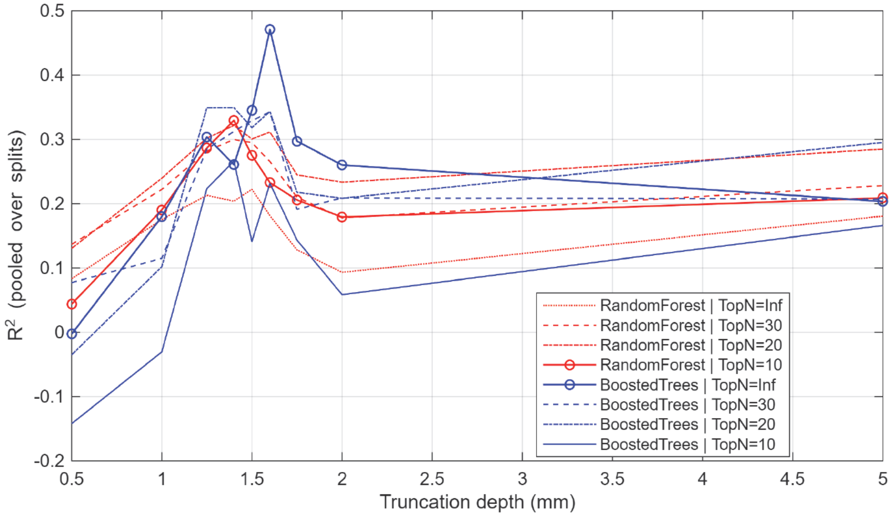

Figure 4 Point-level COMBINED dataset (n = 135): pooled out-of-sample R^2 vs truncation depth d for LOSO (leave-one-surface-out)

Figures 5 and 6 show the surface-aggregated experiment, where the three point measurements were averaged per surface (n = 45). Aggregation reduces within-surface variability by averaging the three measurements per surface. Under groupKfold (figure 5), RF achieved a maximum pooled R^2 of 0.32 (TopN = 10 at d = 1.5 mm) and GBT reached 0.29 (TopN = ∞ at d = 1.6 mm). Under LOSO (figure 6), the best pooled R^2 reached 0.42 for RF (TopN = 10 at d = 1.5 mm) and reached 0.40 for GBT (TopN = 10 at d = 1.6 mm).

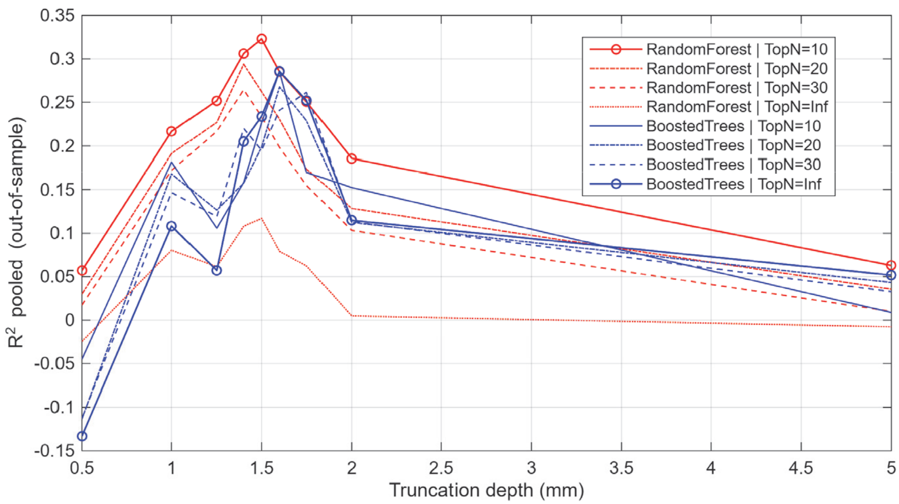

Figure 5 Surface-aggregated COMBINED dataset (n = 45): pooled out-of-sample R^2 vs truncation depth d for groupKfold

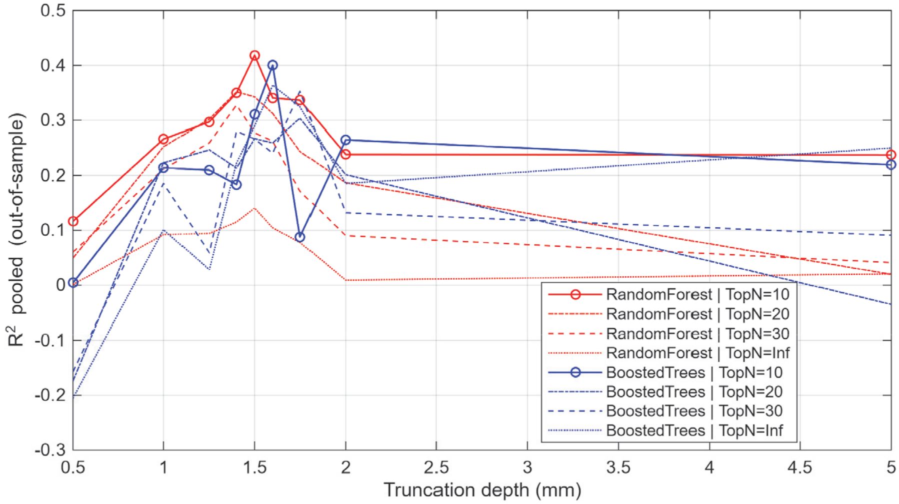

Figure 6 Surface-aggregated COMBINED dataset (n = 45): pooled out-of-sample R^2 vs truncation depth d for LOSO (Note: TopN = ∞ (referenced as Inf in figures) means that all predictors (after the pre-filtering procedure) were used in the model)

Across both point-level (n = 135) and surface-aggregated (n = 45) experiments, model performance consistently peaked at truncation depths around d ≈ 1.4 to 1.6 mm. This repeatable optimum suggests that the pendulum friction response appears to be primarily governed by the upper part of the 3D texture, while very shallow truncation removes informative asperities and deeper truncation introduces larger-scale macro-geometry less relevant to the low-speed BPT interaction. Surface aggregation does not necessarily increase the peak pooled R^2 compared with the best point-level runs, however, it reduces withinsurface variability and can improve robustness to local noise and sampling location.

From a practical perspective, the non-contact Static Road Scanner (SRS) enables repeatable acquisition of high-resolution texture without direct surface contact, which can improve operator safety and supports dense spatial sampling. Such texture measurements may help prioritize sections for further investigation and provide a scalable basis for building larger texture-friction datasets. Nevertheless, predictive power remains moderate under strict group-aware validation, and additional data (more surfaces, repeated campaigns, and broader friction states) are required to confirm generalization and to develop more robust models. Despite moderate pooled R^2 values (≈0.31 to 0.47), the absolute prediction error remains nonnegligible (RMSE on the order of ~10 PTV units), indicating that textureonly prediction with the current dataset size and variability provides limited practical accuracy and should be treated as a baseline for larger future datasets.

For reference only, random holdout and row-wise holdout splits were also evaluated, but they are not reported here because they can yield optimistic estimates when correlated measurements from the same surface appear in both training and testing. While this paper focuses on tree-based ensembles, the modelling framework is compatible with other nonlinear regressors (e.g. SVR or neural networks). Future work should additionally evaluate stricter generalization protocols (e.g. site-wise validation) and investigate whether the descriptor set can be further reduced without loss of accuracy to support operational deployment.

Overall, the study supports the hypothesis that a depthlimited representation of the load-bearing texture contains a higher fraction of frictionrelevant information, with d ≈ 1.5 mm emerging as a pragmatic default for future descriptor extraction in similar datasets.

4 Conclusion

This work evaluated the feasibility of predicting BPT friction (PTV) from 3D-scanner-derived texture descriptors using tree-based machine learning. Using group-aware validation, the best pooled out-of-sample R^2 values consistently occurred at truncation depths d ≈ 1.4–1.6 mm. For point-level data (n = 135), pooled R^2 reached up to ≈0.31 (groupKfold) and ≈0.47 (LOSO). A surface-aggregated variant (averaging three measurements per surface) yielded comparable predictive performance (pooled R^2 up to ≈0.32 in groupKfold and ≈0.42 in LOSO). Feature pre-selection (TopN≈10–30) generally benefited RF and, in several cases, also improved GBT. Across all evaluation schemes, the optimal truncation depth repeatedly occurred around 1.4–1.6 mm. Therefore, d ≈ 1.5 mm is proposed as a pragmatic default for PTV descriptor extraction in similar forthcoming larger datasets. Future work will focus on expanding the dataset (more surfaces and friction states), reporting additional error metrics (e.g. RMSE and MAE), and analysing the stability of important predictors across splits to support a more interpretable 'final' descriptor subset.

Acknowledgements

Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V04-00041.

References

[1] Kováč, M., Brna, M., Decký, M.: Pavement Friction Prediction Using 3D Texture Parameters, Coatings, 11 (2021) 10, 1180, DOI: 10.3390/coatings11101180

[2] Forster, S.W.: Pavement Microtexture and Its Relation to Skid Resistance, Transportation Research Record, 1215 (1989), pp. 151–164

[3] Hu, Y., Sun, Z., Han, Y., Li, W., Pei, L.: Evaluate Pavement Skid Resistance Performance Based on Bayesian-LightGBM Using 3D Surface Macrotexture Data, Materials, 15 (2022), 5275, DOI: 10.3390/ma15155275

[4] Yang, G., Chen, K.-T., Wang, K., Li, J., Zou, Y.: Field Study of Asphalt Pavement Texture and Skid Resistance under Traffic Polishing Using 0.01 mm 3D Images, Lubricants, 12 (2024), 256, DOI: 10.3390/lubricants12070256

[5] Ding, S., Wang, K.C.P., Yang, E., Zhan, Y.: Influence of Effective Texture Depth on Pavement Friction Based on 3D Texture Area, Construction and Building Materials, 287 (2021), 123002, DOI: 10.1016/j.conbuildmat.2021.123002

[6] Breiman, L.: Random Forests, Machine Learning, 45 (2001), pp. 5–32, DOI: 10.1023/A:1010933404324

[7] Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, 29 (2001) 5, pp. 1189–1232, DOI: 10.1214/aos/1013203451

[8] ISO 16610-61:2015: Geometrical product specifications (GPS) – Filtration, Part 61: Linear areal filters – Gaussian filters

[9] [9] ISO 25178-2:2021: Geometrical product specifications (GPS) – Surface texture: Areal, Part 2: Terms, definitions and surface texture parameters