Coupling the Road Construction Process Quality Indicators into Product Quality Indicators

EngD Thesis

to obtain the degree of
Engineering Doctorate (EngD) at the University of Twente,
on the authority of the rector magnificus,
prof.dr.ir. A. Veldkamp
on account of the decision of the graduation committee,
to be defended
on Wednesday the 21 of June 2023 at 14 hours

by

Ir. Qinshuo Shen

born on the 14 March 1996
in Hubei, China

This EngD Thesis has been approved by:

Thesis Supervisor: Prof. Dr. Ir. A. Dorée

Co-supervisor(s): Dr. Ir. S. R. Miller, Dr. Ir. F. Vahdatikhaki

Graduation committee:

Chairman:

Thesis supervisor(s): Prof. Dr. Ir. A. Dorée

Co-supervisor(s): Dr. Ir. S. R. Mille
Dr. Ir. F. Vahdatikhaki
S. Drok
Dr. Ir. F. Bijleveld
Ir. E. van der Velde

Summary

Road infrastructure is vital for socioeconomic development. However, road construction companies face intense competition due to changes in contracting schemes that demand longer guarantee periods. Traditionally, road construction practices heavily rely on experience-based decision-making, leading to variability and poor pavement quality. To address this, contractors must adopt effective construction process quality control to meet functional requirements and reduce operational variability.

Conventional road construction quality control practices are problematic and insufficient. These practices focus on end results and performance-related specifications, neglecting the connection between process quality and product quality. Consequently, they only verify if on-site operational strategies meet functional requirements, without optimizing activities to prevent road failure. This approach fails to identify which specific activities in previous phases contributed to inadequate mechanical properties, hindering improvement in future projects. To address this, a shift is needed towards explicit and process-oriented quality control is necessary, while understanding the relationship between construction process quality and resulting pavement quality is crucial for making this transition successful.

The complexity and non-linearity of the road construction system make it challenging to investigate the relationship between road construction process quality and product quality. One approach to tackle this challenge is to utilize data-driven techniques, such as Machine Learning (ML), because of the compelling capability of ML in revealing the hidden patterns in the data.

On these premises, this research aims to develop a data-driven method to systematically uncover the underlying correlations between road construction process quality and product quality.

Following the concept of design cycle, the methodology was defined. The problem context was investigated concerned with the theoretical framework and societal embedding, through literature review and stakeholder analysis respectively. Subsequently, the requirement engineering was performed based on the stakeholders' needs to generate functional requirements.

The input-output structure of the datasets required for the ML model development was identified, including input variables such as the quality indicator of the on-site operational strategies, weather conditions, mixture type, and auxiliary parameters in the operational phase

of the pavement (such as traffic intensity and climate condition). More specifically, the quality indicator of the on-site operational strategies will be represented by the Effective Compaction Rate (ECR), indicating to which extent the compaction meets the requirements considering the compaction temperature windows and target number of roller passes. The outputs of the identified data structure include density degree, residual lifespan, and IRI, covering the short- and long-term perspectives of the pavement product quality.

A Genetic-Algorithm-based ML model development method was designed. Given the research context, the regression problem will be applied to the ML model development. Specifically, Random Forest (RF) was selected as the ML algorithm due to its promising performance, the capability of overcoming overfitting issues, and interpretability. In addition, because of the time-variant nature of the output regarding long-term pavement performance indicators, the time-series regression will also be applied, where Gated Recurrent Unit (GRU) was utilized to tackle the complexities of non-linear regression concerned with time-series data.

For the validation, case studies were conducted. The regression of density degree will be based on the data provided by the Dutch contractor Heijmans, collected from a series of construction projects around the Schiphol Airport. For the regression of residual lifespan and IRI, two Dutch highway sections (A58 and A4) with a total length of 4.1 km, were selected. Based on the collected data, the regression model for the density degree using RF failed to satisfy the corresponding requirements regarding the model performance. For the regression of residual lifespan and IRI, both the RF and GRU were used to develop corresponding models. For residual lifespan, the developed RF model outperformed the GRU model, with an R^2 of 0. 8297, while the regression of IRI shows contradictory results, where the developed GRU model significantly outperformed (R^2 is 0.8284). After interpreting the permutation importance, both cases show that the construction process quality indicator represented by ECR achieved the third highest importance, revealing the rather high correlation between process quality and product quality in asphalt construction.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

PQi	Process Quality improvement
ML	Machine Learning
ECR	Effective Compaction Rate
RF	Random Forest
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
GA	Genetic Algorithm
CSV	Comma-Separated Values
XML	Extensible Markup Language
RWS	Rijkswaterstaat
JSON	JavaScript Object Notation
R^2	R-squared
MSE	Mean Squared Error
MAE	Mean Absolute Error

1. Introduction

This section provides the background of this Engineering Doctorate (EngD) project, together with the discovered problem and possible solutions. This section concludes with the representation of the outline of the entire report.

1.1. Project Background

Road infrastructure plays an integral role in socioeconomic development by undertaking the function of transporting people and goods from one location to another [1]. While there is a steadily high demand for road infrastructure, road construction companies are facing fierce competition, which mainly comes from the changes in the contracting scheme that demands lengthy guarantee periods [2]. Over centuries, road construction practices have been treated intuitively that have heavily relies on experience-based decision-making, particularly during the construction phase. However, road construction is an intricate and complex operation [3] that requires highly dynamic and adaptive operational strategies [3–5]. Consequently, considerable variability is generated, which is widely regarded as the root cause of poor pavement quality [3]. Therefore, to ensure that the constructed pavement is compliant with the functional requirements, it is of great importance for contractors to implement effective construction process quality control and reduce operational variability.

According to the definition of [6], quality control can refer to all the actions, including techniques and activities, applied in the production process to meet the quality requirements. As pointed out by Miller et al. [7], the conventional road construction process quality control practices are problematic and insufficient. This is because most of the quality control methods are designed and conducted from the perspective of the end result, and they operate based on performance-related specifications. These methods greatly neglected the link between process quality and product quality. Consequently, the conventional road construction quality control schemes can only determine if the implemented on-site operational strategies can pass the verifications regarding pavement functional requirements or not, while no actual improvement can be made to optimize these activities to reduce or prevent premature road failure. In other words, when the post-construction sample tests indicate inadequate mechanical properties, it is not possible to identify which activities in the previous phases went wrong. As a result, the contractors are unable to carry this lesson to their future projects and try to improve their strategies based on past projects. That is why it becomes necessary to shift from implicit and outcome-oriented quality control to explicit and process-oriented quality control, where the core is to closely monitor the road construction process using key process parameters. But, to make this transition possible, it is imperative to properly understand the relationship between the construction process quality and the resulting pavement quality.

1.2. Problem Statement and Potential Solution

To address this, Miller proposed the methodology of Process Quality improvement (PQi), which provides a framework to enable the quantitative measurement and evaluation of the variability during the asphalt paving process [8]. Specifically, the concept of the Internet of Things (IoT) is deployed in the construction process, using embedded sensors, to directly monitor and assess the asphalt.

In the past decade, initiated by a Dutch research network called ASPARi [9], great efforts have been put into validating and applying the PQi framework in Dutch road construction projects. While PQi is gradually becoming a baseline approach for assessing construction variability, much less is known about how the process quality can eventually affect the product quality of the as-built asphalt pavement. Although there exists a strong correlation between the effectiveness of the road construction process and the eventual quality of the pavement, the current system cannot explicitly and quantitatively map this correlation. Consequently, the extent to which the adopted operational strategies contribute to pavement quality is still unknown. Given that the primary goal of monitoring and evaluating process quality is to improve the quality of asphalt pavement and optimize the construction process, it is important to have a comprehensive and unambiguous comprehension of the correlations between process quality and product quality in asphalt construction.

The system's complexity and non-linearity render the investigation of the correlations between road construction process quality and product quality a challenging task. One possible approach to deal with this is to use data-driven techniques, such as Machine Learning (ML). Over the past few years, the Architecture, Engineering, and Construction (AEC) industry has seen momentum to adopt ML as a transformative tool. This is not only because of its capability of capturing hidden patterns underlying the data, but also due to the sheer volume of information and data generated, transmitted, processed, and utilized within the industry that can be used to better analyze the processes [10]. In general, ML can be described as an intelligent method that improves its performance by learning from historical data to make predictions [11]. This process mimics human learning, where knowledge is gained from past experiences through induction, summarization, and internalization.

Specifically, in the field of asphalt construction, great attempts were made by developing and utilizing ML-based empirical models to thoroughly investigate the non-linear correlations between various factors and asphalt pavement quality indicators, including the long-term performance [12–15] and in-place verification results (i.e., the mechanical properties of the pavement obtained through in-place or laboratory measurements) [16–18]. These studies considerably leveraged the capabilities of ML in tackling underlying non-linearities, thus facilitating the exploration of the hidden correlations between input and output parameters. However, previous studies in the field have predominantly limited their modeling scope to just one phase of the asphalt construction lifecycle, resulting in the underappreciation and negligence of the significance of the quality of other crucial phases, such as mixture design and construction to the product quality of the asphalt construction. This is specifically problematic because previous studies failed to perceive the quality of asphalt construction as an integral system, meaning the quality of every phase of its lifecycle is not only affected by itself but also tightly influenced by the previous phases as well. Consequently, when revealing the question of what leads to the achieved product quality of asphalt construction, these studies failed to provide a comprehensive profile that takes all the critical elements into account.

1.3. Project Objective and Questions

On these premises, an urgent problem at hand is to fully accomplish the transition from outcome-oriented road construction quality control to process-oriented quality control. This would require an effective scheme to capture key parameters involved in the road construction process, as well as an explicit understanding of the correlations between the construction process quality and the resulting pavement quality. While the former has been covered in the development and implementation of the PQi methodology, much less has been explored regarding the latter.

Therefore, this EngD design project focused on the investigation of correlations between the construction process quality and the resulting pavement quality in order to quantify the impact of the construction process on the obtained quality of the pavement. Specifically, this would require the application of data-driven techniques, particularly ML. This is because of the considerable system complexity and non-linearity, as well as the compelling capability of ML in revealing the hidden patterns in the data.

To sum up, the objective of this project was formulated:

To develop a data-driven method to systematically uncover the underlying correlations between road construction process quality and product quality.

The design project objective stated in the previous section can be achieved by answering the following questions derived from the project context and the generated objective:

1. How to develop the data structure to illustrate the input-output relations regarding the asphalt pavement process and product quality considering the lifecycle of the asphalt construction?
    a. What are the definitions and scopes of asphalt pavement process and product quality, considering the asphalt construction lifecycle?
    b. What are the indicators that can represent the process and product quality of asphalt pavement?
2. How can the ML-based method be developed and used to extract information from the structured dataset?
    a. What are the input-output structures that can capture the potential correlation between process and product quality indicators?
    b. How to solve the integration conflicts between different data sources to prepare the datasets?
    c. What ML algorithms are suitable given the problem context, modelling objectives, and performance?
    d. How to train, test, and optimize the ML models to achieve optimal performance without yielding generality?
    e. How to analyze and interpret the developed ML models to extract knowledge from ML models?
3. What conclusion and recommendations can be made based on the results of the investigation?

Among them, the first question is concerned with how should data be structured regarding variables that should be considered and their relations. To answer this research question is the prerequisite of obtaining valuable output from the ML approach. The second research question focuses on the implementation of ML method, i.e., how different ML models are developed according to different modelling objectives. To answer this question, it is essential to consider the general ML model developing process, including dataset development, selection of ML algorithms, ML model training and validation, hyperparameter optimization, and analysis of the obtained ML models. The last question is about a final assessment of the entire execution of this design project.

1.4. Design Scope

To specify the design focus and limitations, it is essential to explicitly determine the boundary of the design project.

This project is initiated by ASPARi. Founded by the University of Twente, ten of the largest contractors in the Netherlands, and the Directorate General for Public Works and Water Management of the Netherlands (Rijkswaterstaat), ASPARi has been dedicated to collaboratively connecting multiple organizations in the Dutch road construction industry to improve the performance of the asphalt construction process and its related activities [9]. To incorporate the scope of this project into the specific ambitions, strategies, and needs of ASPARi, this project merely focused on the road construction industry in the Netherlands. Therefore, the data collection was also confined to the Dutch road construction industry and within its particular context.

As the final deliverable of the road construction projects, various road types can be defined based on the construction materials (e.g., asphalt roads, concrete roads, gravel roads, etc.). Among them, because of various advantages that it can provide, such as jointless surface, superior smoothness, low degree of wear, low maintenance cost, high durability, and recyclable usage [19,20], the asphalt pavement has been widely used as the major form of the road infrastructure, particularly in the Netherlands. Therefore, to ensure the data availability and quality to successfully implement the data-driven techniques applied in this project, only asphalt roads were focused.

Furthermore, in this study, the definition of the asphalt construction process quality followed the previous work of Bijleveld et al. [3], as the degree to which extent the variability of key process characteristics, i.e., the homogeneity of the pavement temperature and the consistency and effective compaction, is controlled. Therefore, the assessment of variability resulting from a particular operational strategy can be recognized as process quality.

Product quality can be determined as a cluster of general product attributes that can be expected to satisfy the corresponding demands at an acceptable level [21]. According to [22], the quality characteristics of the product can be both tangible and intangible. The former mainly refers to the physical attributes of the product, such as the physical properties of the product, appearance, and performance, while the latter is concerned with the perceived quality, including the related services, information, and supplier characteristics and behaviours [23]. Similarly, according to [24], the intrinsic characteristics of the product can be defined as the internal quality, which is product-based and can either be objective or subjective. The objective internal quality is impartially assessed based on whether the performance of the product can meet the customers' expectations or has a low failure rate, thus enabling the quantitative reflection of this type of product quality [25]. Nonetheless, the subjective internal quality mainly concerns the perception of the intrinsic cue, which can be too subjective to capture. On this premise, in this study, the scope was confined to the identification of tangible and objective product quality. This is also partly due to concerns about the availability and accessibility of corresponding data for developing a well-structured dataset to enable the application of data-driven techniques.

1.5. Report Structure

The rest of the thesis has been organized as follows: after the introduction, Chapter 2 discusses the design methodology. Next, Chapter 3 represents the theoretical framework and the social context, which together illustrate the results of the investigation of the problem context. Chapter 4 covers the system design, representing the detailed procedure of developing the proposed data-driven approach to fulfil the design project objective. Chapter 5 illustrates the case studies conducted for validating the proposed approach and the obtained results. Chapter 6 represents the corresponding discussion. Lastly, Chapter 7 concludes the project.

2. Project Design Methodology

In order to systematically reach a design solution to the identified problem from the previous chapter, it is essential to follow a methodological approach. In this design project, the adopted methodology was referenced from the concept of the design cycle to provide systematic guidance to fulfil the design project objective [26]. According to Wieringa [26], four critical phases are included in the design cycle, namely the problem investigation, system design, system validation, and system implementation, to ensure that the design and improvements can be made after thoroughly investigating the problem context and opportunities. To align the design cycle with the project context, several adjustments were made on the basis of the original design cycle, as illustrated in Figure 1.

Compared to the design cycle proposed by Wieringa [26], the methodology used in this project primarily focused on the first three design phases. Starting with the problem investigation, in this phase, the knowledge context and the social context of the data-driven method under design were explored, using the literature review and stakeholder analysis respectively. The knowledge context of the design provided existing knowledge related to the design questions, while the social context specified the goals that the design should accomplish based on the expectations of stakeholders. The system design phase included requirement engineering, the development process of ML models, and the interpretation of ML models. Lastly, the design was validated to assess the attainment of the desired performance by the designed products.

It is worth noting that the original design cycle is an iterative process allowing continuous optimization of the design through design validation and design implementation [26]. The latter phase aims to implement the design solution in the real-world environment. However, to deploy the design output of this project into the guidance and decision-making enhancement in the actual road construction practices, it might encounter the challenges such as regulatory and legal considerations, increased cost in digital infrastructure (e.g., data storage, computational resources, and hardware infrastructure, etc.), and interpretability and trust. Consequently, the implementation of the design outcome would only stay at the conceptual level without being able to be validated. Given the constraints on the limited time and resources in this project, the implementation of the design was excluded from the project scope.

The rest of this chapter elaborated on the details of each design phase.

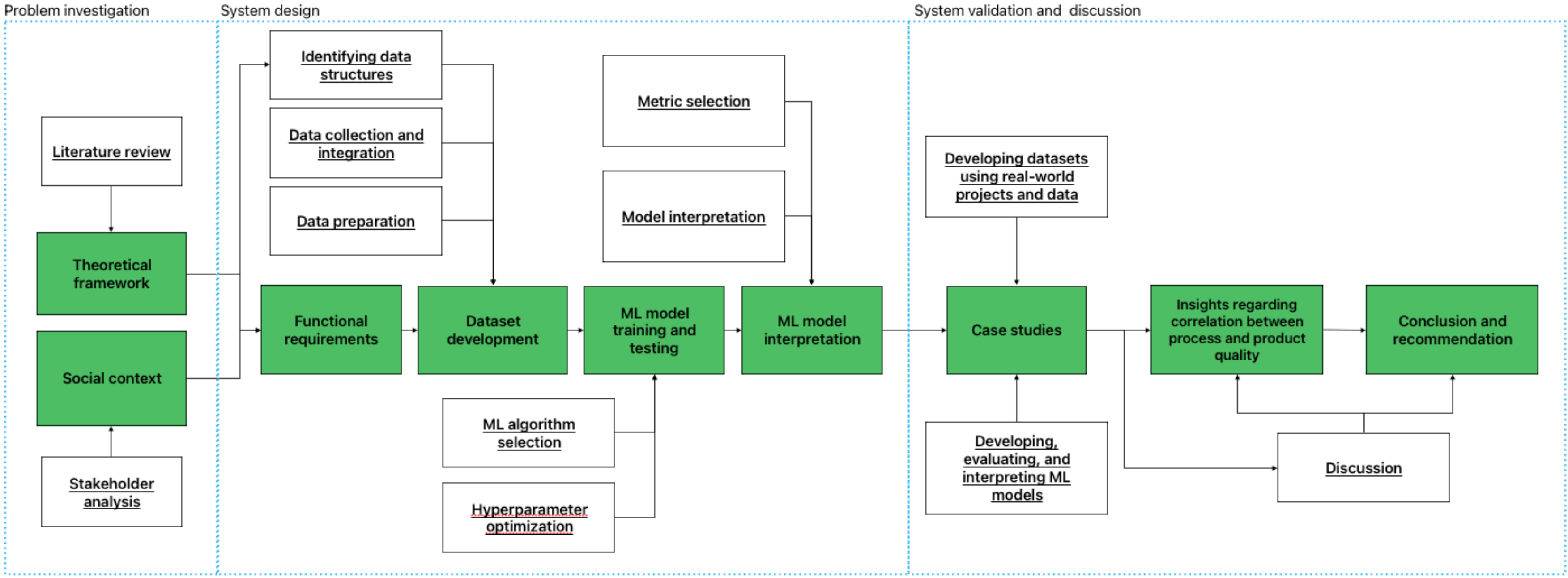


Figure 1. A representation of the detailed design process

2.1. Problem Investigation

In the problem investigation phase, two primary design outputs were obtained, namely the theoretical framework and social context. The former involved exploration and representation of the theories that underlie the design problem achieved through a literature review, while the latter pertained to the interaction between the proposed design and its relevant social environment, including the stakeholders involved and corresponding needs. In the subsequent system design phase, these two outputs served as the foundation for the requirement engineering of the proposed framework, leading to the identification of consistent and testable requirements that ensured the framework meets its intended goals and enabled verification and validation. Moreover, the theoretical framework, provided a formal and explicit model to encode the relationships among the concepts related to quality in asphalt construction and operation, thereby facilitating the formation of the data structure used in the ML model development process.

2.1.1. The Exploration of Knowledge Framework

The theoretical framework of this design project can connect the design questions with the existing knowledge by conducting a critical examination of the relevant literature. This examination resulted in the extraction and definition of key concepts and theories, including a definition of quality composition within the asphalt construction lifecycle, the identification of process and product quality and corresponding indicators, and an overview of the applications of ML in the field of asphalt construction.

To explore the theoretical framework, a literature review was conducted to analyze the existing work on the topic of interest [27]. Figure 2 below represents the general process of the literature review.

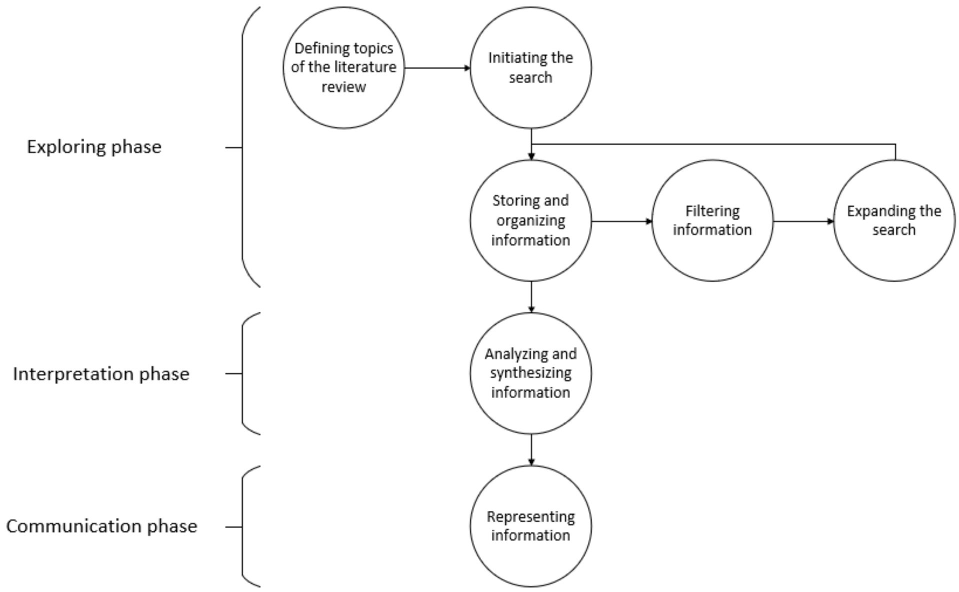


Figure 2. A representation of the literature review process

The literature review process was iterative and contained three phases. The first phase was mainly concerned with searching relevant literature based on identified topics. In order to initiate the literature search and confine the search space of relevant literature, several topics were defined as the keywords for searching the literature, including “asphalt construction process quality”, “asphalt pavement quality”, “asphalt construction quality control”, as well as knowledge concerned with data-driven techniques, such as “data cleansing”, “feature engineering”, “machine learning model development”, “machine learning in asphalt construction”, etc. These topics were then used as inputs in the search process of the corresponding literature. To ensure the quality and comprehensiveness of the extracted information, the types of literature included papers published in journals or conferences with high significance and acknowledgement, dissertations at master or doctorate levels with high quality and novelty, and documents that play an important guiding role in the industry. In addition, Scopus and Google Scholar were selected as the main platforms for the literature search, given that these two databases have indexed an extensive amount of literature covering these defined topics. Then, the pre-selected literature was further examined and filtered. More specifically, the titles, abstracts, and conclusions of the pre-selected literature were assessed to exclude irrelevant or redundant literature regarding certain topics. Subsequently, the pre-defined topics were updated by including keywords with high levels of co-occurrence from the literature after the filtering process, to further expand the search space of more representative and relevant literature. Then, the next iteration started by using the newly updated topics as inputs for the literature search. The entire exploring phase was regarded as completed until no more new topics were identified from the updated literature.

The next phase of the literature review started with extracting key information from the selected literature after conducting the exploring phase, including the key findings, applied methodologies, theoretical frameworks, and any other information that is relevant to the corresponding topic. Then, the extracted information was analyzed by identifying the research patterns and comparing the discrepancies across the literature, thus developing a deep understanding of the existing knowledge on the corresponding. Next, the analyzed information was synthesized by being integrated into cohesive narratives, by identifying the main arguments and theories and organizing them in a logical and coherent manner.

Lastly, the final outcome of the literature review was represented. Based on the analyzed and synthesized information, the theoretical framework was structured.

2.1.2. The Exploration of Social Context

In order to have a better understanding of the social context of the system under design, the social embedding of the proposed assessment framework will be thoroughly explored, which represents the abstract and functional connections between the system under design and other elements that can realize its societal functions.

To discover the social context of the design project, a stakeholder analysis was performed to identify and understand the interests of different organizations towards the project. The stakeholder analysis was initiated with the identification of stakeholders related to the project, by specifying all the individuals and organizations that have interests in or are affected by the project. Internal brainstorming sessions, with participants including the EngD candidate and the supervisory team, were conducted to generate a list of potential stakeholders, considering both the internal and external organizations and individuals that are directly or indirectly connected to the design project.

Next, interactive activities, such as workshops and brainstorming sessions, were organized to collaboratively assess the interests of identified stakeholders and generate their needs. Key stakeholders were invited to participate in these interactive activities, where they were encouraged to share their expectations of the desired project outcome, challenges, and concerns, as well as potential support that can be offered by them. Lastly, insights were captured from the discussions to generate the stakeholders' needs.

2.2. System Design

The system design phase mainly focused on the development of ML models with the alignment of the stakeholders' needs to fulfil the project objective. As shown in Figure *1*, it started with the requirement engineering to generate the requirements for the design output, to explicitly reflect what is the design expected to achieve according to the stakeholders' needs. Then, based on the identified requirements, subsequent activities involved in the system design can be guided and constrained.

The rest of the system design phase was mainly concerned with the development of ML-based regression models to fulfil the design project objective. Figure 3 represents the overview workflow for the developing process of these ML models, including the establishment of datasets, ML model training and testing, and ML model interpretation.

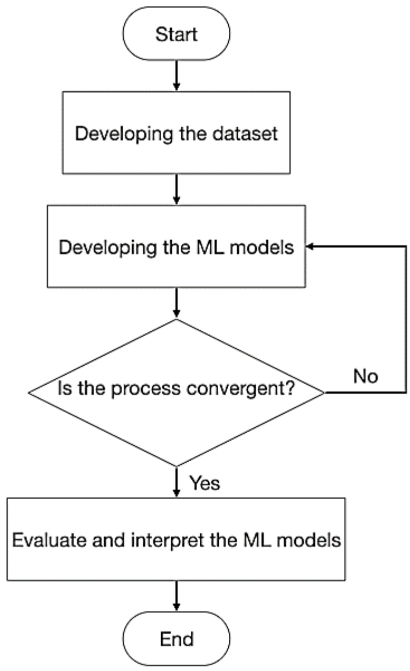


Figure 3. Overview of the proposed ML modelling process

Overall, the process of dataset establishment is a crucial step in the development of ML models, comprising several tasks including the definition of input-output structures aligned with the modelling objectives, data collection, data integration, and data preparation. In the following step, ML models were developed using the established datasets through an iterative process, employing optimization algorithms to obtain optimal configurations based on the evaluation of the ML models on the pre-defined metrics. Lastly, the interpretability of the models was explored by analyzing the contribution of input features to the overall performance, providing deeper insight into the inner workings of the models and a clearer understanding of the relationship between process quality indicators and product quality indicators.

The details of the system design phase were explained as follows.

2.2.1. Requirement Engineering

The main purpose of the system design is to specify the requirements of the system under design and develop the system for the investigated problem. While the identification of stakeholders' needs is concerned with the problem domain, which mainly focuses on the question of "what do stakeholders want", the system needs mainly focus on the solution domain, which answers the question of "what does the system need to do". Therefore, based on the stakeholders' needs and explored knowledge context, system requirements can be identified, which will be used as the guideline for the rest of the system design and validation.

Specifically, the concept of requirements engineering will be applied in the formulation of the system requirements. Based on the outcomes from the stakeholder analysis regarding the stakeholders' needs, firstly, the stakeholders' needs were mapped into various functions of the system under design by linking what do stakeholders need to what functions should the system provide.

Next, the requirements can be formulated. In order to ensure the formulated requirements are consistent, testable, traceable, and complete, and also enable the verification of the requirements, a formulation template, as proposed by [28], will be utilized. The template is shown in Figure 4 below.

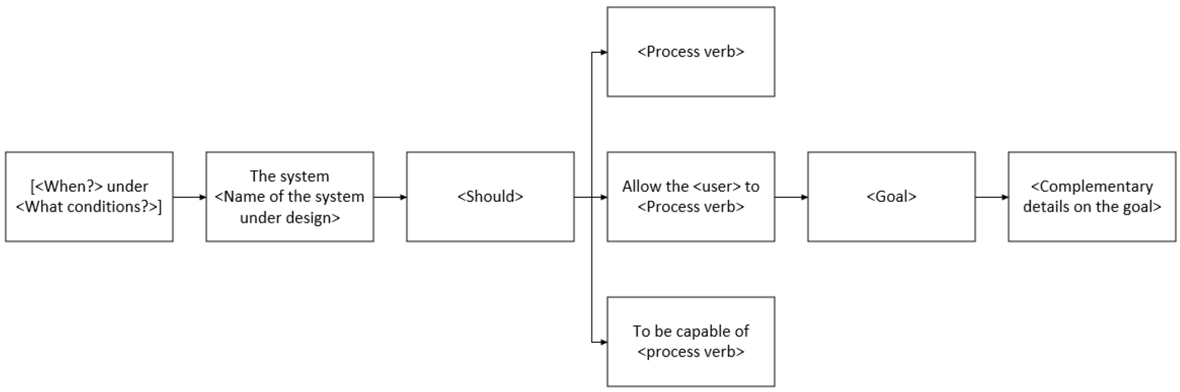


Figure 4. The template of the system requirements formulation [28]

2.2.2. Dataset Development

The first step in ML development is to establish the datasets needed for providing the training and testing data. This step contains the formulation of the data structure, data collection and integration, and data preparation.

2.2.2.1. Data Structure Formulation

In this design project, the primary input for identifying the data structure was the theoretical framework explored in the problem investigation phase, which provided important concepts and relations from the literature. In addition, the pavement lifecycle management ontology developed by Sadeghian [29] was also used as a reference for generating the data structure. Therefore, the target variables, i.e., the outputs that the ML models are expected to predict were first identified, considering whether or not they can be sufficiently representative as the indicators to the pavement product quality. Subsequently, according to different outputs, corresponding input variables, i.e., factors that are likely to influence the identified target variables, were determined by examining the potential cause-effect relationships through the theoretical framework and the pavement lifecycle management ontology [29].

For the identification of both input and output variables, brainstorming sessions and workshops were also held with the participation of domain experts, to collaboratively generate the potential variables. During these interactive sessions, the EngD candidate prepared several pre-selected variables regarding road construction process and product quality indicators, where the participants were encouraged to assess whether these variables were adequate or whether new variables should be considered. In addition, the data availability of the identified variables was also discussed, to determine if it is possible to acquire a certain amount of data to enable the ML model development. Otherwise, variables would be removed if too less data can be collected.

2.2.2.2. Data Collection, Data Integration, and Data Preparation

To efficiently collect data according to the identified input-output structure, a data collection strategy was made. Based on the identified input-output structure, multiple data sources, from which the corresponding data can be extracted, were examined, regarding the data contents, data types, and relations. Subsequently, the accessibility of these data sources was evaluated, considering data owners, data formats (such as CSV, JSON, or XML), and authorization (whether the database is open, or permission is needed). Then, for different data sources, according to their accessibility, different data collection methods were selected, including data request forms, surveys, interviews, etc. To store the collected data with a global structure for mapping heterogeneous data uniformly, a data warehouse was designed based on the pavement lifecycle management ontology [29] and the identified input-output structure.

Because the data collection was concerned with multiple heterogeneous data sources, it is critical to ensure that the collected data with different resolutions can be integrated into the same data structure. Based on the examined data sources and the designed data warehouse, data integration conflicts were identified. To cope with the identified conflicts, the matching fields between data sources were determined, thus resolving the discrepancies.

After collecting data from the identified sources, an exploratory analysis was performed on the collected data by mainly examining the correlations and patterns. According to the results of the performed data exploratory analysis, dimensionality reduction was conducted to remove variables that are highly correlated, to prevent overfitting and underfitting problems. Subsequently, the outliers were detected from the obtained datasets, which were removed eventually. Besides, the data completeness was also assessed, where missing values were completed using statistical characteristics of corresponding variables, such as using mean values. Lastly, the data transformation was also conducted, focusing on encoding the categorical data into numerical data thus enabling the application of ML.

2.2.3. ML Model Training and Testing

The training process of ML models was initiated with the selection of ML algorithms. Based on the problem context of the design project, supervised learning will be applied, in which instances with both the input and output data will be handled [30]. In addition, the development of ML models should fall into the regression problem, where a map between continuous input and output variables is expected to build [31]. In the theoretical framework, several ML algorithms and their characteristics were briefly represented, which was used as the foundation for the ML algorithm selection.

Subsequently, the developed datasets were re-shuffled and randomly divided to support the training and testing processes of the ML models. In the training process, the training sets were used as the input of the machine learning models. However, the re-shuffled and randomly divided training and testing sets could still introduce biases to the predictive performance of the developed ML models based on a single division of training and testing sets, thus being unable to provide consistent assessment regarding the model performance. Therefore, a method named k-fold cross-validation was adopted by further dividing the subset into k non-repeating sections, as shown in Figure 5. During the training process, the models were 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 thus ensuring better utilization of data, and reducing the sensitivity of the models' performances to how the training subset will be further split.

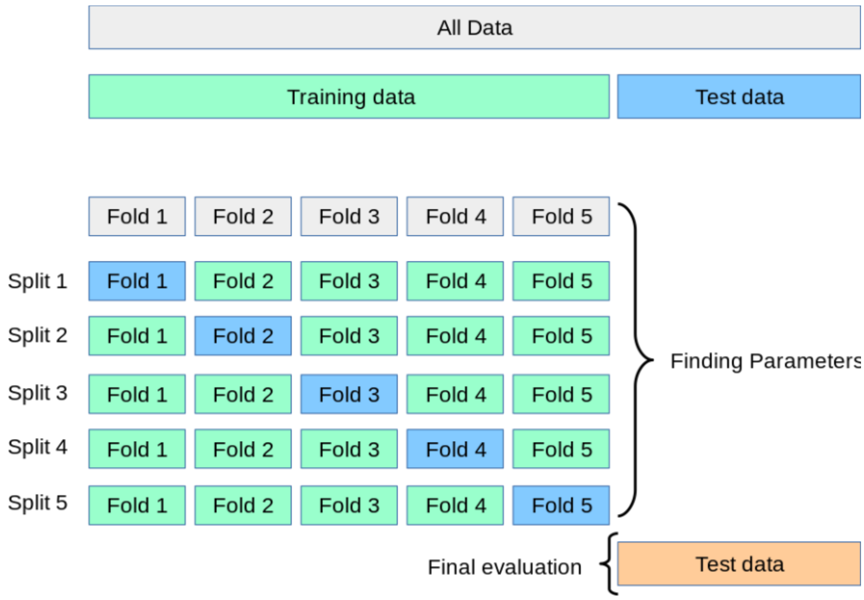


Figure 5. A representation of k-fold cross-validation (k equals to 5) in the ML model training and testing processes [32]

In order to obtain optimal performance of developed machine learning models, it is essential to fine-tune and optimize the model configurations. In this research, optimization focused on the hyperparameters of the models, which are external configurations of the machine learning model, and whose values cannot be estimated and normally define higher concepts within the model. In the system design phase, the hyperparameter optimization algorithm was adopted based on the ease of implementation and capability of finding the global optimum, to iteratively train and evaluate the obtained ML models until the convergence condition was met.

The testing phase is connected with the model evaluation, where the trained models were used to generate predictions given the inputs from the testing sub-dataset. Then, the predictions were compared with the “true values” of the outputs from the testing set based on various metrics, thus obtaining the assessment of the model performance.

2.2.4. ML Model Interpretation

Lastly, in order to explicitly represent the correlations between various process quality indicators and product quality indicators, the obtained regression models were interpreted. However, because of the black-box nature of ML, it is difficult to explicitly explain the internal mechanism of the ML models in terms of how the algorithm results in certain results.

To cope with the difficulty in explaining the ML models, one widely applied approach is the model agnostic method, which solely examines the input and output behaviours of the models instead of inspecting the inner structures and parameters [33]. Therefore, in this project, sensitivity analyses were performed on the developed ML models, to investigate how would the predictions of the output variables change according to the changes regarding the input variables. In other words, the importance of the input variables (features) regarding the impact on the model errors can be examined. Various methods can be identified from the literature for calculating the feature importance, such as impurity-based feature importance and permutation feature importance [34]. Based on the applied ML algorithms, the method that can be applied for all the selected algorithms were applied to provide explicit insights into the models, as well as the hidden correlations between inputs and outputs, i.e., between process and product quality indicators in this study.

2.3. System Validation

It is essential to validate the corresponding design outputs to assess the efficacy of the proposed treatment in addressing the identified issues and fulfilling the established needs and requirements.

This design project endeavours to validate the proposed data-driven method for analyzing the correlations between process and product quality indicators in asphalt construction through the implementation of real-world data via case studies. The initial step was to explore, assess, and analyze the PQi measurements over the past few years. This was followed by the filtering of the measurements and associated archived data based on their completeness. Besides, the projects used in the case studies for modelling the pavement's long-term performance were confined to the surface layer, given that most of the distress takes place on the pavement surface.

Utilizing the identified cases, the system under design functioned as a pipeline to generate the datasets to develop different ML models for different regression tasks. Each ML model was assessed and interpreted. The efficacy of the method was then evaluated according to the alignment of the project outcomes with the requirements established in the requirement engineering phase. Furthermore, the insights generated through model interpretation were compared with existing knowledge and previous findings within the related field, to determine to which extent can the design project address existing knowledge gaps and contribute to a more nuanced understanding of the topic. This comparison also provided insight into the overall robustness and validity of the findings from this design project.

3. Problem Investigation

In this section, outputs of the problem investigation phase will be demonstrated, in terms of the theoretical framework and social context.

3.1. Theoretical Framework

3.1.1. Multiple Aspects of the Quality in the Asphalt Construction Lifecycle

When applying this definition to the context of road construction projects, it is essential to map the inherent characteristics of the road construction project to the scale of its entire lifecycle. This is because the asphalt construction process is highly composite in nature, where the quality of each stage is influenced not only by its own factors but also by the preceding stages [35].

Derived from asphalt construction lifecycle ontology [29], Figure 6 illustrates an overview of the quality of a typical road construction project, with regard to various critical aspects. The pre-construction phase is primarily concerned with the properties of raw materials and the asphalt mixture respectively. For the former, the raw materials mainly include aggregates and bitumen, where the quality of these two types of raw materials plays an integral role in satisfying the mechanical performance of the asphalt mixtures, especially in providing the adhesion, which can be influenced by the chemistry of the aggregates, mixing temperature, and surface texture [36–38]. After determining the composition of different raw materials, the asphalt mixtures can be manufactured from the asphalt plants, whose properties can be reflected from numerous mechanical performance indicators by conducting corresponding type tests. These indicators include the theoretical maximum specific gravity, measured bulk-specific gravity, air void percentage, water sensitivity, dynamic modulus, stiffness, fatigue resistance, rutting depth, resilient modulus, indirect tensile strength, measured stone loss, the bond strength between two adhesive layers, noise reduction, etc. [39–43].

In the construction phase, the quality of the construction operational strategies, i.e., the logistics, paving strategies, and compaction strategies, can be assessed by measuring and evaluating corresponding key construction process characteristics. Verification results, including in-place measurements like nuclear gauge measurements of pavement density and laboratory tests on cores, can be conducted to determine the direct outputs of the on-site operations.

Upon completion of the construction, asphalt pavements are functional in their operations and thus enter the post-construction phase, where external stressors such as traffic intensity and weather can affect their performance. The long-term quality of the as-built pavements is ultimately evaluated through measurements of pavement performance in terms of the severity of various failure modes, such as raveling, cracking, rutting, and roughness.

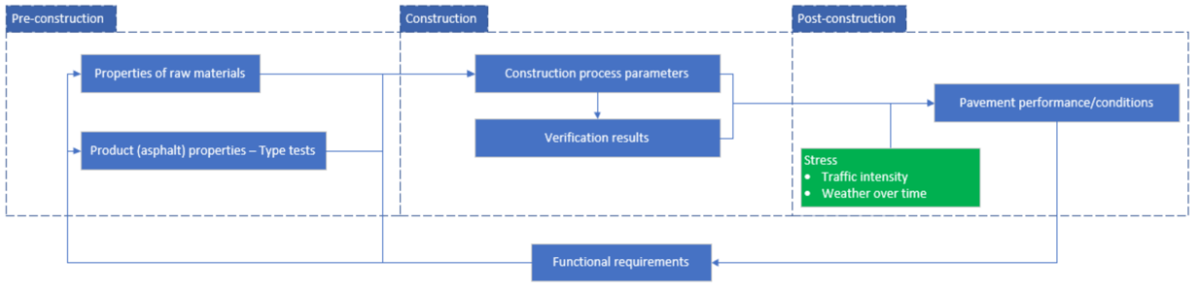


Figure 6. An overview of different aspects of quality within the road construction lifecycle

Based on Figure 6, the various critical aspects pertaining to quality in the asphalt construction lifecycle can be roughly divided into two categories, namely process quality and product quality. The former can refer to the extent to which the process of making an artefact, i.e., the asphalt pavement, has conformed to the standards. Therefore, it includes the direct assessment of the quality of the operations during the construction process, i.e., to which extent the process is well-controlled to reduce variability. Additionally, because the operational strategies are not the sole factor contributing to the variability of the construction process, other factors will also be considered when assessing the process quality, including the ambient conditions and the properties of the input mixtures.

Product quality, on the other hand, refers to how the products satisfy the tangible and intangible requirements, which is positively affected by the process quality indicators. In this design project, the product quality of the asphalt construction will be confined to the tangible quality, given intangible values of the pavement (such as beauty, historical values, emotional responses, etc.) are too abstract to evaluate and model using data-driven methods. In general, product quality encompasses two distinct aspects, including verification results that focus on the in-place or short-term mechanical properties of the pavement, and long-term performance that will be measured regularly during inspections. Similar to the process quality indicators, the product quality will also receive the impact of environmental factors during the operation of pavements, such as climate and traffic intensity. Figure 7 provides the representation of the cause-effect relations between the process and product quality indicators.

The following sections will detail the explanation and identification of the process and product quality and corresponding indicators.

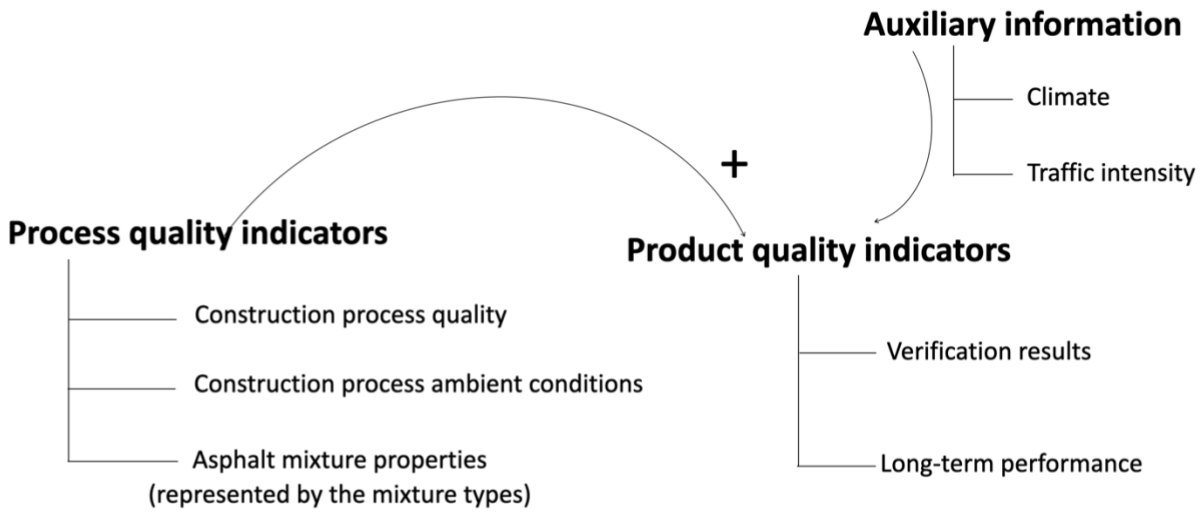


Figure 7. The cause-effect relations between different quality aspects in the asphalt construction lifecycle

3.1.2. Asphalt construction process quality and Relevant Indicators

In this study, the definition of the asphalt construction process quality will follow the previous work of [3], which defines quality as the degree to which the variability of key process characteristics, i.e., the homogeneity of the pavement temperature and the consistency and effective compaction, is controlled. Therefore, the assessment of variability resulting from a particular operational strategy, in conjunction with specific ambient conditions and mixture properties, can be recognized as process quality.

Figure 8 illustrates a holistic picture of the components involved in the asphalt construction process, which can be roughly divided into asphalt mixture production, on-site operations, and transport. Variability in the process can be caused by any of these component classes. Asphalt mixture production refers to the production process of asphalt mixtures, where high variability can be associated with the production, sampling, and testing methods. As for the on-site operations, it is mainly concerned with the operational activities taken on the construction site including the asphalt paving and compaction, which highly rely on craftsmanship in the current practices, and consequently. This dependence causes a high degree of variability [3,7]. Lastly, transport or logistics connects the different stages involved in the asphalt construction phase. However, variability can also be caused during the transport process if the materials are delivered too early or too late, which will affect the delivery and initial temperature of the mixture during the paving process.

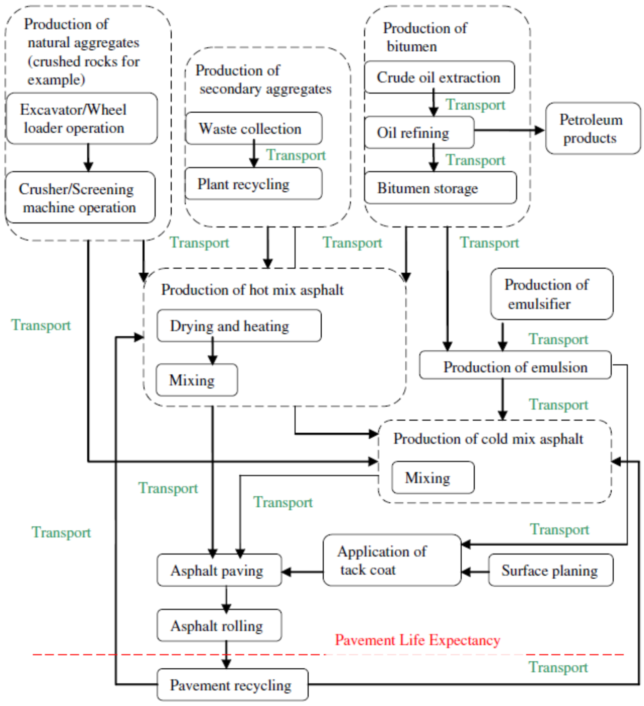


Figure 8. Units involved in the asphalt construction process [44]

The measurable representations of process quality can be regarded as process quality indicators. As illustrated in Figure 8, the on-site operations of asphalt construction mainly include the asphalt paving process and compaction. Therefore, the process quality indicator can be 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) [45–48].

More specifically, the construction process quality is assessed using the Effective Compaction Rate (ECR) index proposed by [49], as indicated in Equation 1.

$$\mathrm{ECR}_{\mathrm{p,k}} = \frac{\mathrm{n}_{\mathrm{p,k}}}{\mathrm{N}} \tag{1}$$

where $\mathrm{n}_{\mathrm{p,k}}$ refers to the number of cells that have received $\pm\mathrm{k}$ passes compared to the target number of passes and at least p% of received passes were within the defined compaction window, and N represents the total number of cells.

However, ECR is only applicable for the assessment of the compaction effectiveness from a macro perspective, such as evaluating different sections in the asphalt pavement. As for the micro perspective, the evaluation of the quality of the construction operation will be scaled down to the centimeter level, such as the PQi methodology. Figure 9 represents the analytical matrix adopted in the PQi measurement, where the concept of cells has been utilized as the basic analytical unit for representing the process and product quality, which is defined by the resolution of collected data due to the update rate and the precision of the infrared camera and GNSS receivers.

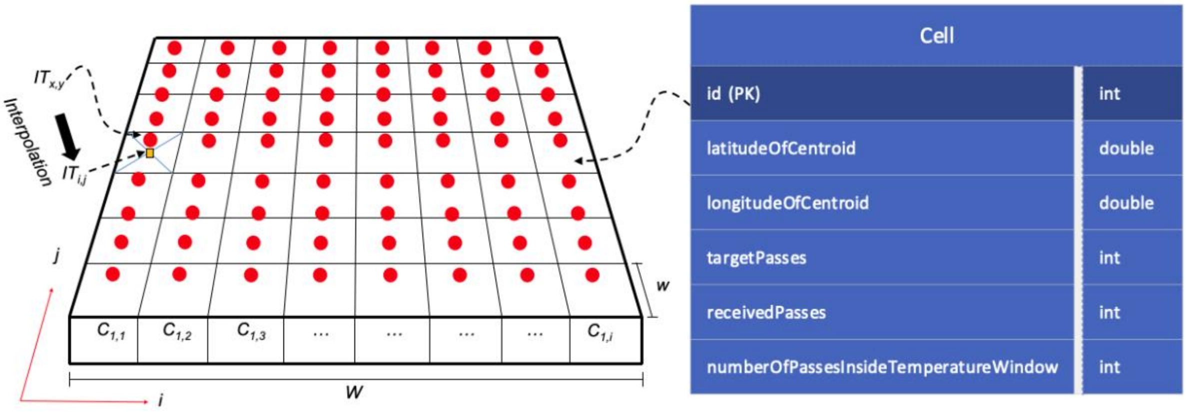


Figure 9. Construction process quality indicators at the cell level

Therefore, when applying this concept to a rather small resolution, i.e., the single cell, an adjustment will be needed. This is particularly essential when aligning the process quality of each cell to the measured mechanical properties such as the density, which are tested on a much smaller scale. According to the definition of ECR, the rate of compaction effectiveness is correlated with two factors, namely (1) the deviation between the achieved and target number of roller passes and (2) the percentage of the received roller passes within the temperature window. Therefore, when applying the concept of the ECR to evaluate the construction process quality on the scale of each cell, these two factors can be used.

In addition, on-site operations can be intricate, whereas multiple factors apart from the exact execution strategies may also jointly influence the process quality, such as the type of asphalt mixture and ambient conditions [21]. Therefore, in the identification of process quality indicators, these two additional factors need to be included.

3.1.3. Asphalt Construction Product Quality and Relevant Indicators

Product quality can be quantitatively characterized by product quality indicators. Given the fact that the ultimate purpose of focusing on product quality is to improve products' capability of satisfying the functional requirements [24,50], it is, therefore, essential to pay attention to the quality indicators of asphalt pavement that can contribute to obtaining a high level of satisfaction. Based on a previous study [51], there is a positive correlation between the performance of products and the level of satisfaction, and the research [52] indicated that a lower failure rate of the products may result in higher satisfaction. Therefore, product quality indicators of the asphalt pavement are defined as quantitative expressions of the tangible and objective attributes of the pavement, related to the performance and resistance to distress from both the short- and long-term perspective.

Based on the given definition, several product quality indicators can be extracted from past studies, which can be categorized into two types, namely indicators related to the in-place properties of the pavement and long-term pavement conditions, as summarized in Table *1*.

Table 1. Identified product quality indicators

Product quality dimensions	**Product quality indicators**	**Definitions**
Pavement properties	Elastic modulus	The elastic (or resilient) modulus of the tested pavement layer [53,54].
	Dynamic modulus	The viscoelastic behaviour of the asphalt layer, with the consideration of time and temperature [55,56].
	Penetration Index (DPI)	The vertical movement of the cone [57].
	Density	The density of the asphalt pavement, which plays an important role in the development of the distress of the pavement, such as rutting [58].
	Stability and flow	The internal friction and cohesion whereby cohesion is a measure of the bitumen binding strength, and internal friction a benchmark of the interlocking and friction resistance of aggregates.
	Aggregate gradation	The particle size distribution of the coarse aggregates.
	Indirect tensile strength	A high tensile strain at failure indicates that a particular asphalt can tolerate higher strains before failing, which means it is more likely to resist cracking than an asphalt with a low tensile strain at failure.
	Air void percentage	The ratio of the air voids contained in the asphalt mixture.
	Layer thickness	The thickness of the asphalt layer.
	Ride quality mean (MRI)	This indicator reflects the pavement's smoothness [59].
Long-term pavement conditions	Pavement roughness	The irregularities of the pavement surface [60]. The condition regarding the roughness can be numerically represented using the International Roughness Index (IRI) [61].
	Cracking	The characteristics and quantification of the type and severity of the surface cracking, including fatigue cracking, block cracking, edge cracking, transverse cracking, longitudinal cracking, etc [62–64].
	Ravelling	The surface distress caused by the shedding of aggregate particles and the loss of asphalt binder due to hardening [64].
	Rutting	The longitudinal surface depression, which may lead to structural failures and the occurrence of hydroplaning [64].
	Shoving	The longitudinal displacement on the pavement [65].
	Patch	The distress that will occur when the original pavement from a certain area has been removed and replaced by either similar or different materials [64].

Pothole	Holes on the pavement surface due to the alligator cracking, localized disintegration, or freeze-thaw cycles [64].
Delamination	Delamination occurs due to the debonding or stripping of the asphalt layer, resulting in the tearing in the surface [66].
Polished aggregate	The exposed coarse aggregate due to the peeling of the surface binder of the pavement [65].
Water pumping	Seeping or ejection of water from beneath the pavement through cracks [65].
Bleeding	Excess asphalt binder occurring on the pavement surface, especially in the wheel paths [65].

3.1.4. An Overview of ML and Corresponding Applications in Asphalt Construction

The relationship between the process and product quality of the asphalt at the lifecycle scale is highly non-linear and too complex to capture and express using conventional statistical methods. To cope with this non-linearity, machine learning (ML) methods can be applied, as they have proven effective in solving similar problems in other domains [67–72]. ML can be roughly explained as an intelligent system that can learn and improve its performance based on historical data to make inferences [11]. As shown in Figure 10, ML mimics the human learning process, where the learning process happens based on the past or experiences through induction, summarization, and internalization. Therefore, when new situations take place in the future, humans can deal with them by utilizing their knowledge gained from the past through the learning process. As for ML, the learning process will mainly rely on the input data and inner algorithms. Based on the problem domains, typical ML problems can be divided into regression problems, classification problems, and clustering. Besides, based on the differences in the given datasets, ML can also be divided into supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, etc. [73].

Ray [74] provided a series of most widely used ML algorithms, which can be briefly classified as gradient descent algorithms, linear regression algorithms, multivariate regression analysis, logistic regression, decision tree, support vector machine, Bayesian learning, Naïve Bayes, K-nearest neighbour, K-means clustering, and back-propagation algorithms. In addition, as the problems that traditional ML algorithms try to solve become more complex, the concept of deep learning was widely used to build "deeper" layers of abstractions from the rather simple and "shallow" ML architectures, thus bringing the concepts of deep learning [75]. Several wide-applied deep learning techniques include Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) [75–78].

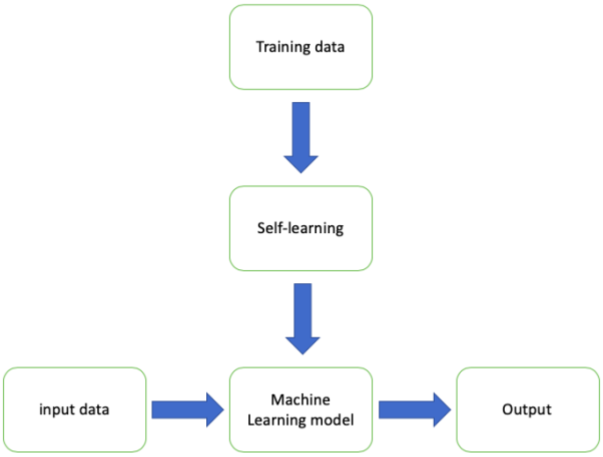


Figure 10. The general framework of the ML approach

In recent decades, the road construction sector has also made several attempts to conform to the trend of the application of ML techniques. One of the most fruitful application domains is developing ML-based predictive models to predict certain pavement condition indicators [14,15,79–81]. In these studies, the developed ML models provided a promising regression of the deterioration of the pavement performance. Furthermore, these studies have given the initiative of applying ML techniques in facilitating the preventive maintenance of road infrastructure. Due to the satisfying predictive performance of the ML models, these models can potentially steer the decision-making process for preventive road pavement maintenance. However, as briefly mentioned in the introduction, most of the aforementioned studies only investigated the correlation between pavement performance and parameters in the operation phase, i.e., traffic intensity, climate condition, road age, and historical inspections. However, the quality of the pavement condition in the operation phase will also be affected by the quality of the previous phases. Neglecting the importance of the construction process quality in the regression modeling will, consequently, not only hinder the model generality and performance but also the possibility of improving the current quality assurance and control practices.

When it comes to the pavement material properties and performance tests, ML techniques were also widely applied in studies for predicting mechanical performance indicators, including air void content [82], viscoelastic behaviour [83], dynamic modulus [84–87], rutting depth [88,89], and indirect tensile strength (ITS) [89]. These applications can provide the possibility for obtaining efficient and reliable predictions of asphalt mechanical properties. This is essential because, in the current quality assurance schemes, tests regarding the properties of pavement normally will be conducted only on the asphalt cores drilled from the paved layers, which cannot cover the entire paved surface and will cause variability in the measurements. Also, these studies lacked a systematic definition and evaluation of the construction process quality. Consequently, it is still unknown to what extent the variability generated from the adopted on-site strategies affects the mechanical properties of the asphalt pavement.

3.2. Stakeholder Analysis

In order to have a better understanding of the social context of the system under design, the social embedding of the proposed framework was thoroughly explored.

Following the description in Chapter 2.1.2, the stakeholder analysis started with the identification of stakeholders through internal brainstorming sessions participated by the EngD candidate and the supervisory team. In this study, stakeholders are defined as any individual or organization "having a vested interest in the decision process and either directly affecting or being affected by its resolution", according to [90].

Based on the outputs from the organized brainstorming sessions, five stakeholders were identified, as shown in Table *2*. Among them, site managers and quality inspectors were persons associated with the contractors, focusing on the on-site operations regarding various construction activities and being involved in quality control. The asset managers, who are mainly responsible for the maintenance planning and pavement condition monitoring during the operational phase of the pavements, can be associated with both contractors and public clients. In addition, the ASPARi network has also been determined as the stakeholder, because as the initiator of this design project, ASPARi was tightly involved in the design process of the project by providing the necessary support in terms of supervising, knowledge, and data. Lastly, the EngD candidate was identified as the final stakeholder because the EngD candidate is directly responsible for the design project and is expected to accomplish the design objective within a limited time. The connections between the identified stakeholders with the method under design have also been explored, as summarized in Table *2*.

A workshop [91] was then organized, with participants including the EngD candidate, the supervisory team, and representatives of the identified stakeholders. During the workshop, multiple needs of the stakeholders were derived, as indicated in Table *2*.

Table 2. The identification of stakeholders and their needs

Stakeholder	Connections with the design	Needs
Site managers	• Providing construction data of road construction projects. • Incorporating the design outcome into the quality control scheme to enhance decision-making regarding operational strategies.	• To improve efficiency and effectiveness in making on-site operational strategies. • The ease of use and interpretability of the data-driven models. • Quantifying the impact of process quality on product quality.
Quality inspectors	• Providing quality inspection data of road construction projects. • Using the project outcome to improve the quality inspection strategy.	• An accurate non-destructive method for measuring pavement asphalt mechanical properties. • Being able to trace back to the construction process quality when the undesired quality of certain areas of the pavement is inspected. • Based on the assessment of the process quality, the pavement quality can be obtained which can cover the entire pavement.
Asset managers	• Providing pavement condition inspection data. • Using the project outcome to improve the maintenance planning and monitoring strategy.	• To reduce the lifecycle cost of asphalt construction regarding maintenance. • To provide more efficient and effective maintenance strategies. • The ease of use and interpretability of the data-driven models.
ASPARi network	• Providing data on historical PQi projects. • Providing supervision. • Providing opportunities to contact practitioners in the road construction industry.	• Inter- and intra-organizational integration of data, processes, materials, and organizations through the entire road construction lifecycle.

		• To obtain an explicit understanding of the process and product quality of asphalt pavement.
EngD candidate	• Developer of the design project.	• Accomplish the design objective.

4. System Design

4.1. Requirements Engineering

The identified stakeholders' needs from the previous chapter were translated into various functions that the project outcome is expected to provide. The identified functions were illustrated in the form of the function tree in Figure 11.

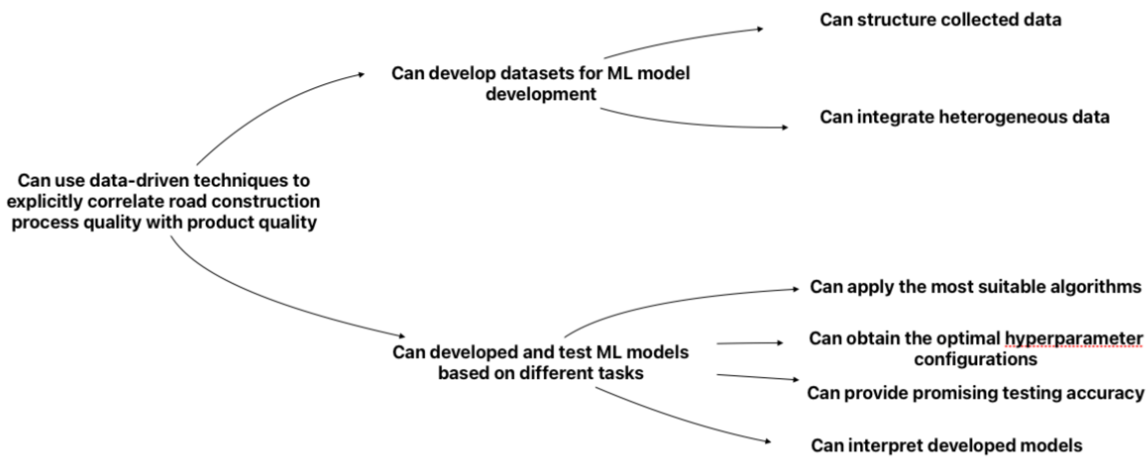


Figure 11. The function tree of the data-driven method under design

Then, each function was allocated with one or more conditions to provide further clarification on the degree to which the function adequately meets the expectations of respective stakeholders. Subsequently, several functional requirements were identified as represented in **Error! Not a valid bookmark self-reference.**.

Table 3. The identified requirements

Requirements
1. The system must explicitly include variables that can assess the quality of the construction on-site operations.
2. The prediction must take into account the parameters in the operation phase as input features.
3. The development of the ML models must use the construction process data collected following PQi methodology.

4. The dataset development must clearly represent the data integration process.
5. The ML model development must include optimization.
6. The derived ML model must accurately predict pavement product quality indicators with R^2 no less than 0.8.
7. The derived ML model must overcome the overfitting problem where the difference between training R^2 and testing R^2 should be less than 0.15.
8. The system must be interpretable by indicating which features are most important in driving the models' predictions.
9. The obtained results from model interpretation must be compared with findings from previous studies.

Among the identified requirements, requirements 1 to 4 are concerned with the dataset development that is needed for the ML model. In addition, requirement 5 is concerned with finding the optimal configuration in the ML model development process to ensure the best regression performance. Furthermore, requirements 6 and 7 specify the requirements for the predictive performance of obtained ML models. These requirements can serve as metrics to directly evaluate whether or not the developed ML models can address the issue of finding the correlation between process and product quality indicators. Last but not least, requirements 8 and 9 mainly focus on the ML model interpretability, thus allowing the researcher to generate insights to answer the formulated design project question.

4.2. The ML model development

The identified requirements from the previous chapter provided a basis for aligning the design with stakeholders' needs, and guiding design decisions by serving as the guideline. In Chapter 5.2, the ML model development was elaborated, according to the process described from Chapter 2.2.2 to Chapter 2.2.4.

4.2.1. Dataset Development

The ML model development was initiated by constructing the datasets, following Chapter 2.2.2, including the formulation of data structure, data collection and integration, and data preparation.

4.2.1.1. Data Structure

The dataset development started with the formulation of the data structure regarding considered input and output variables. The theoretical framework provided the information needed for identifying these variables, including relevant indicators for road construction process and product quality. To better assess these identified indicators considering the context and scope of the project, a workshop [91] was used to collaboratively explore potential variables with the key stakeholders.

It has been determined in the workshop that the consideration of output variables, concerned with product quality, should cover indicators about both the mechanical properties and long-term performance of the pavements. Among all the identified indicators regarding the mechanical properties of the pavements, density is widely perceived as the most representative and prominent parameter to determine the result verification [92–94]. Additionally, when the asphalt pavement reaches the desired density in the appropriate compaction temperature window, the optimal states for other mechanical properties, such as stiffness, fatigue characteristics, resistance to permanent deformation, and moisture resistance, can also be obtained [93]. Therefore, it is of great significance to consider density as the major target for the development of ML models, given its importance in representing the pavement quality and tight connection to the construction process.

For the long-term performance of the asphalt pavement, Table *1* shows a series of indicators, which are mainly represented by the evaluation of the severity of the distress modes. There are several methods to represent the severity of the distress modes. The most direct way is by using geometric definitions, such as the length, width, depth, and size of the impacted area. However, in the practices of the ML application, considering all the geometric definitions of certain distress as model outputs may significantly increase the storage cost and modeling difficulty, due to the increased attribute dimensions. Besides, some data regarding these physical representations of the failure modes are highly confidential, which increases the difficulty in collecting the necessary data. Therefore, it is a common practice in the literature to apply indicators such as Pavement Condition Index (PCI) and Pavement Quality or Performance

Index (PQI) to combine the assessment of multiple failure types into one evaluation formula [14,15,95].

In the Netherlands, pavement conditions of the majority of the road network are inspected, monitored, and managed by the Rijkswaterstaat (RWS), using the model and corresponding database contained in the IVON (Informatiesysteem Verhardings Onderhoud) program. Developed by RWS, IVON has been used as the benchmark for Dutch highway operations to enhance maintenance decision-making the during past decades. Specifically, the IVON program conducts independent inspections of its concerning failure types, including raveling, cracking, rutting, roughness, bearing capacity, and so on, and evaluates the severity of the pavement performance degradation to each failure type to support the decision-making regarding the year of intervention. Lastly, the condition of the road is comprehensively assessed based on the severity of the various modes of failure and the corresponding year of intervention, which will be demonstrated using the overall intervention year. When the shortest expected residual life (directly calculated from the nearest intervention years) of a certain failure type is longer than 5 years, then according to the average lifespan of the mixture characteristics, bearing soils capacities, and environmental and traffic impacts, the lifespan evaluation and intervention plan will be calibrated by an empirical model named IVONLANG.

Another critical indicator that demonstrates the overall pavement performance is the international roughness index (IRI), which quantitatively defines the pavement roughness by using the ratio between accumulated vehicle suspension motion and the traveled distance within the same time window [96]. This indicator can reflect the roughness of the pavement, which is essential to the driving comfort and safety of road users thus also representing the serviceability of the pavement. On these premises, this study also extracted the corresponding IRI data through RWS's inspections as a supplement to the investigation between the process quality and the long-term performance of the pavements in terms of roughness

On these premises, this design project aims to investigate the correlations between the process quality indicators and the pavement density, intervention years obtained from the IVON model, and International Roughness Index (IRI) as regression outputs. Utilizing the theoretical framework established through the identification and analysis of product quality indicators, these variables have been deemed highly indicative of pavement quality from both short- and long-term perspectives and will serve as the focus for the development of ML models.

Furthermore, in an effort to optimize the suitability of the selected outputs for integration into the ML modelling, several modifications will be implemented. With regards to the density, the density degree, i.e., the ratio between the measured density and the target density for the specific asphalt mixture, will be used as an alternative for the direct measurement of density. This is because using density degree can avoid issues such as the imbalance in the data distribution arising from the fact that different asphalt mixtures possess distinct target densities. To elaborate, when processing the density data, the existence of different clusters in the output data can greatly impact the performance evaluation in regression tasks. For instance, when R^2 is used as the metric, the clusters in the output data distribution can result in a huge total sum of squares, therefore, the calculated R^2 can always be high, while the actual ML model could be overfitting. Conversely, the density degree, being the ratio of the measured density to the target density, normalizes the data and eliminates the issue of imbalanced distribution.

In addition to this, the output concerning the intervention years will also be represented as the remained lifespan of the pavement. This is because the intervention years are represented by calendar years in the IVON program, however, it may not be intuitive to interpret the deterioration of pavement performance if the construction years are unshown.

The rest of this section will illustrate the development of the datasets according to the regression of each three of the above-mentioned outputs.

Regression of Density Degree

The determination of input features in the data structure applied to the regression of density degree will be based on the theoretical framework. As demonstrated in the theoretical framework, the evaluation of the quality of the construction operation can be represented in both the macro and micro perspectives. From the macro perspective, the road pavement will be divided into several sections with a certain interval regarding the length. In each section, ECR will be used as the indicator for the overall quality of the construction operation. As for the micro perspective, the evaluation of the quality of the construction operation will be scaled down to the centimetre level, such as the PQi methodology.

Because for the regression of density degree, the output data of the dataset should be provided from the core testing results, which is also at a centimetre level. Therefore, instead of directly using ECR, the evaluation can be performed and represented by the deviation between the target and the actual number of roller passes, and the proportion of the received roller passes that fall within the compaction temperature window.

When it comes to the ambient conditions measured during the construction process, the corresponding data were recorded using the weather stations during the PQi practices, including the temperature, wind speed, pressure, humidity, and precipitation. However, because these weather conditions were registered in a time-series manner, it is necessary to utilize the statistics of the weather conditions of each project to represent the characteristics and trends, while reducing the dimensionality of the dataset. In this design project, for each type of weather condition, the mean and standard deviation will be calculated from each project and used as indicators.

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 design project 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.

Finally, Table *4* below represents the identified input-output structure for the regression of density degree.

Table 4. The summary of the identified data structure for the regression of density degree

	Variable	Description
Input	Roller pass deviation	The deviation of the received roller passes from the target roller passes.
	Percentage of the roller passes within temperature window	The percentage of the number of the received roller passes that are within the compaction temperature window.
	Mixture type	The type of the asphalt mixture.
	Temperature - mean	The mean value of the temperature during the construction process.
	Temperature - standard deviation	The standard deviation value of the temperature during the construction process.

	Wind speed – mean	The mean value of the wind speed during the construction process.
	Wind speed – standard deviation	The standard deviation value of the wind speed during the construction process.
	Humidity – mean	The mean value of the humidity during the construction process.
	Humidity – standard deviation	The standard deviation value of the humidity during the construction process.
	Pressure – mean	The mean value of the pressure during the construction process.
	Pressure – standard deviation	The standard deviation value of the pressure during the construction process.
	Precipitation – mean	The mean value of the precipitation during the construction process.
	Precipitation – standard deviation	The standard deviation value of the precipitation during the construction process.
Output	Density degree	The ratio between the measured density of the asphalt pavement and the target density of the used asphalt mixture.

Regression of Residual Lifespan and IRI

For the regression of residual lifespan and IRI, the input variables for the dataset can directly use those identified for the regression of density degree. However, the inspections and evaluations regarding pavement performance are normally conducted at the hectometer level, meaning the data will be registered based on the inspection and analysis every 100 meters on the corresponding highways. Therefore, the calculated ECR will be used instead of roller pass deviations and percentages of the roller passes within temperature windows.

Additionally, the impact of the operation of the pavements during its service, such as the traffic intensity and weather, can also significantly influence the condition and performance degradation of the pavements [81,95]. For the traffic loads, in the work conducted by [95], the EngD candidate considered the average daily traffic and average daily truck traffic to reflect the 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 the climate conditions during the operational phase of the pavements, as pointed out by [97], temperature variation and moisture change can lead to a great impact on the material properties of the pavement structure, while studies conducted by [98] also indicated that the freeze-thaw cycles during the pavement operation are also a considerable cause to the increase of the performance deterioration process. Therefore, this design project followed the study by [95], to extract climate data including average annual temperature, average annual precipitation, and the number of freeze-thaw cycles. It is worth noting that this design project utilized the definition of [95] to count the number of times the temperature drops from freezing to thawed states as the number of freeze-thaw cycles.

As for the labels in the dataset, although the intervention years and IRI are often evaluated and calculated during regular inspections, they are time-variant metrics. To account for the complexity of changes in road usage, this study considered a rolling time window of one year instead of a long-term average of road use. This approach allows for a more accurate evaluation of the road's condition by considering the most recent measurement and the amount of use the road has seen since then. Table *5* below summarizes the identified input-output structures for both regression outputs.

Table 5. The summary of the identified data structure for the regression of residual lifespan

	Variable	Description
Input	ECR	The deviation of the received roller passes from the target roller passes.
	Mixture type	The type of the asphalt mixture.
	Temperature - mean	The mean value of the temperature during the construction process.
	Temperature - standard deviation	The standard deviation value of the temperature during the construction process.
	Wind speed – mean	The mean value of the wind speed during the construction process.
	Wind speed – standard deviation	The standard deviation value of the wind speed during the construction process.

	Humidity – mean	The mean value of the humidity during the construction process.
	Humidity – standard deviation	The standard deviation value of the humidity during the construction process.
	Pressure – mean	The mean value of the pressure during the construction process.
	Pressure – standard deviation	The standard deviation value of the pressure during the construction process.
	Precipitation – mean	The mean value of the precipitation during the construction process.
	Precipitation – standard deviation	The standard deviation value of the precipitation during the construction process.
	Heavy truck intensity per workday	The mean intensity of the heavy trucks of a certain road section on the workday.
	Medium truck intensity per workday	The mean intensity of the medium trucks of a certain road section on the workday.
	Passenger car intensity per workday	The mean intensity of the passenger vehicles of a certain road section on the workday.
	Annual mean temperature	The mean value of the annual temperature of the road section.
	Annual mean precipitation	The mean value of the annual precipitation of the road section.
	Annual freeze-thaw cycle	The mean value of the annual number of freeze-thaw cycles of the road section.
	Age	The age of the pavement compared to the construction year.
	Residual life-1	The residual lifespan from the previous year.
	IRI-1	The IRI from the previous year.
Output	Residual lifespan	The remaining years until the intervention will be performed, which can be calculated directly from the intervention years.

IRI	The International Roughness Index, which quantitively reflects the roughness of the pavement.

4.2.1.2. Data Collection, Data Integration, and Data Preparation

Data Collection

According to Chapter 4.2.1.1, several databases were identified as the sources for enabling the data collection, including the database of ASPARi network archiving data from historical PQi measurements, the IVON database of RWS containing inspection data regarding pavement conditions, and several public databases providing data regarding traffic intensities and climate. These databases were represented in this chapter, in terms of general data classes, attributes, relations, and formats. Specifically, the relations were illustrated using the concept of private and foreign keys, where the private keys provide unique indexes to a certain data class, while foreign keys provide links to private keys of other data classes [99]. As for the data formats, which is mainly concerned with the types of data, Table 6 below defines the identified formats and descriptions from these databases.

Table 6. Data types and corresponding descriptions

Data types	Description
int	An integer.
double	A normal-size floating point number with specified total number of digits.
varchar	A string containing letters, numbers, and special characters, with varied length.
date	A date in the format of YYYY-MM-DD (e.g., 2023-03-14).

Among these databases, the ASPARi database stores and manages data collected through PQi measurements that are mainly concerned with process quality indicators and other data associated with project characteristics. A relational database has been developed to store and process the data collected following the PQi methodology to support the operator guidance system [49]. The information covered in this relational database can be categorized into three types, namely project information, construction process characteristics, and analytical results. Among these types, the project information includes data such as the contractor's name, the client's name, the site location, the mixture type, and the construction date. The construction process characteristics are mainly concerned with the measurements using the PQi methodology, including the asphalt temperature during the paving and compaction processes, the compaction passes, and the ambient conditions. Lastly, the analytical results encompass real-time feedback provided to the operators, such as the current state of machinery in terms of location and speed, as well as subsequent analytical results, such as the ECR. This database can provide direct data on ECR, roller pass deviation, percentage of roller passes within the temperature window, and various ambient weather conditions, to establish the dataset.

However, the database developed for the purpose of an operator guidance system contains data that is outside the scope of this project. Therefore, a simplified database was extracted from the original database developed in [49] according to the input-output structure identified in the previous chapter, as shown in Figure 12. This simplified database primarily focused on the process quality indicators on the scales of cells and sections respectively. Besides, general project information and ambient conditions were also included.

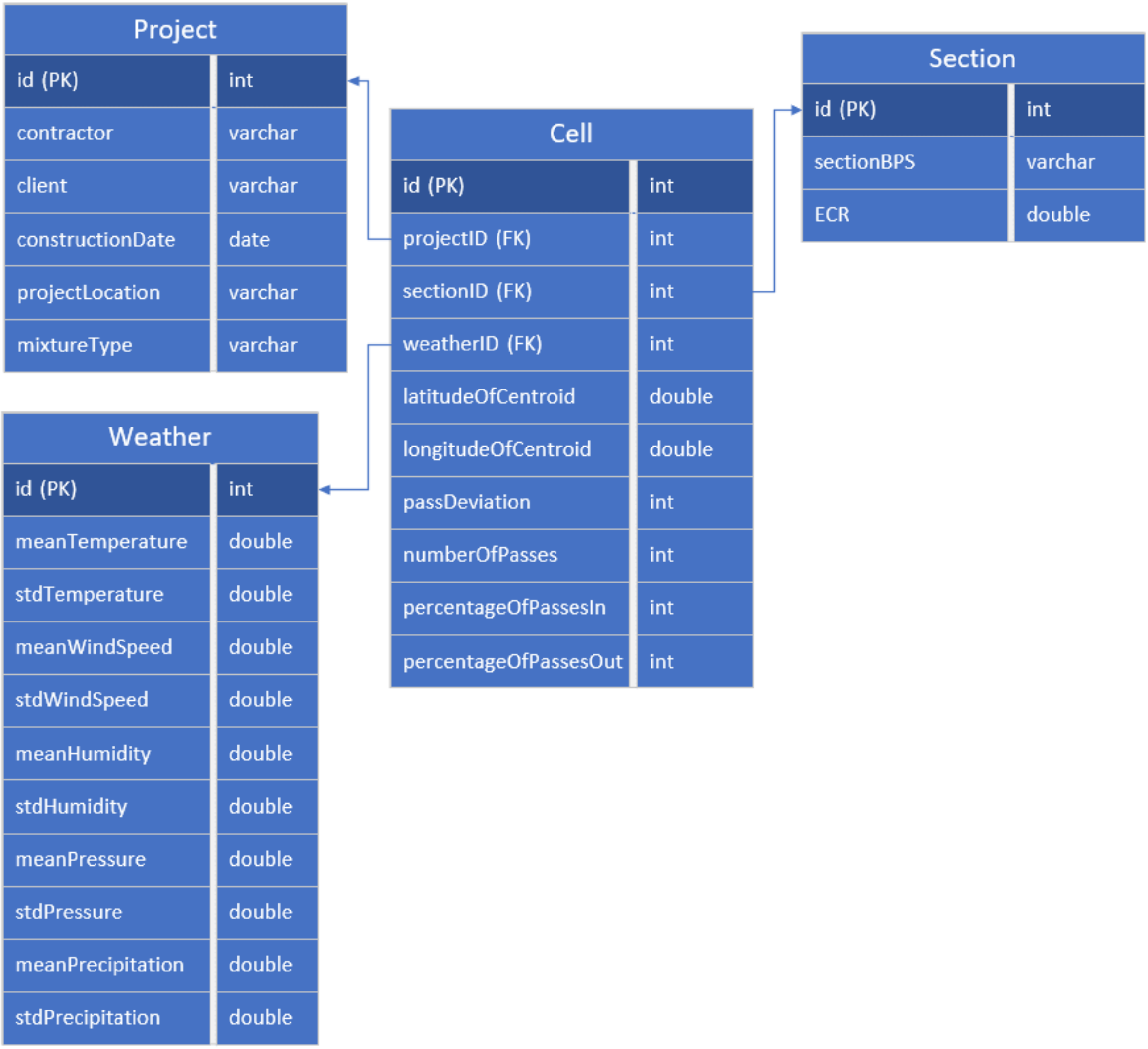


Figure 12. The simplified relational database for data extracted from the PQi measurements

When it comes to the data for product quality indicators, the data collection of pavement density is normally conducted by contractors using either in-place measurements or laboratory tests on the asphalt cores drilled from the pavements. The in-place measurements are typically non-destructive because they do not require the direct extraction of the pavement samples, while the latter is widely considered as the most accurate method for providing the properties of the asphalt pavement [100]. Therefore, in this study, the data concerned with the density measurement was confined to the test results obtained from the laboratory on the core samples. Figure 13 represents the corresponding data structure for this type of data that can be acquired from the contractors, where the attributes "*coreLongitude*" and "*coreLatitude*" refer to the coordinates of the centroid of the extracted cores based on the WGS 84 system.

CoreDensityMeasurements	
id (PK)	int
contractor	varchar
projectLocation	varchar
compactionDegree	double
constructionDate	date
coreLongitude	double
coreLatitude	double

Figure 13. The data structure of the core density measurements

Data related to the pavement performance were retrieved from the IVON database and yearly inspectional records for IRI held by RWS. Figure 14 provides an overview of the data structure of the information that can be requested from RWS.

IVONDatabase	
id (PK)	int
sectionBPS	varchar
constructionDate	int
inspectionYear	int
IRI	double
IVONInterventionYear	int

Figure 14. The data structure of the pavement condition data

It is worth noting that the coordination system adopted in this database does not follow the WGS 84 system. Instead, a descriptive coordination system is applied, named Beschrijvende Plaatsaanduiding Systematiek (BPS) [101]. The BPS system determines the exact location on the national highway network, based on properties including:

- Road type (e.g., national roads, provincial roads, municipal roads, etc.).
- Road number (i.e., the number that uniquely identifies a road of a particular type.).
- The relative distance to the nearest hectometer board.
- The indication on the nearest hectometer board (i.e., the longitudinal distance of the location of the hectometer board to the starting point of the road).

- The type of the traffic-carrying track (*baan* in Dutch) (e.g., main carriageway, connecting road, roundabout, etc.).
- The transverse position of the concerned traffic-carrying track in relation to the road orientation line (e.g., left, right, middle, etc.).
- The type of traffic lane (*strook* in Dutch) (e.g., normal traffic lanes, exit lane, emergency lane, etc.)
- The indication of the position in the transverse direction of a traffic lane in relation to other lanes of the same type on the same traffic-carrying track with respect to the road orientation line (e.g., left, right, middle, etc.).

For the data needed for information during the operational phase of the pavements, two databases were used to extract the required data. For the traffic intensity, data were extracted from the database provided by RWS named INtensiteit op WEgVAkken (INWEVA), as represented in Figure 15. This database covers the entire Dutch highway network and registers the historical traffic intensity of each hectometer section with specified BPS locations. Besides, in this design project, the source for providing the aforementioned weather data is the dataset derived from the Koninklijk Nederlands Meteorologisch Instituut (KNMI), as shown in Figure 16. Therefore, based on the closest weather station of KNMI to the target pavement section, the corresponding weather data can be obtained.

trafficLoading	
id (PK)	int
sectionBPS	varchar
measurementYear	int
heavyTruckPerday	double
mediumTruckPerworkDay	double
passengerCarPerWorkDay	double

Figure 15. The data structure of the traffic intensity data during the operational phase

operationalWeather	
id (PK)	int
sectionBPS	varchar
measurementYear	int
annualMeanTemperature	double
annualMeanPrecipitation	double
annualFreezeThrawCycle	int

Figure 16. The data structure of the weather data during the operational phase

The accessibility of these identified data sources was evaluated. Table 7 below represents the result. Owned by the ASPARi network, the PQi database was assigned with high accessibility. Although permission is needed, because the EngD candidate is affiliated with the ASPARi network, the permission was automatically granted. Besides, because of the relational database as represented in [49], the data extraction could be done by directly executing SQL (Structured Query Language) queries to streamline the collection process. For the data regarding the asphalt density measurements, which are collected and managed by the contractors, the accessibility is low. This is because the contractors heavily rely on ad-hoc data measurements and registration in the pavement quality inspection practices. This resulted in the measured data being managed in a disorganized structure and stored in non-uniform formats (varying from CSV files to paper documentation). Consequently, it led to a lengthy process to extract the needed data from this source. Pavement condition data owned by RWS has medium accessibility, because although permission is needed, the IVON database can ease the corresponding data collection process. Lastly, both traffic intensity data and weather data during pavement operation were allocated with high accessibility, primarily because they are open to public, thus being able to be directly extracted from the corresponding websites.

Table 7. The result of the accessibility evaluation of identified data sources

Data sources	**Data owner**	**Data format**	**Authorization**	**Accessibility**
PQi database	ASPARi network	Database tables	Permission needed	High
Density measurements	Contractors	CSV, paper documentation	Permission needed	Low
Pavement conditions	RWS	Database tables	Permission needed	Medium
Traffic intensity	RWS	CSV	Open data	High

Weather data during pavement operation	KNMI	CSV	Open data	High

The data collection strategy was made accordingly. Firstly, the historical road construction projects were identified, where detailed PQi data must be available. Corresponding data, as described by the structure represented in Figure 12. The simplified relational database for data extracted from the PQi measurements, were retrieved. Then, meetings were arranged between the EngD candidate and data owners from the contractors, to jointly locate the same projects as identified in the previous step and extracted data from all the available documents. Simultaneously, data collection requests were sent to RWS data owners with specified BPS locations of the projects identified in the previous step. Then, the corresponding queries can be made by specialists from RWS to directly provide the EngD candidate with the data described by the structure demonstrated in Figure 14. Lastly, based on these specified BPS locations, data regarding the traffic intensity and weather were directly retrieved from the corresponding websites.

A data warehouse was designed according to the identified input-output structures from the previous chapter and the determined data sources, as shown in Figure 17. This data warehouse functioned as the repository for storing the collected data, which also represents the multidimensional relationships between data to enable the selection and grouping of data [102].

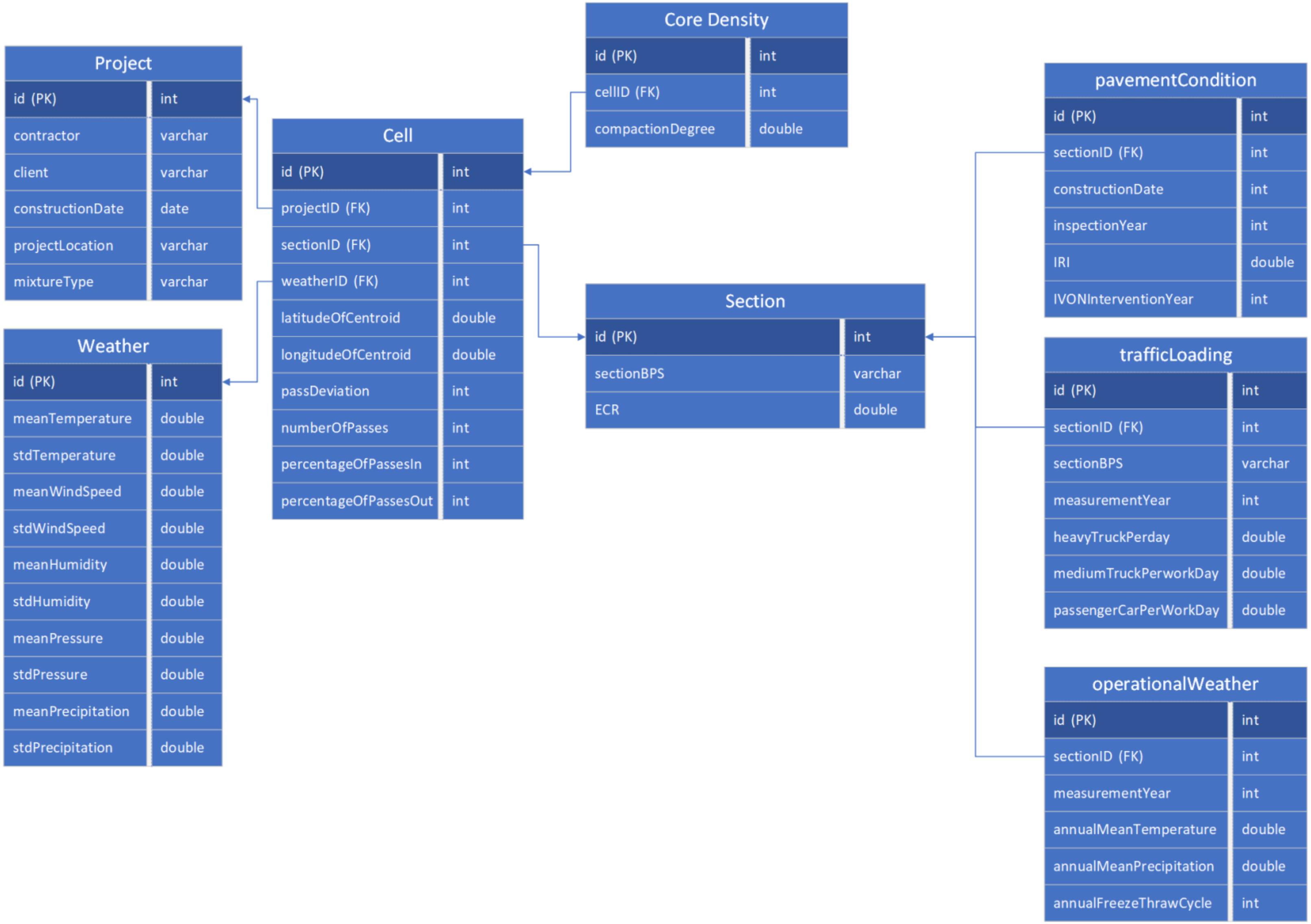


Figure 17. The star schema of the data warehouse for structuring the collected data

Data Integration

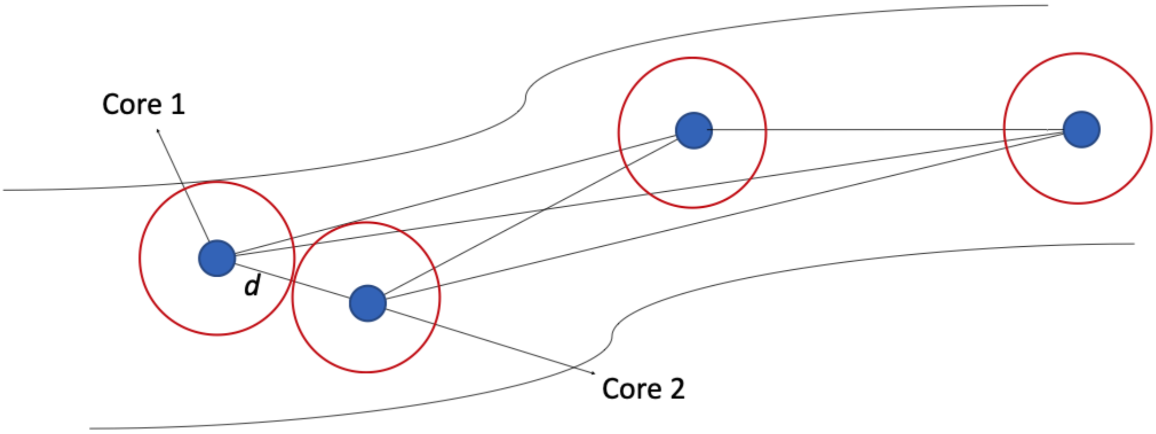
The identified data sources indicate the original form of the data, while the data warehouse describes the eventual structure after integrating the data from multiple sources. In order to seamlessly consolidate and combine the data extracted from heterogeneous sources, it is essential to determine the scheme of data integration.

According to Figure 17, matching fields between tables were found. Specifically, between the "*Cell*" table and tables of "*IVONDatabase*", "*trafficLoading*", and "*operationalWeather*", the matching field was identified as the *sectionBPS*. This is because, for all the cells registered in the data warehouse, corresponding BPS coordinates can be found. These BPS coordinates of different cells can then be aligned with the same BPS coordinates registered in the tables of "*IVONDatabase*", "*trafficLoading*", and "*operationalWeather*".

When aligning the "*Cell*" table and "*CoreDensityMeasurements*" table, the integration conflict was found. Several matching fields between these two tables were identified, including the contractor's name, the data when the project took place, the general location of the project, and coordinates of centroid of the cells and drilled cores. However, coordinates of centroid of the cells and drilled cores are unlikely to overlap between their respective locations. Therefore, it can only be assumed that the process quality of a cell (i.e., in a rather local area on the pavement) can represent the process quality of the core, when the distance between them is distinguishably close.

Therefore, it is essential to find a distance range for each core to include a certain number of cells within the range, which are not only sufficiently representative to reflect the process quality of the core but also adequately unique to avoid the situation where multiple cores share the information of the same cells. The latter is crucial in the application of data-driven techniques to maintain the singularity of the input-output mapping, i.e., one input will only lead to one output.

Figure 18 provides the algorithm adopted in this design project to align the coordinates of the cores with the cells. Initially, the distance between every pair of cores was calculated iteratively, and the smallest distance was then utilized as the diameter of a circular-shaped area to identify the cells that will be covered, as the representation of the process quality of the target cores. However, in case all the cores are generally far from each other, it is necessary to introduce an upper limit of the radius of the circle in Figure 18.



Algorithm: Aligning cores with cells in the same project

Input: Total number of cores N, an array of cores $P_{core} = \{P_1, \dots, P_N\}$,
total number of cells M , an array of cells $p_{cell} = \{p_1, \dots, p_M\}$

Set $smallestDistance = +\infty$
For i from 1 to $N - 1$
  For j from $i + 1$ to N
    Calculate $distance$ between P_i and P_j
    If $distance < smallestDistance$
      $smallestDistance = distance$
    Else
      pass
    End
  End
End
Set $r = {}^{smallestDistance}/_{2}$
Set array $A = Null$
For k from 1 to M
  For q from 1 to N
    Calculate $distance$ between p_k and P_q
    If $distance \leq r$
      A.*append*($[p_k , P_q]$)
    Else
      pass
    End
  End
End
return A

Output: An array with pairs of aligned core and cell

Figure 18. The approach to align the locations of cores with cells

Data Preparation

Lastly, data preparation was conducted. According to different regression tasks, different datasets were derived from the data warehouse following the data structure indicated in Chapter 4.2.1.1.

Then correlations between variables were determined. Specifically, the Pearson correlations coefficient between every two input variables was calculated, using the following equation [103]:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{2}$$

where r refers to the Pearson correlations coefficient, x_i and y_i refer to the values of variables x and y in a sample respectively, and $\bar{x}$ and $\bar{y}$ refer to the mean values of variables x and y in a sample respectively.

The value of r indicates to which extent two variables are linearly correlated. The value of r ranges from -1 to 1. The closer the calculated r to 1, the higher the positive correlation can be determined between the two variables, while when two variables are correlated inversely, the calculated r will be near -1. Therefore, based on the calculated Pearson correlation coefficient, variables that are highly correlated were defined, and only one of them was kept as the input feature for the ML model development to reduce the dimensionality of the dataset to prevent both overfitting (when the dataset is too complex and contains too much noise) and underfitting problems (when the dataset is too small to capture the patterns within the high dimensionality).

Next, the datasets were cleaned by examining the outliers. In this study, the method named interquartile range method (IRM) was applied, where the outliers were defined and removed when the quartile range is less than 1.5 times the upper quantile or the quartile range is 1.5 times larger than the lower quantile.

After the dataset was cleaned, the categorical values within the dataset were transformed into numerical values. This is because of the inherent limitations of various ML algorithms, that meaningful information cannot be extracted from learning data with categorical values. Two methods were used for converting categorical values, namely target coding, and label encoding. Unlike other encoding methods, such as dummy coding or one-hot encoding which will increase the dimensionality of the data structure significantly, target encoding reshapes the categorical features by encoding them based on their effects on the target outputs [104]. However, when the categorical variables are ordinal or the number of categories in one variable is too large or too small, the target encoding might be ineffective. Therefore, the label encoding was used to simply use one integer to replace one certain category [105].

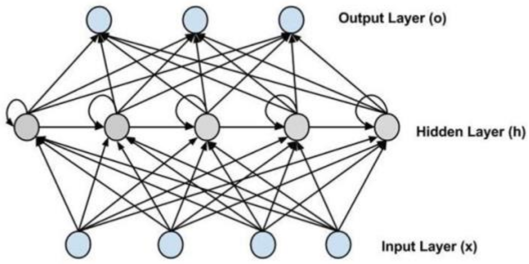
4.2.2. ML Model Training and Testing

4.2.2.1. ML algorithm selection

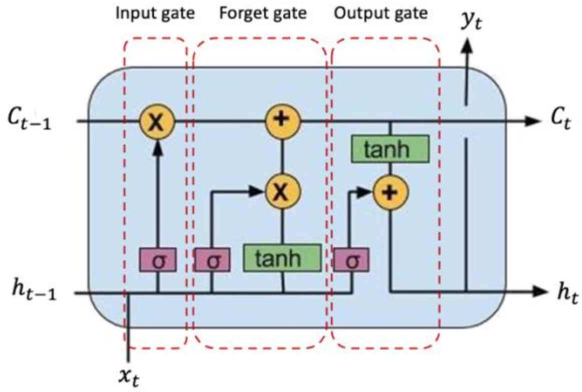
Based on the problem context of the research, in this study, two types of the ML algorithm were selected, namely random forest (RF) and recurrent neural network (RNN). Brief introductions of the two algorithms were given below.

As a widely applied tree-based ML algorithm, RF can solve both the regression and classification problems on the basis of ensembled decision trees, to advance the performance. More specifically, in a typical structure of RF, the training set will be divided into several replicates using the bootstrapping methods and trained by randomized decision trees, and lastly, the final classification or regression can be made by voting to the best predictors [106]. 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 as the conventional classification and regression trees (CART) have demonstrated.

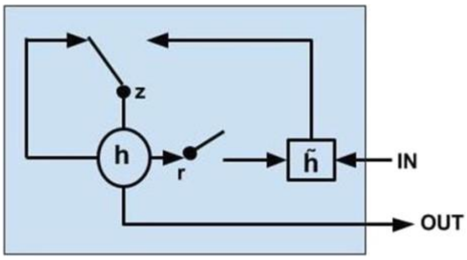
Compared to RF, RNN is dedicated to tackling the complexities of non-linear regression concerned with time-series data. RNN can describe the dynamic behaviour of time-series data, by circulating states in the networks. However, the conventional architecture of RNN soon showed its limits, because of the problems such as the gradient's vanishing and explosion, and the difficulty in learning long-term patterns. In order to tackle the aforementioned challenges, long short-term memory (LSTM) and gated recurrent unit (GRU) were developed and introduced as extensions of conventional RNN [107]. Figure 19 gives a representation of the three types of RNN architectures.



(a). Conventional RNN architecture



(b). LSTM memory cell



(c). GRU cell

Figure 19. RNN architectures [108]

Figure 19. (b) illustrates the general architecture of LSTM, where ⊗ refers to the multiplication operator, ⊕ refers to the sum operator, σ refers to the sigmoid activation function, $tanh$ refers to the hyperbolic tangent activation function, x_t refers to the current input data, h_{t-1} refers to previous hidden layer data, C_{t-1} refers to previous memory cell unit, and y_t refers to the output.

As demonstrated in this figure, LSTM constructs different gates as the replacement for the units in the hidden layer of the conventional RNN architecture, shown in Figure 19. (a). Three gates, namely an input gate, an output gate, and a forget gate, are included in a memory cell of the LSTM, and process the current input data x_t and previous hidden layer data h_{t-1} [109]. In all the gates, a sigmoid activation function is included and decides whether the gate will be open (when the output of σ is 1) or closed (when the output of σ is 0). The input gate additionally contains a multiplication operator, which controls the flow of inputs to the rest of the network [110]. The forget gate allows the transmission of output information from the previous neuron to the next neuron, based on the weights of the output information. Finally, the output gate determines which information from the cell state will be transformed into the current hidden layer data h_t.

When it comes to GRU, Figure 19. (c) indicates the basic architecture. On the basis of LSTM, GRU combined the input and forget gates from the original LSTM architecture into an update gate (denoted as z), which determines which memory will be kept in the cell. Besides, compared to LSTM, GRU directly utilizes a reset gate (denoted as r) to process previous hidden layer data. The reset gate can then decide whether the current state will be integrated with previous information [108].

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. Besides, research has found that GRU has comparable to or even surpassed the performance of LSTM [111]. Therefore, in this design project, GRU will be used.

However, when applying GRU, 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 reshaped into vectors using affine transformation as the internal state of the GRU architecture. This transformed initial state was then added to the hidden state of the GRU when calculating the output [112–114]. In addition, to further tackle the complexities and non-linearities of the problems, a hybrid network was applied by adding more dense layers behind the GRU layer, thus increasing the depth of the network to boost its performance [14,15].

4.2.2.2. GA-based ML Modeling Process

After determining the ML algorithms, the actual development process of ML models was executed, as represented in Figure 20.

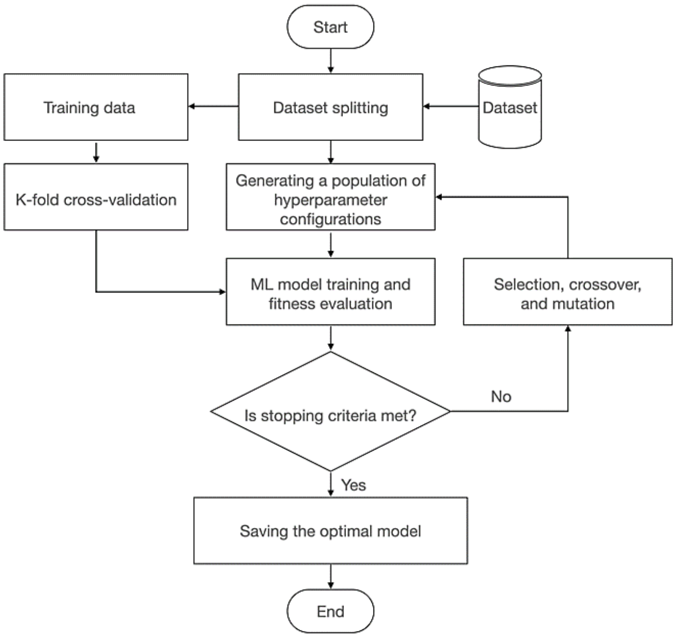


Figure 20. The GA-based hyperparameter optimization framework

Each developed dataset was first divided into 80% training and 20% testing subsets. Subsequently, the k-fold cross-validation was applied. Considering the trade-off in terms of computational time and accuracy, the value of k was set to 10.

To obtain optimal performance of developed ML models, it is essential to fine-tune and optimize the model configurations. Table 8 presents the list of optimized hyperparameters from RF and GRU.

Table 8. 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_features	The number of features that will counted for the best split.
	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.

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 design project used GA-based optimization of the ML models as proposed in the literature [34].

As shown in Figure 20, at the beginning of the optimization process, a random set of hyperparameter arrays was generated and used to develop the first generation of ML models. The performance of each model was assessed and ranked, where the best models were identified. By applying crossover and mutation on the top-ranking solutions, the subsequent generation of models was obtained. The optimization process would continue until the stopping criteria were met.

Finally, the developed ML models were validated to evaluate the performance of data outside the range of the training set.

In this validation process, 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.

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

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{4}$$

$$ASE = \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{n} \tag{5}$$

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, R^2 was also used as the fitness function in the GA-based model optimization process to represent the fitness of each examined chromosome.

4.2.3. ML Model Interpretation

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, 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 for each regression task.

To ensure that the feature importance interpretation can be applied to both ML algorithms, this design project 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 was obtained.

5. System Validation

To validate the proposed framework, case studies were conducted. Following the system design demonstrated in the previous chapter, data were collected to develop datasets required for the ML modeling. The initial step was to explore, assess, and analyze the PQi measurements over the past few years. This was followed by the filtering of the measurements and associated archived data based on their completeness. Besides, the projects used in the case studies for modelling the pavement's long-term performance were confined to the surface layer, given that most of the distress takes place on the pavement surface.

Specifically, the regression of density degree was based on the data provided by the Dutch contractor Heijmans [115], collected from a series of construction projects around the Schiphol Airport. For the regression of residual lifespan and IRI, two Dutch highway sections (A58 and A4) with a total length of 4.1 km were selected.

After the data collection, the data integration was conducted. It is worth mentioning that for aligning the coordinates of drilled cores in the case study provided by Heijmans with the cells, the calculated radius for the circle for covering the cells was 0.3 m. Table 9, Table 10, and Table 11 provide examples of the collected raw data respectively, before applying the data exploratory analysis and data cleansing. In total, 197 data samples were collected for the regression task of density degree, from the projects provided by Heijmans. In addition, 156 data samples were included in the dataset for the regression of residual lifespan, and 62 data samples were used for the regression of IRI.

Table 9. An example of the developed dataset for the regression of density degree

Samples	**Input features**													**Density degree**
	Roller Pass Deviation	Percentage of Roller Passes in Temperature Window	Temp. Mean	Temp. Std	Preci. Mean	Preci. Std	Wind Speed Mean	Wind Speed Std	Humidity Mean	Humidity Std	Press. Mean	Press. Std	Mixture	
1	-0.4333	0.4852	8.7167	1.1040	0.0167	0.0372	24.2500	5.1498	0.8696	0.0352	1013	1.4158	38367	1.0055
2	-0.5916	0.3161	8.7167	1.1040	0.0167	0.0372	24.2500	5.1498	0.8696	0.0352	1013	1.4158	38367	0.9827
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 10. An example of the developed dataset for the regression of residual lifespan

Samples	**Input features**																				**Residual lifespan**
	ECR	Mixture	Age	Temp. Mean	Temp. Std	Preci. Mean	Preci. Std	Wind Speed Mean	Wind Speed Std	Humidity Mean	Humidity Std	Press. Mean	Press. Std	Annual Mean Temp.	Annual Mean Preci.	Annual Freeze/Thaw Cycles	Passenger Cars/Day	Med. Trucks/Day	Heavy Trucks/Day	Residual lifespan-1	
1	0.3725	ZOAB-1	1	21.8260	4.3129	0	0	0.7100	0.5422	0.25	0.4274	1024	0.9018	11.7195	1.5619	60	33039	2843	3419	14	13
2	0.2955	ZOAB-1	1	21.8260	4.3129	0	0	0.7100	0.5422	0.25	0.4274	1024	0.9018	11.7195	1.5619	60	33039	2843	3419	14	13
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 11. An example of the developed dataset for the regression of IRI

Samples	**Input features**																				**IRI**
	ECR	Mixture	Age	Temp. Mean	Temp. Std	Preci. Mean	Preci. Std	Wind Speed Mean	Wind Speed Std	Humidity Mean	Humidity Std	Press. Mean	Press. Std	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	21.8260	4.3129	0	0	0.7100	0.5422	0.25	0.4274	1024	0.9018	11.7195	1.5619	55	47318	3406	3476	1.32	0.93
2	0.1372	ZOAB-1	3	21.8260	4.3129	0	0	0.7100	0.5422	0.25	0.4274	1024	0.9018	11.7195	1.5619	55	47318	3406	3476	1.20	0.64
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

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 12. In this case study, the number of generations is the stopping criterion.

Table 12. Pre-defined GA parameters

GA parameters	Description	Value
Population size	The number of individuals that will be evaluated and selected	100
Offspring size	The number of individuals which will be re-generated after one iteration	100
Crossover rate	The probability of two random individuals after selection replacing their gene fragments	0.8
Mutation rate	The possibility of the occurrence of mutation	0.2
Number of generations	The number of iterations in the GA process	100

5.1 Results

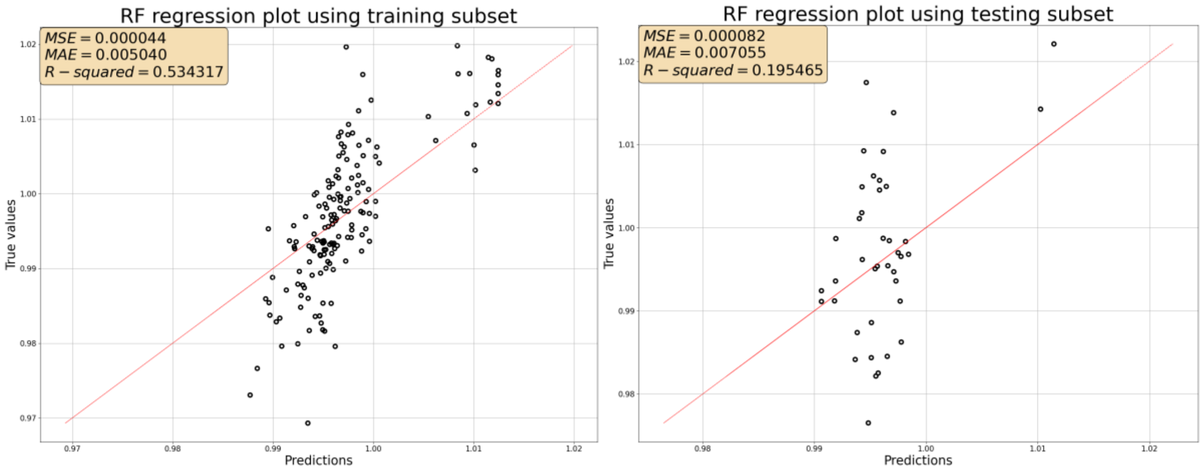
Regression of density degree

For the regression of density regression, only RF was used to build the ML model given the non-time-variant nature of the model output. Based on the dataset represented in Table 9, the target encoding was performed on the variable of asphalt mixtures.

Table 13 presents the model evaluation in terms of MSE, MAE, and R^2. Also, Figure 21 demonstrates the regression plots of the obtained model, including both the training and testing processes. Additionally, Table 14 also summarizes the results of the optimization results of the hyperparameters.

Table 13. The summary of the results of the model validations for the regression of density degree

Model	R^2	MSE	MAE
RF	0.1954	0.0001	0.0071



(a). Performance of RF model on training data (b). Performance of RF model on training data

Figure 21. Regression plots of the developed model for the regression of density degree

Table 14. The summary of the optimization results regarding hyperparameter configurations for the regression of density degree

ML algorithm	Hyperparameter	Value
RF	n_estimators	100
	Max_features	0.05
	max_depth	None
	min_samples_split	10
	min_samples_leaf	2

Based on Table 13 and Figure 21, it is clear that the developed RF model suffered from the underfitting problem, given the poor performance achieved during the training process. In addition, there is also a significant difference between the model evaluations in the training process and testing process respectively. Besides, as shown in Figure 21. (b), it has confirmed poor fitting performance of the resulting model.

When a machine learning model underfitted, it cannot capture the underlying relationships. As a result, the feature importance computed from an underfit model will not bring any value, and will mislead the interpretation. Therefore, for the regression of density degree, the feature importance is not interpreted.

Regression of residual lifespan

For the regression of residual lifespan, both the RF and GRU algorithms were used to develop corresponding models. It is worth noting that because only two projects were selected in the case study, meaning there will only be two scenarios concerning the ambient conditions during the construction process. Because this binary scenario also applies to the mixture type, therefore, the mixture type and all the ambient conditioning variables will be completely correlated. Therefore, all the variables regarding ambient data during the construction phase were removed, because their information can be represented by the mixture type.

Table 15 presents the evaluations of two obtained models in terms of MSE, MAE, and R^2. Also, Figure 22 demonstrates the regression plots of the obtained model, including both the training and testing processes. Additionally, Table 16 summarizes the results of the optimization of the hyperparameters.

The results shown in Table 15 indicate that the two developed models have close predictive performance, however, the RF model slightly outperforms the GRU model, with an R^2 of 0.8297 compared to 0.8172.

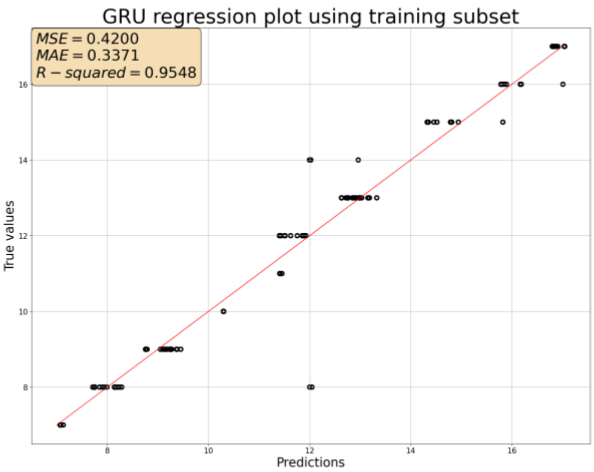
Therefore, for the ML model interpretation and further analysis, the RF model was used. The permutation feature importance was applied, and the results are shown in Figure 23. According to the result, it is clear that the residual lifespan assessed from the previous year has the highest importance among the other features. Besides, the average annual temperature also shows considerably higher importance, while the rest of the features are of comparable importance. It is worth noting that ECR, which represents the overall evaluation of the quality of the construction on-site operations, has a relatively higher contribution to the model's predictive performance than the rest.

Table 15. The summary of the results of the model validations for the regression of residual lifespan

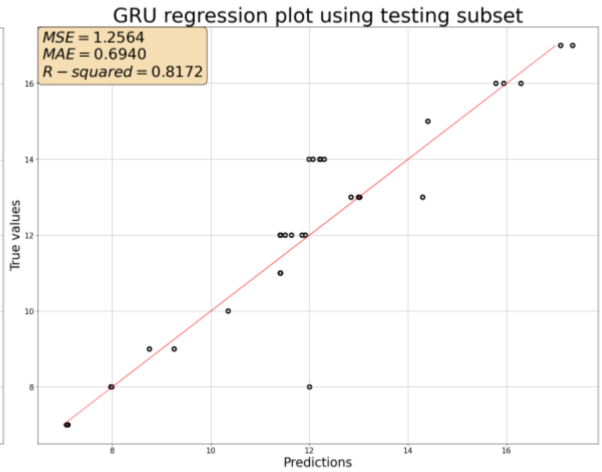
Model	$\mathbf{R^2}$	**MSE**	**MAE**
RF	0.8297	1.3540	0.6358
GRU	0.8172	1.2564	0.6949

(a). Performance of RF model on training data

(b). Performance of RF model on training data



(c). Performance of GRU model on training data



(d). Performance of GRU model on training data

Figure 22. Regression plots of the developed models for the regression of residual lifespan

Table 16. The summary of the optimization results regarding hyperparameter configurations for the regression of residual lifespan

ML algorithm	Hyperparameter	Value
RF	n_estimators	100
	Max_features	0.65
	max_depth	None
	min_samples_split	6
	min_samples_leaf	4
GRU	n_layers	0
	units	12
	epochs	120

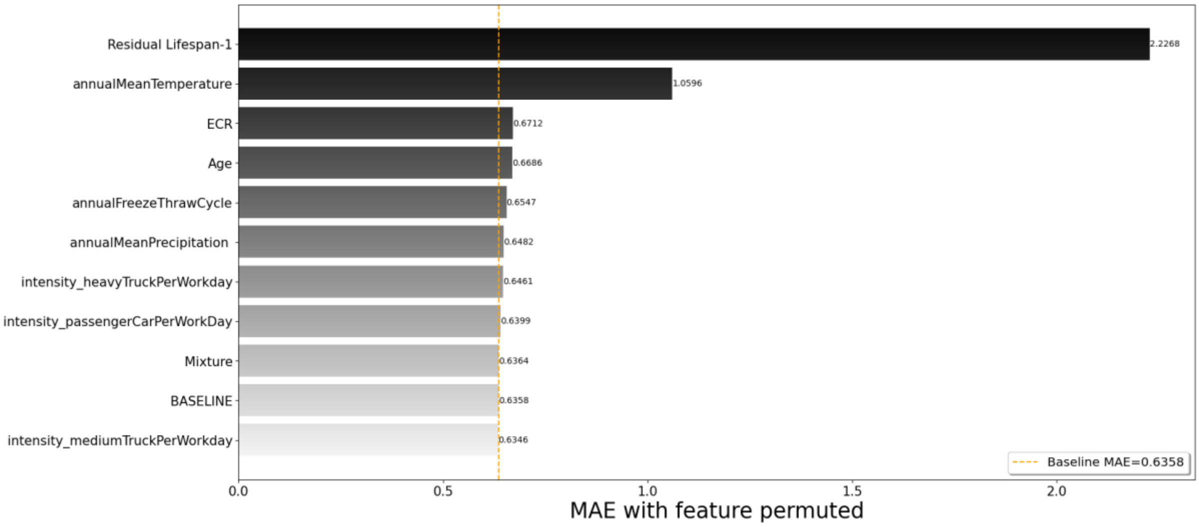


Figure 23. The permutation feature importance of the RF model for the regression of residual lifespan

Regression of IRI

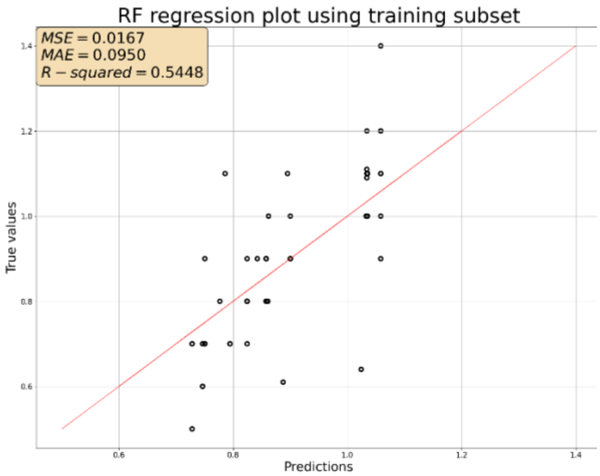
As for the regression of IRI, because its dataset shares the same structure and contents considering the input features, therefore, same as the regression of residual lifespan, all the variables regarding ambient data during the construction phase will be removed.

Similarly, two ML algorithms were selected and used, namely RF and GRU. Table 17 presents the comparison of the two models in terms of MSE, MAE, and R^2. Also, Figure 24 demonstrates the regression plots of each model, including both the training and testing processes. Additionally, Table 18 also summarizes the results of the optimization of the hyperparameters.

Based on Table 17 and Figure 24, it is clear that the developed GRU model significantly outperformed the RF model, where the latter is considerably underfitting. 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.

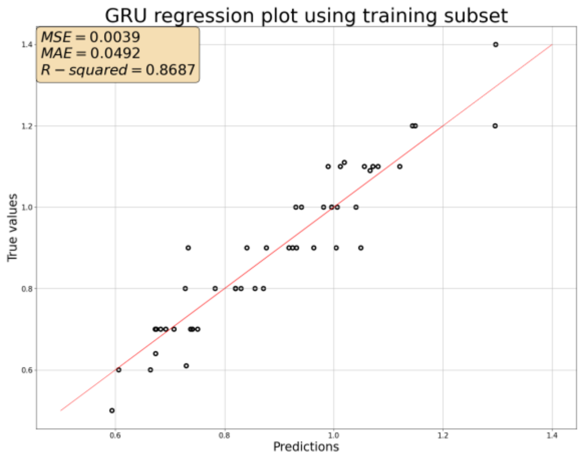
Table 17. The summary of the results of the model validations for the regression of IRI

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

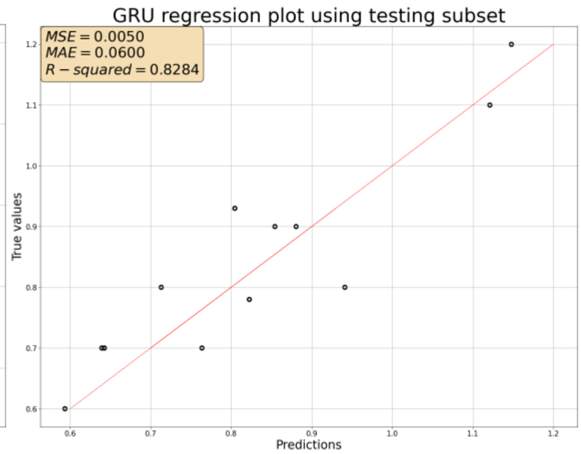


(a). Performance of RF model on training data

(b). Performance of RF model on training data



(c). Performance of GRU model on training data



(d). Performance of GRU model on training data

Figure 24. Regression plots of the developed models for the regression of IRI

Table 18. The summary of the optimization results regarding hyperparameter configurations for the regression of IRI

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 *25* 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 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.

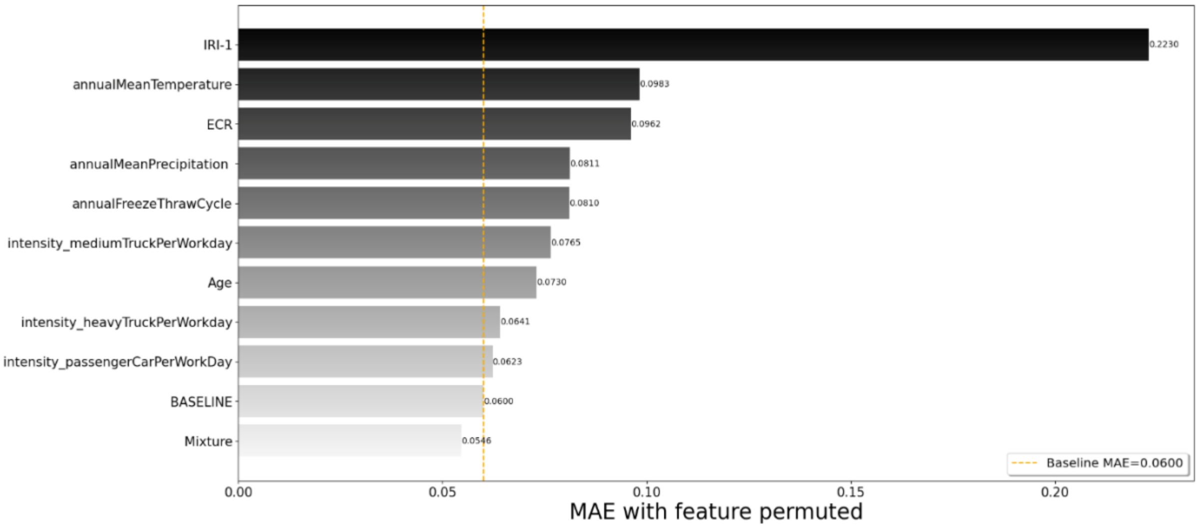


Figure 25. The permutation feature importance of the GRU model for the regression of IRI

6. Discussion

This section demonstrates the analyses and interpretation of the findings based on the undertaken validation of the system. Firstly, several major findings were summarized and contextualized by referring to the theoretical framework explored in this project, as well as other previous studies from the literature. Lastly, the significance of the study was highlighted.

6.1. Discussion on the Results of System Validation

For different modelling tasks, this design project has identified different data structures. However, it is crucial to acknowledge the trade-off made between the comprehensiveness of the mapping of the variables and challenges including the data availability, data completeness, and data quality. In this design project, the development of critical variables of the data structures was initiated by the development of a theoretical framework and the ontology proposed by [29], which encompassed the necessary concepts and theories. However, the identified parameters may not comprehensively reflect the properties of asphalt mixture, and their impact on the construction process quality and pavement quality. It is worth noting that 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. Therefore, a consolidated parameter would be sufficient to cover design-related characteristics in this project. However, this approach would simply assume that the same mixture types will always have the same properties, which neglects that variability can also be generated during the production process of the asphalt mixtures. Consequently, it would be ideal to expand the scope of the identification of key parameters. However, this will also increase the difficulty of data collection, because the expanded dimensionality of the data structure would require more data to ensure the predictive performance of the ML models, while the availability of the data will significantly influence the data collection process.

Overall, the RF model developed for the regression of density degree failed to provide promising performance. The performance achieved in the training and testing process is significantly poor, proving the model is underfitting. Referring to the corresponding requirements, no effective conclusion can be drawn regarding the process quality indicators and the density degree through the case study conducted in this project. However, this deviates from the consensus from the literature that the density of the asphalt pavement should be highly correlated to the compaction effectiveness [49]. The possible reason could be that the determination regarding the target number of roller passes and compaction temperature windows was incorrect, owing to the lack of a clear understanding of the importance of these two parameters to the overall construction process quality, as well as a lack of effective measuring methods among the contractors. By conducting this design project, it is hoped that the struggle of developing a promising ML model regarding pavement mechanical properties can emphasize the importance of acquiring accurate data regarding these two parameters.

On the other hand, the obtained ML models for the regression of long-term pavement performance achieved promising results, in spite of using a small dataset. For the regression of residual lifespan, the developed RF and GRU models all obtained satisfying performance being checked with the requirements, where the RF model (R^2 is 0.8297) slightly outperforms the GRU model (R^2 is 0.8172). When predicting the IRI, the regression model using GRU achieved significantly better performance (R^2 is 0.8284) compared to the RF model (R^2 is 0.5498). The outcomes from the two regression tasks are contradictory to each other. 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. This has been confirmed by the case of the regression of residual lifespan. However, a previous study suggested that the high reliability of the model can be achieved even with small datasets [7]. 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. 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 for the regression of IRI suffered from the overfitting problem with this small amount of data. This could potentially mean that the data used in this study for the regression of IRI is insufficient to support conventional ML algorithms, such as RF. On the other hand, the developed GRU model showed outstanding capability regarding feature extraction. However, it is still necessary to enlarge the dataset in the future, to achieve more generable ML models.

Lastly, for both regression tasks regarding long-term pavement performance, the previous measurements of the pavement performance outranked the other features. This is in line with various studies which also applied time-series regression [15,81], which further emphasizes the temporal dependence of the previous measurements and their significant influence on the future values. Representing the construction process quality, ECR ranked third in both cases, 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 [116].

6.2. Discussions on the Implications for the Industry

During the process of accomplishing this design project, the major obstacle was collecting and integrating data from various organizations to develop datasets with sufficient sizes and adequate quality. The struggle regarding data collection implies poor data management in the road construction industry. Currently, there is an absence of standardization in the industry, in terms of data formats, protocols, and terminology. Consequently, data owned by different organizations are significantly inconsistent, leading to conflicts when consolidating different data sources. To address this issue, it is essential to acknowledge the pivotal role of a global ontology for pavement lifecycle management, which can provide a standardized representation of data using its defined concepts and relations. This can significantly ease the data integration process by eliminating syntactic and semantic conflicts among the data extracted from different sources.

Asphalt is a highly complex material, where the quality product in each phase of the lifecycle is influenced by various factors and also how the previous phases unfolded [35]. 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, which can be regarded as a preliminary attempt to integrate data, organizations, and processes through the entire road construction lifecycle. This would be beneficial to 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 adoption of digital technologies and modernization of the asphalt construction process, thus improving and optimizing the planning, and implementation of explicit and science-based operational strategies. This would lead to increased productivity, and will eventually result in the reduction of the variability during the construction process and improve the pavement quality. Besides, for asset managers, having a clear understanding of the relationships between process and product quality can advance the preventive maintenance practice, enabling the accurate detection and prevention of potential distresses given the specific conditions across the entire lifecycle.

The explicit correlation between process and product quality can also enhance the practices regarding quality inspections in the context of quality control. Conventionally, the pavement quality is inspected as the verification of the construction outcome. Therefore, the primary purpose of performing quality inspections is to test the pavement quality to pass the functional requirements. However, no matter the test results are good or not, the contractors cannot link the obtained pavement quality with their on-site operations. Consequently, when the post-construction sample tests indicate inadequate mechanical properties, they are unable to determine what went wrong during the construction phase and prevent the same operation in the future projects. With an explicit understanding regarding the relationship between the construction process quality and the resulting pavement quality, it is possible to achieve the transition from implicit and outcome-oriented quality control to explicit and process-oriented quality control. Thereby, it is possible to trace back the adopted on-site operations according to the tested pavement quality, thus enabling active and continuous learning and improvement.

The explicit quantification of the impact of road construction process quality on the resulting pavement quality can also be a supplement to the roller operator guidance system. Due to the isolated view of the current operator guidance system, the guidance is provided to operators solely with the direct evaluation of the compaction effectiveness based on the data collected during the construction process. However, although there exists a strong correlation between the compaction effectiveness and the eventual quality of the pavement, the current system cannot explicitly and quantitatively map the quality of the compaction operations into the achieved quality of the construction results. Consequently, to what extent the adopted operational strategies can affect pavement quality is still unknown. The guidance provided to operators may lack precision in terms of understanding how their actions influence the final quality. These imply although operator guidance system has successfully achieved certain intra-process integration, there is an absence of inter-process integration, i.e., the integrations through the entire lifecycle of the road construction, considering data, processes, technologies, and organizations. Hence, by incorporating the explicit quantification of the impact of road construction process quality on the resulting pavement quality into the framework of the operator guidance system, it can facilitate the aforementioned integration, thus optimizing road construction practices and improving quality control.

7. Conclusion and future work

7.1. Conclusion

This design project aimed to investigate the correlations between asphalt construction process quality and product quality, using data-driven techniques. In this design project, based on the investigated theoretical framework and the social context, a data-driven method was proposed and described, which can systematically reveal the hidden correlations between process and product quality indicators through the development and interpretation of ML models.

In this design project, the input-output structure of the datasets required for the ML model development was identified, including input variables such as quality indicators of the on-site operational strategies, weather conditions, mixture type, and auxiliary parameters in the operational phase of the pavement (such as traffic intensity and climate condition). The outputs of the identified data structure include density degree, residual lifespan, and IRI, covering the short- and long-term perspectives of the pavement product quality.

A GA-based ML model development method was designed, where RF and GRU were selected as the ML algorithms. For the validation, case studies were conducted. Based on the collected data, the regression model for the density degree using RF failed to satisfy the corresponding requirements regarding the model performance. For the regression of residual lifespan and IRI, both the RF and GRU were used to develop corresponding models. For residual lifespan, the developed RF model outperformed the GRU model, with an R^2 of 0. 8297, while the regression of IRI shows contradictory results, where the developed GRU model significantly outperformed (R^2 is 0.8284). After interpreting the permutation importance, both cases show that ECR achieved the third highest importance, revealing the rather high correlation between process quality and product quality in asphalt construction.

Referring to the design project questions, overall, the proposed data-driven method and corresponding results obtained through case studies can answer all the formulated design project questions. When it comes to the derived requirements, the developed regression models for two selected long-term performance indicators, following the proposed method, have satisfied all the requirements, in terms of the accuracy, generality, and interpretability. However, the regression model for the density degree failed to provide desirable predictive performance.

7.2. Future work

For future work, because the case study in this design project was performed on a small dataset, the EngD candidate would like to expand the scope of the dataset, because it is believed that a larger dataset with improved data quality can enhance the ML model performance, particularly for the regression of density degree. Besides, the presented design project only focused on a confined selection of product quality indicators, 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.

Besides, the validation of the design project is confined to internal validity by conducting case studies and comparing the results with previous studies. In order to rigorously validate the methodology and outputs, it is critical to expand the validating method to test its concurrent validity in the future. This can be done in the future by comparing the results obtained from the ML predictions with other empirical models or direct measurements of the target product quality indicators, under the same pre-defined input variables concerned with the process quality indicators.

References

[1] A. Hartmann, F.Y.Y. Ling, Value creation of road infrastructure networks: A structural equation approach, Journal of Traffic and Transportation Engineering (English Edition). 3 (2016) 28–36. https://doi.org/10.1016/j.jtte.2015.09.003.

[2] 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).

[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] R. Kuenzel, J. Teizer, M. Mueller, A. Blickle, SmartSite: Intelligent and autonomous environments, machinery, and processes to realize smart road construction projects, Autom Constr. 71 (2016) 21–33. https://doi.org/10.1016/J.AUTCON.2016.03.012.

[5] B.K. Krishnamurthy, H.P. Tserng, R.L. Schmitt, J.S. Russell, H.U. Bahia, A.S. Hanna, AutoPave: towards an automated paving system for asphalt pavement compaction operations, Autom Constr. 8 (1998) 165–180. https://doi.org/10.1016/S0926-5805(98)00060-0.

[6] International Organization for Standardization, ISO 9000:2015: Quality management systems - Fundamentals and vocabulary, (n.d.). https://asq.org/quality-press/display-item?item=T1039&utm_source=qaqc&utm_medium=webpage&utm_campaign=qaqc&utm_id=LAQ (accessed May 10, 2023).

[7] S.R. Miller, A.G. Dorée, A framework for monitoring asphalt mix temperature during construction using Statistical Process Control charts, in: The 11th Conference on Asphalt Pavements for Southern Africa, Capsa, Sun City, South Africa, 2015: pp. 16–19. https://doi.org/https://trid.trb.org/view/1404786.

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

[9] HOME | Aspari, (n.d.). https://en.aspari.nl/ (accessed May 22, 2023).

[10] S.S. Chaurasia, S. Verma, Strategic determinants of big data analytics in the aec sector: A multi-perspective framework, Construction Economics and Building. 20 (2020) 63–81. https://doi.org/10.5130/AJCEB.v20i4.6649.

[11] E. Alpaydin, Introduction to Machine Learning, MIT Press, London, UK, 2010.

[12] S. Amarasiri, B. Muhunthan, Evaluating Cracking Deterioration of Preventive Maintenance–Treated Pavements Using Machine Learning, Journal of Transportation Engineering, Part B: Pavements. 148 (2022) 04022014. https://doi.org/10.1061/JPEODX.0000354.

[13] M. Shariatfar, Y.C. Lee, K. Choi, M. Kim, Effects of flooding on pavement performance: a machine learning-based network-level assessment, Sustain Resilient Infrastruct. 7 (2022) 695–714. https://doi.org/10.1080/23789689.2021.2017736.

[14] 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.

[15] 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.

[16] W. Zhang, A. Raza Khan, S. Yoon, J. Lee, R. Zhang, K. Zeng, Investigation of the correlations between the field pavement in-place density and the intelligent compaction measure value (ICMV) of asphalt layers, Constr Build Mater. 292 (2021) 123439. https://doi.org/10.1016/J.CONBUILDMAT.2021.123439.

[17] N. Baldo, M. Miani, F. Rondinella, J. Valentin, P. Vackcová, E. Manthos, Stiffness Data of High-Modulus Asphalt Concretes for Road Pavements: Predictive Modeling by Machine-Learning, Coatings 2022, Vol. 12, Page 54. 12 (2022) 54. https://doi.org/10.3390/COATINGS12010054.

[18] S.S. Todkar, V. Baltazart, A. Ihamouten, X. Dérobert, D. Guilbert, One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data, J Appl Geophy. 192 (2021) 104392. https://doi.org/10.1016/J.JAPPGEO.2021.104392.

[19] Y. Huang, Z. Qiao, H. Zhang, Evaluation of an economy-technology-green development system for asphalt pavement construction in China based on synergetics, J Clean Prod. (2020) 125132. https://doi.org/10.1016/j.jclepro.2020.125132.

[20] Y. Zhao, L. Feng, Y. Sun, K. Song, Historical Analysis of Urban Public Transportation Development in Modern Tianjin (1902-1949), International Planning History Society Proceedings. 18 (2018) 1036–1047. https://doi.org/https://doi.org/10.7480/iphs.2018.1.2750.

[21] O. Safonova, L. Tatarnikova, Assessment of the competitiveness of industrial companies and methods for assessing the quality of construction products, in: A. Zheltenkov, A. Mottaeva (Eds.), E3S Web of Conferences, EDP Sciences, 2020: p. 09027. https://doi.org/10.1051/e3sconf/202016409027.

[22] T. Levitt, Marketing success through differentiation—of anything, in: Harv Bus Rev, 1980: pp. 83–91.

[23] R.M. Toivonen, Product quality and value from consumer perspective-An application to wooden products, J For Econ. 18 (2012) 157–173. https://doi.org/10.1016/j.jfe.2011.12.004.

[24] F.-J. Molina-Castillo, R.J. Calantone, M.A. Stanko, J.-L. Munuera-Aleman, Product Quality as a Formative Index: Evaluating an Alternative Measurement Approach *****, Journal of Product Innovation Management. 30 (2013) 380–398. https://doi.org/10.1111/j.1540-5885.2012.01005.x.

[25] S. Curkovic, S.K. Vickery, C. Droge, An empirical analysis of the competitive dimensions of quality performance in the automotive supply industry, International Journal of Operations and Production Management. 20 (2000) 386–403. https://doi.org/10.1108/01443570010308121.

[26] R.J. Wieringa, Design science methodology: For information systems and software engineering, Springer-Verlag Berlin Heidelberg, 2014. https://doi.org/10.1007/978-3-662-43839-8.

[27] A.J. Onwuegbuzie, R. Frels, Seven Steps to a Comprehensive Literature Review : A Multimodal and Cultural Approach, SAGE Publications Ltd, 2016.

[28] J.-L. Boulanger, Requirements Management, in: Certifiable Software Applications 3, Elsevier, 2018: pp. 7–27. https://doi.org/10.1016/B978-1-78548-119-2.50002-9.

[29] M. Sadeghian, Develop an Ontology for Pavement Lifecycle Management, EngD thesis, University of Twente, 2023.

[30] K. Worden, G. Manson, The application of machine learning to structural health monitoring, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 365 (2007) 515–537. https://doi.org/10.1098/rsta.2006.1938.

[31] V. Cherkassky, F. Mulier, Learning from Data: Concepts, Theory, and Methods: Second Edition, John Wiley and Sons, 2006. https://doi.org/10.1002/9780470140529.

[32] 3.1. Cross-validation: evaluating estimator performance — scikit-learn 1.2.2 documentation, (n.d.). https://scikit-learn.org/stable/modules/cross_validation.html (accessed June 12, 2023).

[33] D. Baehrens, S. Harmeling, M. Kawanabe, K. Hansen KHANSEN, C. Edward Rasmussen, How to Explain Individual Classification Decisions Timon Schroeter * Klaus-Robert M ¨ uller, Journal of Machine Learning Research. 11 (2010) 1803–1831.

[34] 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.

[35] 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.

[36] A. Cala, S. Caro, Y. Rojas-Agramonte, Aggregate properties affecting adhesive quality in asphalt mixtures, Advances in Materials and Pavement Performance Prediction II - Contributions to the 2nd International Conference on Advances in Materials and Pavement Performance Prediction, AM3P 2020. (2020) 16–19. https://doi.org/10.1201/9781003027362-4.

[37] A. Cala, S. Caro, Predictive quantitative model for assessing the asphalt-aggregate adhesion quality based on aggregate chemistry, Road Materials and Pavement Design. 23 (2022) 1523–1543. https://doi.org/10.1080/14680629.2021.1900896.

[38] G.D. Airey, A.C. Collop, S.E. Zoorob, R.C. Elliott, The influence of aggregate, filler and bitumen on asphalt mixture moisture damage, Constr Build Mater. 22 (2008) 2015–2024. https://doi.org/10.1016/J.CONBUILDMAT.2007.07.009.

[39] H. Zhang, Z. Liu, X. Meng, Noise reduction characteristics of asphalt pavement based on indoor simulation tests, Constr Build Mater. 215 (2019) 285–297. https://doi.org/10.1016/J.CONBUILDMAT.2019.04.220.

[40] Z. Qian, Q. Lu, Design and laboratory evaluation of small particle porous epoxy asphalt surface mixture for roadway pavements, Constr Build Mater. 77 (2015) 110–116. https://doi.org/10.1016/J.CONBUILDMAT.2014.12.056.

[41] G. Shafabakhsh, M. Taghipoor, M. Sadeghnejad, S.A. Tahami, Evaluating the effect of additives on improving asphalt mixtures fatigue behavior, Constr Build Mater. 90 (2015) 59–67. https://doi.org/10.1016/J.CONBUILDMAT.2015.04.046.

[42] F. Yin, J. Garita, A. Taylor, R. West, Refining the indirect tensile (IDT) Nflex Factor test to evaluate cracking resistance of asphalt mixtures for mix design and quality assurance, Constr Build Mater. 172 (2018) 396–405. https://doi.org/10.1016/J.CONBUILDMAT.2018.03.251.

[43] S. Angelone, M.C. Casaux, F. Martinez, Implementation of a static strength test for evaluating the rutting resistance of asphalt mixtures and its application for quality controls, in: Asphalt Pavements - Proceedings of the International Conference on Asphalt Pavements, ISAP 2014, Taylor and Francis - Balkema, 2014: pp. 679–688. https://doi.org/10.1201/b17219-86.

[44] Y. Huang, R. Bird, O. Heidrich, Development of a life cycle assessment tool for construction and maintenance of asphalt pavements, J Clean Prod. 17 (2009) 283–296. https://doi.org/10.1016/j.jclepro.2008.06.005.

[45] G.K. Chang, K. Mohanraj, W.A. Stone, D.J. Oesch, V. (Lee) Gallivan, Leveraging Intelligent Compaction and Thermal Profiling Technologies to Improve Asphalt Pavement Construction Quality: A Case Study, Transportation Research Record: Journal of the Transportation Research Board. 2672 (2018) 48–56. https://doi.org/10.1177/0361198118758285.

[46] D. Liu, M. Lin, S. Li, Real-Time Quality Monitoring and Control of Highway Compaction, Autom Constr. 62 (2016) 114–123. https://doi.org/10.1016/j.autcon.2015.11.007.

[47] C. Newcomer, J. Withrow, R.E. Sturgill, G.B. Dadi, R.E. Sturgill, Towards an Automated Asphalt Paving Construction Inspection Operation, in: Mutis I., Hartmann T. (Eds) Advances in Informatics and Computing in Civil and Construction Engineering, Springer, Cham, 2019: pp. 593–600. https://doi.org/10.1007/978-3-030-00220-6_71.

[48] F. Bijleveld, S.R. Miller, A.H. de Bondt, A.G. Doree, Too hot to handle, too cold to control - influence of compaction temparature on the mechanical properties of asphalt, in: E. Beuving, P. Dewez, G. Malkoc, M. Souther (Eds.), E. Beuving, P. Dewez, G. Malkoc, & M. Southern (Eds.), Proceedings of the 5th Eurasphalt and Eurobitume Congress, 13-15 June 2012, Foundation Eurasphalt, Istanbul, Turkey, 2012: pp. 1–11. https://research.utwente.nl/en/publications/too-hot-to-handle-too-cold-to-control-influence-of-compaction-tem (accessed April 12, 2021).

[49] 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.

[50] N.A. Morgan, D.W. Vorhies, Product quality alignment and business unit performance, Journal of Product Innovation Management. 18 (2001) 396–407. https://doi.org/10.1111/1540-5885.1860396.

[51] R. Sebastianelli, N. Tamimi, How product quality dimensions relate to defining quality, International Journal of Quality and Reliability Management. 19 (2002) 442–453. https://doi.org/10.1108/02656710210421599.

[52] J. Lemmink, H. Kasper, Competitive Reactions to Product Quality Improvements in Industrial Markets, Eur J Mark. 28 (1994) 50–68. https://doi.org/10.1108/03090569410074255.

[53] Y. Huang, manual-pavement-analysis-and-design-2nd-edition-huang, Prentice Hall, Upper Saddle River, NJ, 1993.

[54] H. Wang, M. Li, P. Szary, X. Hu, Structural assessment of asphalt pavement condition using backcalculated modulus and field data, Constr Build Mater. 211 (2019) 943–951. https://doi.org/10.1016/j.conbuildmat.2019.03.250.

[55] J.T. Lin, Y. Xiao, Microstructure and performance characteristics of cold recycled asphalt mixtures, in: Eco-Efficient Pavement Construction Materials, Elsevier Inc., 2020: pp. 51–76. https://doi.org/10.1016/B978-0-12-818981-8.00004-7.

[56] C. Olidis, D. Hein, Guide for the Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures Materials Characterization Is your Agency Ready?, in: 2004 Annual Conference of the Transportation Association of Canada, 2004.

[57] M. Sabouri, An investigation on perpetual asphalt pavements in Minnesota, International Journal of Pavement Research and Technology. 13 (2020) 247–254. https://doi.org/10.1007/s42947-020-0149-2.

[58] T. Xu, X. Huang, Investigation into causes of in-place rutting in asphalt pavement, Constr Build Mater. 28 (2012) 525–530. https://doi.org/10.1016/j.conbuildmat.2011.09.007.

[59] G.K. Chang, K. Mohanraj, W.A. Stone, D.J. Oesch, V.(Lee,) Gallivan, Leveraging Intelligent Compaction and Thermal Profiling Technologies to Improve Asphalt Pavement Construction Quality: A Case Study, Transp Res Rec. 2672 (2018) 48–56. https://doi.org/10.1177/0361198118758285.

[60] S. Islam, W.G. Buttlar, Effect of Pavement Roughness on User Costs, Transportation Research Record: Journal of the Transportation Research Board. 2285 (2012) 47–55. https://doi.org/10.3141/2285-06.

[61] A.T. Papagiannakis, B. Raveendran, International Standards Organization-Compatible Index for Pavement Roughness, Transportation Research Record: Journal of the Transportation Research Board. 1643 (1998) 110–115. https://doi.org/10.3141/1643-14.

[62] H. Wang, B. Behnia, W.G. Buttlar, H. Reis, Development of two-dimensional micromechanical, viscoelastic, and heterogeneous-based models for the study of block cracking in asphalt pavements, Constr Build Mater. 244 (2020) 118146. https://doi.org/10.1016/j.conbuildmat.2020.118146.

[63] Y. Zhao, M. Alae, G. Fu, Investigation of mechanisms of top-down fatigue cracking of asphalt pavement, Road Materials and Pavement Design. 19 (2018) 1436–1447. https://doi.org/10.1080/14680629.2017.1303394.

[64] 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.

[65] J.S. Miller, W.Y. Bellinger, Distress Identification Manual for the Long-Term Pavement Performance Program, United States. Federal Highway Administration. Office of Infrastructure Research and Development, 2003.

[66] K. Hoegh, L. Khazanovich, K. Maser, N. Tran, Evaluation of Ultrasonic Technique for Detecting Delamination in Asphalt Pavements, Transportation Research Record: Journal of the Transportation Research Board. 2306 (2012) 105–110. https://doi.org/10.3141/2306-12.

[67] S.K. Dasari, A. Cheddad, P. Andersson, Predictive modelling to support sensitivity analysis for robust design in aerospace engineering, Structural and Multidisciplinary Optimization. 61 (2020) 2177–2192. https://doi.org/10.1007/S00158-019-02467-5.

[68] S.K. Dasari, A. Cheddad, P. Andersson, Random Forest Surrogate Models to Support Design Space Exploration in Aerospace Use-Case, IFIP Adv Inf Commun Technol. 559 (2019) 532–544. https://doi.org/10.1007/978-3-030-19823-7_45.

[69] K. Prudviraj, S. Deshmukh, R.K. Tripathy, K. Supradeepan, P. Tandon, P.K. Jha, Machine Learning-based Approach for the Prediction of an Orifice size of Aerospace Vehicle RCS Thrusters during Cold Flow Calibration, 2021 IEEE 6th International Conference on Computing, Communication and Automation, ICCCA 2021. (2021) 455–459. https://doi.org/10.1109/ICCCA52192.2021.9666377.

[70] M. Botvinick, S. Ritter, J.X. Wang, Z. Kurth-Nelson, C. Blundell, D. Hassabis, Reinforcement Learning, Fast and Slow, Trends Cogn Sci. 23 (2019) 408–422. https://doi.org/10.1016/J.TICS.2019.02.006.

[71] S. Singaravel, J. Suykens, P. Geyer, Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction,

Advanced Engineering Informatics. 38 (2018) 81–90. https://doi.org/10.1016/J.AEI.2018.06.004.

[72] C. Fan, F. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms, Appl Energy. 195 (2017) 222–233. https://doi.org/10.1016/J.APENERGY.2017.03.064.

[73] S.B. Kotsiantis, Supervised Machine Learning: A Review of Classification Techniques, 2007.

[74] S. Ray, A Quick Review of Machine Learning Algorithms, Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019. (2019) 35–39. https://doi.org/10.1109/COMITCON.2019.8862451.

[75] T.D. Akinosho, L.O. Oyedele, M. Bilal, A.O. Ajayi, M.D. Delgado, O.O. Akinade, A.A. Ahmed, Deep learning in the construction industry: A review of present status and future innovations, Journal of Building Engineering. 32 (2020) 101827. https://doi.org/10.1016/j.jobe.2020.101827.

[76] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, Pattern Recognit. 77 (2018) 354–377. https://doi.org/10.1016/J.PATCOG.2017.10.013.

[77] A.K. Langroodi, F. Vahdatikhaki, A. Doree, Activity recognition of construction equipment using fractional random forest, Autom Constr. 122 (2021) 103465. https://doi.org/10.1016/J.AUTCON.2020.103465.

[78] S. Qinshuo, A machine-learning-based approach for the design optimization of wind turbine foundations, University of Twente, 2020.

[79] N. Sholevar, A. Golroo, S.R. Esfahani, Machine learning techniques for pavement condition evaluation, Autom Constr. 136 (2022) 104190. https://doi.org/10.1016/J.AUTCON.2022.104190.

[80] Y.C. (James) Tsai, Y. Zhao, B. Pop-Stefanov, A. Chatterjee, Automatically detect and classify asphalt pavement raveling severity using 3D technology and machine learning, International Journal of Pavement Research and Technology. 14 (2021) 487–495. https://doi.org/10.1007/S42947-020-0138-5/METRICS.

[81] 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.

[82] N. Zavrtanik, J. Prosen, M. Tušar, G. Turk, The use of artificial neural networks for modeling air void content in aggregate mixture, Autom Constr. 63 (2016) 155–161. https://doi.org/10.1016/J.AUTCON.2015.12.009.

[83] A. Sadat Hosseini, P. Hajikarimi, M. Gandomi, F. Moghadas Nejad, A.H. Gandomi, Optimized machine learning approaches for the prediction of viscoelastic behavior of modified asphalt binders, Constr Build Mater. 299 (2021) 124264. https://doi.org/10.1016/J.CONBUILDMAT.2021.124264.

[84] H. Gong, Y. Sun, Y. Dong, B. Han, P. Polaczyk, W. Hu, B. Huang, Improved estimation of dynamic modulus for hot mix asphalt using deep learning, Constr Build Mater. 263 (2020) 119912. https://doi.org/10.1016/J.CONBUILDMAT.2020.119912.

[85] G.S. Moussa, M. Owais, Pre-trained deep learning for hot-mix asphalt dynamic modulus prediction with laboratory effort reduction, Constr Build Mater. 265 (2020) 120239. https://doi.org/10.1016/J.CONBUILDMAT.2020.120239.

[86] J. Barugahare, A.N. Amirkhanian, F. Xiao, S.N. Amirkhanian, Predicting the dynamic modulus of hot mix asphalt mixtures using bagged trees ensemble, Constr Build Mater. 260 (2020) 120468. https://doi.org/10.1016/J.CONBUILDMAT.2020.120468.

[87] Y. Ali, F. Hussain, M. Irfan, A.S. Buller, An eXtreme Gradient Boosting model for predicting dynamic modulus of asphalt concrete mixtures, Constr Build Mater. 295 (2021) 123642. https://doi.org/10.1016/J.CONBUILDMAT.2021.123642.

[88] H. Majidifard, B. Jahangiri, P. Rath, L. Urra Contreras, W.G. Buttlar, A.H. Alavi, Developing a prediction model for rutting depth of asphalt mixtures using gene expression programming, Constr Build Mater. 267 (2021) 120543. https://doi.org/10.1016/J.CONBUILDMAT.2020.120543.

[89] S. Rahman, A. Bhasin, A. Smit, Exploring the use of machine learning to predict metrics related to asphalt mixture performance, Constr Build Mater. 295 (2021) 123585. https://doi.org/10.1016/J.CONBUILDMAT.2021.123585.

[90] V. Ferretti, From stakeholders analysis to cognitive mapping and Multi-Attribute Value Theory: An integrated approach for policy support, Eur J Oper Res. 253 (2016) 524–541. https://doi.org/10.1016/j.ejor.2016.02.054.

[91] Discover Functionality of the Machine Learning Decision Support System | Aspari, (n.d.). https://www.aspari.nl/events-1/discover-functionality-of-the-machine-learning-decision-support-system (accessed March 29, 2023).

[92] S.R. Miller, T. Hartmann, A.G. Dorée, Measuring and visualizing hot mix asphalt concrete paving operations, Autom Constr. 20 (2011) 474–481. https://doi.org/10.1016/J.AUTCON.2010.11.015.

[93] D.S. Decker, State-of-the-Practice for Cold-Weather Compaction of Hot-Mix Asphalt Pavements, Transportation Research Circular. (2006).

[94] H.L. ter Huerne, Compaction of asphalt road pavements: Using finite elements and critical state theory, University of Twente, 2004. https://doi.org/10.2/JQUERY.MIN.JS.

[95] 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).

[96] M.W. Sayers, T.D. Gillespie, C.A.V. Queiroz, The international road roughness experiment: a basis for establishing a standard scale for road roughness measurements, Transp Res Rec. (1986) 76–85.

[97] 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.

[98] B.S. Smith, Design and Construction of Pavements in Cold Regions: State of Design and Construction of Pavements in Cold Regions: State of the Practice the Practice, Brigham Young University , 2006. https://scholarsarchive.byu.edu/etd (accessed October 18, 2022).

[99] C.S. Jensen, T.B. Pedersen, C. Thomsen, Multidimensional Databases and Data Warehousing, Morgan & Claypool, 2010. https://doi.org/10.2200/S00299ED1V01Y201009DTM009.

[100] M. Ameri, M. Kashani Novin, B. Yousefi, Comparison of the field measurements of asphalt concrete densities obtained by ground-penetrating radar, pavement quality indicator and the borehole coring methods, Http://Dx.Doi.Org/10.1080/14680629.2014.909874. 15 (2014) 759–773. https://doi.org/10.1080/14680629.2014.909874.

[101] J. Driessen ... et al, Dienst Weg- en Waterbouwkunde (Delft), Beschrijvende plaatsaanduiding systematiek, (1994).

[102] C.S. Jansen, T.B. Pedersen, C. Thomsen, Fundamental Concepts, in: Multidimensional Databases and Data Warehousing, Morgan & Claypool, 2010: pp. 7–32.

[103] J.J. Berman, Understanding Your Data, Data Simplification. (2016) 135–187. https://doi.org/10.1016/B978-0-12-803781-2.00004-7.

[104] P. Rodríguez, M.A. Bautista, J. Gonzàlez, S. Escalera, Beyond one-hot encoding: Lower dimensional target embedding, Image Vis Comput. 75 (2018) 21–31. https://doi.org/10.1016/J.IMAVIS.2018.04.004.

[105] P. Rodríguez, M.A. Bautista, J. Gonzàlez, S. Escalera, Beyond one-hot encoding: Lower dimensional target embedding, Image Vis Comput. 75 (2018) 21–31. https://doi.org/10.1016/J.IMAVIS.2018.04.004.

[106] D. Kocev, C. Vens, J. Struyf, S. Džeroski, Ensembles of multi-objective decision trees, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 2007: pp. 624–631. https://doi.org/10.1007/978-3-540-74958-5_61.

[107] 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.

[108] D. Zhang, M.R. Kabuka, Combining weather condition data to predict traffic flow: a GRU-based deep learning approach, IET Intelligent Transport Systems. 12 (2018) 578–585. https://doi.org/10.1049/IET-ITS.2017.0313.

[109] J. Chen, H. Jing, Y. Chang, Q. Liu, Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process, Reliab Eng Syst Saf. 185 (2019) 372–382. https://doi.org/10.1016/J.RESS.2019.01.006.

[110] F. Shahid, A. Zameer, M. Muneeb, Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM, Chaos Solitons Fractals. 140 (2020) 110212. https://doi.org/10.1016/J.CHAOS.2020.110212.

[111] 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.

[112] 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.

[113] 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.

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

[115] Homepage Dutch | Heijmans NV, (n.d.). https://www.heijmans.nl/nl/ (accessed February 10, 2023).

[116] R.W. Perera, C. Byrum, S.D. Kohn, Inc. Soil and Materials Engineers, Investigation of Development of Pavement Roughness, (1998). https://rosap.ntl.bts.gov/view/dot/42761 (accessed January 25, 2023).