Machine learning-aided rheological prediction models of asphalt binders based on chemical properties

F. Zhang[1], D. Wang[1,2], Y. Sun[1], A.C. Falchetto[1,3]
[1]*Department of Civil Engineering, Aalto University, Finland*
[2]*Department of Civil Engineering, University of Ottawa, Canada*
[3]*Department of Civil Environmental and Architectural Engineering, University of Padova, Italy*

ABSTRACT: This work aims to provide rapid rheological characterization of asphalt binders through their chemical properties based on advanced machine learning tools. With this objective, Fourier transform infrared spectroscopy (FTIR) and dynamic shear rheometer (DSR) are adopted to measure the chemical and rheological properties. Results indicate that the raw six FTIR features can be reduced to two principal components (PC 1 and PC 2), and the variance and role of PC 1 are more significant than PC 2. Multiple linear regression models can predict the phase angle accurately but not for modulus. Gaussian process regression model with higher R^2 and lower RMSE values can accurately predict both modulus and phase angle.

Keywords: asphalt binders; rheological properties; chemical composition; machine learning

1 INTRODUCTION

Rheology of asphalt binders is the study of their flow and deformation behavior under various temperature and loading conditions (Zhang et al., 2024). It is a crucial aspect of understanding the performance characteristics of asphalt materials used in pavements. Asphalt binders exhibit viscoelastic properties, meaning they show both viscous (liquid-like) and elastic (solid-like) responses depending on the temperature and rate of loading. At high temperatures, they behave more like viscous fluids, while at low temperatures, they act more like elastic solids. The rheological properties of asphalt binders directly influence the durability and performance of road pavements, impacting resistance to deformation (rutting), cracking, and fatigue. Accurate rheological analysis helps in selecting and modifying binders to meet specific climatic and traffic demands, ensuring better long-term performance of asphalt pavements.

Measuring the rheology of asphalt binders can be a time-consuming process due to the complex and detailed analyses required to fully characterize their viscoelastic behavior. Rheological testing often involves conducting multiple assessments, such as dynamic shear rheometer (DSR) tests and bending beam rheometer (BBR) tests, across a range of temperatures and loading frequencies. These tests are designed to simulate the real-world performance of asphalt binders under different traffic and climatic conditions, which requires precise sample preparation, conditioning, and repeated measurements to ensure reliability. Additionally, each test can take significant time due to the need for temperature equilibration, application of controlled stress or strain, and data collection. The time investment is necessary to capture the binder's behavior over short-term (high traffic speed) and long-term (slow-moving or stationary loads) performance. While these processes provide invaluable insights into the binder's potential durability and suitability for specific applications, they require a considerable commitment of time and resources (Wang et al., 2022).

The chemical properties of asphalt binders play a fundamental role in determining their rheological behavior and overall performance (Shan et al., 2023). Asphalt binders are composed of complex mixtures of hydrocarbons, including asphaltenes, resins, saturates, and aromatics (Salehfard et al., 2024), each contributing differently to their physical characteristics. The balance between these chemical constituents influences the binder's response to temperature changes and mechanical stress. For instance, asphaltenes contribute to the stiffness and elasticity of the binder (Ilyin and Yadykova, 2024), while lighter fractions, such as saturates and aromatics, provide fluidity and flexibility (Shan et al., 2024). The interactions between these components affect the viscoelastic nature of the binder, dictating how it responds to high temperatures (resisting rutting) and low temperatures (resisting cracking). Understanding the chemical composition and interactions within asphalt binders is essential for modifying their properties to achieve

desired performance characteristics, especially under varying traffic loads and environmental conditions.

Machine learning (ML) offers superior predictive accuracy over traditional statistical models like regression, which struggle with nonlinear interactions in asphalt characterization. ML techniques, including decision trees and neural networks, identify key chemical parameters influencing rheology, enhancing model interpretability and efficiency. The integration of ML with Fourier-transform infrared (FTIR) spectroscopy marks a significant advancement in asphalt binder analysis. FTIR identifies chemical compositions, while ML algorithms establish correlations between these compositions and rheological behavior, enabling predictive models that reduce reliance on extensive physical testing. This approach improves efficiency in predicting critical properties like stiffness and temperature susceptibility, facilitating optimized binder formulation and material selection. Despite prior reliance on empirical models, limited studies have systematically applied ML to predict asphalt rheology using chemical composition data. This study benchmarks ML against conventional models, demonstrating its advantages in data-driven pavement design.

2 MATERIALS AND METHODS

2.1 *Materials*

The materials used in this research include 18 fresh asphalt binders. In order to improve the applicability of the model, these binders are provided by different suppliers with different penetrations.

2.2 *DSR tests*

In this research, the frequency sweep of asphalt binders is performed via a DSR machine. The frequency is selected from 0.1 Hz to 20 Hz. The tests are carried out at three different temperature ranges: low-temperature ranges (-30 °C to 6 °C), mid-temperature ranges (-4 °C to 40 °C), and high-temperature ranges (28 °C to 76 °C). The corresponding sample plate is 25 mm in diameter for high temperatures, 8 mm in diameter for medium temperatures, and 4 mm in diameter for low temperatures.

2.3 *FTIR tests*

The FTIR test was employed to measure the chemical composition of different asphalt binders. The measuring wave numbers are from 400 cm^{-1} to 4000 cm^{-1}. FTIR can identify chemical composition by determining absorbance from input and transmitted light. It also can quickly and non-destructively characterize the chemical composition of materials.

2.4 *Machine learning models*

Principal component analysis (PCA) is a method to reduce the dimensionality of the data. PCA can avoid the limitations of over-counting by finding a new coordinate system and reducing the dimensionality from n to k ($k < n$).

Since the multiple linear regression (MLR) model has the advantage of simplicity and practicality, it is used in this paper to predict the rheological properties of asphalt binders. In this model, the number of independent variables is greater than or equal to two; hence, it is called multiple regression. In fact, the combination of numerous independent variables is more accurate and realistic than single variable prediction (Hu et al., 2015).

Gaussian Process Regression (GPR) model is a more complex and powerful machine learning model that can be used to represent the distribution of the function. GPR can establish the model with infinite dimensions (Schulz et al., 2018). Each input point is associated with a random variable, and their joint distribution can be modeled as multiple Gaussian distributions. This model is a non-parametric, Bayesian, supervised learning method (Schulz et al., 2018).

3 RESULTS AND DISCUSSION

3.1 *Principal component analysis*

The difference in chemical bond areas can distinguish different binders in this research. Considering the contribution of all of these as input variables would make the calculation much more complex. PCA method is applied to reduce the dimensionality. It can be found that the dimensionality of binders from FTIR results is six, so the extracted original principal component number is six. The eigenvalues and cumulative variance, as shown in Figure 1, can be generated after PCA analysis. A larger eigenvalue indicates that the principal component it represents has the largest contribution among all principal components. It can be observed that the first two principal components have the largest eigenvalues of 3.03 and 1.24. The results of PCA can be considered acceptable when the cumulative variance of the principal components is greater than 60% (Margaritis et al., 2020, Paranhos and Petter, 2013). The cumulative variance of the first two principal components in Figure 2 is 71.0% (above 60%); therefore, PC 1 and PC 2 can represent all original features of binders. The description of PCs is shown in Equations 1 and 2. It can be found that all original features are positively correlated with the distribution of PC 1, with the vibration of C-H and CH_2 having the highest correlation and the vibration of S=O having the lowest correlation. However, for PC 2, the vibration of S=O has the highest correlation, while C-H has the lowest value. The vibrations of S=O, CH3, and CH2 positively correlate

with PC 2, and vibrations of C-H, CH_2/CH_3, and C=C/C=O have a negative correlation.

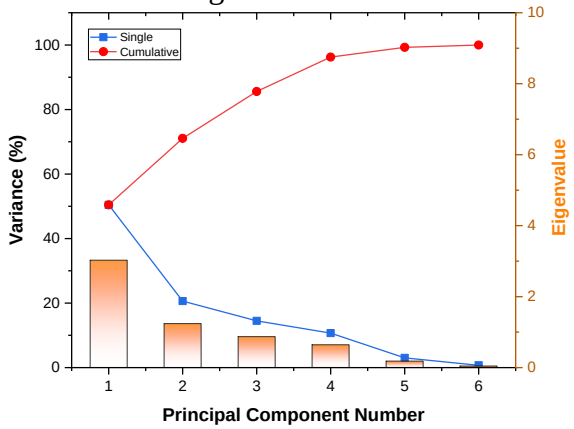

Figure 1. Variance and eigenvalue of each component.

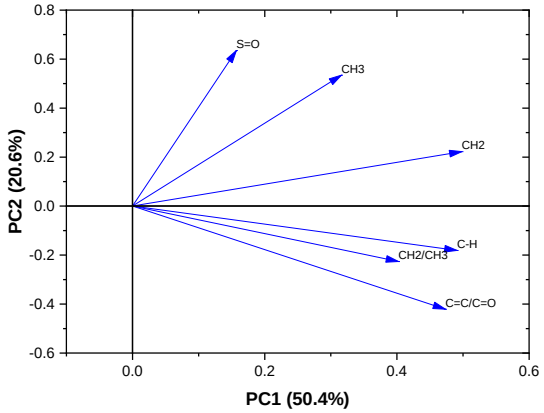

Figure 2. Projection of the original features onto the system.

$$PC\ 1 = 0.491(C-H) + 0.157(S=O) + 0.316(CH_3) + 0.498(CH_2) + 0.473(C=C \text{ or } C=O) + 0.402(CH_2 \text{ or } CH_3) \quad (1)$$

$$PC\ 2 = -0.181(C-H) + 0.635(S=O) + 0.534(CH_3) + 0.222(CH_2) - 0.421(C=C \text{ or } C=O) - 0.227(CH_2 \text{ or } CH_3) \quad (2)$$

3.2 *Development of prediction model*

In this section, PC1, PC2, temperature, and frequency were selected as input variables, ensuring that both compositional and environmental factors were considered in the prediction. The dataset was split into training (70%) and validation (30%) sets. Performance Evaluation: The predictive accuracy of the model was assessed using R^2 and RMSE, demonstrating the effectiveness of GPR in capturing complex relationships between chemical composition, temperature, and rheological behavior.

3.2.1 *Multiple linear regression*

MLR, as the simplest regression model, has the advantage of fast computation speed with low computer configurations. Figure 3 is the predicted modulus of binders through the MLR model. A total of 7200 data by simultaneously considering temperature and frequency. However, the model is unsatisfactory for modulus prediction, with the R^2 of 0.5488 and 0.5447 and RMSE of 11.12. Hence, it is necessary to apply a more advanced and powerful model to predict modulus. Figure 4 is the δ results of the binders through the MLR model. The δ prediction is more accurate than the modulus. The R^2 of the training and testing set is above 0.9, and small RMSE values in Figure 4 reveal that the prediction of δ is reliable.

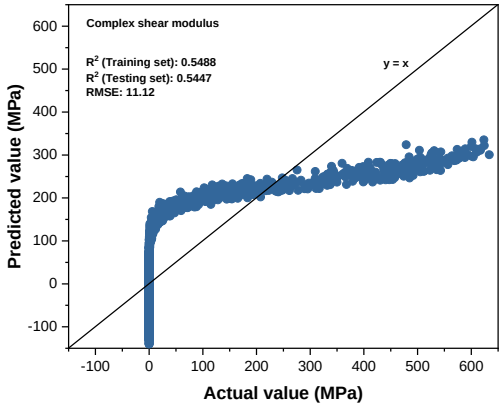

Figure 3. MLR model of |G*|.

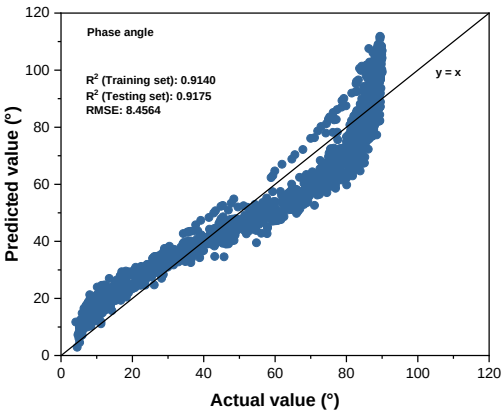

Figure 4. MLR model of δ.

3.2.2 *Gaussian Process Regression*

Since MLR cannot provide sufficiently accurate prediction $|G^*|$, a more advanced GPR model was introduced to achieve this purpose. Figure 5 shows the GPR predictions of binders affected by temperature and frequency. It showed a promising correlation in Fig. 5. The R^2 values are above 0.98, and the RMSE is 1.9370 in this model. Regarding δ, the GPR model also exhibited desirable results in Figure 6. The predicted values in Figure 6 are almost equal to the actual values; its R^2 is above 0.98, and the RMSE value is 3.1672. This indicates that the GPR model can achieve the purpose of predicting δ.

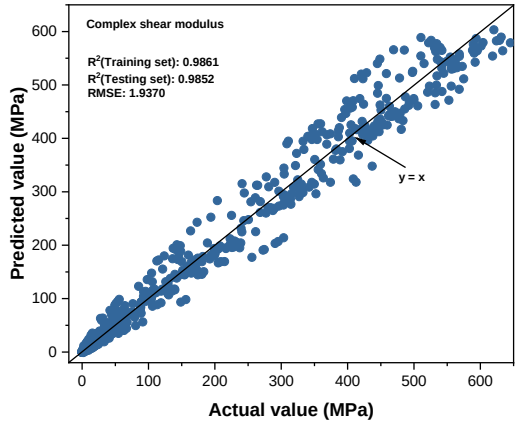

Figure 5. GPR model of |G*|.

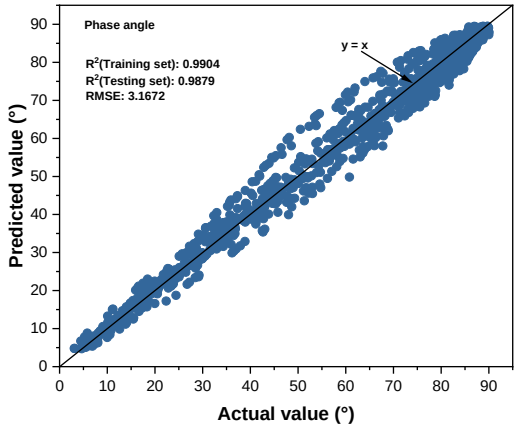

Figure 6. GPR model of δ.

3.3 *Discussion*

The sample type and dataset are not enough. Thus, in the next phase of this research, it is recommended that the analysis and the prediction models be extended to modified asphalt binders with a big dataset, especially those with chemical modification. In addition, test of additional regression models, such as Ridge Regression, Support Vector Regression, or Random Forest Regression, are needed to ensure the best model.

4 CONCLUSIONS

- The original chemical features of binders can be reduced to two principal components.
- The multiple linear regression (MLR) model can predict the phase angle of binders but not the modulus.
- On the other hand, the Gaussian process regression (GPR) model enables the prediction of both modulus and phase angle.

The neat asphalt binders in this work have the same type of chemical composition; only the concentration is different. Thus, in the next phase of this research, it is recommended that the analysis and the prediction models be extended to modified asphalt binders (more chemical composition types), especially those with chemical modification.

REFERENCES

Hu, X., Madden, L. V., Edwards, S. & Xu, X. (**2015**). Combining models is more likely to give better predictions than single models. *Phytopathology*, 105**,** 1174-1182.

Ilyin, S. O. & Yadykova, A. Y. (**2024**). Eco-friendly bitumen binders from heavy crude oil and a relaxation approach to predicting their resistance to rutting and cracking. *Journal of Cleaner Production*, 434**,** 139942.

Margaritis, A., Soenen, H., Fransen, E., Pipintakos, G., Jacobs, G. & Blom, J. (**2020**). Identification of ageing state clusters of reclaimed asphalt binders using principal component analysis (PCA) and hierarchical cluster analysis (HCA) based on chemo-rheological parameters. *Construction and Building Materials*, 244**,** 118276.

Paranhos, R. S. & Petter, C. O. (**2013**). Multivariate data analysis applied in Hot-Mix asphalt plants. *Resources, conservation and recycling*, 73**,** 1-10.

Salehfard, R., Yeganeh, S., Dalmazzo, D., Underwood, B. S. & Santagata, E. (**2024**). Linking Chemical Structure to the Linear and Nonlinear Properties of Asphalt Binders. *Transportation Research Record***,** 03611981241244793.

Schulz, E., Speekenbrink, M. & Krause, A. (**2018**). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of mathematical psychology*, 85**,** 1-16.

Shan, B., Cao, X., Hao, Z., Wu, Y., Yang, X. & Li, X. (**2024**). Experimental and Molecular Dynamics Simulations Investigating the Flexibility and Compatibility of Epoxidized Soybean Oil–Reinforced Epoxy Asphalt. *Journal of Materials in Civil Engineering*, 36**,** 04024327.

Shan, L., Wang, Y., Liu, S., Qi, X. & Wang, J. (**2023**). Establishment of correlation model between compositions and dynamic viscoelastic properties of asphalt binder based on machine learning. *Construction and Building Materials*, 364**,** 129902.

Wang, D., Baliello, A., Poulikakos, L., Vasconcelos, K., Kakar, M. R., Giancontieri, G., Pasquini, E., Porot, L., Tušar, M. & Riccardi, C. (**2022**). Rheological properties of asphalt binder modified with waste polyethylene: An interlaboratory research from the RILEM TC WMR. *Resources, Conservation and Recycling*, 186**,** 106564.

Zhang, F., Cannone Falchetto, A., Yuan, D., Wang, W., Wang, D. & Sun, Y. (**2024**). Research on performance variations of different asphalt binders results from microwave heating during freeze-thaw cycles. *Construction and Building Materials*, 448**,** 138280.