A Novel Approach to Coke Strength Prediction Using Self Organizing Maps

L. North[1], K. Blackmore[2], K. Nesbitt[2], K. Hockings[3], and M. Mahoney[1]
[1]Centre for Ironmaking Materials Research, The University of Newcastle, Callaghan, NSW, Australia
[2]School of Electrical Engineering and Computing, The University of Newcastle, Callaghan, NSW, Australia
[3]Coal Technical Marketing, BHP Billiton, Brisbane, QLD, Australia

Abstract - *Vitrinite reflectance is an important property of coal, and can be used in the prediction of the quality of coke derived from coal blends. However, limited methods that incorporate the associated reflectance distributions within prediction models exist. We present a novel method of classifying vitrinite reflectance distributions using self organizing maps. This approach is shown to provide representative and repeatable distributions. In comparison to approaches that capture only the spread of the distribution, the work presented here is able to capture the often multimodal nature of coal blends. Implementation of the resulting distributions within regression based prediction models showed statistically significant improvement to model prediction, furthering the understanding and control of coal blending. More generally, the work presented here may inform other application domains where the general shape of data distributions needs to be incorporated as a single attribute into simulation models.*

Keywords: Self Organizing Maps; Pattern Recognition; Data Distributions; Coke Prediction

1 Introduction

1.1 Coal, coke, and their properties

Metallurgical coke is a heterogeneous, porous, brittle solid, used primarily in iron production. The strength and porosity of this material are critical to the successful operation of the reactor vessel [1]. Coke is derived from the pyrolysis of particular types of coal in the absence of air [2]. Coal is formed from the chemical and physical transformation of plant material over millions of years [3]. This process results in a heterogeneous material, which can be considered as a mixture of reactive or fusible, and inert components [4]. Of particular interest to this study is vitrinite, one of the reactive components. During the conversion of coal to coke, the reactive components soften. At higher temperatures, these reactive components resolidify to form the binder which holds the coke together [5]. Traditionally, the vitrinite is characterized by its ability to reflect light, which is related to its chemical composition, particularly the degree of aromatization. Individual grains of coal reflect different amounts of light [6], and in a single coal, measurement of the reflectance of many grains of coal results in an approximately Gaussian distribution of vitrinite reflectance [7]. Different coals will have a unique vitrinite reflectance distribution as a result of their individual formation conditions. Vitrinite reflectance is related to the behavior of the coal upon heating, particular the swelling and binding abilities [5]. This behavior in turn, has implications on coke properties and hence coke performance, which is typically reported as coke strength after reaction (CSR) in more recent generations of prediction models [5].

1.2 Prediction of coke properties

Due to the complex nature of metallurgical coke, the understanding and prediction of its properties from the parent coals used in the coking blend presents as an interesting materials challenge. At present, there is no singular method from which the properties of coke derived from coals of any global coal basin can be predicted with reasonable accuracy [5]. Furthermore, prediction abilities of existing models diminish as coal blends increase in complexity [5]. The majority of these prediction methods rely on bulk, aggregated coal measures, such as, mean maximum vitrinite reflectance and thermoplastic properties as model inputs, which may not adequately capture the nuances arising from the heterogeneity of the coals. Hence, rather than attempting to build the 'perfect' prediction model from existing inputs, this work considers the extraction of additional information from the distributions of coal quality values. This allows the distinct characteristics of the distribution, rather than their singular averaged value, to be used as an input to prediction models.

1.3 Concepts of vitrinite reflectance and coal blending decisions

The mean vitrinite reflectance ($R_{o,max}$) is a commonly used input to predictive models of coke strength after reaction (CSR) [5, 8-10]. With the usage popularized by Schapiro and co-workers in the 1960's [4, 11], the concept of vitrinoid types, or V-Groups, is the breakdown of the measured vitrinite reflectance distribution into cumulative bins of 0.1% reflectance, resulting in a histogram of the distribution of coal reflectance over the measured spectrum. Half V-Groups are a more detailed characterization of the distribution, where the

bin size used is 0.05%. In coke making practice, a single peak within the histogram, representing completely overlapping ranges of reflectance, is preferred by some manufacturers over a multimodal vitrinite distribution [12, 13]. For some, this profile is mandated regardless of whether desired aggregate properties (for instance mean reflectance, volatile matter, and thermoplastic properties) are met [12].

Whilst the concept of V-Groups is an important factor in industrial blend preparation, the application of V-Groups for the determination of coke quality models is limited. Schapiro et al. [4] utilize the V-Group as part of their strength prediction model, implying that for each V-Group there is an optimum ratio of inerts to reactive material. This graphical approach has been criticized by some investigators for various reasons, notably the limited data set from which it was derived [14], and for the poor applicability beyond Appalachian coal measures [15-17]. A number of investigators [9, 18-20] consider an aggregate value of a portion of the distribution in their prediction models of CSR, in effect creating a weighted Ro,max. Attempts have been made to analyze the distribution behavior statistically [20-22], with the results used to infer details of blending. Bukharkina et al. [20] consider the standard deviation of the distribution as an attribute in their regression, finding it more important than the $R_{o,max}$ value, but achieving a poor overall correlation coefficient. Similarly, Stankevich et al. [22] utilize the standard deviation, reporting good correlation in their prediction of the coke Micum indices for the limited selection of coals for which the model was developed. Whilst the standard deviation may capture information about the spread of the distribution, it does not necessarily capture the modality of the distribution, which may be a key feature in the determination of the behavior of the resulting coke. Despite the role of $R_{o,max}$ in blending decisions, in the available literature, the effects on, and prediction of, coke properties resulting from the profile of vitrinite reflectance have not been well explored.

1.4 Emerging methods of information extraction

Data mining presents as an increasingly used method for both extracting information and predicting behavior within large datasets, with applications in (but certainly not limited to) medicine, advertising, and manufacturing [23 ,24]. Within competitive industries such ironmaking, documented implementation and development of such techniques has been limited. A key limiting factor in the adoption of data mining in this domain is the availability of data sets that contain both the variety of coals and number of instances required to develop a robust, and generalizable model. Despite this, the use of data mining techniques is growing within the study of coal and coke behavior, with a recent shift towards the prediction of coking coal behavior using "black box" artificial neural network approaches based on influential parameters identified in regression studies [9, 25-28]. In terms of inferring additional information from measurements, the complicated environment of the blast furnace has shown to be a breeding ground for more novel applications of data mining techniques [29-32].

Of particular interest to this study is the work of Martín et al. [33], who apply a sub-class of the neural network category of data mining, the Kohonen self organizing map (SOM) [34], to identify patterns within furnace top temperature probe measurements taken from the above burden probes. Self organizing maps are an interactive non-linear pattern recognition technique, relying on a process known as vector quantization. In essence, this technique groups like features together, which is differentiated from traditional clustering approaches (including K-means, nearest neighbour techniques) through the production of a map of prototype groups that spatially represents the cluster similarity. From the SOM patterns identified by Martín et al. [33] within the blast furnace data set, undesirable changes in the descent of material through the furnace were pre-empted, helping to maintain operating stability.

In the context of vitrinite reflectance data, the SOM technique presents as a novel method of classifying blend distributions, extending upon the method used by Martín et al. [33], and utilizing the resulting distributions to infer the influence of such distributions on the behavior of coking coal blends. This study presents the proof of concept for such an approach, with the implications on blending decisions discussed in a future paper. Section 2 of the paper develops the modelling framework including the data pre-processing and model tuning. Section 3 details the application of the model to filtered data, followed by the application to data that is offset to a centralized value (with these scenarios defined in Section 2). Conclusions drawn from this proof of concept study are discussed in Section 4.

2 General Modelling Framework

A flow diagram displaying the modelling phase is shown in Fig. 1.

2.1 Data selection and preparation

This modelling and analysis was completed using MATLAB Release 2015b, Neural Network Toolbox 8.4, and Statistics and Machine Learning Toolbox 10.1 [35].

2.1.1 Data source

The data used in this study is half V-Group data, with bin sizes of 0.05% reflectance, taken from pilot coking oven experimental results. The dataset has 401 observations for coking of blends of Australian and some overseas coals.

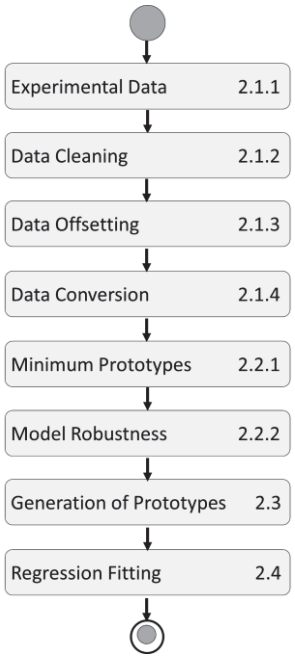


Fig. 1. Flow diagram of modelling framework

2.1.2 Data cleaning

The dataset was manually cleaned to remove instances with missing values, and instances reported as full V-Groups (bin size 0.1% reflectance). The removal of full V-Groups was considered necessary to maximize the information extracted from the data set, as the use of the full V-Group data would require the aggregation and associated loss of information of the half V-Group data. As the V-Group data is a closed data set; that is, the cumulative value should add to 100, a distribution with a cumulative value differing from this value may be erroneous. Hence, groups where the cumulative frequency was not between 99 and 101 were removed. As only blend behavior was considered in this study, blends of more than one distinct component were included. Further filtering removed blends where one single coal represented 90% or more of the blend, so that the behavior of blends could be distinguished from single coals.

2.1.3 Data offsetting - modification of data set to eliminate influence of $R_{o,max}$

Due to the previously known significance of the mean-max vitrinite reflectance on the strength properties of coke, we model a scenario where the reflectance distribution was offset to a central reflectance value based on the instance mean-max reflectance (see Fig. 2), with this model herein referred to as the "offset model". This allowed examination of the influence of the blending profile alone on coke strength. In order to offset the data, the number of steps between the sample $R_{o,max}$ and the selected central $R_{o,max}$ value was determined, with the instance vector padded with this number of zeros at the start or the end, such that the newly offset instance had the central $R_{o,max}$.

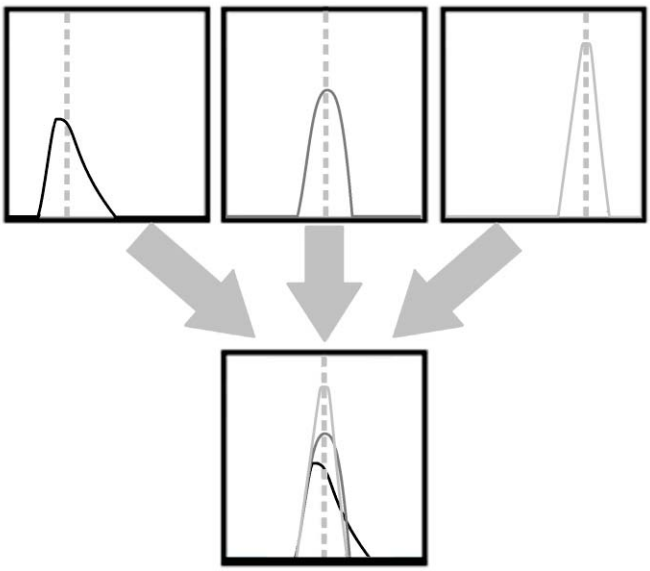


Fig. 2. Offsetting of the V-Group histogram, where the mean maximum reflectance value shown as the dotted line is centralized for each case

2.1.4 Data conversion

The authors conducted preliminary testing to evaluate the suitability of the approach. This preliminary testing indicated that the use of the raw half V-Group data in the SOM resulted in prototype groupings that overrepresented the cases with strong single peaks (i.e. where a single V-Group had a comparatively high value), and appeared to reduce the representation of cases where the vitrinite reflectance was distributed over a wide range or where the value was close to zero. This meant that the differences between blends with small changes in composition were not adequately represented. To reduce this sort of biasing, the raw data was transformed to a binary two-dimensional matrix by conversion to an image of a line plot of the data with a line thickness of three pixels. This in effect normalized the data to the range [0,1] whilst maintaining the importance of the spatial distribution of features along the V-Group reflectance histogram.

2.2 Selection of tuning parameters

In this research, the SOM neural network structure consists of an input layer matching the size of the converted data, directly mapped to an output layer. The size of the output layer was determined through the iterative testing of various sizes of the SOM to verify that the derived prototype structures have sufficient resolution and statistical accuracy [36]. The topology was set as a grid, and default parameters were used for cover steps (100), initial neighborhood size (3), training epochs (200), and distance function ('linkdist' [35]).

2.2.1 Determination of minimum prototypes

The minimum number of prototypes was determined using the "elbow method" [37] to analyze the reconstruction error (sum of the mean square error, SSE) as a function of the number of clusters.

The reconstruction error plot (see Fig. 3) was determined using the entire filtered dataset at various grid dimensions. An iterative approach to collect the sum of squares error (SSE) for each number of prototypes was used, with the grid maintained as approximately a square, to a maximum grid size of 6x6. For instance, six prototypes would be defined as a 3x2 or 2x3 grid, rather than a 6x1 grid.

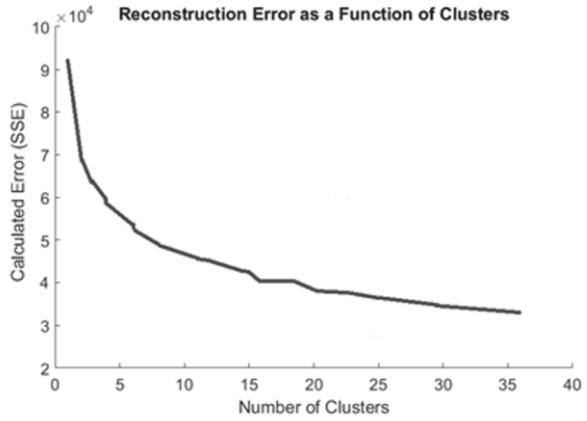


Fig. 3 V Group SOM reconstruction error as a function of the number of clusters

2.2.2 Determination of model robustness

In order to assess the validity of model parameters, the robustness of the model was determined. Using grid dimensions based on the reconstruction error plot, the robustness of the parameters selected was assessed using a 10-fold Monte Carlo cross validation [38], with the data portioned into a 60% train and 40% test split. Some variations in the ordering of prototypes occurred between folds, thus equivalent prototype vectors were matched between each of the folds, and the prototypes were renumbered accordingly. The test portion of the data was assigned to the corresponding closest prototype vector for each of the folds. As there is no 'ground truth' to validate against in this case, the 'correct' assignment of each instance was assumed to be the prototype vector the instance was most frequently assigned to. In cases where the highest frequency was the same between two clusters, the correct cluster was assigned to the lower number (i.e. if cluster 4 and cluster 5 both had 2 allocations, cluster 4 would be assigned as the correct cluster.). These assumptions allowed the generation of a confusion matrix [39], from which the consistency of allocation could be assessed. Provided that the prototype vectors were similar between cases, and the classification consistency was greater than 80%, the selected parameters were determined to be of adequate robustness. The final structure resulting from this process and the cross validation results can be found in Section 3.1.

2.3 Generation of prototypes and grouping of instances

Once suitable modelling parameters were established, a single SOM was generated using the entire cleaned dataset (n = 401). Each instance was assigned to the closest prototype, and a new attribute named "Cluster" that identified the closest prototype was added to each record in the original dataset.

2.4 Generation of regression models

2.4.1 Regression model for original dataset

A regression model was generated for the original dataset, ignoring the newly added attribute. Regression analysis was selected such that comparison could be made with the prediction models of other investigators. The parameters selected for the regression were commonly used within the available literature for similar regressions, capturing rank, rheology, petrology, and chemistry [40]. These parameters were the base-10 logarithm of the maximum fluidity, mean maximum reflectance, modified basicity index, as well as their squared values. A linear regression model was fitted, with parameters insignificant at the 5% level excluded from the results.

2.4.2 Regression model for grouped dataset

A regression model was fitted as described in Section 2.4.1, however the newly assigned Cluster attribute was forced in to the model as a categorical variable. This allowed the importance of the V-Group attribute on the regression model to be assessed.

3 Results and Discussion

3.1 Assessment of tuning parameters

Using the method described in Section 2.2, the model set up was tuned for both the raw model and the offset model. From the reconstruction error plot shown in Fig. 3, the approximate minimum number of prototypes was determined to be approximately six. Nine prototypes arranged in a 3x3 grid was trialed. This selection was made on the basis of capturing the minimum number of prototypes, whilst ensuring a broader selection of possible distributions was included in a favored grid layout. The validity of this assumption was tested through the 10-fold cross validation step. It is observed that the clusters are visually quite robust; that is that they are repeated consistently between the validation folds. The prototype behavior between different validation folds is consistent, as detailed in Table 1. Due to selecting more than the optimum number of clusters, there is more similarity between some of the clusters between folds, resulting in increased false allocation within the confusion matrix.

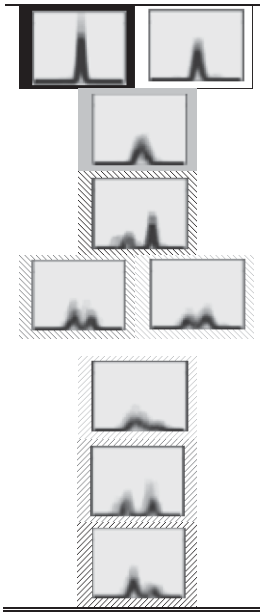
The resulting confusion matrix (Table 2) shows that the 10-fold classification correctly allocates 72.9% of cases in the offset data set. The low correct allocation frequency (72.9%) in the offset data case in Table 2 is due to the similarity

between vectors 1, 2, 4 and 5. It is likely that some of these vectors would be merged if a lower number of prototypes was selected as an input parameter. However, for the purposes of this study, it was anticipated that the additional similar groupings would not have a negative impact on the regression results.

3.2 Implementation of the SOM on offset data

A single self organizing map was trained and tested on the complete offset data set using the process described in Section 2, yielding the prototype vectors shown in Fig. 4. The prototype vectors are shaded based on their behavior in the model robustness study. Fig. 5 shows the relationship between these vectors. As identified in the robustness testing, vectors 1, 2, 4 and 5 are similar to each other, reflected by the small weight distances in Fig. 5.

Table 1 Comparison of vectors between folds



Vector 9 and **Vector 6** represent strong single reflectance behaviour (i.e. the blending of similar reflectance coals)

Vector 3 represents a more distributed single peak

Vector 8 represents a single reflectance coal, with a small inclusion of low reflectance coals

Vector 4 and **Vector 5** represent a bimodal distribution, with the individual coal distributions touching and slightly overlapping in some cases

Vector 1 represents a multimodal distribution that overlaps completely

Vector 7 represents a bimodal distribution with distinct separation between their reflectance

Vector 2 represents bimodal distribution with some overlap, merging to a distributed single peak in some folds

Table 2 Confusion matrix for the 10-fold cross validation of V-Group distribution clusters

		Predicted As								
		1	2	3	4	5	6	7	8	9
Correct Allocation	1	128	0	0	0	11	0	1	12	0
	2	0	122	0	0	5	0	0	0	0
	3	0	0	134	0	1	20	0	0	0
	4	0	0	0	294	30	5	34	17	9
	5	10	5	0	10	149	18	8	3	0
	6	0	0	12	1	5	105	0	0	10
	7	0	0	0	10	4	0	228	5	0
	8	6	0	0	3	0	0	6	63	5
	9	0	0	0	7	0	9	1	8	86

Black are correctly classified values, light grey are misclassification frequency less than 20, dark grey are misclassification frequency greater than or equal to 20

3.2.1 Original regression

As a point of comparison, a linear regression was implemented on the filtered data set, using the method described in Section 2.4.1. The results of this regression are shown in Table 3. Whilst the results of this regression are by no means best of class (r^2=0.698 compared to typical values above 0.8), the limited filtering, restriction to blend only data, and breadth of coals included was anticipated to detrimentally affect the regression fitting. Nevertheless, as a benchmark to assess the impact of the vitrinite distribution clustering, this regression is deemed adequate.

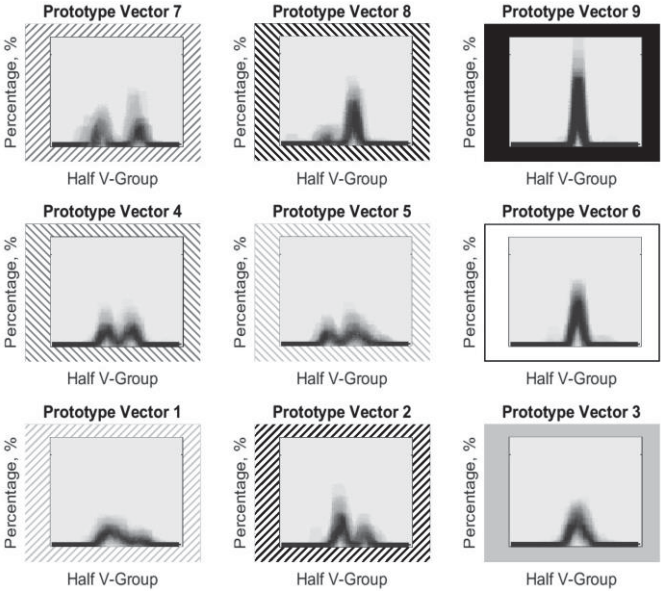


Fig. 4 Prototype vectors for the offset data set

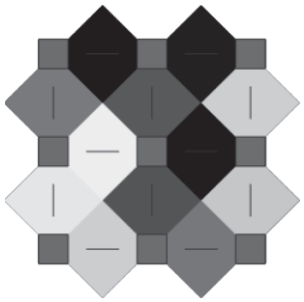


Fig. 5 SOM neighbor weight distances. Nodes are represented as squares, and the shading of diamonds represents the distance between neighbors, where the darker the shading, the more dissimilar the vectors are

Table 3 Original regression parameters

Attribute	Coefficient	Standard Error	p Value
Constant	-XX	20.448	0.017
logMF	XX	2.090	0.000
$R_{o,max}$	XX	31.305	0.000
MBI	-XX	0.737	0.000
$logMF^2$	-XX	0.531	0.000
$R_{o,max}^2$	-XX	12.000	0.030
r^2	0.690		
r^2, *adj*	0.686		

Coefficients removed for confidentiality.

3.2.2 Modified regression including offset data prototypes

The regression was modified to include the prototype groups, the results of which are shown in Table 4. Whilst the $R_{o,max}^2$ term is appears insignificant at the 5% significance level, given the marginal deviation from this significance level and to allow comparison between the original and modified regressions, the decision was made to retain this term within the regression. Although the regression in this offset case shows a small increase to the r^2 value, the prototype vectors are statistically significant at the 10% significance level (with the exception of prototype 6). Examining the coefficients of the prototype vectors, sorted in Table 5, and their associated averaged maximum fluidity, mean maximum reflectance, and modified basicity index, there is no discernible trend. This suggests that either the prototypes are capturing information about the blending profile, or are representative of another attribute not included in the regression.

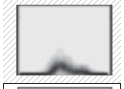
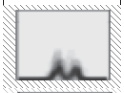
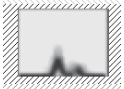
Comparison of the coefficients of the prototype vectors and the average of other measured coal parameters including grind parameters, proximate and ultimate analysis, and softening and dilation parameters, yields no apparent relationship. This again, suggests that the prototypes are capturing some otherwise unmeasured parameter, a non-linear combination of parameters not considered in the model, or behavior associated with the blend structure. The possible explanations for this behavior are to be studied in future work.

Table 4 Offset data modified regression parameters

Attribute	Coefficient	Standard Error	p Value
Constant	-XX	20.795	0.022
logMF	XX	2.217	0.000
$R_{o,max}$	XX	32.026	0.001
MBI	-XX	0.755	0.000
$logMF^2$	-XX	0.558	0.000
$R_{o,max}^2$	-XX	12.316	0.055
Vector 2	XX	1.562	0.023
Vector 3	XX	1.368	0.002
Vector 4	XX	1.339	0.093
Vector 5	XX	1.456	0.093
Vector 6	XX	1.383	0.886
Vector 7	XX	1.622	0.011
Vector 8	XX	1.645	0.011
Vector 9	XX	1.442	0.003
r^2	0.707		
r^2, *adj*	0.697		

Coefficients removed for confidentiality.

Table 5 Offset data vectors, re-ordered by coefficient

Re-Ordered Vector	Group Average $R_{o,max}$	Group Average logMF	Group Average MBI	Image
Vector 1	1.19	2.25	1.48	
Vector 6	1.25	2.07	1.51	
Vector 4	1.26	2.09	1.11	
Vector 5	1.23	2.32	1.55	
Vector 2	1.18	2.40	1.31	
Vector 7	1.27	1.62	1.46	
Vector 8	1.32	1.96	1.39	
Vector 3	1.29	2.11	1.25	
Vector 9	1.28	2.01	1.55	

4 Conclusion

This study has provided proof of concept on the application of self organizing maps in the classification of vitrinite reflectance distributions. The distributions determined by the self organizing map were determined to be representative of the underlying data. The use of offset data in modelling distributions was shown to capture representative distributions of the blends. When the distribution classification from the SOM, was used as a parameter in a regression model it resulted in improved model correlation, compared to models that excluded this parameter. It is noted, however, that this model is limited by the underlying data it is based on, which is biased towards blends featuring Australian coals. As previously noted, there is limited external data available in this area due to the closed nature of the industry. This makes it difficult for benchmarking activities to be undertaken to further validate the effectiveness of the technique. Regardless, it is desirable that further benchmarking against alternative classification techniques is undertaken to further validate the proposed approach.

Despite these limitations, this model presents as a promising approach for capturing the underlying vitrinite distribution behavior, and for inferring the implications of blending decisions on coke quality. The approach may also be broadly relevant in other domains where information contained within the distribution of values is important. Future work shall study the applicability of this model to the blend properties of Australian and international coals.

5 Acknowledgement

The authors acknowledge the constructive advice from Richard Roest, Hannah Lomas, and Kim Colyvas (The University of Newcastle). The Australian Coal Association Research Program is gratefully acknowledged for their financial support of this project.

6 References

[1] A. K. Biswas, “Principles of Blast Furnace Ironmaking: Theory and Practice”. Brisbane: Cootha Publishing, 1981.

[2] R. Loison, P. Foch, and A. Boyer, “Coke: Quality and Production”. Elsevier Science, 2014.

[3] J. G. Speight, “The Chemistry and Technology of Coal, Third Edition”. Boca Raton, Florida: CRC Press, 2012.

[4] N. Schapiro, R. J. Gray, and G. Eusner, "Recent developments in coal petrography," Blast Furnace, Coke Oven and Raw Materials Committee, Proc. 20, pp. 89-112, 1961.

[5] M. A. Díez, R. Alvarez, and C. Barriocanal, "Coal for metallurgical coke production: predictions of coke quality and future requirements for cokemaking," International Journal of Coal Geology, vol. 50, pp. 389-412, 2002.

[6] I. e. Suárez-Ruiz and J. C. e. Crelling, “Applied Coal Petrology: The Role of Petrology in Coal Utilization”. Burlington, Massachusetts: Academic Press, 2008.

[7] M. B. Underwood, “Thermal Evolution of the Tertiary Shimanto Belt, Southwest Japan: An Example of Ridge-trench Interaction”. Geological Society of America, 1993.

[8] A. S. Artser and V. A. Vents, "Prediction of high-temperature properties of coke," Koks i Khimiya, issue 11, pp. 13-19, 2001.

[9] P. S. Dash, M. Guha, D. Chakraborty, and P. K. Banerjee, "Prediction of Coke CSR from Coal Blend Characteristics using Various Techniques: A Comparative Evaluation," International Journal of Coal Preparation and Utilization, vol. 32, pp. 169-192, 2012.

[10] B. D. Ryan and J. T. Price, "British Columbia Geological Survey Geological Fieldwork 1992," 1992.

[11] N. Schapiro and R. Gray, "The use of coal petrography in coke making," Journal of the Institute of Fuel, vol. 11, pp. 234-242, 1964.

[12] N. Choudhury, P. Boral, and S. Hazra, "Coal Petrography for Prediction of Coke Quality of Blended Coals," in International Seminar on Coal Science & Technology, 2005, p. 224.

[13] B. Yao, "The Function Reflectance Distribution Map in Guiding Coal Blend for Cokemaking (In Chinese)," Fuel & Chemical Processes, vol. 39, pp. 11-19, 2008.

[14] R. J. Gray and C. M. Bowling, "Petrographic prediction of coking properties for the Curragh coals of Australia," International Journal of Coal Geology, vol. 27, pp. 279-298, 1995.

[15] B. Ryan, J. F. Grandsden, and J. Price, "Fluidity of Western Canadian Coals and its Relationship to Other Coal and Coke Properties," Geological Fieldwork 1997, 1998.

[16] C. Coin and A. Broome, "Coke quality prediction from pilot scale ovens and plant data," in International Coal Conference, 1997, pp. 325-333.

[17] J. F. Gransden, J. G. Jorgensen, N. Manery, J. T. Price, and N. J. Ramey, "Applications of microscopy to coke making," International Journal of Coal Geology, vol. 19, pp. 77-107, 1991.

[18] P. S. Dash, M. Guha, D. Chakraborty, S. Krishnan, and P. Banerjee, "Application of Various Techniques for Predicting Coke CSR from Coal Blend Properties through Laboratory Scale Experiments," Tata Search, vol. 1, pp. 89-100, 2005.

[19] A. Suresh, T. Ray, P. S. Dash, and P. K. Banerjee, "Prediction of coke quality using adaptive neurofuzzy inference system," Ironmaking & Steelmaking, vol. 39, pp. 363-369, 2012.

[20] T. V. Bukharkina, V. L. Luk’yanov, O. Y. Sal’nikova, and E. E. Sokolovskaya, "Optimizing the batch composition at OAO Moskoks," Coke and Chemistry, vol. 55, pp. 138-141, 2012.

[21] D. Pearson, "Probability Analysis of Blended Coking Coal," International Journal of Coal Geology, vol. 19, pp. 109-119, 1991.

[22] A. Stankevich, N. Chegodaeva, V. Vens, and A. Cheremiskina, "Optimization of the Coal-Blend Composition and Prediction of Coke Quality Based on Chemico-Petrographic Parameters," Coke and Chemistry, pp. 11-17, 1998.

[23] U. M. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery: an overview," in Advances in knowledge discovery and data mining, M. F. Usama, P.-S. Gregory, S. Padhraic, and U. Ramasamy, ed: American Association for Artificial Intelligence, 1996, pp. 1-34.

[24] H.-H. Tsai, "Global data mining: An empirical study of current trends, future forecasts and technology diffusions," Expert Systems with Applications, vol. 39, pp. 8172-8181, 2012.

[25] Q. A. Cui, Z. He, and F. X. Cui, "Prediction model of coke quality based on support vector machines," Huagong Zidonghua Ji Yibiao/Control and Instruments in Chemical Industry, vol. 33, pp. 28-31, 2006.

[26] J. Deng, X.-z. Lai, M. Wu, and W.-h. Cao, "Intelligent optimization method for coal blending based on neural network and simulated annealing algorithm [J]," Metallurgical Industry Automation, vol. 3, pp. 19-23, 2007.

[27] S. Maharana, B. Biswas, A. Ganguly, and A. Kumar, "Artificial neural network prediction for coke strength after reaction and data analysis," World Academy of Science, Engineering and Technology, pp. 4-9, 2010.

[28] L. Qing, X. Keming, H. Xiaogang, and X. Jun, "Prediction of coke mechanical strength based on support vector machines," in Computer Communication Control and Automation (3CA), 2010 International Symposium on, 2010, pp. 35-37.

[29] W. Chen, B. X. Wang, and Y. Chen, "Prediction for the sulfur content in pig iron of blast furnace by combining artificial neural network with genetic algorithm," in Advanced Materials Research vol. 143-144, ed, 2011, pp. 1137-1142.

[30] L. Jian, Y. Song, S. Shen, Y. Wang, and H. Yin, "Adaptive Least Squares Support Vector Machine Predictor for Blast Furnace Ironmaking Process," ISIJ International, vol. 55, pp. 845-850, 2015.

[31] H. Saxen and F. Pettersson, "Nonlinear Prediction of the Hot Metal Silicon Content in the Blast Furnace," ISIJ International, vol. 47, pp. 1732-1737, 2007.

[32] H. Saxen and F. Pettersson, "A Data Mining Method Applied to a Metallurgical Process," in Computational Intelligence and Data Mining (CIDM) - IEEE Symposium, 2007, pp. 368-375.

[33] D. Martín, J. Mochón, L. F. Verdeja, R. Barea, P. Rusek, and J. Jiménez, "Above Burden Temperature Data Probes Interpretation to Prevent Malfunction of Blast Furnaces - Part 2: Factory Applications," Steel Research International, vol. 80, pp. 194-201, 2009.

[34] T. Kohonen, "Self-organized formation of topologically correct feature maps," Biological Cybernetics, vol. 43, pp. 59-69, 1982.

[35] The MathWorks Inc., "MATLAB, Neural Network Toolbox 8,4, and Statistics and machine Learning Toolbox 10.1. Release 2015b,". Natick, Massachusetts, United States, 2015.

[36] T. Kohonen, “MATLAB Implementations and Applications of the Self-Organizing Map”. Aalto, Finland: Aalto University, School of Science, 2014.

[37] E. Alpaydin, “Introduction to Machine Learning”. Cambridge, Massachusetts: MIT Press, 2014.

[38] Q.-S. Xu, Y.-Z. Liang, and Y.-P. Du, "Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration," Journal of Chemometrics, vol. 18, pp. 112-120, 2004.

[39] J. Han, J. Pei, and M. Kamber, “Data Mining: Concepts and Techniques”. Waltham, Massachusetts: Elsevier Science, 2011.

[40] H. P. Tiwari, P. K. Banerjee, and V. K. Saxena, "A novel technique for assessing the coking potential of coals/coal blends for non-recovery coke making process," Fuel, vol. 107, pp. 615-622, 2013.