

BITUMEN CONTENT ESTIMATION OF ATHABASCA OIL SAND FROM BROAD BAND INFRARED REFLECTANCE SPECTRA

B. Rivard,^{1*} D. Lyder,² J. Feng,¹ A. Gallie,³ E. Cloutis,⁴ P. Dougan,⁵ S. Gonzalez,⁵ D. Cox⁶ and M.G. Lipsett⁷

1. Earth Observation Systems Laboratory, Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E3
2. Alberta Environment, Edmonton, Alberta, Canada T5J 1G4
3. Department of Earth Sciences, Laurentian University, Sudbury, Ontario, Canada P3E 2C6
4. Department of Geography, University of Winnipeg, Winnipeg, Canada
5. Syncrude Canada Ltd, 9421-17 Ave., Edmonton, Canada T6N 1H4
6. Suncor Energy, Inc., P.O. Box 4001, Fort McMurray, Alberta, Canada T9H 3E3
7. Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G8

Oil sand is a mixture of quartz grains, clay minerals, bitumen, water, and minor accessory minerals. There is a need in oil sands mining operations for a robust method to estimate total bitumen content in real time; and so modelling of the total bitumen content (TBC) in Athabasca oil sands of Western Canada was undertaken on the basis of hyperspectral reflectance spectra. A selection of different bitumen, water, and clay mineral spectral features (3.0–30.0 μm) was used to develop broad-band TBC predictive models that have good accuracy, with less than 1.5% error with respect to laboratory methods of bitumen assay. These models are also robust, in that they are independent of mine location. Simple broad band models, based upon previously identified Gaussian features or wavelet features, provide an incremental improvement over the currently deployed industry two-band ratio model. An improved two-band model was also developed, which makes use of a combination of the same two bands but normalised to their mean. A wavelet-based, broad-band model comprised of indices and five bands, where the bands are normalised to the mean of the bands, adequately addresses the influence of water, clay, and textural variation on selected bitumen features. This five-band model appears to produce the most robust estimator of TBC, with a dispersion of ~ 1.1 –1.5%, which can be applied to different sites within a mine and to different mines without additional tuning or calibration, as evidenced by regression slopes of 0.99–1.0 for modelling, validation, and blind data sets.

Keywords: oil sand, bitumen, infrared, reflectance spectra

INTRODUCTION

Oil sand is a mixture of quartz grains, clay minerals, bitumen (a mixture of heavy hydrocarbons), water, and minor accessory minerals (Bichard, 1987). Typically oil recovery in surface-mineable Athabasca oil sands deposits involves excavation of the ore, which is then mixed with heated water and process additives to encourage separation (Schramm et al., 2000; Fong et al., 2004). The ore deposits are not homogeneous and the ore displays considerable variability in clay, bitumen, and fines (Bichard, 1987; Hepler and His, 1989), which impact bitumen recovery.

High-accuracy ($\leq 0.1\%$) determination of total bitumen content (TBC) by % weight (%wt) in oil sand ore is conducted by traditional analytical approaches, which are time consuming (typ-

*Author to whom correspondence may be addressed.
E-mail address: benoit.rivard@ualberta.ca
Can. J. Chem. Eng. 88:830–838, 2010
© 2010 Canadian Society for Chemical Engineering
DOI 10.1002/cjce.20343
Published online 17 June 2010 in Wiley Online Library
(wileyonlinelibrary.com).

ically the analysis requires on the order of several hours to be completed) and destructive in nature (Yang et al., 1988; Cloutis, 1989) but of value to define mine models of the ore body. This knowledge is supplemented by field estimates of TBC by geologists and by a broad-band analysis of reflected light (a simple ratio of a peak bitumen absorption to continuum) in the near infrared region (NIR: 1.0–2.5 μm) (Thompson, 1984) of the oil sand as it is transported on conveyors to be processed. The broadband method reported by Thompson (1984) implemented as a commercial unit some 20 years ago, and later modified and known as the Wright and Wright method, had an accuracy in the estimation of TBC of $\pm\sim 3\%$ based on a small sample. The accuracy currently achieved commercially with a two-wavelength instrument is in the order of 1–1.5% though little has been published on this.

Bitumen content is used for reserve estimation and geological block modelling. Current methods based on core data give only approximate estimates of the bitumen delivered to a separation plant, which is why conveyor-based instruments for ore grade estimation were developed. A more accurate estimate of bitumen content into an extraction plant would allow for a more accurate material balance of the bitumen species and thus an improved estimate of recovery. Hence, there is a need for a robust method to estimate total bitumen content in real time.

In a previous study (Lyder et al., 2010) we examined the selection of different bitumen (or other component) spectral features that had been based upon Gaussian fitting and wavelet analysis (Torrence and Compo, 1998) of hyperspectral data in the 3.0–30.0 μm range towards the development of a more robust and accurate model of TBC. Given the impracticality of determining TBC from hyperspectral data in real time, the current study aimed to use the features identified in Lyder et al. (2010) to guide the development of broad band predictive models that not only improve the accuracy of the estimate of TBC but also are robust to changes in ore location and nonbitumen components in ore. Improvements would need to minimise the current requirement for site specific calibration (Shaw and Kratochil, 1990) and account for the influence of non-bitumen factors, for example, water content and texture, which change the shape of the spectra (Thompson, 1984; Shaw and Kratochil, 1990; Lyder et al., 2010).

Specifically, such models must take into consideration two types of variability:

1. Variation due to viewing geometry/illumination and sample texture. Samples from different mines and different sites in mines will possess different textural features, such as clumpiness, which affect the extent of shadowing and overall brightness. In addition, changes in particle size impact overall reflectance (Clark, 1999) and spectral contrast (Hunt and Salisbury, 1970; Salisbury et al., 1991; Salisbury, 1993). In Lyder et al. (2010), where hyperspectral data was analysed, this variation was implicitly removed as part of the continuum via Gaussian or wavelet analysis; but in a broad band system the variation would have to be removed by other means. This work examines the mean normalisation of each band comprising the model to the mean value of all of the selected bands (mean normalised). For each spectrum, this approach should reduce the absolute variability in these bands by scaling them to their mean value.
2. Variation in the continuum due to water and clays. As noted above, the sample texture and illumination/viewing system affects the continuum, but the contribution to bitumen features from water and/or clays must also be accounted for. Two ratios are considered in this study: A simple ratio of a

mean normalised oil band to a mean normalised water band and a simple ratio of a second mean normalised oil band to a mean normalised clay band. As an alternative approach, we also examine the use of normalised indices, that is, the difference between the mean normalised reflectance for an oil feature and that of a water feature divided by their sum, and the difference between the mean normalised reflectance for a second oil feature and that of a clay feature divided by their sum.

With these considerations, linear models have been developed that estimate TBC with good accuracy, that is, $\sigma < 1.5\%$, and that are robust, that is, independent of mine location. Results are presented for a relatively small sample population to contrast the merits of models based on features derived from Gaussian as opposed to wavelet analysis. A larger sample population is then used to focus on features derived from wavelet analysis, and to build a five-band model relating spectral features to bitumen content. The performance of this five-band model is compared to a modern two-band model derived from the Wright and Wright method as well as an improved version of this two-band model that makes use of mean normalised bands.

EXPERIMENTAL

Sample Suites

Models were examined that were derived from two sample collections. In Collection One, 51 samples were available that were broken down into three suites: a model suite of 21 samples used to construct Gaussian and wavelet models, and a validation suite of 15 samples and a blind suite of 15 samples used to validate the model. The model suite comprised samples from multiple mines, while the samples of the blind suite and validation suite were obtained from a separate mine not included in the model suite. In this study, the samples that were scanned had been comilled and dry-mixed to homogenise the ore. Actual ore will exhibit some variability, especially for ore that is more heterogeneous. Routine laboratory analyses were conducted to measure the bitumen, water and solid mass fraction in each sample. True bitumen content values were found by Dean-Stark analysis from samples taken from the same ore used for spectral analysis (Bulmer and Starr, 1979). The ore samples are typically homogenised by dry mixing before taking the analysis sample, so as to reduce the possibility of subsampling errors. TBC ranged from 3.43% to 16.26% with water content varying from 0.56% to 11.82%. The distribution of TBC was fairly uniform over the sample range for the model suite and blind suite but clustered into a more bimodal distribution for the validation suite (at $\sim 7\%$ and 15% TBC).

Collection Two had 85 new samples from multiple locations as a modelling suite used to construct only a wavelet model, with the 51 samples of Collection One used as a validation suite. This almost tripled the number of samples, while maintaining the diversity of mine provenance to further test the general applicability, that is, site independence of the derived bitumen estimation models.

Spectral Measurement

Reflectance spectra were acquired with a Bomen MB102 Fourier Transform Infrared (FTIR) spectrometer equipped with a Mercury/Cadmium/Telluride (MCT) detector. The spectral coverage was set between 1.667 μm (6000 cm^{-1}) and 20 000 μm (500 cm^{-1}) with a spectral resolution of 8 cm^{-1} . The instantaneous field of

view was a circle with a diameter of 20 mm. Thirty-two consecutive scans for a given location were averaged to provide one measurement. The reflectance spectra were obtained from the ratio of each measurement to that of an illuminated diffuse gold panel of known reflectance taken with the same geometry. The final spectra considered in the modelling were generated by averaging six measured spectra, each obtained at a different location on the sample. The reproducibility of a measurement was assessed by collecting seven measurements of the infragold panel collected at an interval of 5 min within a period of 35 min. The rms error over the entire spectral region is between 0.06% and 1.09% reflectance, absolute, with a mean value of 0.45%. Additional processing was required to convert the hyperspectral data into suitable broad band data. This was accomplished by convolving the spectra with Gaussians of unity amplitude with the appropriate centre and width.

For the Gaussian analysis the fitting of Collection One spectra was performed using the commercial package Peakfit, Version 4.12 (SeaSolve Software, Inc., Framingham, MA). Spectra were converted to absorbance as Peakfit is designed to fit to spectral peaks rather than troughs. Gaussian fitting to the spectra was done in wavelength space conducted over the range 2.23–2.60 μm . Figure 1 shows the average absorption spectra for the suites of model, validation, and blind data. Peakfit was allowed to make an initial guess as to the number of peaks and peak locations on the basis of inflection points within the spectra for a given continuum type, for example, constant, linear or quadratic, and using Savitsky–Golay smoothing (Savitsky and Golay, 1964) and amplitude rejection threshold. The amplitude rejection threshold for a given smoothing window serves as a threshold above which Peakfit will consider peaks as being real and is used by Peakfit when selecting portions of the spectra to which it fits the continuum. From a visual inspection of this initial fit to the spectra, groups of peaks (peaks that were similar in shape but slightly different in location) were selected that were common to all spectra as a template. A second Gaussian fitting (to all spectra) was then undertaken. The peaks allowed were then constrained in number to match the number of peaks identified in the preliminary fitting with the peak locations now fixed to the average locations for all spectra, while peak amplitudes and widths and continuum types were allowed to vary for each spectra. For each peak fitted with a Gaussian for a given continuum type, estimates were found of the Gaussian full-width at half-maximum (FWHM), the Gaussian area as determined by numerical integration (Area_{int}) and by an analytical solution (Area_{ana}), and the Gaussian amplitude (Amp). These parameters were then exploited to derive the Gaussian-based, broad band models.

Wavelet analysis was conducted on the spectra of both collections in order to improve the detection of weak spectral features. Each reflectance spectrum in wavelength is represented as a sum of similar wave-like functions (wavelets) that are equivalent to the variation within a window of varying size, or scale, shifted over the entire spectrum (Figure 2). The wavelet power then describes the strength of correlation between the wavelet shape and the spectrum at a given scale, which in turn may be correlated to the quantity of interest, that is, TBC. For the wavelet-based broad band models the feature centres corresponded to the locations of peak correlation between the wavelet power and TBC while the width is determined by the width where the correlation is significant. Table 1 provides a summary of the relevant feature parameters for both the Gaussian-based and wavelet-based broad band models and the associated correlation thresholds.

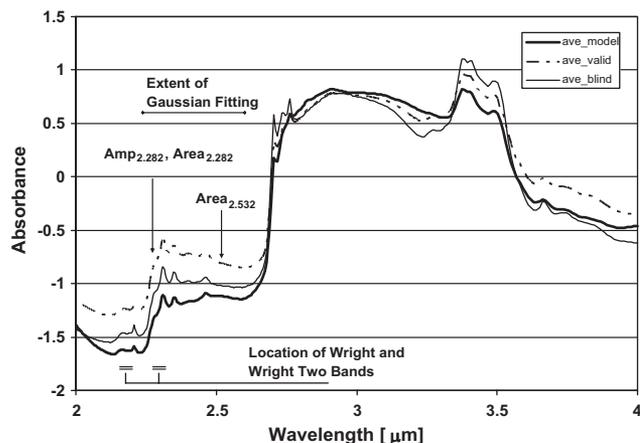


Figure 1. Average absorption spectra of model suite (21 samples), validation suite (15 samples) and blind suite (15 samples) comprising Collection One. The range over which Gaussian fitting was attempted as well as the Gaussian peak parameters used in the modelling are shown. For comparison the location and width of the Wright and Wright two-band spectrometer are also shown.

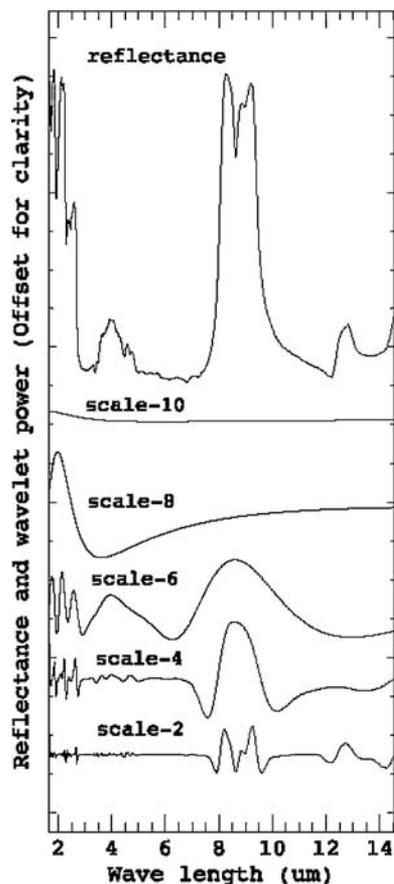


Figure 2. Example of the decomposition of a reflectance spectrum for the model sample C712 (16.26% TBC) into the wavelet power. Lower valued scales represent the variation of the spectrum over small scales. Larger scales represent the large structure usually identified with the baseline or continuum.

Table 1. Individual detectable features

Feature ID	Feature centre (μm)	Physical interpretation	Width (μm)
Wavelet: O1	1.710	Bitumen	0.160
Wavelet: O2	1.754	Bitumen	0.240
Wavelet: W1	2.054	Water	0.320
Wavelet: C1	2.171	Clay	0.070
Wavelet: C2	2.210	Clay	0.034
Wavelet: O3	2.274	Bitumen	0.061
Wavelet: O4	2.396	Bitumen	0.162
Wavelet: W2	2.770	Water	0.157
Wavelet: O5	3.725	Bitumen	0.118
Gaussian	2.282	Clay, bitumen	0.038
Gaussian	2.532	Atmosphere, bitumen	0.072

For wavelet analysis, features have power correlated/anti-correlated to modelling suite TBC values and $R > 0.85$ for bitumen features and > 0.80 for water and clay features. O, oil; W, water; C, clay.

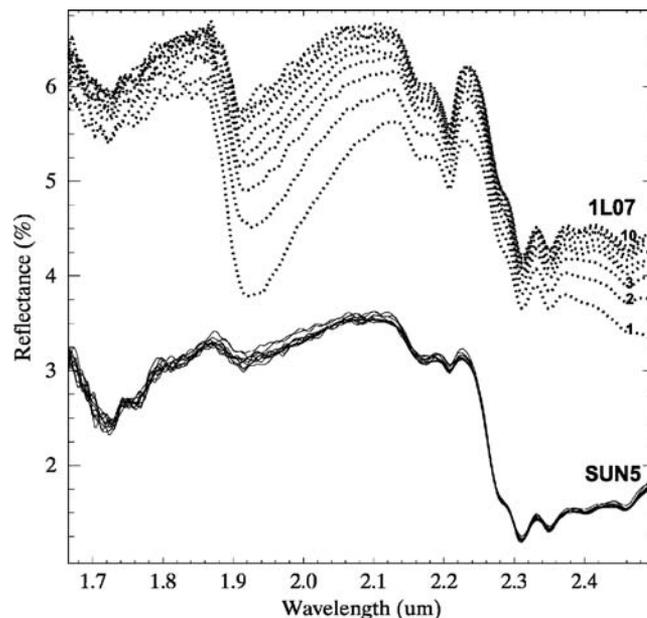


Figure 3. Drying experiment for sample SUN5 (1.99% water) and sample #1L07 (7.07% water). Spectra are labelled from 1 (0 min elapsed) to 10 (45 min elapsed) and illustrate the effect of drying.

ANALYSIS AND RESULTS

Effect of Liquid Water

Figure 3 displays spectra collected from two samples (1L07, 7.07% water and SUN5, 1.99% water) exposed to ambient air and measured at time intervals of 5 min for 45 min. These data illustrate the spectral influence of water on oil sand spectra. The most substantial spectral changes occur early in the drying process (e.g., first 5 min) and are most noticeable for the sample with lower bitumen content (1L07). For sample 1L07 the mean reflectance of the 1.67–2.5 μm region increases from 4.93% (1st spectrum) to 5.89% (10th spectrum). The magnitude of the change is determined by the water content in the sample, the rate of evaporation and the bitumen content. Samples with higher bitumen content display less change in the spectrum during drying. For sample 1L07, the shapes of the spectra are similar except near 1.91 and 2.21 μm where the feature depth is affected respectively by liquid water and clays.

Results for Dataset Collection One (51 Samples)

Construction of simple Gaussian-based broad band models

Broad band models were constructed from the Gaussian analysis (peak centre, amplitude, and width) of hyperspectral data for absorption features at 2.282 μm (clay and bitumen) and 2.532 μm (bitumen) (Table 1) (Lyder et al., 2010). Figure 4 shows the linear models of estimated TBC obtained from these two features using two variations of optimization: the L2 norm or simple least-squares optimization and the L1 norm or minimisation of greatest outlier optimization. Table 2 summarises the relevant statistics for these models. For all three data sets the general trends are in agreement but grossly overestimate low TBC values and underestimate high TBC values. There also remains a considerable offset in the fit between the three data sets. The L1 norm optimization does appear to be marginally superior to the L2 norm optimization in the sense that the mean offset (μ) for both the validation and blind suites are lower.

Construction of simple wavelet-based broad band models

In an analogous manner a simple broad band solution (no mean normalisation) was based on three features at 2.274, 2.396, and 3.725 μm identified from the wavelet analysis of the hyperspectral data (Table 1). Figure 5 is a plot of the estimated TBC obtained with the resulting models (L2 norm—left panel, L1 norm—right panel). As before, Table 2 provides the relevant statistics for these models. Using these features there is improvement in the estimated TBC values compared to that derived using the Gaussian features, particularly with respect to the mean offset (μ). Given the additional information available to this wavelet-based model, it is reasonable to expect it to perform better as a classifier than the simple Gaussian-based model. However, there is a significant difference in the trend of the blind data suggesting the solution may not be robust.

Construction of normalised wavelet-based broad band models

Given the better performance of the wavelet-based model further Gaussian-based modelling was not pursued. To address variations in the continuum due to variable factors such as variable sample content in water and clays, a new model was developed based on five wavelet-based mean-normalised bands. As a first step, the reflectance in five bands (C2, O3, O4, W2, and O5—Table 1) that included a clay band (C2), a water band (W2) and three oil bands ThTT were each normalised by the mean of all five bands. Then, two indices were formed to account for the influence of clay and water, respectively, on the oil reflectance. These are defined as:

$$\text{IND}_{2.274} = \frac{R_{2.210} - R_{2.274}}{R_{2.210} + R_{2.274}} \quad (1)$$

and

$$\text{IND}_{3.725} = \frac{R_{3.725} - R_{2.770}}{R_{3.725} + R_{2.770}} \quad (2)$$

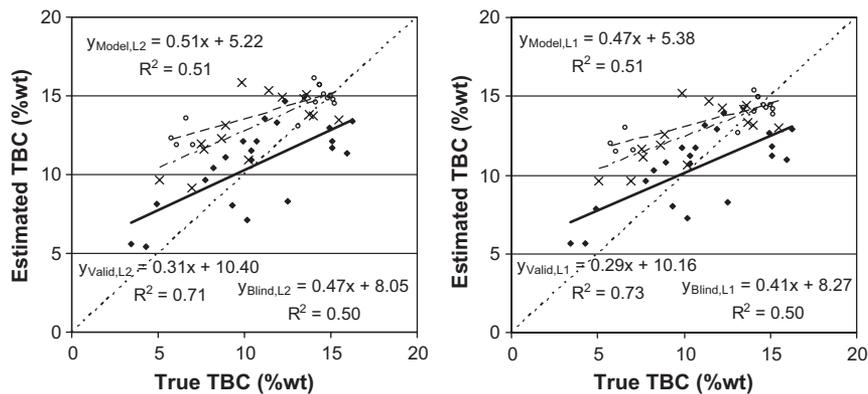


Figure 4. Simple Gaussian two-band model (2.282 and 2.532 μm). Here the fit is determined using the classical least-squares criteria of normalising the distribution of the residuals (left L2 norm optimization) and minimising the greatest residual (right L1 norm optimization). The solid line represents the fit to the model set (diamonds), the long-dash line to the validation set (circles), the long-short-dash line the blind set (crosses) and the short-dash line the line of perfect prediction.

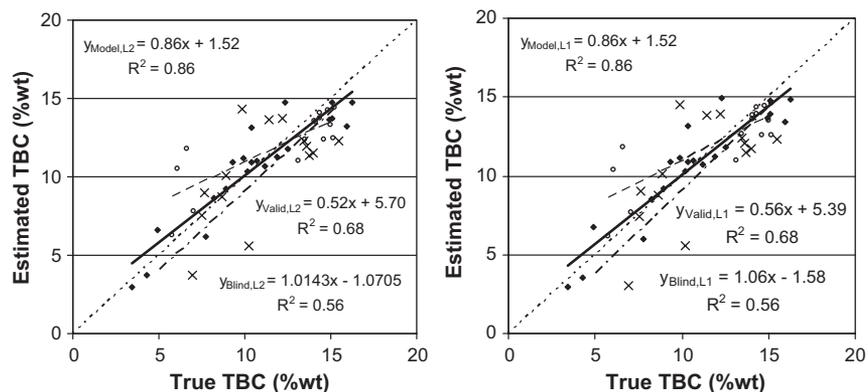


Figure 5. Simple wavelet three-band model (2.274, 2.396 and 3.725 μm), using L2 norm (left) and L1 (right) norm optimizations. The solid line represents the fit to the model (diamonds), the long-dash line to the validation set (circles), the long-short-dash line the blind set (crosses) and the short-dash line the line of perfect prediction.

$\text{IND}_{2.274}$ measures the change in the mean-normalised reflectance of the oil feature at 2.274 μm (O3) relative to the mean-normalised reflectance of the clay absorption at 2.210 μm (C2). $\text{IND}_{3.725}$ measures the change in the mean-normalised reflectance of the oil feature at 3.725 μm (O5) relative to the water absorption at 2.770 μm (W2). These two indices, and the mean-normalised reflectance ($R_{2.396}$) at 2.396 μm (O4, the broadest bitumen absorption feature in the NIR), are all correlated to bitumen content. Since these indices are independent of one another, they may be used to form a linear model to estimate TBC.

The results of this fit are shown in Figure 6. The statistics for the normalised wavelet-based 5 band broad band models are provided in Table 3. Using the L2 norm (Figure 6, left) the model shows a good correlation between the estimated and true TBC

but there remains an offset between the blind and validation estimates and the model estimates for high TBC values. Using the L1 norm (Figure 6, right) removes this discrepancy and produces an excellent solution ($\sigma \sim 1.5\%$ TBC) that is robust and superior to the non-normalised Gaussian-based and wavelet-based models.

Discussion based on dataset Collection One (51 samples)

From Figure 4 it is clear that the simple Gaussian-based broad band models provide a predictive capability of TBC, that is, similar for all three data suites though with significant dispersion (σ). The Gaussian broad band models, however do perform more poorly relative to hyperspectral Gaussian models ($\sigma \sim 1.2\text{--}1.5\%$; Lyder et

Table 2. Summary of statistics for simple two-band Gaussian-based and simple three-band wavelet-based models

Broad band base	Optimization	Model		Validation		Blind	
		μ	σ	μ	σ	μ	σ
Gaussian	L1	0.24	2.59	-1.44	2.73	-2.06	2.23
	L2	0.00	2.58	-2.00	2.66	-2.47	2.20
Wavelet	L1	-0.01	1.39	0.04	2.72	0.90	2.96
	L2	0.00	1.39	0.15	2.23	0.92	2.82

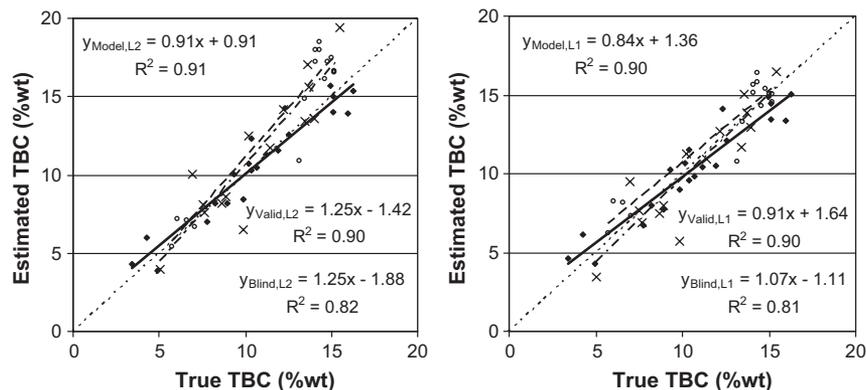


Figure 6. Calculation of TBC using indices constructed from wavelet-based five broad bands normalised to the mean value of the five bands using the L2 norm (left) and L1 (right) norm optimizations. The solid line represents the fit to the model (diamonds), the long-dash line to the validation set (circles), the long-short-dash line the blind set (crosses) and the short-dash line the line of perfect prediction.

al., 2010). The broad band models may not be as accurate in modelling the continuum and will be susceptible to contamination by non-bitumen features.

In regards to the simple wavelet-based model (Figure 5) the explicit use of a clay band ($2.210 \mu\text{m}$) and water band ($3.725 \mu\text{m}$) provide a direct measurement of these quantities and their influence on the reflectance spectra. However, the high dispersion in the estimated TBC values of the blind suite and potential bias in the estimated TBC values of the validation suite suggests that at least one of the bands is not as simple as proposed and there may exist some overlap in the selected broad bands with features from other minerals present.

The five-band model (Figure 6) was based upon the combination of five mean normalised wavelet-based broad bands seems to be effective in removing the influence of clays using the $IND_{2.274}$ index and the influence of water using the $IND_{3.725}$ index, if judged by the reduced dispersion in the L1 and L2 norm errors ($\sim 1.2\text{--}1.5\%$), in comparison to the simple wavelet-based model. It is interesting to note that only the L1 norm appears to produce a robust model in the sense that the estimated TBC values follow the same trend for all data suites. The L2 norm model does appear to produce similar dispersion for all three data suites but the validation and blind TBC values are overestimated for high TBC. If the errors produced in the estimated values of TBC were normally distributed and of equal weight, which they are not, then the L2 norm would be expected to be the superior model. The L1 norm optimization serves to reduce the influence of outliers on the model by varying the weight given to the model samples. Samples with low TBC that produce faint bitumen features may prove to be difficult to model once the bitumen spectral feature is diluted by the broad band sampling. Samples with high TBC may be difficult to model if the bitumen features are quite broad and the complete spectral response is not adequately captured by the

broad band sampling. The L1 norm will attempt to compensate for these differences by varying the weight given to the samples in the modelling suite while the L2 norm will try and distribute the error equally to all of the samples.

Results for Collection Two (136 Samples)

The wavelet-based five-band model established with Collection One made use of two features beyond $2.5 \mu\text{m}$ (2.77 and $3.725 \mu\text{m}$). Using Collection Two we aimed to develop a model based solely on bands located in the short wave infrared region and thus make use of a single detector technology for the potential implementation of a monitoring system. We then conducted a comparison with a two-band model derived from the Wright and Wright method in use in industry today and devised an improved version of this two-band model.

Normalised wavelet based five-band broad band model

The five band model applied to Collection One made use of three bitumen features (O3, O4, O5), one clay feature (C2) and one water feature (W2) (Table 1). Features O5 and W2 lie beyond $2.5 \mu\text{m}$ and were replaced by a bitumen feature at $1.754 \mu\text{m}$ (O2) and a water feature at $2.054 \mu\text{m}$ (W1), respectively. The bitumen feature at $1.710 \mu\text{m}$ (O1) was not considered because of the poorer signal-to-noise level of reflectance spectra. C2 captures the absorption feature of kaolinite and illite clays, both of which are present in the Athabasca oil sands (Cloutis et al., 1995). The reflectance value of each was then normalised to the mean of the five bands to address variations in the continuum.

Three spectral indices were constructed from the normalised bands to measure the strength of the O2 ($1.754 \mu\text{m}$), O3

Table 3. Summary of statistics for normalised five-band wavelet-based models

Broad band base	Optimization	Model		Validation		Blind	
		μ	σ	μ	σ	μ	σ
Normalised	L1	0.34	1.17	-0.47	1.20	0.34	1.61
	L2	0.00	1.08	-1.60	1.77	-0.78	1.96

The bold entry indicates the best model found in this study.

Table 4. Broad band models and their performance

Model	Bands	Slope ^a	R ² range	Model		Validation	
				μ _{Model}	σ _{Model}	μ _{Valid}	σ _{Valid}
Normalised five-band	5	0.99–1.00	0.71–0.93	0.00	1.04	−0.33	1.52
Two-band (Ratio)	2	0.97–1.08	0.69–0.75	0.00	1.81	−0.92	2.02
Improved two-band models							
Model 1	2	0.98–1.01	0.80–0.85	0.00	1.40	−0.15	1.47
Model 2	2	0.97–1.01	0.73–0.78	0.00	1.67	−0.39	1.73

^a Slope of the trend between predicted and observed TBC, which is expected to be 1.0 for an unbiased estimation.

(2.274 μm), and O4 (2.396) bitumen absorption features:

$$\text{IND}_{2.274} = \frac{R_{2.210} - R_{2.274}}{R_{2.210} + R_{2.274}} \quad (3)$$

$$\text{IND}_{1.754} = \frac{R_{2.054} - R_{1.754}}{R_{2.054} + R_{1.754}} \quad (4)$$

and,

$$\text{Ratio}_{2.396} = \frac{R_{2.210}}{R_{2.396}} \quad (5)$$

IND_{1.754} measures the change of reflectance at 1.754 μm (O2 bitumen feature) relative to the water absorption at 2.054 μm (W1). IND_{2.274} measures the change of reflectance at 2.274 μm (O3 bitumen feature) relative to the clay absorption at 2.210 μm (C2). R_{2.396} measures the change of reflectance at 2.394 μm (O3 bitumen feature) relative to the clay absorption at 2.210 μm (C2). These three indices were combined into a single L1 optimized model using the new 85 samples as the modelling suite and applied to the 51 sample validation suite (essentially Collection One). Results (Table 4 and Figure 7) show that this model generates a mean error of up to 0.33% and a standard deviation of up to 1.52%. The slope of the data trend (R² = 0.71) between predicted and observed TBC is 1.00 indicating that this model is unbiased.

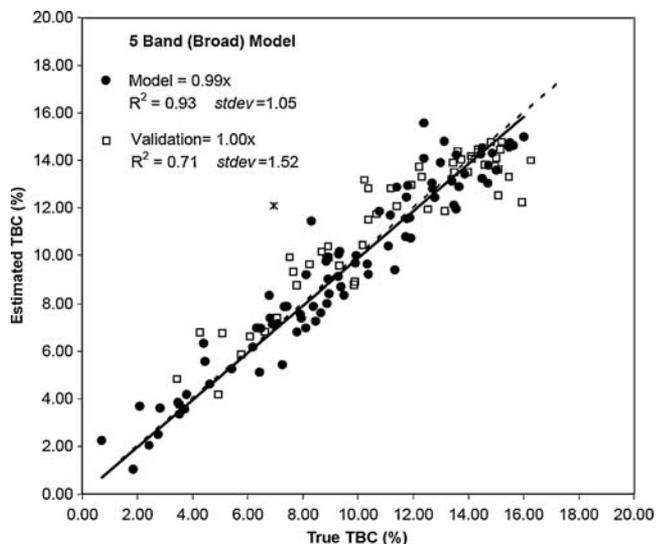


Figure 7. Performance of the wavelet-based normalised five-band broad band model derived from Collection Two. The X label refers to the validation sample with highest error (#1B07) which has an estimated TBC of 12.09% and a true TBC of 6.94%.

Two band ratio and improvements

Comparison of two band ratio with normalised five-band model. The industry has made use of a reflectance ratio of two bands centred at 2.22 and 2.33 μm (Ratio = $R_{2.22}/R_{2.33}$) with a FWHM of 70 nm for each band. The 2.33 μm feature is a strong bitumen feature and the 2.22 μm band lies on the wing of a clay feature. When this method is applied to the Collection Two suites, the correlation between predicted and observed TBC is relatively strong with R² ranging from 0.69 to 0.75 (Table 4) for the model and validation suites, which suggests this ratio is a fairly robust predictor of TBC. However, the agreement in the regression slope for the modelling and validation suite (0.97–1.08) is not as good when compared with the wavelet-based normalised five-band model (0.99–1.0). The mean error and standard deviation are also larger (−0.92, 2.02) than for the normalised five-band model (−0.33, 1.52). Hence, the simple two-band (Ratio) is neither as precise nor as robust as the wavelet-based normalised five-band model.

Comparison with improved two band model. Given the merits of the mean normalisation to minimise variability in the continuum we examined the applicability of this method to the two-band (Ratio). We calculated the mean normalised reflectance at 2.22 and 2.33 μm and the difference of these bands ($D^n = R_{2.22}^n - R_{2.33}^n$) to estimate the strength of the 2.33 μm feature relative to the 2.22 μm feature. From these variables, two new two-band ratio models were derived using an L1 optimization of the 85 model samples:

$$\text{TBC} = 7.03 + 2.03 \times R_{2.22} - 3.89 \times R_{2.33} + 11.56 \times D^n, \quad (6)$$

$$\text{TBC} = 12.25 - 12.55 \times \text{Ratio} + 41.21 \times D^n. \quad (7)$$

Both models make use of the difference of the normalised reflectance but model 1 (Equation 6) also uses the absolute reflectance at 2.22 and 2.33 μm while model 2 uses the two-band (Ratio) of absolute reflectance values. The performance of these two models on the 51 validation samples is shown in Figure 8 and described in Table 4. Model 1 outperforms model 2 with a lower mean error (0.15 vs. 0.39), lower standard deviation of errors (1.47 vs. 1.73) and higher R² (0.80–0.85 vs. 0.73–0.78). Both models shows an improvement over the two band (Ratio) but not when compared to the wavelet-based normalised five-band model.

The use of normalised bands in models 1 and 2 does appear to offer a better correction for the effects of water and clay than the standard two-band (Ratio). A significant benefit of this improvement is that it can be implemented on data acquired by existing two band spectrometers through software modifications without requiring hardware modifications and implementations of new bands.

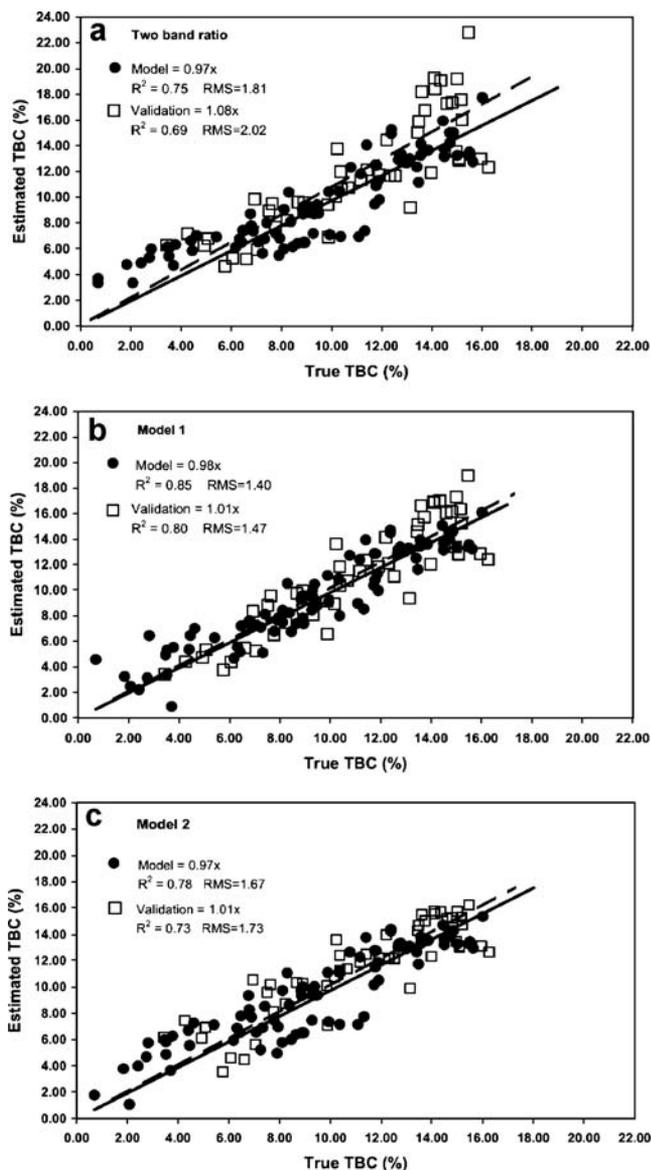


Figure 8. Performance of two-band (Ratio) method (a) and potential improvements based upon two different two-band models using a mean normalised variable (b) and (c).

CONCLUSIONS

In this study we presented broad band predictive models of total bitumen content of oil sands based on band centres and widths guided by our previous analysis of the hyperspectral NIR and SWIR reflectance spectra (Lyder et al., 2010). The results presented here show that simple broad band models based upon previously identified Gaussian features or wavelet features provide an incremental improvement over the currently deployed industry two-band ratio model. We were able to provide improved two-band models that made use of a combination of the same two bands but now normalised to their mean. A wavelet-based broad-band five-band model comprised of indices and bands, where the bands were normalised to the mean of the bands, was adequate to address the influence of water, clay and textural variation on selected bitumen features. This five-band model makes use of the L1 norm optimization to vary the weight given to the model samples to adequately compensate for limitations in the spectral

response to TBC at the extreme values of TBC. This model appears to produce the most robust estimator of TBC with a dispersion of ~ 1.1 – 1.5% which can be applied to different sites within a mine and to different mines without additional tuning or calibration as illustrated by regression slopes of 0.99–1.0 for all data sets.

True bitumen content is not the only factor in predicting the effectiveness of aqueous separation processes. Drilling programs for geological block modelling and mine planning include assaying for both bitumen and fines (usually defined as sub $44\ \mu\text{m}$ particles, regardless of composition and surface activity, although some laboratories will also test for surface activity using methods such as methylene blue). Future work will consider estimation of parameters for the clay fines fraction, to provide additional on-line features for operational control.

ACKNOWLEDGEMENTS

Syncrude, Suncor and the GEOmatics for Informed Decisions (Geoide) Network of Centres of Excellence of Canada contributed financially to the support of this research.

END NOTES

Absorption = $\ln(1/\text{Reflectance})$.

REFERENCES

- Bichard, J. A., "Oil Sand Composition and behavior Research," in "AOSTRA Technical Publication Series 4," Chapter 1, Alberta Oil Sands Technology and Research; Authority, Edmonton, Canada (1987), pp. 6–62.
- Bulmer, J. T. and J. Starr, "Syncrude Analytical Methods for Oil Sand and Bitumen Processing in Methods 2.9 and 2.10: Classical and Syncrude Methods for the Determination of Bitumen," Water and Solids Content of Froth, AOSTRA (1979).
- Clark, R. N., "Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy," in A. N. Rencz, Ed., "Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences," John Wiley and Sons, New York (1999), pp. 3–58.
- Cloutis, E. A., "Spectral Reflectance Properties of Hydrocarbons: Remote Sensing Applications," *Science* 245, 165–168 (1989).
- Cloutis, E. A., Gaffey, M. and Mosiow, T., "Characterization of minerals in oil sands by reflectance spectroscopy," *Fuel* 74, 874–879 (1995).
- Fong, N., S. Ng, K. H. Chung, Y. Tu, Z. Li, B. D. Sparks and L. S. Kotlyar, "Bitumen Recovery From Model Systems Using a Warm Slurry Extraction Process: Effects of Oil Sands Components and Process Water Chemistry," *Fuel* 83, 1865–1880 (2004).
- Hepler, L. G. and C. His, "AOSTRA Technical Handbook on Oil Sands, Bitumen's, and Heavy Oils," in AOSTRA Technical Publication Series 6, Chapter 2, Alberta Oil Sands Technology and Research Authority, Edmonton, Canada (1989), pp. 14–20.
- Hunt, G. R. and J. W. Salisbury, "Visible and Near-Infrared Spectra of Minerals and Rocks: I. Silicate Minerals," *Mod. Geol.* 1, 283–300 (1970).
- Lyder, D., J. Feng, B. Rivard, A. Gallie and E. Cloutis, "Remote Bitumen Estimation of Athabasca Oil Sand From Hyperspectral Infrared Reflectance Spectra Using Gaussian Singlets and Derivative of Gaussian Wavelets," *Fuel* 89, 760–767 (2010).

- Salisbury, J. W., "Mid-Infrared Spectroscopy Laboratory Data: Laboratory Data, Chapter 4," in C. Pieters and P. Englert, Eds., "Remote Geochemical Analysis: Elemental and Mineralogical Composition," Cambridge University Press, Cambridge (1993), pp. 79–98.
- Salisbury, J. W., L. S. Walter, N. Vergo and D. M. D'Aria, "Infrared (2.1–25 μm) Spectra of Minerals," Johns Hopkins University Press, Baltimore, Maryland (1991), pp. 294.
- Savitsky, A. and M. J. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Anal. Chem.* **36**, 1627–1639 (1964).
- Schramm, L. L., E. N. Stasiuk, M. MacKinnon, "Surfactants in Athabasca Oil Sands Slurry Conditioning, Flotation Recovery, and Tailings Processes," in L. L. Schramm, Ed., "Surfactants: Fundamentals and Applications in the Petroleum Industry," Cambridge University Press, Cambridge, UK (2000).
- Shaw, R. and B. Kratochil, "Near Infrared Diffuse Reflectance Analysis of Athabasca Oil Sand," *Anal. Chem.* **62**, 167–174 (1990).
- Thompson, G. R., "Method and Apparatus for On-Line Monitoring of Bitumen Content in Tar Sand," Canadian Patent 1139702 (1984).
- Torrence, C. and G. P. Compo, "A Practical Guide to Wavelet Analysis," *Bull. Am. Meteorol. Soc.* **79**, 61–78 (1998).
- Yang, P. W., H. H. Mantsch, L. S. Kotlyar and J. R. Woods, "Characterization of Athabasca tar Sand Malthenes by Diffuse Reflectance Infrared Spectrometry," *Energy Fuels* **2**, 26–31 (1988).

Manuscript received May 7, 2009; revised manuscript received November 26, 2009; accepted for publication November 30, 2009.