

**Chemo-rheological characterization of bitumen ageing based on multivariate analysis combined with variable selection methods**

Ma, L.; Varveri, A.; Erkens, S.

**DOI**

[10.1201/9781003402541-13](https://doi.org/10.1201/9781003402541-13)

**Publication date**

2024

**Document Version**

Final published version

**Published in**

Bituminous Mixtures and Pavements VIII

**Citation (APA)**

Ma, L., Varveri, A., & Erkens, S. (2024). Chemo-rheological characterization of bitumen ageing based on multivariate analysis combined with variable selection methods. In *Bituminous Mixtures and Pavements VIII* (pp. 106-113). CRC Press. <https://doi.org/10.1201/9781003402541-13>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)  
as part of the Taverne amendment.**

More information about this copyright law amendment  
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:  
the publisher is the copyright holder of this work and the  
author uses the Dutch legislation to make this work public.

# Chemo-rheological characterization of bitumen ageing based on multivariate analysis combined with variable selection methods

L. Ma, A. Varveri & S. Erkens

*Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands*

**ABSTRACT:** This study aims to correlate the chemical and rheological properties of bitumen at different ageing states and understand the chemical mechanisms of bitumen degradation due to ageing. The relationship between Fourier transform infrared (FTIR) spectral data and rheological results is investigated using partial least squares (PLS) regression integrated with two variable selection methods. The spectral region of  $1800 - 800 \text{ cm}^{-1}$  is identified as the most informative for accurate estimation of the rheological properties of bitumen. Variable selection methods, particularly moving windows (MW), improve the prediction accuracy of the regression models.

## 1 INTRODUCTION

Oxidative ageing is one of the main contributors for pavement degradation (Petersen 2009). Oxygen in the atmosphere can diffuse into bitumen film (Ma *et al.* 2021). The presence of oxygen within bitumen, in combination with high temperature and/or UV radiation, leads to oxidative ageing and causes an increase in stiffness and brittleness of bitumen, as well as a decrease in adhesion between bitumen and aggregates (Aguiar-Moya *et al.* 2015).

Many studies have been focusing on the evaluation of ageing kinetics and their effects on bitumen properties (Herrington 2012; Tauste *et al.* 2018). Various test protocols, parameters, and indices are proposed to describe the chemical changes during the ageing process and the resulting changes in the physical, rheological, and mechanical properties. Ageing effects can be integrated into the performance modeling and analysis of asphalt pavement to provide more accurate evaluation, modeling, and prediction of its long-term performance. This requires correlating the ageing-caused chemical changes to the mechanical degradation of bitumen, and further to the deterioration of pavement performance (Nobakht *et al.* 2020; Pipintakos *et al.* 2022; Redelius and Soenen 2015).

A single evaluation index has been demonstrated to be inadequate in evaluating the chemical and mechanical properties of bitumen (Pipintakos *et al.* 2022), as well as the chemo-mechanical relationship, especially when targeting at the generalization of all types of bitumen. Bitumen from different crude oil sources exhibits significant differences in their chemical components and micro-structures (Petersen 2009). The evolving use of polymer modifiers, rejuvenators, and additives further adds to this complexity. Instead of solely relying on a single index, the combination of multiple indices seems to be a promising approach to study the chemo-mechanical relationship in a general way. Despite all the efforts in chemo-mechanical characterization of aged bitumen (Ma *et al.* 2021; Tauste *et al.* 2018; Weigel and Stephan 2017). More work is still needed to achieve an accurate prediction of the mechanical properties of bitumen and pavement performance due to the complicated chemical composition of bitumen from different sources, the viscoelastic properties, the

interactions with aggregates, as well as their physicochemical changes due to environmental factors.

To obtain more accurate evaluation of the chemo-mechanical relationship, except for deriving multiple parameters from different tests, it is possible to utilize more variables from one test with promising quality. Fourier transform infrared spectroscopy (FTIR) can quickly and effectively identify the multiple chemical components of bitumen, which makes it an important tool in evaluating chemical properties of bituminous materials (Jing *et al.* 2019; Mirwald *et al.* 2020). Specifically, sulfoxide peaks at  $1030\text{ cm}^{-1}$  and carbonyl peaks at  $1700\text{ cm}^{-1}$  are two main parameters used to evaluate ageing levels of bitumen. However, additional peaks and spectral regions in the FTIR spectra are also informative of the chemical changes in bitumen due to ageing. To analyze the complete FTIR spectra, advanced tools such as principal component analysis (PCA), linear discriminant analysis (LDA), and partial least squares (PLS) are needed to deal with high-dimensional datasets. Studies have been conducted to identify bitumen types and ageing states (Goosen and Jenkins 2022; Ma *et al.* 2023), and to predict its SARA fractions and rheological properties (Aske *et al.* 2001; Siroma *et al.* 2021, Mohammadi *et al.* 2021; Weigel and Stephan 2017) based on multivariate analysis. With the increasing number of variables, numerous variable selection techniques such as genetic algorithm (GA), interval PLS, and moving windows (MW) have been developed to improve the quality of multivariate analysis (Mehmood *et al.* 2012; Mohammadi *et al.* 2021; Yun *et al.* 2019).

This paper aims to predict rheological properties of bituminous binders based on PLS regression of FTIR spectral data, and to understand the chemical mechanisms underlying the degradation of bitumen properties during ageing process. Variable selection approaches were utilized to improve the prediction accuracy and to capture important chemical components.

## 2 MATERIALS AND METHODS

### 2.1 Materials

To investigate the chemo-rheological relationship across a wide range, bituminous binders derived from three different crude oil sources and with varying penetration grades (PEN) were applied. Additionally, some of these were modified with Styrene-Butadiene-Styrene (SBS). In total, eight different types of binders were prepared. Specifically, binders Q460 (PEN 40/60), Q710 (PEN 70/100), and QPMB (modified by SBS) were obtained from source Q; T710 (PEN 70/100), T1015 (PEN 100/150), and T1622 (PEN 160/220) were derived from source T; and V710 (PEN 70/100) and VPMB (modified by SBS) were sourced from V.

Each sample underwent four distinct ageing processes: one short-term oven ageing and three long-term Pressure Ageing Vessel (PAV) ageing for durations of 20, 40, and 80 hours. Initially, samples of each bitumen type were heated to  $140^{\circ}\text{C}$  for five minutes and was then distributed among four 140mm-diameter containers, each containing precisely  $50.0 \pm 0.5$  grams of bitumen. These containers were then put in a  $140^{\circ}\text{C}$  oven for an additional five minutes to form a uniform 3.2 mm film. Short-term ageing followed the thin film oven test (TFOT) protocol, involving five hours in an oven set at  $163^{\circ}\text{C}$ . Based on short-term ageing, the three long-term ageing conditions were conducted in the PAV at 2.1 MPa and  $100^{\circ}\text{C}$  for the specified durations.

### 2.2 Test methods

FTIR was utilized to measure the chemical properties of all binder samples. Three separate measurements were performed for each specimen. The measured wavenumber spanned from  $600\text{ cm}^{-1}$  to  $4000\text{ cm}^{-1}$ , with a resolution of  $1\text{ cm}^{-1}$ , and 32 scans were used for each test. The spectral bands between  $2400\text{--}1900\text{ cm}^{-1}$  and  $4000\text{--}3700\text{ cm}^{-1}$  were excluded due to

their limited chemical information. Readers are referred to Ref. (Ma *et al.* 2023) for detailed discussion of test results.

The frequency sweep test with strain-controlled mode was carried out at temperatures of 0–80°C. Different plate sizes and gap distances were utilized based on the testing temperatures: an 8 mm plate with a 2 mm gap was used for temperatures between 0 and 40°C, while a 20 mm plate with a 1 mm gap was used for temperatures ranging from 50 to 80°C. Before the frequency sweep, an amplitude sweep was conducted at both 0°C and 50°C to identify the maximum strain in the linear viscoelastic (LVE) range. The chosen strain amplitudes were 0.1% for temperatures ranging from 0 to 40°C, and 0.5% for temperatures between 50 and 80°C.

The master curves were constructed combining Christensen-Anderson-Marasteanu (CAM) model and Williams-Landel-Ferry (WLF) time-temperature superposition principle (TTSP) (Yusoff *et al.* 2011). Four rheological parameters were computed based on derived master curves, i.e., crossover frequency ( $f_c$ ), crossover modulus ( $G_c$ ), Glover-Rowe (G-R) parameter, and rheological index ( $R$ ). The  $f_c$  and  $G_c$  are the frequency and complex modulus at which the phase angle equals to 45°. The G-R parameter can be used to evaluate age-induced fatigue cracking of bituminous materials (Mensching *et al.* 2015), with a higher G-R value relating to a higher susceptibility to fatigue cracking. The  $R$  describes the transition from elastic behavior to steady flow state (Pipintakos *et al.* 2022).

### 2.3 Partial least squares regression

The initial spectral data are complex, high-dimensional, and potentially subject to noise originating from instrument limitation or operational process. Therefore, data preprocessing (DP) is essential for conducting accurate multivariate analysis concerning the ageing behavior in bitumen. The preprocessing methods utilized include standard normal variate (SNV), Savitzky-Golay (SG) (Rinnan *et al.* 2009), and outliers detection. The SNV normalization is developed to reduce the deviation of baseline shifts and scattering effects in data sets. The Savitzky-Golay (SG) method consists of two primary functions, i.e. curve smoothing and derivative calculation. The Q-residuals and Hotelling's T2 were used to detect outliers. Q-residuals are the sum of squares of residuals and describe how well the fitted results conform to the model. A larger Q-residual can be observed for outliers compared to the rest samples. Hotelling's T2 is the F-distribution of Mahalanobis distances from the center of a data group. The outliers were identified based on Q-residuals and Hotelling's T2 with a confidence level of 95% (Morais *et al.* 2020).

Before Partial Least Squares (PLS) regression, the preprocessing was conducted for the spectral data to increase regression accuracy. The SNV method was initially applied to correct wavenumber scaling and background effects. Subsequently, the SG method with a window size of 10, a 5-degree polynomial, and a 1st derivative was then used to smooth the FTIR curve. After this, the spectra with a resolution of 1 cm<sup>-1</sup> were merged into a resolution of 5 cm<sup>-1</sup>. Therefore 520 spectra variables were ultimately considered for further multivariate analysis.

PLS regression is a powerful technique for modeling and analyzing large, multivariate, and collinear data sets. In this method, a data matrix of variables  $X$  and a data matrix of responses  $Y$  are decomposed into their respective scores  $T$  and  $U$  and loadings  $P$  and  $Q$ , such that  $X = T \times P^T$  and  $Y = U \times Q^T$ . The covariance between the scores of independent variables and responses can be maximized by adjusting the decomposition directions (i.e., loadings) using algorithms such as the non-linear iterative partial least squares (NIPALS) (Westerhuis *et al.* 1998). This allows the transformed  $Y$  to be predicted as a function of transformed  $X$ . As a result, the original  $Y$  can be described as

$$Y = \beta X \quad (1)$$

where  $\beta$  is the coefficient matrix.

Before regression analysis, preprocessed bitumen samples were randomly divided into calibration and validation subsets in the ratio of 7:3. Model training was initially implemented using the calibration subset. Subsequently, to quantify the robustness of the regression models and avoid overfitting, a 5-fold cross-validation method was utilized for the calibration subset. The trained model was then used on the validation data set.

The prediction quality of regression models was evaluated by the root mean square errors of calibration (RMSEC), cross validation (RMSECV), and prediction (RMSEP), the determination coefficients of calibration ( $R_c^2$ ), cross-validation ( $R_{cv}^2$ ) and prediction ( $R_p^2$ ), the residual prediction deviation (RPD), and the average bias:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (2)$$

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

$$RPD = \frac{SD}{RMSE} \quad (4)$$

$$Bias = \frac{\sum_{i=1}^N \hat{y}_i - y_i}{N} \quad (5)$$

where  $y_i$  and  $\hat{y}_i$  represent the measured and predicted values for sample  $i$ ,  $\bar{y}$  is the mean value of  $y_i$ ,  $N$  is the number of samples for calibration, cross-validation or prediction, and SD is the standard deviation of measured values for  $N$  samples. An RPD value higher than 2 implies a good accuracy of the model, and models with RPD larger than 3 are considered having excellent performance (Douglas *et al.* 2018). The  $R^2$  and bias indicate the linearity of regression models, i.e., the deviations from random distribution of residues and the existence of systematic errors.

#### 2.4 Variable selection methods

Two variable selection approaches, i.e., moving windows (MW) and simulated annealing (SA) were used together with PLS regression to identify important features contributing to the regression model. The MW selection method allows the selection of a narrower region than the full FTIR spectrum, which is more informative for the prediction of bitumen properties. For MW, the spectral width in the range of  $50 \text{ cm}^{-1}$  to  $500 \text{ cm}^{-1}$  and a wave-number moving step of  $50 \text{ cm}^{-1}$  for each spectral width were used, which correspond to a window size of 10–100 variables and a moving step of 10 variables. For each window size, a series of sub-windows were generated by moving the window through the whole spectra with a moving step of 10. The optimal window size and window region were assessed by the root mean square error of cross-validation. Simulated annealing is a stochastic optimization technique used to identify the most informative variables relevant to the targets by exploring various combinations of variables. For FTIR data set, the number of variables ranging from 10 to 200 was compared. The maximum iteration time was defined as 150. Trial regression modeling results regarding crossover modulus showed that the prediction accuracy had limited improvement after 150 iterations. For complete details on variable selection methods, readers are referred to our previous work (Ma *et al.* 2023).

### 3 RESULTS AND DISCUSSION

#### 3.1 Modeling performance

The performance of PLS regression modeling for various rheological parameters using different variable selection methods is presented in Table 1. The  $R^2$  values of four rheological properties are mostly above 0.9, demonstrating the high quality of the trained regression model. The modeling without variable selection presents lower accuracy for calibration dataset and better performance for validation dataset, as indicated by RMSECV and  $R_{cv}^2$ . In contrast, after SA variable selection, the  $R_{cv}^2$  significantly increases, and RMSECV is reduced, which demonstrates increased model performance for calibration dataset. However, this compromises the prediction accuracy for validation dataset. This indicates that applying SA to PLS regression increases the risk of overfitting, and thus the prediction for unknown samples is impaired. Among three variable selection methods, i.e., SM, MW, and SA, MW presents overall the best performance, with high prediction accuracy for both calibration and validation datasets. The comparison between predicted and measured crossover frequency and crossover modulus based on integrated MW-PLS modeling is shown in Figure 1. A good precision is observed for both rheological properties and for both calibration and validation datasets. The high  $R^2$  and the low bias reveal the high linear correlation between FTIR spectra and rheological properties, also verified by the random distribution of calibration and validation data points around identity line. Most RPD values for four rheological properties are greater than 3, demonstrating the high performance of these models.

The FTIR data usually presents high collinearities. As an example, the bands ranging from 2965 to 2864  $\text{cm}^{-1}$  are informative of the antisymmetric and symmetric C-H stretching vibration of  $\text{CH}_2$  and  $\text{CH}_3$  groups while the spectral regions of 1485 – 1357  $\text{cm}^{-1}$  are linked to the bending vibration of  $-\text{CH}_2$  and  $-\text{CH}_3$  groups (Castro and Vazquez 2009). Both ranges describe the aliphatic structure in bitumen. Through MW approach, a narrower spectral region is selected which has the highest correlation with targets. Compared to the full spectrum, the selected region involves much less variables, which also means less noisy information. The high predictive ability of the selection region suggests that the chemical information in this region is sufficient to predict these rheological properties. More variables can bring into similar information but also more noise. Consequently, PLS modeling with MW shows the best performance for all samples.

Table 1. PLS regression accuracy using different variable selection (VS) methods. “SM” refers to the PLS regression without variable selection.

Parameters	VS	$R_c^2$	RMSEC	$R_{cv}^2$	RMSECV	$R_p^2$	RMSEP	RPD	bias
$f_c$	SM	0.975	0.179	0.896	0.366	0.970	0.214	5.911	0.038
	MW	0.992	0.107	0.957	0.247	0.969	0.213	5.780	0.044
	SA	0.998	0.059	0.992	0.107	0.936	0.305	3.954	0.004
$G_c$	SM	0.961	0.061	0.924	0.084	0.915	0.100	3.761	0.042
	MW	0.990	0.030	0.948	0.069	0.880	0.102	2.905	-0.011
	SA	0.995	0.021	0.982	0.041	0.844	0.116	2.534	0.004
G-R	SM	0.982	0.141	0.854	0.397	0.921	0.283	3.680	0.074
	MW	0.993	0.095	0.947	0.263	0.981	0.152	7.293	-0.013
	SA	0.995	0.085	0.981	0.159	0.928	0.296	3.746	0.02
R	SM	0.970	0.074	0.916	0.123	0.959	0.091	4.937	0.004
	MW	0.985	0.044	0.949	0.083	0.934	0.093	4.105	-0.030
	SA	0.991	0.036	0.977	0.055	0.831	0.148	2.821	-0.075

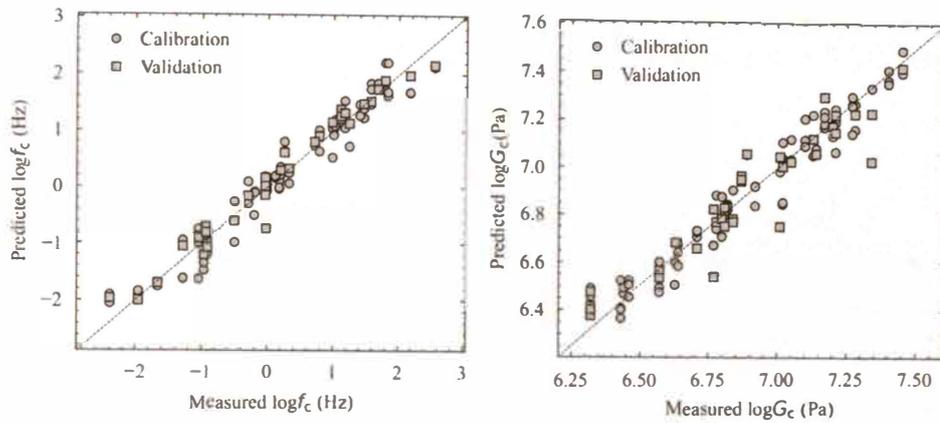


Figure 1. Predicted versus measured crossover frequency and crossover modulus utilizing PLS regression of FTIR spectra.

### 3.2 Chemical characterization of rheological properties

Figure 2 shows the regression coefficients for all variables derived from the PLS regression using complete FTIR spectra. The identified wavenumber regions highly correlated to the prediction of  $f_c$  are mainly related to carbonyl functional groups, aliphatic  $-\text{CH}_3$ , and aromatic  $-\text{CH}$ . For  $G_c$ , the important chemical groups are  $-\text{CH}_3$ ,  $-\text{CH}_2$ , carbonyl group, sulf-oxide group, and aromatic  $-\text{CH}$ . The  $G-R$  is strongly correlated with carbonyl group, aromatic  $\text{C}=\text{C}$ , and aromatic  $-\text{CH}$ . The carbonyl group,  $-\text{CH}_3$ ,  $-\text{CH}_2$ , and aromatic  $-\text{CH}$  are primary chemical groups contributing to the  $R$  (Ma *et al.* 2023). Apart from these groups, the spectral region near  $920\text{ cm}^{-1}$  is also identified as the most relevant parameter for all four rheological properties. According to the spectra of bitumen at different ageing levels (Ma *et al.* 2023), the intensity of this region decreases with increasing ageing level. Given that peak overlapping occurs for FTIR spectra (Asemani and Rabbani 2020), this decreased intensity may be caused by the peaks around it, which changes significantly with ageing or sources. This observation underlines the importance of applying spectra decomposition to the analysis of FTIR data. Moreover, it should be noted that, due to the collinearity inherent in FTIR spectra, some peaks that represent similar chemical information to those with high coefficients may not be recognized. In general, the spectral regions of  $1700\text{--}1400\text{ cm}^{-1}$  and  $1100\text{--}600\text{ cm}^{-1}$  are highly related to the prediction of rheological properties.

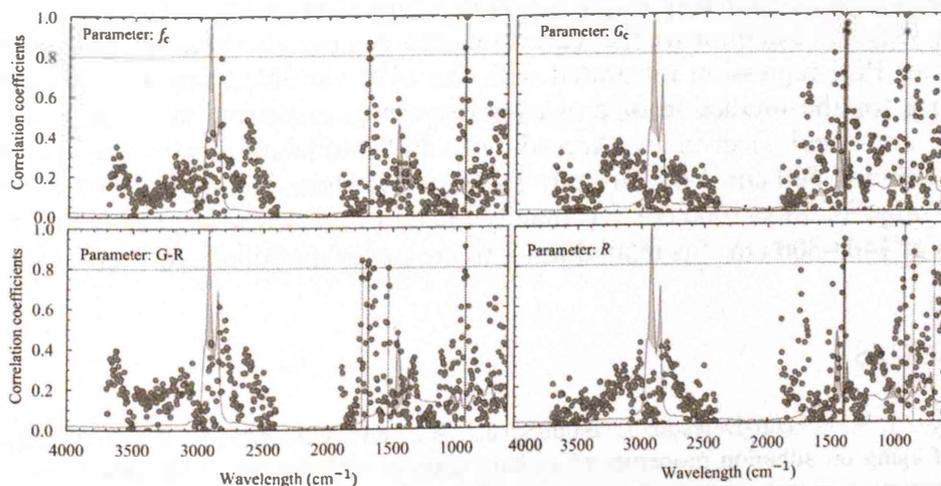


Figure 2. Normalized absolute regression coefficients obtained from PLS regression using full FTIR spectra. Vertical dotted lines are plotted for coefficients higher than 0.8.

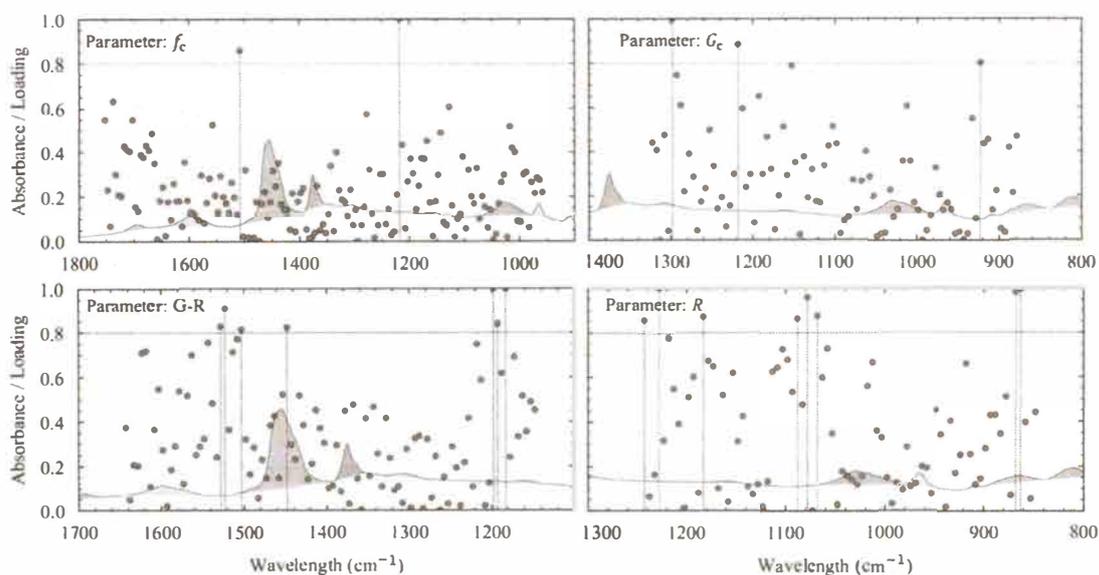


Figure 3. Normalized absolute regression coefficients derived from PLS regression integrated with MW variable selection.

Figure 3 shows the regression coefficients for the PLS regression using optimized wavenumber ranges as input variables. For  $f_c$ ,  $G_c$ , G-R and R, the optimized wavenumber ranges are 1750–950  $\text{cm}^{-1}$ , 1320–880  $\text{cm}^{-1}$ , 1640–1150  $\text{cm}^{-1}$ , and 1240–840  $\text{cm}^{-1}$ , respectively. The selected spectral regions for  $f_c$  and G-R are similar while  $G_c$  and R obtain best prediction at similar spectral regions. The region between ca. 1800–1000  $\text{cm}^{-1}$  is mainly linked to polar functional groups, aliphatic -CH<sub>3</sub> and -CH<sub>2</sub>, and aromatic C=C bonds. The spectral region between ca. 1400–800  $\text{cm}^{-1}$  involves polar functional groups and aromatic -CH group. This difference suggests that  $f_c$  and  $G_c$  depend partly on different chemical changes during ageing process.

#### 4 CONCLUSIONS

This paper investigated the chemo-rheological relationship and predicted rheological properties through PLS regression and variable selection approaches. The PLS regression overall demonstrates high performance in predicting rheological properties of bitumen. Utilization of variable selection can improve the regression accuracy but also hold the risk of overfitting. Despite that, PLS regression integrated with the MW variable selection presents the best performance for the prediction of crossover frequency, crossover modulus, Glover-Rowe parameter, and rheological index. According to PLS modeling results, the spectral region between 1800 and 600  $\text{cm}^{-1}$  are strongly correlated to these rheological properties, among which the range of 1800–1000  $\text{cm}^{-1}$  is mainly linked to crossover frequency and G-R while the region at 1400–800  $\text{cm}^{-1}$  is related more to crossover modulus and rheological index.

#### REFERENCES

- Aguiar-Moya, J. P., Salazar-Delgado, J., Baldi-Sevilla, A., Leiva-Villacorta, F. & Loria-Salazar, L. 2015. Effect of aging on adhesion properties of asphalt mixtures with the use of bitumen bond strength and surface energy measurement tests. *Transportation Research Record*, 57–65.
- Asemani, M. & Rabbani, A. R. 2020. Detailed FTIR spectroscopy characterization of crude oil extracted asphaltene: Curve resolve of overlapping bands. *Journal of Petroleum Science and Engineering*, 185.

- Aske, N., Kallevik, H. & Sjoblom, J. 2001. Determination of saturate, aromatic, resin, and asphaltenic (SARA) components in crude oils by means of Infrared and Near-Infrared spectroscopy. *Energy & Fuels*, 15, 1304–1312.
- Castro, L. V. & Vazquez, F. 2009. Fractionation and characterization of Mexican crude oils. *Energy & Fuels*.
- Goosen, E. S. & Jenkins, K. J. 2022. The viscoelastic transition: a tool to determine equivalent binder grade. *Materials and Structures*, 55.
- Herrington, P. R. 2012. Diffusion and reaction of oxygen in bitumen films. *Fuel*, 94, 86–92.
- Jing, R. X., Varveri, A., Liu, X. Y., Scarpas, A. & Erkens, S. 2019. Ageing effect on chemo-mechanics of bitumen. *Road Materials and Pavement Design*.
- Ma, L., Varveri, A., Jing, R. & Erkens, S. 2021. Comprehensive review on the transport and reaction of oxygen and moisture towards coupled oxidative ageing and moisture damage of bitumen. *Construction and Building Materials*, 283.
- Ma, L., Varveri, A., Jing, R. & Erkens, S. 2023. Chemical characterisation of bitumen type and ageing state based on FTIR spectroscopy and discriminant analysis integrated with variable selection methods. *Road Materials and Pavement Design*, 24, 506–520.
- Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. 2012. A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62–69.
- Mensching, D. J., Rowe, G. M., Daniel, J. S. & Bennert, T. 2015. Exploring low-temperature performance in Black Space. *Road Materials and Pavement Design*, 16, 230–253.
- Mirwald, J., Werkovits, S., Camargo, I., Maschauer, D., Hofko, B. & Grothe, H. 2020. Understanding bitumen ageing by investigation of its polarity fractions. *Construction and Building Materials*, 250.
- Mohammadi, M., Khorrami, M. K., Vatani, A., Ghasemzadeh, H., Vatanparast, H., Bahramian, A. & Fallah, A. 2021. Genetic algorithm based support vector machine regression for prediction of SARA analysis in crude oil samples using ATR-FTIR spectroscopy. *Spectrochim Acta A Mol Biomol Spectrosc*, 245, 118945.
- Morais, C. L. M., Lima, K. M. G., Singh, M. & Martin, F. L. 2020. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. *Nat Protoc*, 15, 2143–2162.
- Nobakht, M., Zhang, D., Sakhaeifar, M. S. & Lytton, R. L. 2020. Characterization of the adhesive and cohesive moisture damage for asphalt concrete. *Construction and Building Materials*, 247.
- Petersen, J. C. 2009. A review of the fundamentals of asphalt oxidation chemical, physicochemical, physical property, and durability relationships. *Transport Research Circular Number E-C140*.
- Pipintakos, G., Lommaert, C., Varveri, A. & Van Den Bergh, W. 2022. Do chemistry and rheology follow the same laboratory ageing trends in bitumen? *Materials and Structures*, 55.
- Redelius, P. & Soenen, H. 2015. Relation between bitumen chemistry and performance. *Fuel*, 140, 34–43.
- Rinnan, Å., Berg, F. V. D. & Engelsen, S. B. 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28, 1201–1222.
- Siroma, R. S., Nguyen, M. L., Hornych, P., Lorino, T. & Chailleux, E. 2021. Clustering aged bitumens through multivariate statistical analyses using phase angle master curve. *Road Materials and Pavement Design*.
- Tauste, R., Moreno-Navarro, F., Sol-Sánchez, M. & Rubio-Gámez, M. C. 2018. Understanding the bitumen ageing phenomenon: A review. *Construction and Building Materials*, 192, 593–609.
- Weigel, S. & Stephan, D. 2017. The prediction of bitumen properties based on FTIR and multivariate analysis methods. *Fuel*, 208, 655–661.
- Westerhuis, J. A., Kourti, T. & Macgregor, J. F. 1998. Analysis of multiblock and hierarchical PCA and PLS models. *Journal of Chemometrics*, 12, 301–321.
- Yun, Y.-H., Li, H.-D., Deng, B.-C. & Cao, D.-S. 2019. An overview of variable selection methods in multivariate analysis of near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 113, 102–115.
- Yusoff, N. I. M., Shaw, M. T. & Airey, G. D. 2011. Modelling the linear viscoelastic rheological properties of bituminous binders. *Construction and Building Materials*, 25, 2171–2189.