minerals

MDPI

Article

Near-Infrared Spectroscopy of Limestone Ore for CaO Estimation under Dry and Wet Conditions

Sungchan Oh [1], Chang-Uk Hyun [2,*] and Hyeong-Dong Park [3]

[1] Seoul National University Research Institute of Energy and Resources, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea; scoh@snu.ac.kr
[2] Korea Polar Research Institute, 26 Songdomirae-ro, Yeonsu-gu, Incheon 21990, Korea;
[3] Department of Energy Systems Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea; hpark@snu.ac.kr
* Correspondence: changuk.hyun@gmail.com; Tel.: +82-32-760-5348

Received: 16 September 2017; Accepted: 11 October 2017; Published: 13 October 2017

Abstract: Quantitative analysis of CaO in limestone mining is mandatory, not only for ore exploration, but also for grade control. A partial least squares regression (PLSR) CaO estimation technique was developed for limestone mining. The proposed near-infrared spectroscopy (NIR)-based method uses reflectance spectra of the rock sample surface in the visible to short-wave infrared wavelength regions (350–2500 nm (4000–28,571 cm^{-1})) without the need to crush and pulverize the rock samples. The root mean square (RMS) error of CaO estimation using limestone ore fragment was 1.2%. The CaO content estimated by the PLSR method was used to predict average CaO content of composite samples with a sample size of 15, which resulted in an RMS error of 0.3%. The prediction accuracy with moisture on sample surfaces was also examined to find out if the NIR-based method showed a similar RMS error. Results suggest that the NIR technique can be used as a rapid assaying method in limestone workings with or without the presence of groundwater.

Keywords: NIR spectroscopy; partial least squares regression; limestone; moisture effect

1. Introduction

Limestone is a common sedimentary rock primarily composed of calcium carbonate ($CaCO_3$) with lesser amounts of calcium magnesium carbonate ($CaMg(CO_3)_2$). Depending on the various mineralogical and chemical compositions of limestone, the extracted ore is used as a source material for a wide range of products such as cement, lime, glass, metallurgical flux, filler and extender. Because the product manufacturing process requires an appropriate chemical composition of limestone, grade control is an important component in the extraction stage of limestone mining.

Chemical analysis of limestone is frequently conducted by X-ray fluorescence (XRF), more specifically, by energy-dispersive XRF, in limestone mining and the cement industry. Detection limits of XRF are element-specific and technique-specific (energy-dispersive or wavelength-dispersive), however, in general, XRF can accurately measure a wide range of elements with detection limits within ppm levels. The sample preparation of XRF is comprised of grinding and pelletizing using a pressed powder method or grinding and fusing with lithium tetraborate ($Li_2B_4O_7$) for the glass bead method [1]. Results of XRF are usually reported by element name, e.g., Ca and Mg, whereas the results of the glass bead method can also be presented in oxide form, e.g., CaO and MgO, due to the release of CO_2 from carbonate minerals during fused bead production. It should be noted that CaO is not direct chromophore in the visible and infrared spectra, and, in this paper, CaO was intended to represent calcite quantity in limestone. The CaO concentration is usually reported in the range of 0–56 in percent oxide units (%), and its limit is c.a. 0.001%. Although accurate measurements can be made using XRF, the disadvantage is the considerable amount of time required for the analysis,

specifically in sample preparation. To reduce analysis time, a handheld XRF (HHXRF) tool was developed for greater versatility. Commercial HHXRF (Bruker S1 series, Olympus Delta series, etc.) analyzers are energy dispersive and designed for compactness and energy efficiency. However, obtaining reliable data from naturally broken rock is difficult due to X-ray attenuation on irregular surface geometry and the particle size effect on sample homogeneity, which is a general disadvantage of common XRF analysis. Moreover, HHXRF usually requires at least 30–60 s for each measurement [2].

Spectral analyzing techniques for CaO in limestone ore using near-infrared (NIR) spectroscopy and partial least squares regression (PLSR) can be used to improve the efficiency of chemical analysis in mining. The use of the NIR spectrum, i.e., the wavelength range from 350 to 2500 nm (4000–28,571 cm^{-1}), is proposed for this study for two reasons. First, carbonate minerals show characteristic absorption features in this region, and the amount of carbonate affects the intensity of the absorption feature [3,4]. Second, the use of a portable spectrometer enables the measuring of a single location in a few seconds, and the measurement can be performed using either a rock powder or solid rock sample.

Studies using the absorption features in NIR spectroscopy have been conducted, but they were not intended for grade checking or for immediate use. It was demonstrated that this approach was applicable for estimating sulfide ore grade [5]; however, the spectral measurements were performed on unweathered rock surfaces with reasonably homogeneous mineralogy. The chemical content in iron ore was estimated by analyzing the absorption properties of a synthetic powder mixture, which required grinding and sieving [6]. More recently, rock samples that were cut in half were used for sensor-based sorting [7].

More practical approaches have suggested the possibility of using NIR spectroscopy of rock surfaces in the determination of carbonate rock chemistry. Wavelength position of carbonate absorption feature was used to determine the chemical composition of Ca and Mg content from eight measured locations of four carbonate rocks [8]. Another study exploited absorption strength of carbonate feature to estimate CaO and Al_2O_3 for chemical quality control of Portland cement-grade limestone, in which 15 rock samples were used to generate a prediction model per each rock type, i.e., dark-gray, light-gray and dolomitic limestone [9].

A mathematical approach of spectral unmixing is another important method in hyperspectral imaging where spectral mixture of multiple endmembers should be considered. The number of spectral endmembers and their abundances can be estimated through spectral unmixing. An overview of the main techniques was thoroughly summarized by the authors of [10]. A recent method based on synthetic endmembers has been proposed to quantify the mineral abundances [11].

A data-driven estimation using PLSR was recently applied in this field. PLSR is used for modeling the relationship between independent and dependent variables by projecting them onto low-dimensional spaces as in principal component analysis (PCA) [12]. This approach has been successfully applied in the field of chemometrics, especially when the number of variables is greater than the observations, and the data are noisy and collinear. PLSR was implemented to quantify gangue mineral in copper ore with particle sizes ranging from 0.1 to 50 mm [13]. Likewise, serpentine content was measured in nickel laterite ores by using synthetic mineral mixtures [14].

This study aims to develop a rapid analyzing technique for CaO content in limestone ore using NIR reflectance spectroscopy and PLSR. We also examine the prediction accuracy, with or without water, on limestone surface to find out if the suggested method can be applicable in the presence of groundwater.

2. Materials and Methods

2.1. Limestone Ore Samples

The Donghae limestone mine operated by Daesung MDI is located in the eastern part of Gangwon Province in South Korea, where the geology primarily consist of the Cambro-Ordovician Joseon Supergroup in which limestone is predominant [15]. The excavated ore is microcrystalline or

fine-grained limestone with crystal size of 0.01–0.3 mm, and it shows blackish to pale gray color. X-ray diffraction (XRD) analysis showed that the main reason for the lower CaO content of the ore is the presence of reddish clay (comprised of kaolinite, illite and quartz) that infills fractures. The infilled rock fractures are few meters to tens of meter long and the range of joint frequency is c.a. 2–5 per meter.

The mined ore is supplied to Posco Inc. (Pohang, South Korea) as metallurgical flux for ironmaking. In the ironmaking process, silicate impurities derived from the clay minerals form blast-furnace slag, which decreases the overall efficiency of the process. To avoid this, limestone used in blast-furnace has specific chemical compositions, i.e., CaO + MgO > 54.0%, SiO_2 < 0.1% [16]. The chemical content of limestone in the study mine mostly satisfies the above conditions when CaO is over 51% along with average 3.7% MgO in ores. In this regard, Posco Inc. provides monetary incentives for ore delivered with a CaO content greater than 51%.

Limestone ore samples were collected at 100 locations at the study mine to capture the mineralogical variability and corresponding chemical composition of Pungchon limestone. For each location, ore fragments of approximately 12 cm diameter, i.e., 0.7–1.0 kg, were taken randomly from mine face or ore stock. From visual inspection of color, texture and mineral size of rocks, it was almost impossible to differentiate between lower-grade and higher-grade ore. The CaO content of the samples was measured by XRF analysis using a Rigaku Supermini 200 (Rigaku, Tokyo, Japan) and applying the glass bead method. The range of CaO content based on the XRF measurements was 48.1–53.9%.

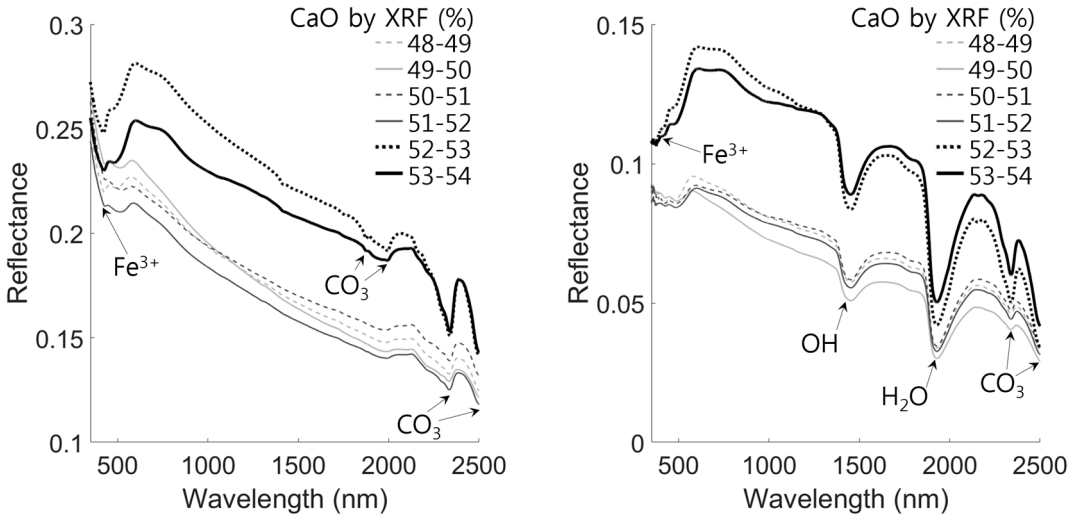
2.2. NIR Spectra of Ore Samples

NIR spectra of ore samples were measured using an ASD FieldSpec 3 portable spectrometer (Analytical Spectral Devices, Inc., Boulder, CO, USA). This instrument is configured in the spectral range of 350 to 2500 nm (4000–28,571 cm^{-1}), with a spectral resolution of 3 nm at 700 nm (14,286 cm^{-1}) and 10 nm at 1400 (7143 cm^{-1}) and 2100 nm (4762 cm^{-1}), with a sampling interval of 1.4 nm in the 350–1050 nm (9524-28,571 cm^{-1}) range and 2 nm in the 1000–2500 nm (4000–10,000 cm^{-1}) range. The measured spectrum is interpolated by 1 nm intervals by an appropriate cubic spline function [17] in order to obtain 2151 reflectance values in the 350–2500 nm (4000–28,571 cm^{-1}) wavelength region.

Limestone spectra were measured in dry and wet surface conditions with consideration of underground mining conditions where groundwater can exist. For dry conditions, samples were dried at room temperature, and, for wet conditions, rocks were soaked in tap water for 10 s. Surface clay was not cleaned or shaken-off in any process step. No additional treatment on rock surfaces was applied, e.g., polishing or grinding, to conserve natural rock surface. Ten spectra were obtained from each sample, and the measuring locations were distributed as consistently as possible. Each spectrum was an averaged spectrum of 30 measurements. A white reference standard made of Spectralon (polytetrafluoroethylene, PTFE) was used for calibration for every 10 measurements. The ASD contact probe was used to measure the spectral reflectance of limestone surfaces with a constant spot size of 20 mm. This contact probe has an internal halogen light source with a fixed viewing geometry, so bias due to illumination and measuring geometry is minimized.

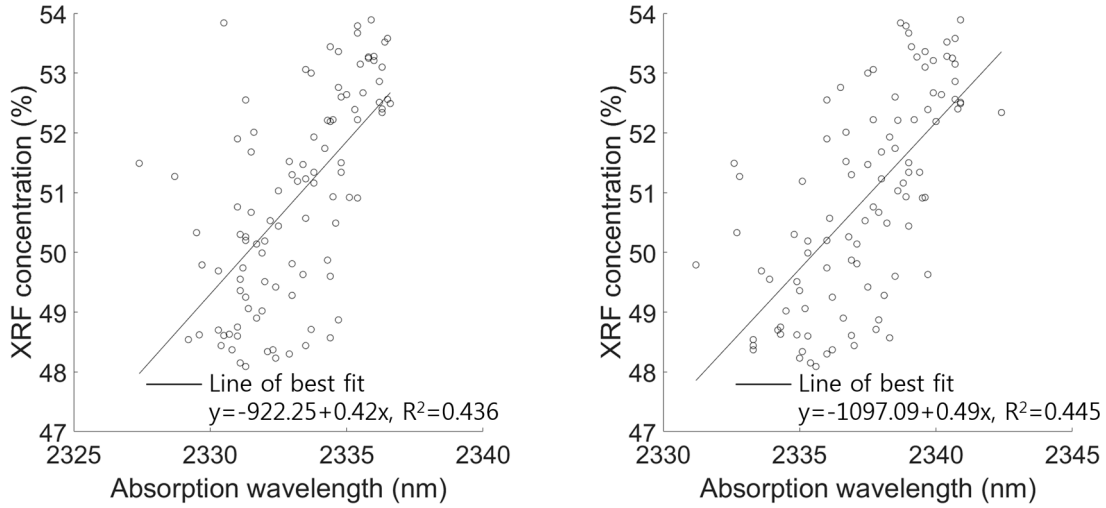
Limestone ore spectra in this study showed characteristic absorption features of calcite and clay minerals (Figure 1; [4,18]). Calcite features at 1850, 2000, 2350 and 2500 nm (5405, 5000, 4255 and 4000 cm^{-1}) are commonly observed, and additional three features at 1944, 2041 and 2219 nm (5145, 4900 and 4505 cm^{-1}) that might be anharmonically coupled to a lattice mode in calcite exist in the NIR spectrum of limestone [4,19]. In this study, calcite features at 2350 and 2500 nm (4255 and 4000 cm^{-1}) were observed from NIR spectra in both moisture conditions, but other calcite features at 1850, 2000 nm (5405 and 5000 cm^{-1}) were only seen from dry spectra. This is because a broad water feature at 1900 nm (5263 cm^{-1}) suppressed calcite absorptions at 1850 and 2000 nm (5405 and 5000 cm^{-1}). Calcite absorption at 1944, 2041 and 2219 nm (5145, 4900 and 4505 cm^{-1}) was subtle. Absorption in 350–700 nm (14,286–28,571 cm^{-1}) was due to ferric ion in illite, and this feature was observed regardless of moisture conditions. For rough prediction of CaO from spectral features, we plotted absorption strength, wavelength and area with respect to XRF CaO, but none of the results showed

a coefficient of determination over 0.45. Figure 2 shows regression results of the average absorption wavelength of 2350 nm (4255 cm^{-1}) feature against CaO content.



(**a**) (**b**)

Figure 1. Average spectra of all limestone ore samples discriminated by CaO content when the rock surface is (**a**) dry and (**b**) wet.



(**a**) (**b**)

Figure 2. Correlation between the absorption wavelength of 2350 nm (4255 cm^{-1}) calcite absorption feature and CaO content when the rock surface is (**a**) dry and (**b**) wet. The coefficient of determination was the largest among regression results from various absorption characteristics of carbonate spectra.

2.3. Partial Least Squares Regression (PLSR)

PLSR is a generalization of a multiple linear regression by which the relationship between predictor (X) and observation (Y) data matrices are modeled [12]. This method projects data matrices to new individual spaces so that the covariance of two projected variables is maximized. PLSR is an iterative process where X, Y score vectors are extracted along the multidimensional variance direction, and the original X, Y matrices are deflated by the information that corresponds to the latest extracted vectors. The repetition number of data extraction corresponds to the number of PLS

components. One of the main outputs of PLSR is beta coefficients. Using a linear combination of beta coefficients and input spectrum added by a constant, a predicted Y is obtained [20].

To select predictor variables that correlate with observations, the variable importance in the projection (VIP) was used. VIP score of *j*-th waveband is calculated by:

$$VIP_j = \sqrt{p \sum_{k=1}^{h} (b_k^2 t_k^t t_k (w_{jk}/\|w_k\|)^2) / \sum_{k=1}^{h} b_k^2 t_k^t t_k} \quad (1)$$

where p is the number of variables; h is the number of components; t_k is the k-th column vector of the X score matrix; b_k is the scalar that explains the relation between Y and t_k; w_{jk} is the loading weight of the j-th wavelength in the k-th PLS component; w_k is the k-th column vector of the weight matrix (Equation (1); [20,21]). Since the wavebands with a higher VIP better explain Y, a more generalized model is created by excluding data with a smaller VIP value. It is common to use the average of the VIP scores for the threshold, i.e., 1.00. However, it has been suggested that a threshold between 0.83 and 1.21 is acceptable [21].

Prior to the PLS application, the spectral data were preprocessed in three steps. The first step is to determine the representative spectrum for each sample. We applied several methods, including arithmetic, geometric, and harmonic means as well as maximum and minimum of the reflectance value for the wavebands, and concluded that the geometric mean was the best approach. Thus, the geometric means of ten spectra were calculated for each rock sample. The second step is log-transformation of ore spectra, which exaggerates the spectral difference between 0 and 1. Lastly, centering and scaling was applied to the spectral data [12].

2.4. Model Selection of PLSR

To optimize the PLSR model parameters, the total CV method was used [12]. The total CV uses k-fold cross validation where k ranges from 5 to 9. K-fold cross validation randomly divides entire data into k subsets where k-1 sets are used for training the PLSR model and the remaining 1 set for testing. This process is repeated k times, changing the test set each time. The difference between measured and predicted Y is calculated by predicted residual sum of squares (PRESS; Equation (2)) which is collected from k parallel models. In Equation (2), y_i is the observed value (XRF content), $\hat{y}_i$ is the predicted value by PLSR and N is the number of observations. The number of PLS components for the best model is chosen with the smallest PRESS/($N - A - 1$) where N and A are numbers of observations and components, respectively. The standard error of cross validation (SECV) can also be used as a measure to compare candidate models with a changing number of PLS components [22]. SECV is the root mean square of the average error, thus choosing the number of PLS components by PRESS/($N - A - 1$) or SECV gives similar results.

$$\text{PRESS} = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (2)$$

To determine the best PLSR model for this study, we conducted a total CV method with 9-fold cross validation. PRESS/($N - A - 1$) was calculated with a different number of PLS components from 3–6. Additionally, we also performed dimensionality reduction using the VIP threshold. VIP thresholds applied for this study ranged from 0.85 to 1.20 with a 0.05 step. After determining the best PLSR model that has the smallest PRESS/($N - A - 1$) value, root mean square (RMS) was calculated to present accuracy of prediction (Equation (3)). Variables of this equation are the same as in Equation (2).

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

2.5. Assaying CaO Content in Bulk Limestone Ore

To estimate the prediction accuracy of bulk limestone ore, we generated hypothetical composite samples and calculated average CaO content based on XRF and PLSR results. The size of the composite sample was assigned to be 15 so that the weight of the hypothetical sample generally matches that collected by the in situ sampling protocol.

A total of 1000 hypothetical composite samples were generated from 100 limestone samples where constituent rock was randomly chosen with repetition. True and predicted CaO content of composite samples were calculated by arithmetic mean of CaO by XRF and PLSR assuming that the weight of each rock fragment is substantially same. The average CaO content by XRF and PLSR was displayed as a scatter plot, and the RMS error was calculated.

3. Results

3.1. CaO Estimation of the Single Limestone Sample

A 9-fold cross validation was applied for PLSR model selection for dry and wet limestone spectra. The training and test sets for the cross validation were randomly generated but commonly used in this process with different numbers of PLS components and VIP thresholds. Figure 3 shows calculation results of PRESS/($N - A - 1$) and optimal model parameters (Table 1).

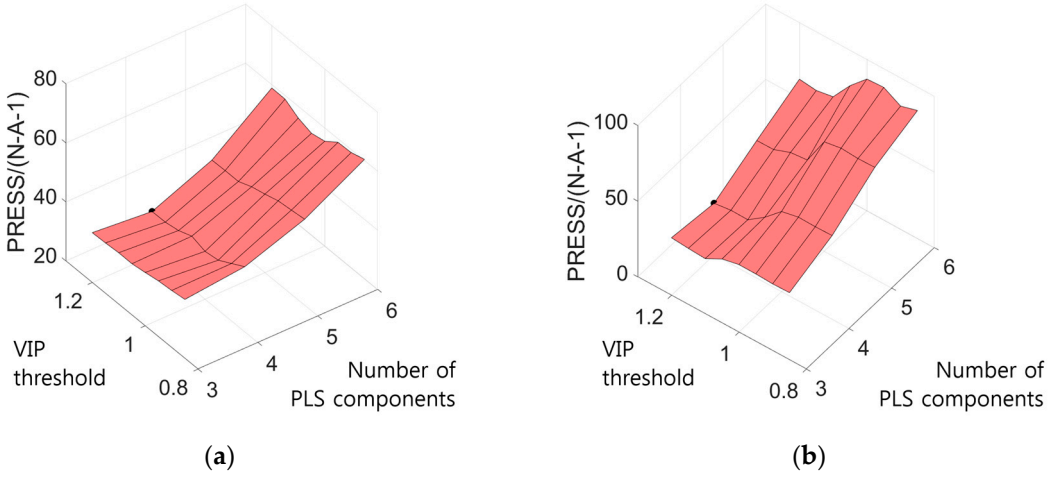


Figure 3. PRESS/($N - A - 1$) of the Partial Least Squares Regression (PLSR) model generated from (**a**) dry and (**b**) wet limestone spectra. The predictive ability is maximized when PRESS/($N - A - 1$) is lowest (black dot above surface).

Table 1. Optimal model parameters for PLSR obtained from cross validation.

Surface Moisture State	Optimal Number of PLS Components	Optimal Variable Importance in the Projection (VIP) Threshold
Dry condition	4	1.20
Wet condition	4	1.20

The optimal model parameters in Table 1 were used to build PLSR models for dry and wet limestone spectra. The CaO predicted by PLSR showed a correlational trend with XRF CaO for both moisture conditions (Figure 4) where error measure (RMS errors) and goodness-of-fit (R squared) in the two models showed differences of less than 0.01% and 0.005, respectively (Figure 4, Table 2).

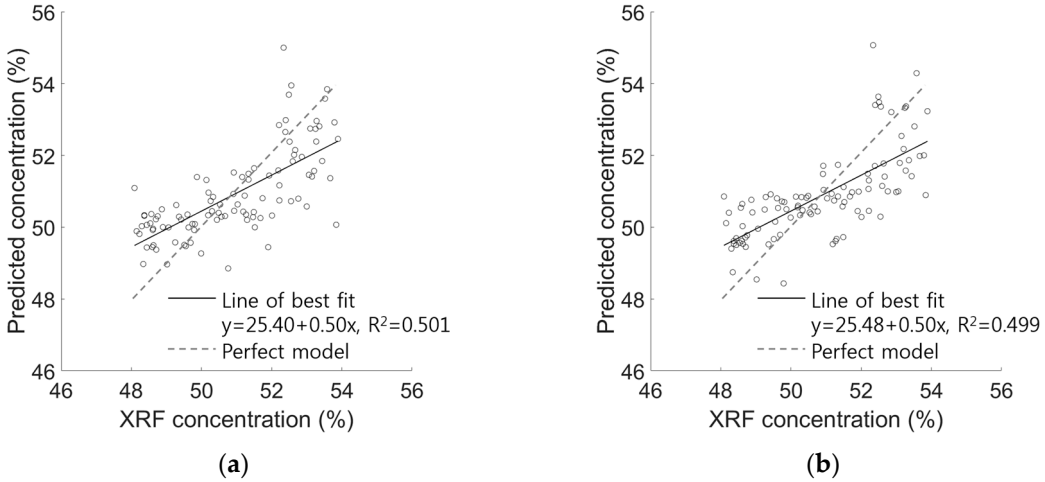


Figure 4. PRESS/($N - A - 1$) of the PLSR model generated from (**a**) dry and (**b**) wet limestone spectra. The predictive ability is maximized when PRESS/($N - A - 1$) is lowest (black dot above surface).

As for the testing of the PLS-based method which requires splitting data into training and testing sets, the number of samples, i.e., 100, was not sufficient. Therefore, we instead conducted 10-fold cross validation. Figure 5 is a collection of test results from 10 cross-validation models. The overall trend of test results was similar to the model generated by all samples (Figure 4). The average RMS errors of test sets for dry and wet spectra were 1.30% and 1.29%, respectively, which are slightly larger than the RMS error of the whole data model.

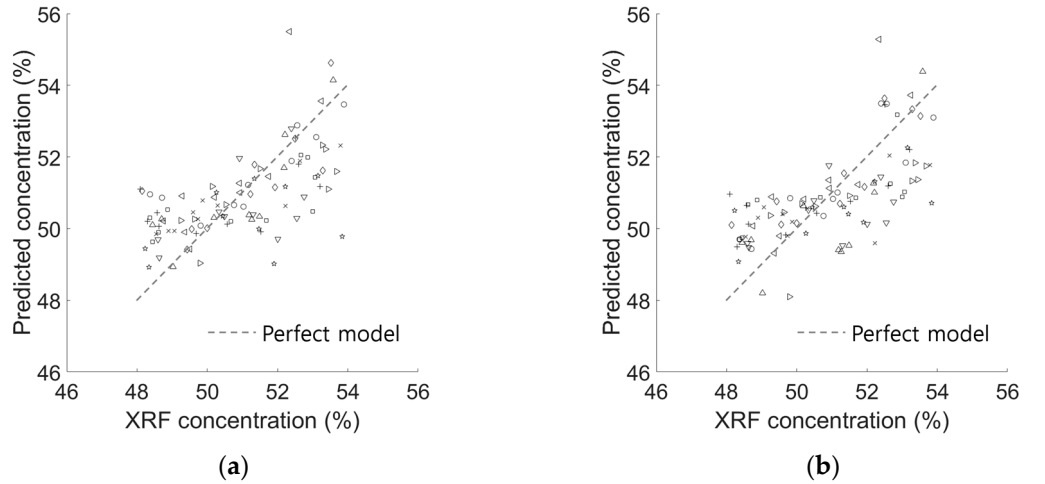


Figure 5. Test results of CaO prediction in the single limestone sample estimated by 10-fold cross validation using (**a**) dry and (**b**) wet limestone spectra. Each test set used in the cross validation was represented by different symbols. Distribution pattern of cross-validated results is similar to that of the PLSR model generated by whole sample data (Figure 4). The average root mean square (RMS) error of cross-validated models was slightly bigger than the whole data model, but less than 0.1%.

3.2. CaO Estimation of the Composite Limestone Sample

From the prediction results of the whole data model (Figure 4), hypothetical composite samples were generated (Figure 6). The slope and offset of the regression line were similar to Figure 4 as with the coefficient of determination. RMS errors using hypothetical composite samples were 0.311% and 0.310% for dry and wet conditions (Table 2), respectively, which suggests estimation of CaO using bulk ore is more reliable than that of single rock fragments.

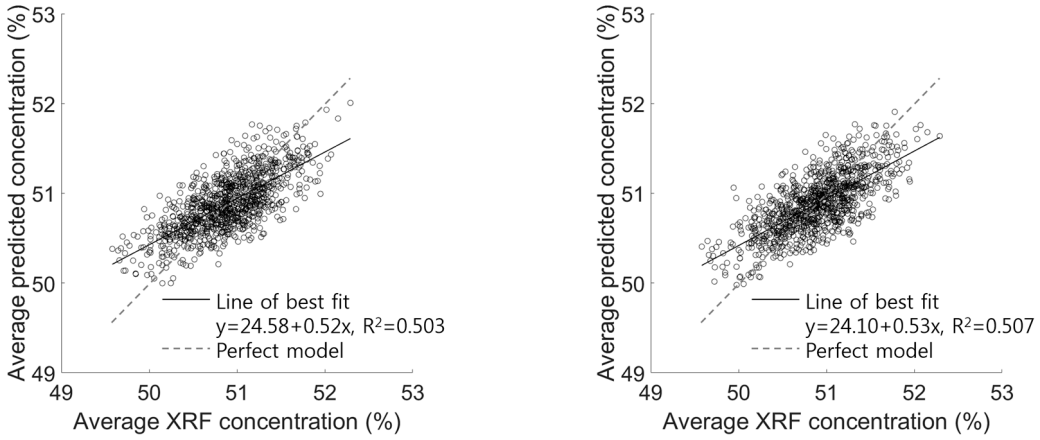


Figure 6. Predictive models of CaO in the hypothetical composite sample under (**a**) dry and (**b**) wet surface conditions. Regression lines for two distributions were similar to the corresponding results of single limestone ore (Figure 4), but predicted values were distributed closer to best fit lines.

Table 2. RMS error (%) of CaO prediction by PLSR with or without the presence of water on the rock surface. RMS error of the single limestone sample was obtained from the prediction results of 100 single ore fragments (Figure 4), and that of the composite limestone sample was calculated from the average CaO values of 1000 hypothetical samples (Figure 6).

Prediction Accuracy	Dry Conditions	Wet Conditions
RMS of the single limestone sample (%)	1.220	1.222
RMS of the composite limestone sample (%)	0.311	0.310

4. Discussion

The PRESS/($N - A - 1$) calculated from dry and wet limestone spectra both had the lowest values when the number of PLS components and VIP threshold were 4 and 1.20, respectively (Figure 3, Table 1). From Figure 3, a common tendency was found that PRESS/($N - A - 1$) slightly changes until the number of components becomes four and rapidly increases afterwards. A declining trend of PRESS/($N - A - 1$) against the VIP threshold was evident for wet limestone spectra, which suggests that dimensionality reduction with a higher VIP threshold had a more positive effect on prediction accuracy than the adverse effect of model simplification. For dry spectra, this trend was less apparent but still valid in that the lowest PRESS/($N - A - 1$) was observed with the VIP threshold at its upper bound.

As for the overall relationship between input limestone spectra and CaO concentration, beta coefficients of whole data models were examined (Figure 7). Wavelengths not correlated with $CaCO_3$ content were excluded by VIP.

Beta coefficients of the prediction models showed two convex shapes near 2350 and 2500 nm (4255 and 4000 cm^{-1}) in common. This indicates that the ore sample with higher CaO concentration had stronger absorption in these bands while other wavelengths including calcite bands at 2120–2160 (4630–4717 cm^{-1}), 1970–2000 (5000–5076 cm^{-1}) and 1850–1870 nm (5348–5405 cm^{-1}) were less effective for the prediction. Additionally, beta coefficients of dry spectra showed an extra concave feature at 400–500 nm (20,000–25,000 cm^{-1}). This was attributed to the fact that the majority of the limestone spectra with red clay had convex features at 400–500 nm (20,000–25,000 cm^{-1}), but purer limestone with smaller clay content had a weaker absorption feature in the range. Data in 400–500 nm (20,000–25,000 cm^{-1}) were not used for CaO prediction of wet limestone, which means that the spectral characteristics of this band were less correlated with CaO content in this case. As for moisture effect on the NIR spectrum, water suppressed the overall spectrum in 350–2500 nm (4000–28,571 cm^{-1}) and caused two broad absorption features at 1400 and 1900 nm (7143 and 5263 cm^{-1}; Figure 2), but the

statistical trend of spectral suppression and water absorption did not correlate with CaO content and only spectral data related to calcite or clay mineral were exclusively used for the prediction.

By using optimal PLSR model parameters in this study, a moderate to strong correlation trend was observed from Figure 4. However, large variance and a couple of outliers reduced the linearity and accuracy of the prediction model. With coefficients of determination c.a. 0.5, it was found that predicting CaO content from rock spectra is still challenging compared to previous results. Compared to the studies where absorption characteristics of rock spectra were used [8,9], our study used CaO content that represents the volume of rock samples instead of surface CaO content, so the prediction could be more sensitive to the measuring location of NIR spectra. Additionally, clay coatings on rock surfaces, which were geological characteristics of limestone in the study area, could also disturb spectral characteristics of carbonate minerals. Nevertheless, this study revealed that PLSR with the VIP threshold method could increase the reliability of prediction compared to estimating CaO solely depending on absorption characteristics, and this was valid regardless of surface moisture conditions (Figures 2 and 4).

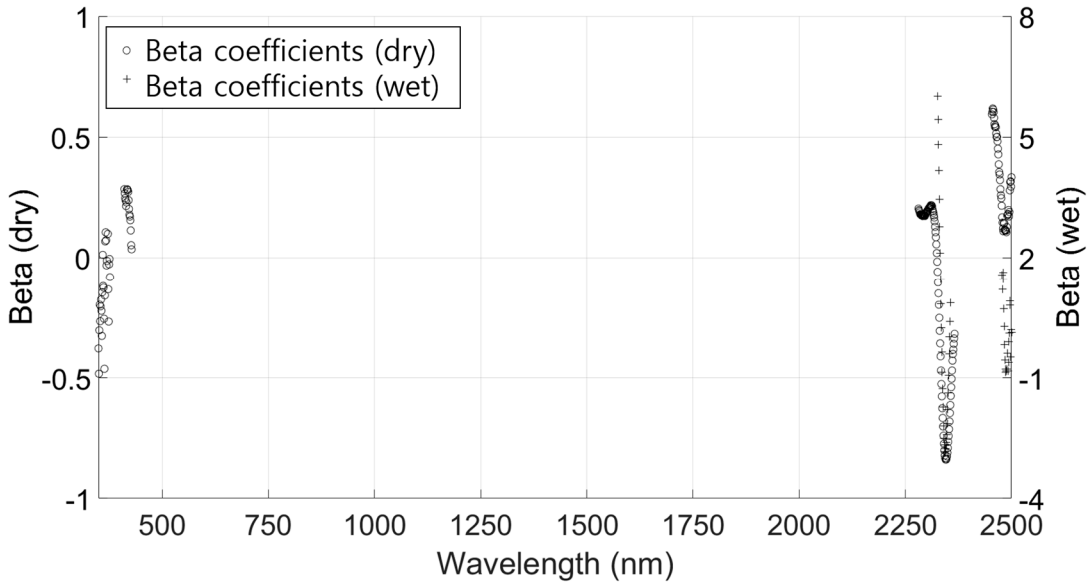


Figure 7. Beta coefficients of the PLSR model using dry and wet limestone spectra. The above results indicate that spectral data with a calcite absorption pattern at 2350 and 2500 nm (4255 and 4000 cm^{-1}) were used to predict CaO in PLSR models regardless of moisture conditions. A concave spectral pattern in 400–500 nm (20,000–25,000 cm^{-1}) was only exploited for the prediction when the rock surface was dry.

5. Conclusions

We developed and tested a PLS-based technique for the estimation of CaO content using NIR spectra of limestone ore samples. We also showed that water on rock surfaces can decrease the prediction accuracy of the suggested method. The procedure is comprised of rock sampling, spectrum measurements, data preprocessing and PLSR. In order to obtain a reliable estimate, appropriate numbers of PLS components with VIP thresholds were selected by cross validation. The PLSR of rock spectra predicted CaO content with RMS errors of 1.2% under dry or wet surface conditions. These results could not be achieved by using the absorption properties, e.g., absorption strength, area or wavelength. Since average CaO content of bulk ore is more important for decision making in the study mine, we also generated hypothetical composite samples to calculate average CaO content. The accuracy of prediction was 0.3% in RMS errors in both moisture conditions.

Although the approach was useful to estimate CaO content from limestone rock spectra, there were some limitations. Reflectance spectroscopy only penetrates a few microns of the rock surface. Hence, the CaO content of the estimated ore is surface specific. Measured spectra can also be distorted

by the inhomogeneity of rock affected by crystal size, and the existence of a calcite vein or clay coating with nonlinear mixing effects. Geometry is another important element that causes deviation in rock spectra. More robust sampling protocols and preprocessing methods are required to reduce the above effects, i.e., randomizing the measuring location on rock surfaces, excluding the spectrum with large standard deviation compared to others, using derivatives of spectra, etc.

Data-driven methods including PLSR have the inherent limitation that training data of predictor (NIR spectra) and response variables (XRF content) are necessary for prediction. This also means that training data and a prediction model have to be generated case by case. The process of finding optimal parameters for the prediction method is mandatory to avoid overfitting or underfitting. Most of the time, the regression process of a data-driven method is more time consuming than the direct regression method based on absorption feature characteristics. Therefore, use of a data-driven method is recommended when regression methods of absorption characteristics do not work well but a prediction model is still required.

A distinct difference between this study and previous NIR approaches lies in that our study revealed that the PLSR-based prediction method can also be used with rock spectra and the method can improve the reliability of prediction compared to making predictions based solely on absorption characteristics. It was also revealed that these two findings were valid regardless of the surface moisture conditions of limestone ores. The described technique can be used directly for grade checking in a working rock face as well as for bulk separation of excavated ore. This method can provide immediate results of CaO content of rock while saving time and energy associated with conventional chemical analysis. The suggested method can also be applied for the development of NIR spectroscopy-based systems that could be used for rapid ore grade characterization on conveyor belt systems.

Acknowledgments: This work was supported and funded by the Korea Institute of Energy Technology Evaluation and Planning (Grant No. 2013T100100021) through grant funded by the Korean Government.

Author Contributions: Sungchan Oh and Hyeong-Dong Park conceived and designed the experiments; Sungchan Oh performed the experiments and analyzed the results; Sungchan Oh and Chang-Uk Hyun were involved in the manuscript preparation; The whole paper was finally checked by Chang-Uk Hyun and Hyeong-Dong Park. All authors read and approved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wheeler, B.D. Analysis of limestones and dolomites by X-ray fluorescence. *Rigaku J.* **1999**, *16*, 16–25.
2. Simandl, G.; Paradis, S.; Stone, R.; Fajber, R.; Kressall, R.; Grattan, K.; Crozier, J.; Simandl, L. Applicability of handheld X-ray fluorescence spectrometry in the exploration and development of carbonatite-related niobium deposits: A case study of the Aley Carbonatite, British Columbia, Canada. *Geochem. Explor. Environ. Anal.* **2014**, *14*, 211–221.
3. Van der Meer, F. Spectral reflectance of carbonate mineral mixtures and bidirectional reflectance theory: Quantitative analysis techniques for application in remote sensing. *Remote Sens. Rev.* **1995**, *13*, 67–94.
4. Clark, R.N. Spectroscopy of rocks and minerals, and principles of spectroscopy. In *Manual of Remote Sensing*; Renz, A.N., Ed.; John Wiley and Sons: New York, NY, USA, 1999; Volume 3, pp. 3–58.
5. Gallie, E.; McArdle, S.; Rivard, B.; Francis, H. Estimating sulphide ore grade in broken rock using visible/infrared hyperspectral reflectance spectra. *Int. J. Remote Sens.* **2002**, *23*, 2229–2246.
6. Magendran, T.; Sanjeevi, S.; Bhattacharya, A.K.; Surada, S. Hyperspectral radiometry to estimate the grades of iron ores of Noamundi, India—A preliminary study. *J. Indian Soc. Remote Sens.* **2011**, *39*, 473–483.
7. Dalm, M.; Buxton, M.W.N.; van Ruitenbeek, F.J.A.; Vonken, J.H.L. Application of near-infrared spectroscopy to sensor based sorting of a porphyry copper ore. *Miner. Eng.* **2014**, *58*, 7–16.
8. Zaini, N.; van der Meer, F.; van der Werff, H. Determination of carbonate rock chemistry using laboratory-based hyperspectral imagery. *Remote Sens.* **2014**, *6*, 4149–4172.
9. Zaini. N.; van der Meer, F.; van Ruitenbeek, F.; de Smeth, B.; Amri, F.; Lievens, C. An alternative quality control technique for mineral chemistry analysis of Portland cement-grade limestone using shortwave infrared spectroscopy. *Remote Sens.* **2016**, *8*, 950.

10. Bioucas-Dias, J.M.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral Unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2012**, *5*, 354–479.
11. Schmidt, F.; Legendre, M; Le Mouëlic, S. Minerals detection for hyperspectral images using adapted linear unmixing: LinMin. *Icarus* **2014**, *237*, 61–74.
12. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
13. Goetz, A.F.; Curtiss, B.; Shiley, D.A. Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy. *Miner. Eng.* **2009**, *22*, 490–499.
14. Basile, A.; Hughes, J.; McFarlane, A.J.; Bhargava, S.K. Development of a model for serpentine quantification in nickel laterite minerals by infrared spectroscopy. *Miner. Eng.* **2010**, *23*, 407–412.
15. Choi, D.K.; Chough, S.K. The Cambrian-Ordovician stratigraphy of the Taebaeksan Basin, Korea: A review. *Geosci. J.* **2005**, *9*, 187–214.
16. Oates, J.A.H. *Lime and Limestone: Chemistry and Technology, Production and Uses*; John Wiley & Sons: Weinheim, Germany, 2008; pp. 94–99.
17. Hatchell, D.C. *ASD Technical Guide*, 3rd ed.; Analytical Spectral Devices, Inc.: Longmont, CO, USA, 1999; 11–14.
18. Hunt, G.R.; Salisbury, J.W. Visible and near infrared spectra of minerals and rocks. II. Carbonates. *Mod. Geol.* **1971**, *2*, 23–30.
19. Hopkinson, L; Rutt, K.J. Crystal chemical correlations between the mid and near-infrared in carbonate minerals. *Spectrochim. Acta Part. A Mol. Biomol. Spectrosc.* **2016**, *162*, 105–108.
20. De Jong, S. SIMPLS: An alternative approach to partial least squares regression. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 251–263.
21. Chong, I.-G.; Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112.
22. Burns, D.A.; Ciurczak, E.W. *Handbook of Near-Infrared Analysis*, 2nd ed.; Marcel Dekker: New York, NY, USA, 2001; p. 122.