Coupling asphalt construction process quality into product quality using data-driven methods

Q. Shen[a] , F. Vahdatikhaki[a] , S.R. Miller[a], and A.G. Dorée[a]

[a,] Department of Construction Management and Engineering, University of Twente, the Netherlands
q.shen@utwente.nl, f.vahdatikhaki@utwente.nl, s.r.miller@utwente.nl, a.g.doree@utwente.nl

Abstract –

The long-term quality of the asphalt layer is crucial for maintaining the functionality of roads. Despite extensive research on predicting pavement failure modes and the effect of design and road use on the quality of the asphalt layer, there is limited understanding of how the quality of road construction impacts the long-term quality of asphalt pavement. This paper presents a data-driven approach to studying the impact of construction process quality on the International Roughness Index (IRI) of roads. Two machine learning models (Random Forest and Gated Recurrent Unit) were compared in a case study, with the GRU model (R^2 of 0.8284) outperforming the RF model (R^2 of 0.5498). Results showed that construction process quality was the third most significant factor affecting IRI.

Keywords –

Asphalt construction; construction process quality; international roughness index (IRI); data-driven methods; regression; machine learning

1 Introduction

The demand for road infrastructure and longer guarantee periods has driven the asphalt construction industry to prioritize quality in order to remain competitive [1]. However, the industry is reliant on craftsmanship and experience-based decision-making, making quality assurance a challenge. To address this, the Process Quality Improvement (PQi) methodology was proposed by Miller [2] and has since been widely adopted as part of asphalt construction quality control methods in the Netherlands [3–5]. This methodology uses advanced sensing technologies in an integrated network, the Internet of Things (IoT), to monitor compaction efficiency/consistency and temperature homogeneity during the construction process. The collected data is used to evaluate process quality and provide feedback to contractors. While it is known that construction process quality impacts the long-term quality of asphalt pavement [6], the extent of this impact remains unknown. Further investigation is needed to determine the effect of operational strategies on the quality of the asphalt pavement.

At present, the correlation between the process and product quality of the asphalt pavement is still treated implicitly and intuitively during the construction practices, primarily owing to the system's non-linearities. In the past few years, attempts have been made by various studies to develop data-driven empirical models regarding the long-term performance of asphalt pavement [7–10]. Various data-driven techniques, such as machine learning (ML), have been applied to extract valid patterns and knowledge. Particularly, the pavement condition regarding the roughness, which is represented by International Roughness Index (IRI), has received the most attention. However, although the data-driven techniques have been successfully applied to these studies, investigations in these studies were confined to correlating the pavement's long-term performance with indicators during the operational stage, such as traffic intensities, climate conditions, and previous inspections of the pavement conditions. To provide a more comprehensive understanding of the pavement performance it is essential to take the quality of the construction process into consideration.

On these premises, this research aims to explicitly investigate the correlations between the asphalt construction process quality (PQi data) and the long-term asphalt pavement quality, focusing primarily on IRI. A dataset covering the design, construction, and operation phases of road construction lifecycle was developed using data from two highway sections in the Netherlands, resulting in 62 samples. Considering the general performance of non-linear regression and the time-variant characteristic of the IRI data, Random Forest (RF) and Gated Recurrent Unit (GRU) were selected. The GRU model outperformed the RF model with an R^2 of 0.8284 compared to 0.5498. Permutation feature importance analysis revealed the construction process quality indicator as an important factor for pavement performance.

The remainder of the paper is structured as follows. First, the methodology of the research is presented. This is followed by a case study demonstrating the validation of the proposed method. The paper ends with a discussion and a conclusion.

2 Methodology

Figure 1 provides an overview of the methodology applied in this research. In general, this methodology included the development of the dataset, ML model development, and evaluation and interpretation. Overall, the dataset development phase mainly focused on building a structured dataset that can be used for ML-based analysis. The second phase was dedicated to the development of ML models. An optimization algorithm was used to find the optimal configuration of the ML models. In the last phase, first, the developed ML models were evaluated based on the pre-defined metrics and then interpreted by analyzing the importance of the input features, i.e., how changes within the features influence the overall performance of the models. This can give more insights regarding the inner mechanism of the ML models, as well as a more intuitive representation of how to process quality indicators that contribute to certain product quality indicators. The following will provide a detailed explanation of each step.

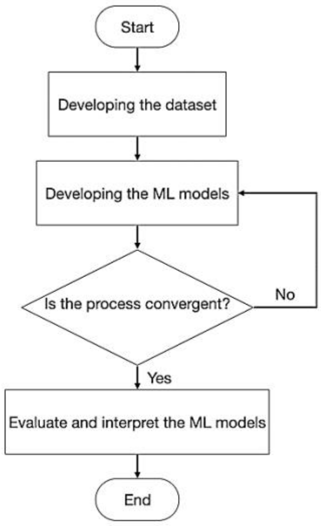


Figure 1. A schematic overview of the proposed modeling process

2.1 Dataset development

The proper input-output structure of a dataset is crucial for successful data mining. This involves understanding the cause-and-effect relationships that connect process quality indicators with the product quality indicator, in this case, the International Roughness Index (IRI). Three distinct types of parameters (i.e., features in the ML nomenclature) were considered in this research, namely design, construction, and operation parameters (each pertaining to a phase in the lifecycle of the road).

As for the design-related parameters, this study considered the type of asphalt mix. While the design characteristics of the asphalt mix play a significant role in the long-term performance of the road, for the purposes of understanding the correlation between construction process quality and long-term pavement quality, a simplification was necessary. Instead of examining each individual design characteristic (such as bitumen content, aggregate type, etc.), only the mix type indicator was used. This choice was made because the focus of this research was not on exploring the individual impact of design characteristics on long-term performance, but rather on ensuring that the design phase was accurately represented in the ML model. This was achieved by consolidating all design parameters into the mix type indicator.

The definition of quality of the asphalt construction process was adopted from the previous work of the authors [3], where the process quality is defined as the degree to which the asphalt layer was compacted sufficiently (i.e., enough compaction passes) at the right temperature (i.e., avoiding the compaction of the asphalt layer when it is too hot or too cold). More specifically, the construction process quality is assessed using the Effective Compaction Rate (ECR) index proposed by [4], as indicated in Equation (1).

$$ECR_{p,k} = \frac{n_{p,k}}{N} \tag{1}$$

where $n_{p,k}$ refers to the number of cells that have received $\pm k$ passes (i.e., tolerance margin) of the target number of passes, and at least $p\%$ of received passes were within the defined compaction window. In this equation, N represents the total number of measurement cells.

In addition, the operation of the pavements during their service, such as the traffic intensity and weather, can also significantly influence the condition and performance of the pavements [8,11]. For the traffic loads, the author considered the average daily traffic and average daily truck traffic to reflect the daily intensity of the investigated highways. Therefore, in this study, the average hourly traffic intensities of three different types were considered, including passenger vehicles, heavy trucks, and medium trucks. When it comes to weather conditions, as pointed out in the literature [12], temperature variation and moisture change can have a great impact on the material properties of the pavement structure. It is shown that freeze-thaw cycles during pavement operation also contribute to the deterioration rate of the asphalt [13]. Therefore, this research considered the average annual temperature, average annual precipitation, and the number of freeze-thaw cycles.

As for the labels in the dataset (i.e., the parameter about which the prediction is to be made), IRI is used as the sole metric. Although IRI is often recorded during regular inspections, it is a time-variant metric. To account for the complexity of changes in road usage, this study considers a rolling time window of 1 year instead of a long-term average of road use. This approach allows for a more accurate evaluation of the roads' condition by considering the most recent IRI measurement and the amount of use the road has seen since then. Table 1 below summarizes the identified input-output structure.

Table 1. The summary of the identified dataset structure

	Variable	Description
Input	ECR	The effective compaction rate, which was described in Equation 1
	Mixture type	The type of the asphalt mixture
	Age	The age of the pavement compared to the construction year
	Heavy truck intensity per workday	The mean intensity of the heavy trucks of a certain road section on the workday from the previous year
	Medium truck intensity per workday	The mean intensity of the medium trucks of a certain road section on the workday from the previous year
	Passenger car intensity per workday	The mean intensity of the passenger vehicles of a certain road section on the workday from the previous year
	Annual mean temperature	The mean value of the annual temperature of the road section from the previous year
	Annual mean precipitation	The mean value of the annual precipitation of the road section from the previous year
	Annual freeze-thaw cycle	The annual number of freeze-thaw cycles of the road section from the previous year
	IRI-1	The IRI value from the previous year
Output	IRI	International Roughness Index, which quantitively reflects the roughness of the pavement

2.2 Machine learning model development

2.2.1 Selecting machine learning algorithms

Based on the problem context of the research, two types of ML algorithms were used, namely random forest (RF) and gated recurrent unit (GRU).

As a widely applied tree-based ML algorithm, random forest (RF) can solve both regression and classification problems using ensembled decision trees. An RF model is built by randomly ensembling various decision trees using the bagging method and obtaining the output(s) by voting [14]. As a powerful ML algorithm, RF can overcome the overfitting problem and improve the robustness against the outliers, without compromising the performance in handling non-linear classification and regression problems. The ensembling technique and bootstrapping also allow RF to achieve potentially good performance on small datasets.

As previously mentioned, regression modelling faces great system non-linearities. To tackle the complexities of non-linear regression and provide better performance with time-series data, GRU was also selected. GRU is a special form of recurrent neural network (RNN) that can describe the dynamic behaviour of time-series data by circulating states in the networks. However, the conventional architecture of RNN soon showed its limits, because of problems such as the gradient's vanishing and explosion, and the difficulty in learning long-term patterns. To tackle the aforementioned challenges, long short-term memory (LSTM) and gated recurrent unit (GRU) was developed and introduced as extensions of conventional RNN [15].

Previous studies found that GRU has comparable or even surpassed the performance of LSTM [16]. In addition, although the structure of the GRU unit is similar to LSTM, the architecture of the GRU cell will require fewer external gating signals. Therefore, fewer parameters are needed, and the training process will be more efficient. Therefore, in this research, GRU was used.

However, when applying GRU, one issue is that apart from the time-variant variables, such as IRI-1, traffic intensities, and climate conditions, other input features including ECR and mixture types cannot be processed by the default GRU layer, because these input features are time-invariant. Therefore, these time-variant features were converted into vectors using affine transformation as the internal state of the GRU architecture. This transformed initial state is then added to the hidden state of the GRU when calculating the output [17–19]. In addition, to further tackle the complexities and non-linearities of the problems, a hybrid network can be used by adding more dense layers behind the GRU layer, thus increasing the depth of the network to boost its performance [9,10].

2.2.2 Hyperparameter optimization

To obtain optimal performance of developed ML models, it is essential to fine-tune and optimize the model configurations. Table 2 presents the list of optimized hyperparameters from RF and GRU.

A widely applied approach for hyperparameter optimization in ML is to use meta-heuristic methods, e.g., genetic algorithm (GA) or particle swarm optimization (PSO). This research used GA-based optimization of the ML models as proposed in the literature [20]. Figure 2 represents the flowchart of the GA-based hyperparameter optimization framework.

The developed dataset was first divided into 80% training and 20% testing subsets. Subsequently, the *k*-fold cross-validation was applied to the training subset by further dividing the subset into k non-repeating sections. During the training process, the model was trained k times, using k-1 sub-training sets each time. The remaining sub-training set was then used for the evaluation of the model. By averaging the evaluation of the model k times, the fitness score was calculated. By using k-fold cross-validation, it can be ensured that all the samples of the training subset are involved in both the training and testing process. Therefore, this method reduces the sensitivity of the models' performances to how the training subset will be further split. Considering the trade-off in terms of computational time and accuracy, the value of k was set to 10.

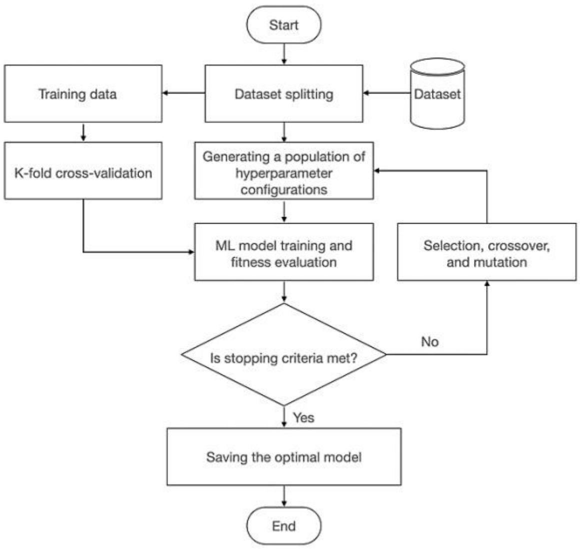


Figure 2. The GA-based hyperparameter optimisation framework

Table 2. The summary of the selected hyperparameters required to be optimized

ML algorithm	Hyperparameter	Description
RF	n_estimators	The number of decision trees in the RF structure.
	max_depth	The allowed maximum depth of each decision tree.
	min_samples_split	The minimum number of samples needed to split an internal node.
	min_samples_leaf	The minimum number of samples required to be at a leaf node.
GRU	n_layers	The number of the hybrid dense layer.
	n_neurons	the number of neurons within each hybrid dense layer.
	units	The number of GRU units.
	epochs	The number of epochs for the model training.

At the beginning of the optimization process, a random set of hyperparameter arrays are generated and used to develop the first generation of ML models. The performances of each model are assessed and through a ranking process, the best models are identified. By applying crossover and mutation on the top-ranking solutions, the subsequent generation of models is generated. The optimization process will continue until the stopping criteria are met.

2.3 Model evaluation and interpretation

Finally, the developed ML models were validated using the data outside the range of the training dataset. Several metrics were used to represent the regression performance of developed ML models, including R-squared (R^2), mean squared error (MSE), and mean absolute error (MAE). The equations of these three metrics were given below.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{3}$$

$$MAE = \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{n} \tag{4}$$

where n refers to the total number of samples, y_i refers to the true value, $\hat{y}_i$ refers to the prediction, and $\bar{y}$ refers to the mean value of the sample.

Additionally, among these metrics, R^2 was also used as the fitness function in the GA-based model optimization process to represent the fitness of each examined chromosome. To explicitly represent the correlations between various process quality indicators and product quality indicators, the obtained regression models were interpreted in the form of feature importance, which was through a sensitivity analysis. The feature importance reflects how important the features are for the regression. Therefore, the feature importance analysis can provide explicit insights into the models, as well as the hidden correlations between inputs and outputs. For this purpose, the model with the highest predictive performance was used. To ensure that the feature importance interpretation can be applied to both ML algorithms, this research adopted the permutation importance as the interpreting approach, which randomly shuffles a certain input feature and re-evaluates the model performance. By comparing the performance changes with the baseline performance, the importance of a certain input feature can be obtained.

3 Case study

To validate the proposed framework, a case study was conducted. Two Dutch highway sections (A58 and A4) with a total length of 4.1 km, were selected. For both sections, PQi measurements and regular IRI inspections were available. The historical PQi measurements (i.e., construction process quality indicator) were retrieved

Table 3. An example of the developed dataset

Samples	Input features										IRI
	ECR	Mixture	Age	Annual Mean Temp.	Annual Mean Preci.	Annual Freeze/Thaw Cycles	Passenger Cars/Day	Med. Trucks/Day	Heavy Trucks/Day	IRI-1	
1	0.1792	ZOAB-1	3	10.7713	2.2784	55	47318	3406	3476	1.32	0.93
2	0.1372	ZOAB-1	3	10.7713	2.2784	55	47318	3406	3476	1.20	0.64
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

from the database of the Dutch research network ASPARi [21]. The IRI data were retrieved from the IVON database and yearly inspectional records held by the Ministry of Infrastructure and Water Management of the Netherlands (Rijkswaterstaat). Regarding the data about the operation phase of the roads, two databases were used. For the traffic intensity, data were extracted from the INtensiteit op WEgVAkken (INWEVA) database of Rijkswaterstaat. This database covers the entire Dutch highway network and registers the historical traffic intensity of each hectometer section with specified BPS locations. Besides, the weather data were derived from the Koninklijk Nederlands Meteorologisch Instituut (KNMI) database. Table 3 provides an example of the developed dataset.

Several parameters for the GA-based hyperparameter optimization were also defined including the size of the population, the size of offspring, crossover and mutation rates, and the total number of generations, as shown in Table 4. In this case study, the number of generations is the stopping criterion.

Table 4. Pre-defined GA-parameters

GA parameter	Value	Description
Population size	100	The total number of individuals contained in one population.
Offspring size	100	The total number of individuals that will be generated after each iteration.
Crossover rate	0.8	The probability of the occurrence of crossover between two individuals.
Mutation rate	0.2	The probability of the occurrence of mutation within one individual.
Generation number	100	The number of iterations

3.1 Results

As explained in section 2.2.1, two ML algorithms were selected and used, namely RF and GRU. Table 5 presents the comparison of the two models in terms of MSE, MAE, and R^2. Also, Figure 3 demonstrates the regression plots of each model, including both the training and testing processes. Additionally, Table 6 also summarizes the results of the optimization of the hyperparameters.

Based on Table 5 and Figure 3, the developed GRU model significantly outperformed the RF model, where the latter is considerably underfitting, due to the high training and testing errors. This underfitting issue will be further discussed in the following section. The GRU model achieved a promising result regarding R^2, with a value of 0.8284. Besides, the errors of the predictions compared to the true values, which are reflected by MSE and MAE, were well-controlled. The results between the training process and testing process are close, meaning that the developed GRU model has reasonable generality.

Table 5. The summary of the results of the model validations

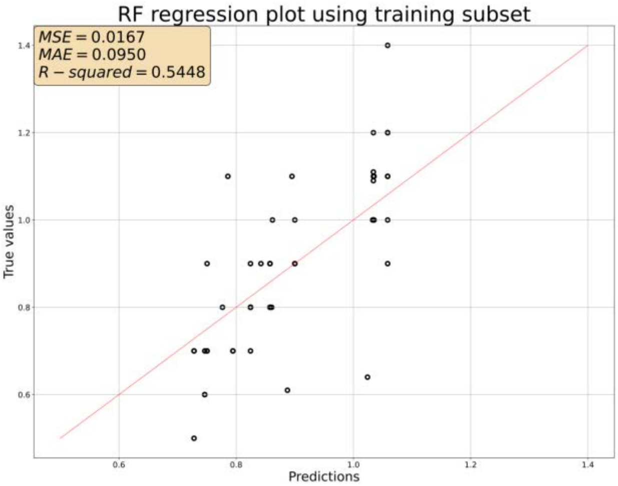
Model	R^2	MSE	MAE
RF	0.5498	0.0123	0.0847
GRU	0.8284	0.0050	0.0600

As mentioned earlier, only the best-performing model (i.e., GRU model) was used for the feature importance. MAE was used as the metric to evaluate the impact of each feature on the overall model performance.

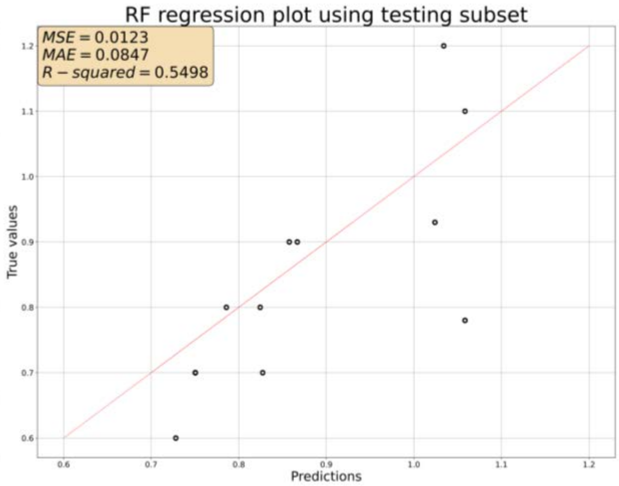
Table 6. The summary of the optimization results regarding hyperparameter configurations

ML algorithm	Hyperparameter	Value
RF	n_estimators	100
	max_depth	None
	min_samples_split	20
	min_samples_leaf	1
GRU	n_layers	2
	n_neurons_first_layer	14
	n_neurons_second_layer	8
	units	18
	epochs	210

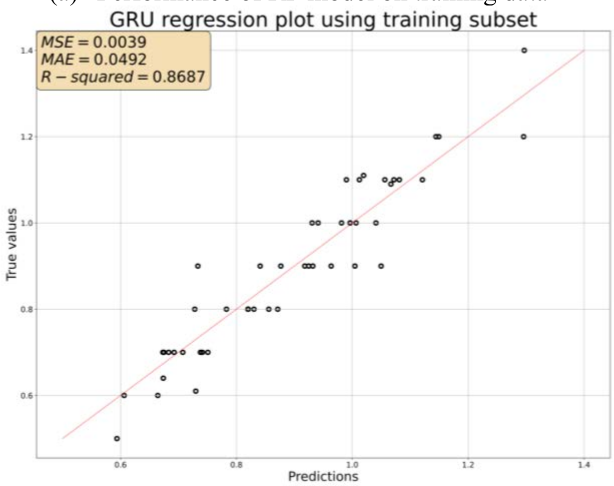
Figure 4 shows the permutation importance of each input feature. Based on the results, the feature IRI-1 has the highest importance. By changing the values of this feature, the model performance reduces dramatically reduced. Compared to other features, features including the mean annual temperature and ECR also have rather higher importance, ranking second and third respectively. The feature importance of the rest of the features is quite lower, while the differences are not considerable. However, the feature representing the characteristics of the mixtures ranks the lowest among all the features.



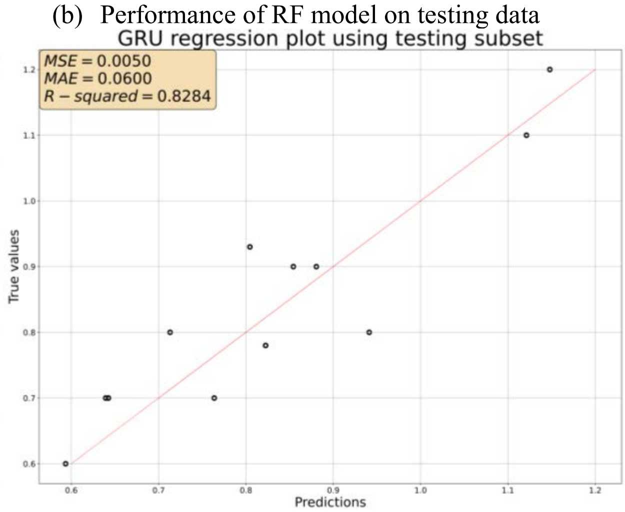
(a) Performance of RF model on training data



(b) Performance of RF model on testing data



(c) Performance of GRU model on training data



(d) Performance of GRU model on testing data

Figure 3. Regression plots of the developed models

4 Discussion

The main contribution of the presented research is to systematically investigate the correlation between process quality and product quality in asphalt construction. Asphalt is a highly complex material, where the quality in each phase of the lifecycle is influenced by various factors and also how the previous phases unfolded [22]. Therefore, this study provides an opportunity to scale up the regression task from the focus on one phase of the asphalt construction lifecycle to multiple phases. This would be significantly beneficial in the highly competitive environment of the asphalt construction sector. For instance, to the contractors, the explicit correlation between process and product quality can eventually help further justify the efforts to improve and optimize the planning and implementation of the on-site operational strategies.

The GRU model performed well despite the small dataset. The size of the dataset is essential in drawing reliable conclusions; however, the consideration of other factors such as the quality of the data and the model's ability to identify significant features and relationships is also crucial to the reliability and validity of the conclusions derived from data analysis.

Essentially, unlike conventional ML algorithms, algorithms such as GRU will have a much higher level of abstraction, thus prone to be greedy to the amount of the data to prevent the overfitting problem. However, a previous study suggests that the high reliability of the model can be achieved even with small datasets [7]. In this study, the R^2 of developed models reached 0.9941 and 0.9893 on two different datasets. Besides, in the presented study, RF was also utilized. Compared to GRU, RF has a rather simpler architecture and less complexity. However, in the represented study, the developed RF model suffered from the underfitting problem with the small amount of data. This could potentially mean that the data used in this study is insufficient to support conventional ML algorithms, such as RF. On the other hand, the developed GRU model showed outstanding capability regarding feature extraction.

Lastly, the previous measurements of the IRI outranked the other features. This is in line with various studies which also applied time-series regression to the IRI data [8,10]. Representing the construction process quality, ECR ranked third, which indicates a rather high impact of construction process quality on product quality. This further highlights the importance of investigating the asphalt product quality from the life-cycle perspective. In addition, the feature representing the properties of the asphalt mixtures ranked the lowest. This is in line with the findings of the previous studies [23].

5 Conclusion and future work

This research aimed to investigate the correlations between asphalt construction process quality and product quality, using data-driven techniques. In this research, the IRI was selected as the output and representation of the pavement product quality indicator.

A GA-based ML model development framework was designed, where RF and GRU were selected as the algorithms. For the validation, a case study was conducted. Based on the collected data, the developed GRU model significantly outperformed the RF model, with an R^2 of 0.8284. After interpreting the permutation importance, ECR achieved the third highest importance, revealing the rather high correlation between process quality and product quality in asphalt construction.

For future work, because the case study in this research was performed on a small dataset, the authors would like to expand the scope of the dataset. Besides, the presented research only focused on the IRI, while in the further study, more product quality indicators concerned with both in-place pavement properties (i.e., density, thickness, etc.) and long-term pavement performance (i.e., raveling, cracking, rutting, etc) can be considered.

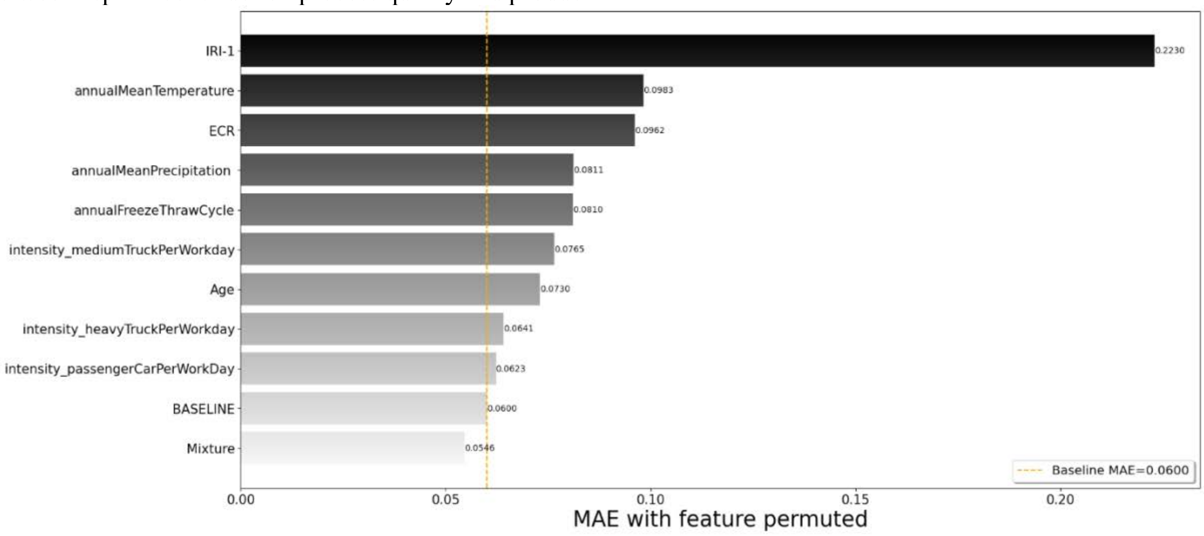


Figure 4. The premutation feature importance

References

[1] S.R. Miller, H.L. ter Huerne, A.G. Doree, Towards understanding asphalt compaction: An action research strategy (in special issue for the IPRC), Built & Human Environment Review. 1 (2008) 11–24. https://research.utwente.nl/en/publications/towards-understanding-asphalt-compaction-an-action-research-strat (accessed February 10, 2021).

[2] S.R. Miller, Hot mix asphalt construction : towards a more professional approach, University of Twente, 2010. https://doi.org/10.3990/1.9789036531283.

[3] F.R. Bijleveld, S.R. Miller, A.G. Dorée, Making Operational Strategies of Asphalt Teams Explicit to Reduce Process Variability, J Constr Eng Manag. 141 (2015) 04015002. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000969.

[4] D. Makarov, F. Vahdatikhaki, S. Miller, A. Jamshidi, A. Dorée, A framework for real-time compaction guidance system based on

compaction priority mapping, Autom Constr. 129 (2021) 103818. https://doi.org/10.1016/J.AUTCON.2021.103818.

[5] D. Makarov, S.R. Miller, F. Vahdatikhaki, A. Doree, A Generic Framework for Automating the Asphalt Construction Process, in: 12th Conference of Asphalt Pavements for Sourthen Africa (CAPSA), Sun City, South Africa, 2019: pp. 1227–1241. https://research.utwente.nl/en/publications/a-generic-framework-for-automating-the-asphalt-construction-proce (accessed April 21, 2021).

[6] S. Sørumgård, G. Sindre, Aspects of process quality, in: Proc. 4th Software Quality Conference, Dundee, Scotland, 1995: pp. 318–326.

[7] M. Mazari, D.D. Rodriguez, Prediction of pavement roughness using a hybrid gene expression programming-neural network technique, Journal of Traffic and Transportation Engineering (English Edition). 3 (2016) 448–455. https://doi.org/10.1016/J.JTTE.2016.09.007.

[8] H. Gong, Y. Sun, X. Shu, B. Huang, Use of random forests regression for predicting IRI of asphalt pavements, Constr Build Mater. 189 (2018) 890–897. https://doi.org/10.1016/J.CONBUILDMAT.2018.09.017.

[9] J. Li, G. Yin, X. Wang, W. Yan, Automated decision making in highway pavement preventive maintenance based on deep learning, Autom Constr. 135 (2022) 104111. https://doi.org/10.1016/J.AUTCON.2021.104111.

[10] J. Li, Z. Zhang, X. Wang, W. Yan, Intelligent decision-making model in preventive maintenance of asphalt pavement based on PSO-GRU neural network, Advanced Engineering Informatics. 51 (2022) 101525. https://doi.org/10.1016/J.AEI.2022.101525.

[11] F. Alharbi, Predicting pavement performance utilizing artificial neural network (ANN) models, Iowa State University, 2018. https://dr.lib.iastate.edu/entities/publication/aef0e177-bebe-4cd4-b7ae-3c82c4a024fb (accessed September 28, 2022).

[12] L. Žiliūte, A. Motiejūnas, R. Kleiziene, G. Gribulis, I. Kravcovas, Temperature and Moisture Variation in Pavement Structures of the Test Road, Transportation Research Procedia. 14 (2016) 778–786. https://doi.org/10.1016/J.TRPRO.2016.05.067.

[13] B.S. Smith, Design and Construction of Pavements in Cold Regions: State of the Practice, Brigham Young University, 2006. https://scholarsarchive.byu.edu/etd/1111/ (accessed March 31, 2023).

[14] L. Breiman, Random forests, Mach Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.

[15] A. Sherstinsky, Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network, Physica D. 404 (2020) 132306. https://doi.org/10.1016/J.PHYSD.2019.132306.

[16] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, (2014). https://doi.org/10.48550/arxiv.1412.3555.

[17] A. Karpathy, L. Fei-Fei, Deep Visual-Semantic Alignments for Generating Image Descriptions, IEEE Trans Pattern Anal Mach Intell. 39 (2014) 664–676. https://doi.org/10.48550/arxiv.1412.2306.

[18] O. Vinyals, A. Toshev, S. Bengio, D. Erhan, Show and Tell: A Neural Image Caption Generator, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 07-12-June-2015 (2014) 3156–3164. https://doi.org/10.48550/arxiv.1411.4555.

[19] philipperemy/cond_rnn: Conditional RNNs for Tensorflow / Keras., (n.d.). https://github.com/philipperemy/cond_rnn (accessed January 24, 2023).

[20] Q. Shen, F. Vahdatikhaki, H. Voordijk, J. van der Gucht, L. van der Meer, Metamodel-based generative design of wind turbine foundations, Autom Constr. 138 (2022) 104233. https://doi.org/10.1016/J.AUTCON.2022.104233.

[21] HOME | Aspari, (n.d.). https://www.aspari.nl/ (accessed January 25, 2023).

[22] C. Han, F. Tang, T. Ma, L. Gu, Z. Tong, Construction quality evaluation of asphalt pavement based on BIM and GIS, Autom Constr. 141 (2022) 104398. https://doi.org/10.1016/J.AUTCON.2022.104398.

[23] Perera, R. W., Byrum, C., Kohn, S. D., & Soil and Materials Engineers, Inc. (1998). Investigation of Development of Pavement Roughness. https://rosap.ntl.bts.gov/view/dot/42761(accessed March 31, 2023).