

AsphaltMine: The asphalt test result database for performance prediction

Mohammad Abbas¹, Martins Zaumanis^{1*}

¹Empa, Swiss Federal Laboratories for Materials Science and Technology, Switzerland

Received: 18 March 2026 / Accepted: 30 September 2026 / Published online: 27 October 2026

© The Author(s) 2026. This article is published with open access and licensed under a Creative Commons Attribution 4.0 International License.

Abstract

The lack of a unified, accessible platform for collecting and sharing standardized asphalt test result data has limited progress in data-driven pavement research. To address this challenge, AsphaltMine (<https://asphaltmine.org>) is being developed as a comprehensive web application platform dedicated to advancing data-driven pavement engineering. It allows users to collect asphalt test results in an organized online database, which can later be used to train machine learning models for predicting asphalt performance. The platform enables laboratories to contribute and retrieve unified laboratory test data for different tests, currently including the Rutting test, Marshall test, ITS, TSRST, and dynamic stiffness test. In addition to mixture test results, users are requested to input the corresponding component material test results and material proportions. The schema for these tests follows the European test standards, allowing standardized parameters to be entered. Users can upload data via an interactive web form or through a unified Excel sheet using a bulk import tool. The results are ultimately stored in a NoSQL database. In this paper, we discuss the potential benefits and challenges of creating such a database as well as describe the system architecture and data model of AsphaltMine, including the curation process and record storage, record structures, and user record management functionalities. We then demonstrate the data analysis and prediction features using the first test result prediction tool that was developed using the collected test results.

Keywords: Database; Web Application; Asphalt Performance Tests; Machine Learning; Prediction Platform.

1 Introduction

The advancement of artificial intelligence, particularly in machine learning, has made material data repositories essential, as these collections provide the training data necessary to develop and improve the accuracy of models to reflect real world physical behavior [1,2]. In construction and infrastructure fields, several initiatives are trying to gather the data in organized datasets, such as Deep Foundation Load Test Database (DFLTD) [3], steel database [4,5], structural dynamics datasets [6], and concrete-related databases [7,8].

Similar to other fields, in pavement engineering there is a driving need for reliable laboratory test data on asphalt mixture performance. This is especially important for advanced mechanical asphalt test methods (sometimes referred to as performance-based test) since the use of these methods can accelerate the development of new, better

asphalt mixtures [9]. Likewise, the application of balanced mix design approaches that use performance-based test methods has been shown to allow designing mixtures that ensure long-term pavement performance even when using innovative component materials [10–12]. The downside of the performance-based test methods, including rutting, cracking, stiffness, fatigue, and moisture susceptibility tests, is the significant effort needed to perform the tests. Gathering the test result data in a database might allow the development of data-driven test result prediction models that enable reducing the needed testing effort while still allowing significant innovation and optimization of mixture recipes. According to the European asphalt pavement association EAPA [13], Europe alone produced over 202 million tons of hot and warm-mix asphalt in 2023. Due to the large amount of asphalt produced, even small improvements in the

*Corresponding Author: E-mail: martins.zaumanis@empa.ch

material performance can lead to significant reductions in cost, resource consumption, and environmental impact.

The potential benefits of pavement databases have been recognized for many years. Several such databases exist globally, most notably the US Federal Highway Administration's Long-Term Pavement Performance (LTPP) [14], which maintains the world's largest pavement performance database, covering asphalt and concrete sections across both the US and Canada. A related project is the Info-Materials [15] that contains asphalt and concrete material research testing results. Similarly, Sweden has an LTPP monitoring project [16]. Another example is the Caltrans Accelerated Pavement Test (CAL/APT) program [17], which developed a relational database that includes laboratory and Heavy Vehicle Simulator (HVS) test results for asphalt and concrete, in addition to project and traffic data. Pavement Management Systems (PMS) databases track network condition indices and maintenance history, such as the US Massachusetts state PMS [18], the municipal street PMS in Sweden [19], and national PMS like in the US [20,21], in Sweden [22] and in Switzerland [23].

While the above-mentioned are clearly needed efforts that have brought many advantages, almost all these databases are closed and owned by a single entity. A crowdsourced database could allow accumulating a large amount of data that is typically not available to single laboratories. A large dataset would then enable the development of reliable data-driven prediction models. The use of unified European standards permits the creation of a cross-country database, thus enabling even small laboratories to contribute to and benefit from the creation of large datasets.

To enable the benefits that a large test result dataset can provide, we are developing AsphaltMine (<https://asphaltmine.org>) [24]. It is a web application for collecting asphalt mixture test results in a structured online database from laboratories that perform the tests according to the European standards. The data will enable training of predictive models for the test results based on mixture recipe and component material properties. AsphaltMine has the following key features:

- It is a crowdsourced database, which enables the combination of data from multiple contributors, resulting in the potential to develop a much larger overall dataset than any single laboratory could produce.
- It will host machine learning predictive tools for asphalt mixture test result prediction. This way, it can be used to virtually test several mixtures before performing physical testing in the laboratory. It will also allow updating mixture recipes when the properties of component materials change, such as in the case of using reclaimed asphalt.
- Once individual test result prediction tools are developed, it will enable creating a tool that allows users to define the desired properties of a mixture and back-calculate the required mixture recipe and the needed materials.

- It will allow users to visualize and statistically characterize the data available in the database.

We have also identified several challenges that arise when trying to gather data for AsphaltMine:

- Even if European Standards are used, users can still define many parameters and choose between multiple testing methods, making the comparison between the tests difficult.
- A wide variety of asphalt types and component material sources are used across Europe, which are not necessarily characterized to the extent that would enable development of accurate universal prediction models (this same problem exists for physics-based modeling of asphalt performance).
- There are large differences in how data is stored and managed in various laboratories, and often manual effort for moving the data and saving it in AsphaltMine is required.
- A certain quality of the uploaded data has to be ensured, which can be challenging considering that the data originates in different laboratories across many countries.
- The confidentiality and legal aspects can prevent laboratories from contributing data.

Some of these challenges can be addressed by increasing the size of the final dataset and performing a deep data analysis before using the data on a major scale. Other challenges could be overcome if practical benefits of the tools are demonstrated, thus justifying the effort required for collecting the data. We will discuss how to overcome the challenges in the paper.

This paper provides an in-depth description of AsphaltMine, including its purpose, structure, and capabilities. First, we describe the system architecture of AsphaltMine, including the data curation process and the internal XML storage. Next, we detail the record structure in AsphaltMine and the rationale for selecting particular performance tests. We then discuss user record management, highlighting the workspace feature that facilitates user collaboration on the platform, as well as other capabilities. Finally, we demonstrate the prediction platform of AsphaltMine and provide a general discussion about the potential benefits and challenges of creating the platform.

2 Objective

The objective of AsphaltMine is to enable the development of data-driven models for asphalt test result prediction. To achieve this, a web-based platform is developed for collecting asphalt test results along with the respective component material proportions and properties in an organized online database.

3 Database System Overview

AsphaltMine, as shown in Figure 1, is designed as an end-to-end pipeline from data collection to a prediction tool. It is a web-based platform where participants upload asphalt mixture composition, component materials, and mix test

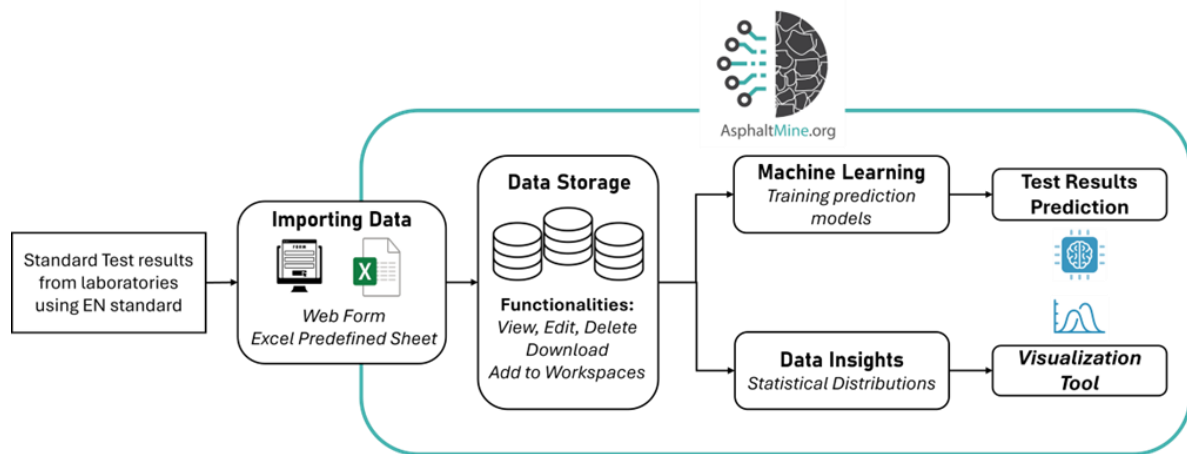


Figure 1. The AsphaltMine Web Application Overview.

results into a central database. This data can then be used to train machine learning models for predicting performance tests, followed by publishing user-friendly prediction tools on the platform. The data can also be used for visualization, co-working, and statistical analysis.

The backend of the platform is built starting from the open-source Configurable Data Curation System (CDCS) developed by NIST [25,26], which is a flexible framework for data repositories that supports custom schemas and collaboration. The AsphaltMine web application runs online using the open-source Django [27], which works with Python.

The data curation and internal data saving process are explained in more detail in the supplementary materials.

4 Records Overview

The XML records in AsphaltMine are based on the XSD files that define the permitted schema or structure of the record. This non-relational, XSD-based structure was selected

because it accommodates evolving standards, allowing the XSD to be updated and existing records to be migrated to a new version. It also provides flexibility, as records vary across materials and users can add custom properties without a fixed limit. The XSD for each test defines the record detailed tree structure for the test. In AsphaltMine, every single record has four main branches as shown in Figure 2, which are the Data Source, Mixture, Results, and Notes branches. More information about the definition of the branches and subbranches can be found in the supplementary materials.

The contributors of the data are acknowledged on the AsphaltMine website, and we request that data users acknowledge them also in any published materials where the data from AsphaltMine is used. The statistics on the number of test results in the database are provided and regularly updated on the website.

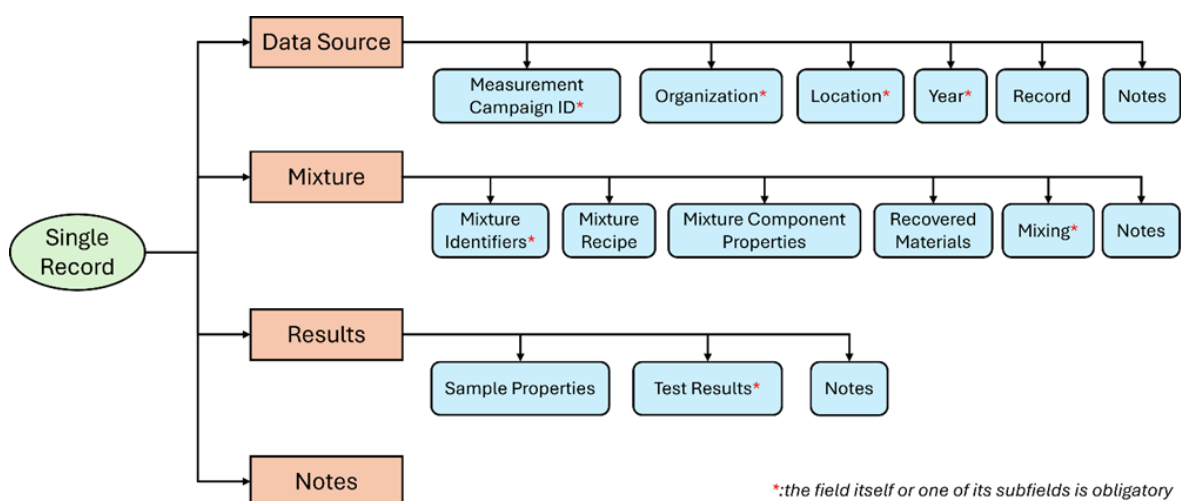


Figure 2. The tree structure of AsphaltMine database records, showing the first two levels.

4.1 Test Methods included in the AsphaltMine

AsphaltMine follows the European test standard methods, and results include fields for each of the items a particular standard requires to report. This approach allows unified data collection. One challenge lies in the fact that often European standards permit user-defined test conditions or provide a selection of alternative methods. These inconsistencies across Europe in some cases make it difficult to use the data for a single model. The performance tests currently included in AsphaltMine are as follows:

- Hamburg Wheel-Tracking Test according to standard EN 12697-22 [28].
- Marshall Test according to EN 12697-34 [29].
- Indirect Tensile Strength (ITS) according to EN 12697-23 [30].
- Thermal Stress Restrained Specimen Test (TSRST) according to EN 12697-46 [31].
- Uniaxial Tension Stress Test (UTST) according to EN 12697-46 [31].
- Stiffness Test according to EN 12697-26 [32].

As of September 2026, AsphaltMine included 3067 Marshall test records, 59 rutting records, 31 ITS records, and 24 stiffness test records. The live number of records and the detailed statistics per test are openly available in the visualization tool at <https://data.asphaltmine.org/visualization/>.

In the future, other test methods can be added based on the need and availability of test results by users. The priority for inclusion in AsphaltMine will be given to test methods for which a significant amount of test results is available and that require significant effort to perform the test (thus justifying the need for a prediction model). Since the database might need to be updated whenever a standard is revised, AsphaltMine's structure supports XSD modification. As a

feature, existing data is either migrated to the new schema or retained in its original format with a version number, depending on compatibility.

4.2 User's Record Management

Users can manage records in the database independently through different functionalities, such as creating workspaces, attaching files, editing records, or downloading them. AsphaltMine includes a quick start guide and a complete manual that guides users through all these functionalities. In addition, the platform supports workspaces, which are repositories that hold records. Workspaces are an important feature that allows teams and groups to share data in an organized manner, such as when collaborating on a project or sharing laboratory data among lab or group members. A chart of the AsphaltMine workspace with its components is shown in Figure 3. More details about the platform user functionalities can be found in the supplementary materials.

AsphaltMine provides three levels of user access:

- 1) Unregistered users: Unregistered users can use the visualization and prediction tools.
- 2) Registered users - read access: Any user can create an account through email verification and receive read access instantly. This allows users to inspect the database, filter by workspace, download openly available datasets or datasets for which they have permission.
- 3) Registered users - read and write access: Registered users can submit a request through their AsphaltMine accounts for additional write access, after which the access is granted. Registered users with write access can upload data.

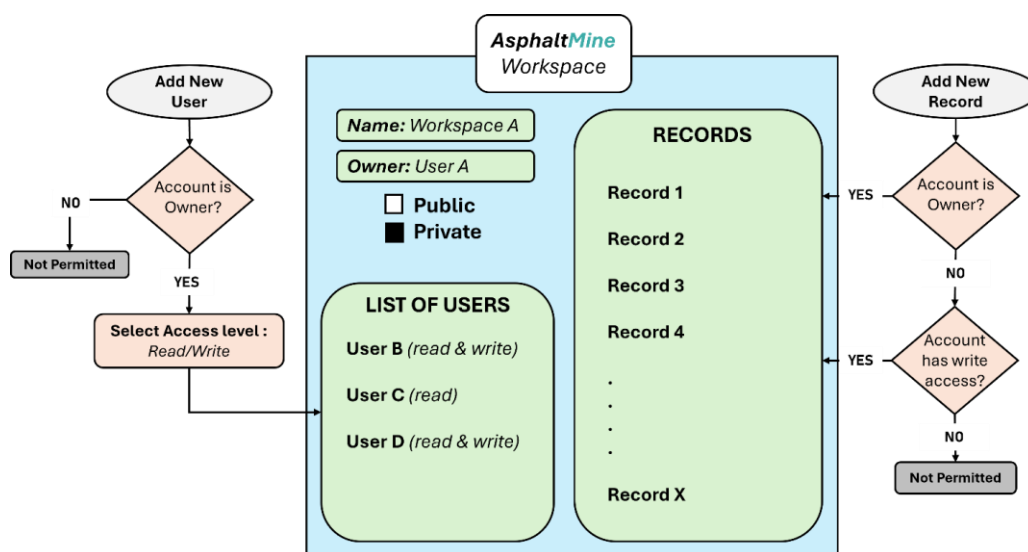


Figure 3. AsphaltMine Workspace structure and the processes of adding new users and records to it.

The tiered access measure provides a basic security and governance mechanism to prevent automated or unauthorized accounts from uploading inappropriate or low-quality data. Once write access is granted, contributors can upload their records and organize them within their own workspaces. The decision to use a dataset for a particular analysis remains the responsibility of the user, and access to a dataset should not be interpreted as independent third-party validation of its quality. Appropriate analysis of the data should be performed, which includes checking its consistency, suitability, variability and documentation for the intended application by the users.

5 AsphaltMine Platforms

The goal of the AsphaltMine platform is to use the gathered data to develop prediction tools for asphalt tests. These tools will be published on the AsphaltMine platform and will enable users to virtually test various asphalt mixture recipes. This process can enable optimizing recipes before performing a physical test in the laboratory. It can also help to adjust the recipe if the material properties change. The mixture tweaking can be especially beneficial when using reclaimed asphalt since the properties of this material change over time depending on the pavement that was milled.

Ultimately, we envision developing a tool that helps users design asphalt mixtures with the desired properties through optimization of the mixture recipes. By extension, this could lead to improved pavement performance, more efficient material use, and optimization of the recipes for sustainability.

5.1 Prediction Platform Description

AsphaltMine prediction platform contains "Performance Prediction" and "Mix Design" tabs as shown in Figure 4. In the "Performance Prediction" tab, each lab test is listed separately, allowing the user to choose the desired test. The "Mix Design" tab will include in the future the ability to select multiple desired properties that a certain mix design needs to achieve, for which the model will try to predict the necessary component materials and their proportions in a backward prediction method.

One example currently available in the AsphaltMine prediction tool is the Marshall test for asphalt, defined by EN

12697-34 [29]. This particular test was selected and included in this paper for first demonstration of the model capabilities because of the large amount of test results available. The full development of the model and its performance evaluation is described in [33].

In Marshall test, the output parameters are the Marshall Stability in kN representing the maximum load applied during compression, and Marshall Flow in mm representing the deformation at that maximum load. The model was built using a Physics-Informed Neural Network (PINN) approach, and it was trained on 2656 test results from Switzerland. The model considers Marshall test experimental reproducibility as defined in EN 12697-34 [29], giving the final prediction output for Marshall Stability and Flow as ranges rather than finite values. According to [33], the prediction interval coverage was 93% for both output parameters and the reproducibility-adjusted R^2 [34] was 0.88 for Marshall Stability and 0.55 for Marshall Flow on the test set, with very similar values obtained through cross-validation, as described in [33].

The input prediction form for Marshall Test in AsphaltMine is shown in Figure 5. The form states the source of the model at the top and the type of results required for the input parameters (materials recovered from the mixture). It shows the input fields needed for the model for prediction, divided into categories (volumetric, binder and composition). In addition, each input field has a tooltip that gives a brief description of the field, and a unit (°C, %, etc.).

It is important to note that although the amount of data used to train the model [33] was more than 2500 records, it was sourced solely from the Swiss Cantons of Zurich and Aargau. Thus, this model is expected to perform well for mixtures produced in these regions, as they share a similar pool of materials to those used in training. However, its performance will likely be worse when applied to mixtures from other regions in Europe, as those datasets may have significant differences in materials and their proportions. Expanding the size and including various material sources in the training set will result in better generalization and improvement of model performance. This is a key reason we encourage other European laboratories to contribute their data to AsphaltMine.

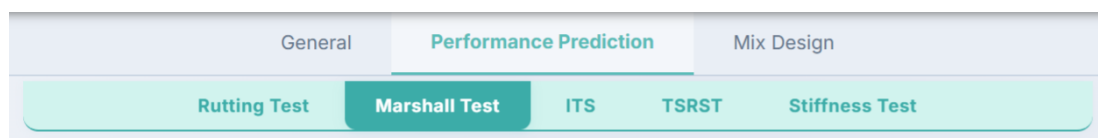


Figure 4. AsphaltMine prediction tool navigator.

General Performance Prediction Mix Design

Rutting Test **Marshall Test** ITS TSRST Stiffness Test

Marshall Test Performance Prediction

Input your parameters to predict Marshall Stability and Flow.
The parameters below are for materials recovered from the mixtures.
For the model scope, check [Assumptions & Conditions](#)
You can clear or randomly fill all fields by clicking the buttons below.

[Random Fill](#) [Clear](#)

VOLUMETRIC PROPERTIES

Maximum Density: Mg/m³

Bulk Density: Mg/m³

BINDER PROPERTIES

Penetration: 0.1mm

Softening Point: °C

COMPOSITION

Binder Content: %

Gradation:

0.063 mm	0.125 mm	0.25 mm	0.5 mm	1 mm	2 mm	4 mm
5.6 mm	8 mm	11.2 mm	16 mm	22.4 mm	31.5 mm	45 mm

[Predict](#)

Figure 5. AsphaltMine Marshall Test prediction form.

5.2 Marshall Test Prediction Examples

In this section, we demonstrate the AsphaltMine Marshall test prediction tool to predict four different test results based on mixture composition. The physical tests for these same mixtures have been performed in Switzerland and enable us to compare the predictions to the tested results. The provided cases serve to demonstrate the functionality of the prediction tool rather than for validating the developed model. For a full description of the model, the reader is referred to [33].

5.2.1 Samples

Four distinct Swiss samples were selected, including two originating from Canton Zurich (denoted Z1 and Z2) and two from Canton Aargau (denoted A1 and A2). The experimentally measured properties of the samples are shown in Table 1. Note that neither of the four samples has been used for training the model in [33].

5.2.2 Prediction Results

AsphaltMine prediction tool is used to predict Marshall Stability and Flow using the properties and composition provided in Table 1. These properties are filled in the form shown in Figure 5, and then the “Predict” button is clicked, after which AsphaltMine performs the calculation in less than 5 seconds and returns a summary of mixture properties and the Marshall test prediction results, as shown in Figure 6 for sample Z1.

As shown in Figure 6, the tool produces the prediction results both as a plot and as numerical values stating the prediction intervals for the Marshall Stability and Flow. Both formats consider the test reproducibility, based on the standard used (EN 12697-34 in this case). During prediction analysis, it is important to consider the test reproducibility, since it means two different labs can get two different results using the exact

Table 1. Marshall test lab results for four different plant-produced mixtures.

Parameters	Samples			
	Z1	Z2	A1	A2
Swiss Canton Source	Zurich	Zurich	Aargau	Aargau
Maximum Density (mg/m ³)	2.478	2.653	2.404	2.529
Bulk Density (mg/m ³)	2.335	2.430	2.075	2.399
Penetration (0.1mm)	46	37	48	23
Softening Point (°C)	76.6	56.8	61.6	69.9
Binder Content (%)	5.47	3.79	6.94	3.94
0.063 mm (%)*	10.5	8.7	12.0	8.0
0.125 mm (%)*	13.8	10.9	13.0	10.0
0.25 mm (%)*	18.1	13.6	13.5	13.2
0.5 mm (%)*	23.9	17.0	14.0	17.4
1 mm (%)*	34.9	21.7	15.4	23.1
2 mm (%)*	54.4	28.9	20.5	31.5
4 mm (%)*	77.5	41.5	95.5	43.4
5.6 mm (%)*	88.0	49.5	100	51.2
8 mm (%)*	98.5	59.6	100	59.6
11.2 mm (%)*	100	71.1	100	66.8
16 mm (%)*	100	84.7	100	84.3
22.4 mm (%)*	100	100	100	98.6
31.5 mm (%)*	100	100	100	100
Marshall Stability (kN)	14.7	12.0	6.5	20.5
Marshall Flow (mm)	3.6	2.6	2.2	2.7

*Percentage passing the corresponding sieve sizes

same sample. This makes the adoption of ranges rather than finite values more accurate to cover the possible values for different predictions.

For all samples, the AsphaltMine prediction results are listed in Table 2. In the table, reproducibility is demonstrated, which refers to the acceptable difference between results obtained for the same sample in two different laboratories. It is described by the standard deviation σ_R and the maximum allowable difference R . The provided example cases were selected to cover different situations and provided the following results:

- For the sample Z1, one can see the experimentally measured Marshall Stability (14.7 kN) and Flow (3.6 mm) are within the prediction limits of the model, as Marshall Stability is in the σ_R range while Marshall Flow is slightly outside σ_R range but within R range.
- For sample Z2, which has high maximum density compared to the records used for training in [33], the measured Marshall Stability (12 kN) and Flow (2.6 mm) were within R and the internal σ_R ranges, respectively.
- For sample A1, both measured Marshall Stability (6.5 kN) and Flow (2.2 mm) are within the internal σ_R range.
- For sample A2, the measured Marshall Stability (20.5 kN) is outside both σ_R and R predicted ranges, while the Marshall Flow (2.7 mm) is in the center of σ_R range. This is expected to a certain extent, since none of the samples that the model was trained on had a Marshall Stability higher than 20 kN [33]. Moreover, the model itself showed underestimated Marshall Stability in higher ranges, which could certainly affect the Marshall Stabilities above 20 kN. This behavior of the model is a

result of insufficient amount of data for higher Marshall Stability ranges, showcasing a need to add more data to AsphaltMine. More variable data would decrease the model bias and increase the prediction accuracy.

The provided four examples show that while the model (like any model) is not perfect and cannot replace laboratory testing, it can be used as a preliminary design tool and for the evaluation of the effect of various mix design parameters on the test results. It can be seen from the reproducibility parameters that the model functions best around the typical requirement range for the Marshall test, since in this range most results were available for training the model. At the same time, the provided examples, especially sample A2, highlight the need for the model user to understand how the models function and their limitations. To enable this, we always provide a link on the prediction tool webpage to documentation explaining the particular model.

5.3 Visualization Tool

The AsphaltMine visualization tool is available to any unregistered user. It provides an interactive overview of the datasets available in the database, as shown in Figure 7. It summarizes the distribution of records by test type, binder grade, year, mixing method, and key mixture and performance properties. Users can view mean values, uncertainty ranges, and record counts, and can select individual charts to inspect detailed distributions.

Upon registration, users gain access to filter data by workspace, mixture type, year, and binder grade, with further data visualizations planned for future development.

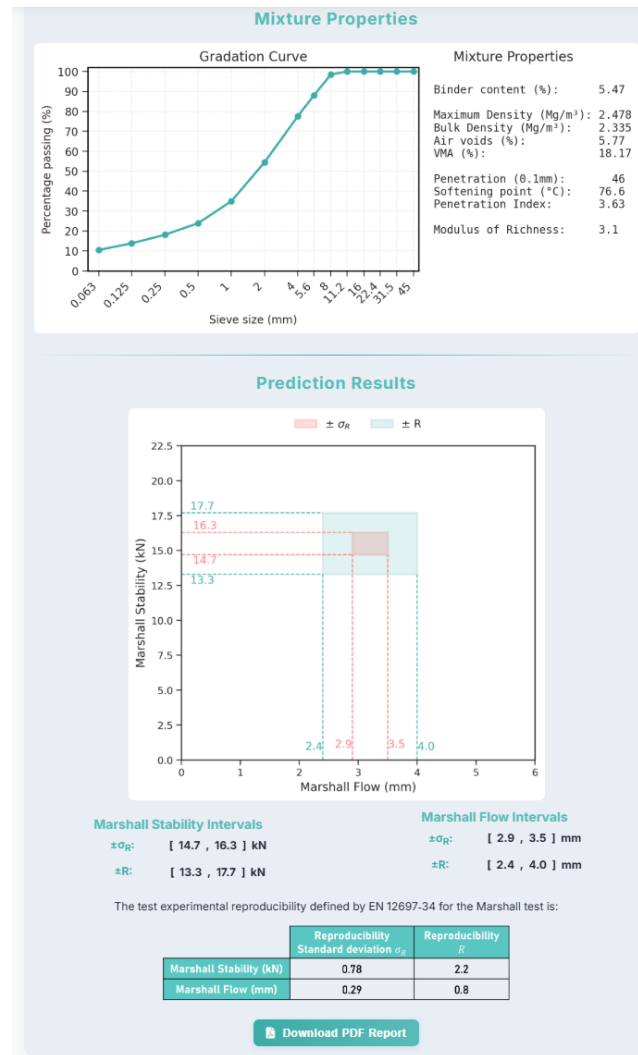


Figure 6. AsphaltMine Marshall Test prediction tool output for Sample Z1 showing the input mixture and the prediction results using EN 12697-34 reproducibility.

Table 2. AsphaltMine Marshall test final prediction output for samples Z1, Z2, A1 and A2.

Output	Marshall Stability (kN)				Marshall Flow (mm)			
Sample	Physical test result	Prediction Range		Tested result is within:	Physical test result	Prediction Range		Tested result is within:
		±σ _R	±R			±σ _R	±R	
Z1	14.7	14.7 to 16.3	13.3 to 17.7	σ _R	3.6	2.9 to 3.5	2.4 to 4.0	R
Z2	12.0	10.0 to 11.6	8.6 to 13.0	R	2.6	2.3 to 2.8	1.8 to 3.4	σ _R
A1	6.5	5.8 to 7.4	4.4 to 8.8	σ _R	2.2	1.7 to 2.3	1.2 to 2.8	σ _R
A2	20.5	17.4 to 18.9	16.0 to 20.4	None	2.7	2.5 to 3.1	2.0 to 3.6	σ _R

This allows users to focus on specific mixture groups or time periods and assess whether enough relevant data are available for a particular analysis. The visualizations of sieve gradation, binder content, RAP content, and other parameters also provide a rapid understanding of the composition and variability of the available mixtures. The visualization platform also includes a live correlation matrix

that reports both Pearson and Spearman correlation coefficients, together with the number of paired records used for each calculation. Selecting a matrix cell opens a scatter plot showing the detailed relationship between the corresponding variables, as illustrated in Figure 8.

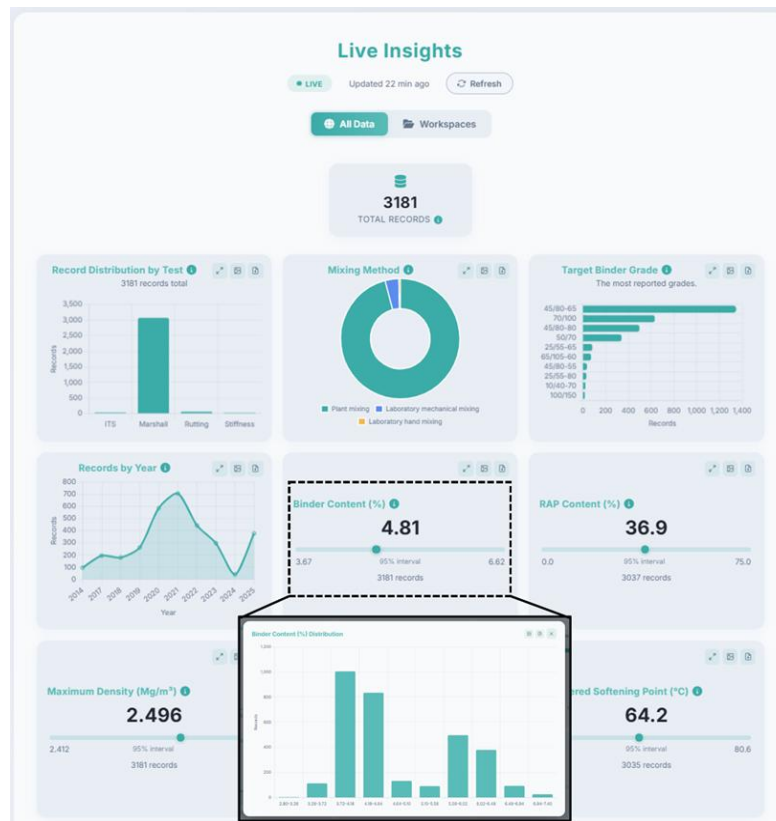


Figure 7. AsphaltMine visualization tool.

This helps users investigate meaningful relationships, distinguish general trends from possible outliers, and evaluate whether a correlation is supported by a sufficient number of records.

The main benefit of the visualization tool is that it enables users to evaluate the suitability of the data for a specific study. Researchers can identify which material types, test results, or regions are well represented and where important data gaps remain. This can help guide the selection of laboratory tests and materials, avoid analyses based on insufficient data, and support the preparation of reliable statistical and machine learning models. Thus, the platform is useful for experimental planning, numerical research, and data-based decision making by both researchers and asphalt industry stakeholders. An additional benefit for the stakeholders is the ability to visualize their own data, for example to evaluate the production consistency or annual variability.

6 Discussion

AsphaltMine is an initiative toward crowdsourced collaborative data collection that potentially can result in gathering enough data for developing tools that predict asphalt test results. It can be conceived that eventually the models help develop better performing, longer lasting, and more sustainable asphalt pavement materials.

6.1 Database

With a wide range of features, and more that could be added in the future, the platform serves as a data repository and a user-friendly tool for machine learning model use.

AsphaltMine allows users to manage and control the access rights to their records. We encourage users to make their data public, even if in an anonymized form, to benefit the research community. The data can be collected from quality control laboratories, asphalt plant operators, and researchers, facilitating the organization of their own data within the workspaces and helping them find additional data already available on AsphaltMine, which is useful for analysis and for comparison.

Another possibility to consider for the future is for the database to become a unified destination for data uploads for international organizations or research journals. This would support open research data principles. Making the data public and available in a unified format, like other data already shared in AsphaltMine [35], facilitates data reuse by the research community. The community-driven nature of the AsphaltMine project means that user feedback will guide its development.

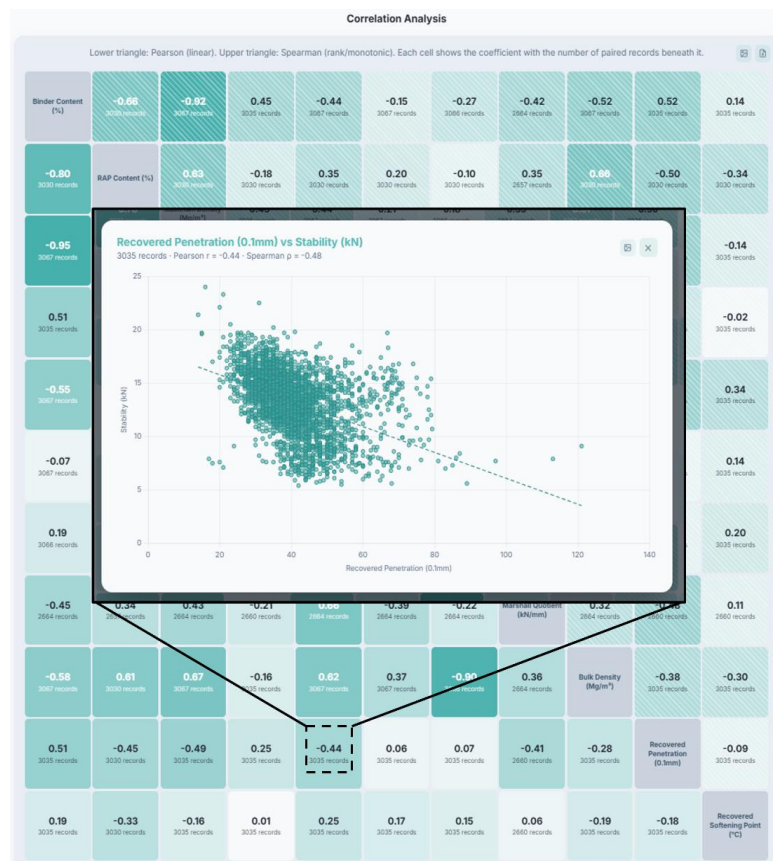


Figure 8. AsphaltMine correlation matrix in the visualization tool.

6.2 Prediction tools

The datasets gathered through AsphaltMine can be used to develop data-driven models that predict test results. Using a prediction tool can potentially aid quality control engineers, asphalt plant operators, and researchers in quickly optimizing mixture recipes, while considerably reducing the number of samples that need to be tested experimentally. The prediction models will need to be continuously updated as new data is added, thereby increasing their accuracy.

The applicability domain of the models will depend on the geographical and material diversity of the training data. Adding data from different regions and mixture types will progressively expand the range of conditions for which reliable predictions can be made.

Ultimately, a tool can be developed to potentially help optimize mixture recipes based on the desired test results. This would allow users to select the desired performance criteria for specific tests, such as the rutting test and TSRST, and then, through a back-calculation process, the tool would determine the optimal mix design for this set of properties, such as aggregate gradation, binder content, and overall mix composition. This would make the selection of the best mix design more efficient.

In the future, the test results prediction could be used to estimate pavement performance and service life. Furthermore, coupling the tools with an emission calculation

tool could allow developing more sustainable asphalt mixtures.

6.3 Challenges and potential solutions

Projects like AsphaltMine face uncertainty and various challenges. It is important to acknowledge them to enable implementing suitable solutions. A list of challenges and potential solutions is shown in Table 3.

7 Conclusions

This paper introduces the AsphaltMine web application, a novel infrastructure for asphalt pavement research in Europe. It provides a database for crowdsourcing asphalt mixture performance tests results that can then be used for developing data-driven test result prediction tools. We encourage wide participation in gathering and using the data.

In this paper, we first introduced the database system in AsphaltMine, its curation mechanism and the internal data saving process. After that, the record structure was detailed, showing the AsphaltMine XML schema definition, XSD and the European test standards used. Next, the user record management was described, including functionalities such as downloading records and workspace features. Finally, the AsphaltMine prediction platform was introduced, in which physical test results were compared to the output of the first developed prediction tool.

The following conclusions can be drawn:

- AsphaltMine addresses the pressing need for performance-based asphalt test data, by facilitating data sharing among laboratories. The initiative aims to accelerate the use of machine learning and AI in pavement engineering to accelerate knowledge gain.
- AsphaltMine encourages pavement researchers and labs to contribute their data to the database. As more participants join, the pool of data will grow, eventually enabling the development of accurate machine learning models that can predict asphalt behavior from mix designs.
- AsphaltMine prediction tools have the potential to reduce the resources needed for testing. By predicting test outcomes, the number of physical mixtures required to select a final performing mix is decreased, saving time and money.
- AsphaltMine visualization tools can help users rapidly assess data availability and quality, identify knowledge gaps, and select suitable datasets for experimental research, statistical analysis, and machine learning application. It also allows to rapidly analyze ones own data.
- AsphaltMine functionalities will continue developing and expanding, including incorporating additional features and collecting more data from various sources across Europe.
- AsphaltMine data structure is flexible and can be modified as needed to include new data fields or features without affecting previously stored records.
- AsphaltMine contributes to open science by encouraging data contributors to make data public and accessible for others to view, download, and analyze.

Table 3. AsphaltMine challenges and potential solutions.

Challenges	Potential Solutions
Confidentiality: Asphalt producers and road owners are often concerned about data privacy and ownership, which may discourage them from sharing mix designs and test results.	• AsphaltMine addresses this by allowing contributors to restrict access to their records, share with select partners, or make them public (for example, after paper publication). • Regardless of its publication status, all data is anonymized during the machine learning process, ensuring that only the model, not the raw data, is shared publicly. • AsphaltMine allows users to store their data in their affiliated entity, instead of the AsphaltMine server.
Data Heterogeneity: Despite standardization in Europe, differences in test procedures exist among contributors. Furthermore, data quality and trustworthiness must be ensured to enable training of reliable models. Finally, depending on the geographical location, and the locally available materials, the types of mixtures across the AsphaltMine data contributors vary significantly.	• Before using any dataset for training a model a deep statistical data analysis must be performed. This approach uncovers insights and trends, allows to identify any outliers or systematic errors stemming from a particular data contributor. Such analysis allows to ensure a certain quality of the data and guides the selection of the most appropriate model for predicting test performance. • The selection process may involve dividing the dataset into groups with different properties and developing separate models for each group. • An ensemble model may be implemented after detailed analysis and based on model accuracy. • The research is an ongoing process, with models being updated as new data is contributed. • The analyses behind each test performance prediction model published on AsphaltMine will be made public.
Manual Labor: There is a wide difference in how the data is stored and managed in the labs, requiring manual effort for data upload in AsphaltMine.	• In addition to the web form, AsphaltMine provides users with a predefined MS Excel spreadsheet designed mainly for users who already have their data stored in tabular format. This makes it easier for laboratories to copy data and then upload it to AsphaltMine. • It is possible to create custom plugins for specific users to automatically import data from multiple standard sheets into the predefined AsphaltMine spreadsheet. • Some manual labor will always be required initially (like converting PDF formats and gathering data). It is important to note that while uploading existing old data may require effort at the beginning, adding new data as it is generated will be much easier. • The data contributors are acknowledged on the AsphaltMine website, and we ask anyone using the data from AsphaltMine to acknowledge the data contributors in the publications. • The benefits of having accurate prediction tools should outweigh the invested effort, thus justifying the effort.

Authorship statement (CRediT)

Mohammad Abbas: Conceptualization, Methodology, Software, Visualization, Formal analysis, Investigation, Data Curation, Validation, Writing - Original Draft. **Martins Zauamanis:** Conceptualization, Methodology, Project administration, Supervision, Resources, Funding Acquisition, Data Curation, Writing - Original Draft

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Statement on the use of AI

During the preparation of this work the authors used Perplexity AI in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgments

The authors would like to thank Mateusz Wyrzykowski, Beat Münch, Michele Griffa and Marko Orešković for their extensive support during the planning of the database.

This work was supported by Swiss National Science Foundation (SNSF) grant No. 200021E_213163, under the WEAVE project No 213163. Project partners include the University of Antwerp, Belgium, funded by FWO, and TU Wien, Austria, funded by FWF.

References

- Himanen L., Geurts A., Foster A. S., Rinke P., Data-Driven Materials Science: Status, Challenges, and Perspectives, *Adv. Sci.*, Volume 6, Article 1900808, (2019).
<https://doi.org/10.1002/adv.201900808>
- Rumor L., Andrade-Campos A., On the need for material model databases: A state-of-the-art review, *Adv. Mech. Eng.*, Volume 14, Article 16878132221130575, (2022).
<https://doi.org/10.1177/16878132221130575>
- Kalavar S., Ealy C., FHWA Deep Foundation Load Test Database, New Technological and Design Developments in Deep Foundations, American Society of Civil Engineers, Denver, Colorado, United States, pp. 192–206, (2000).
[https://doi.org/10.1061/40511\(288\)14](https://doi.org/10.1061/40511(288)14)
- Skidopoulos A., de Castro e Sousa A., Lignos D. G., Experiments and proposed model for residual stresses in hot-rolled wide flange shapes, *J. Constr. Steel Res.*, Volume 210, Article 108069, (2023).
<https://doi.org/10.1016/j.jcsr.2023.108069>
- Skidopoulos A., de Castro e Sousa A., Lignos D. G., Database of Residual Stress Measurements on Hot-rolled Wide Flange Steel Cross Sections, (2023).
<https://doi.org/10.5281/zenodo.7677600>
- EPFL, Open Data, EPFL, n.d.
https://www.epfl.ch/labs/eesd/data_sets/
- Oland C. B., Ferraris C. C., Concrete Materials Database, NIST, Volume 22, pp. 28–33, (2000).
- National Concrete Pavement Technology Center, Concrete Infrastructure Research Database, n.d.
<https://www.cptechcenter.org/concrete-infrastructure-research-database/>
- Zauamanis M., Poulikakos L. D., Partl M. N., Performance-based design of asphalt mixtures and review of key parameters, *Mater. Des.*, Volume 141, pp. 185–201, (2018).
<https://doi.org/10.1016/j.matdes.2017.12.035>
- Li N., Hao P., Yao Y., Zhang C., The implementation of balanced mix design in asphalt materials: A review, *Constr. Build. Mater.*, Volume 402, Article 132919, (2023).
<https://doi.org/10.1016/j.conbuildmat.2023.132919>
- Mogawer W. S., Austerman A. J., Stuart K. D., Abdalfattah I. A., Zhou F., Comprehensive Balanced Mixture Design Protocol: Design Phase through Production Quality Assurance, *Transp. Res. Rec.*, Article 03611981231186422, (2023).
<https://doi.org/10.1177/03611981231186422>
- Sreedhar S., Coleri E., Obaid I. A., Kumar V., Development of a Balanced Mix Design Method in Oregon to Improve Long-Term Pavement Performance, *Transp. Res. Rec.*, Volume 2675, pp. 1121–1137, (2021).
<https://doi.org/10.1177/03611981211032222>
- EAPA, Asphalt in Figures – Key Figures of the European Asphalt Industry, (2025).
- Afsharikia Z., Groeger J., Gardner L., Ostrom B., Rada G., Development and Use of the LTPP Analysis-Ready Materials Dataset, (2023).
- FHWA, InfoMaterials, Federal Highway Administration, (2025).
<https://infomaterials.fhwa.dot.gov/>
- Göransson N.-G., Uppföljning av vägars tekniska tillstånd: Lägesrapport för observationssträckor ingående i det svenska LTPP-projektet till och med december 2019, Statens väg- och transportforskningsinstitut, (2020).
- Roesler J., Monismith C. L., Harvey J. T., Coetzee N. F., Caltrans Accelerated Pavement Test (CAL/APT) Program: Summary Report Six Year Period: 1994–2000, (2000).
- MassDOT, Open Data Portal, (2025).
<https://geodot.mass.gov/pages/open-data-portal/>
- Afridi M. A., Erlingsson S., Sjögren L., Municipal Street Pavement Management Systems in Sweden, *Springer Nature*, pp. 437–446, (2024).
- National Academies of Sciences, Engineering, and Medicine, Pavement Management Systems: Putting Data to Work, The National Academies Press, Washington, DC, (2017).
<https://doi.org/10.17226/24682>
- FHWA (Federal Highway Administration), The Highway Performance Monitoring System (HPMS), (2021).
<https://www.fhwa.dot.gov/policyinformation/hpms.cfm>
- PMSv4, PMSv4, Trafikverket, (2025).
<https://pmsv4.trafikverket.se/>
- OFROU, TRA – Trasse, n.d.
<https://www.astra.admin.ch/astra/fr/home/fachleute/weitere-bereiche/fachanwendungen/tra.html>
- AsphaltMine, AsphaltMine, (2025).
<https://asphaltmine.org/>
- Dima A., Bhaskarla S., Becker C., Brady M., Campbell C., Dessauw P., et al., Informatics Infrastructure for the Materials Genome Initiative, *JOM*, Volume 68, pp. 2053–2064, (2016).
<https://doi.org/10.1007/s11837-016-2000-4>
- Brady K., The Configurable Data Curation System (CDCS), National Institute of Standards and Technology, Gaithersburg, Maryland, (2018).
<https://www.nist.gov/itl/ssd/information-systems-group/configurable-data-curation-system-cdcs>
- Django, Django, (2025).
<https://www.djangoproject.com/>
- EN 12697-22:2020 – Bituminous mixtures – Test methods – Part 22: Wheel tracking, iTeh Standards, (2020).
<https://standards.iteh.ai/catalog/standards/cen/4442b66e-a324-4503-880a-ee4342064693/en-12697-22-2020>
- EN 12697-34:2020 – Bituminous mixtures – Test methods – Part 34: Marshall test, (2020).
- EN 12697-23:2017 – Bituminous mixtures – Test methods – Part 23: Determination of the indirect tensile strength of bituminous specimens, iTeh Standards, (2017).
<https://standards.iteh.ai/catalog/standards/cen/4920c3b0-a083-4d95-acb7-1c9a3fc1e0c5/en-12697-23-2017>
- EN 12697-46:2020 – Bituminous mixtures – Test methods – Part 46: Low temperature cracking and properties by uniaxial tension tests, iTeh Standards, (2020).
<https://standards.iteh.ai/catalog/standards/cen/ef35c3b0-17de-4666-acaf-b4a17ed38ecc/en-12697-46-2020>
- EN 12697-26:2012 – Bituminous mixtures – Test methods for hot mix asphalt – Part 26: Stiffness, iTeh Standards, (2012).
<https://standards.iteh.ai/catalog/standards/cen/6e19a3f7-d3aa-4906-b631-882b28846217/en-12697-26-2012>

-
- [33] Abbas M., Zaumanis M., Methodology for Physics-Informed Neural Networks for materials test result prediction: Case study on the Marshall test for asphalt, *Case Stud. Constr. Mater.*, Volume 24, Article e05829, (2026).
<https://doi.org/10.1016/j.cscm.2026.e05829>
- [34] Abbas M., Zaumanis M., Uncertainty-aware metrics for evaluating machine learning regression models in materials testing, *Measurement*, Volume 290, Article 122931, (2026).
<https://doi.org/10.1016/j.measurement.2026.122931>
- [35] Abbas M., Zaumanis M., Paperna O., Using machine learning and statistics for mapping BTSV boundaries from penetration and softening point requirements, *Int. J. Pavement Eng.*, Volume 26, Article 2571739, (2025).
<https://doi.org/10.1080/10298436.2025.2571739>