

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

An Empirical Model-based Method for Signal Restoration of SWIR in ASD Field Spectroradiometry

Chinsu Lin, Khongor Tsogt, and Chein-I Chang

Abstract

ASD spectroradiometer field measurements in the SWIR water absorption region are sometimes problematic due to atmospheric effects such as moisture in the air. The reduced signal-to-noise ratio (SNR) in these wavebands makes it difficult to diagnose water stress in tree foliage using spectroscopy. This paper investigates the SNR issue in the 1,350 to 1,410 nm waveband using laboratory-based experiments and practical field measurements of mountainous tree foliage spectra. With laboratory spectra data, three empirical signal models along with a Gaussian bias model were tested and validated using noisy field spectra data. Results demonstrate that a combination of either a logistic or sigmoid signal model coupled with a Gaussian bias (residual) model (LOGGM and SIGGM complex signal models) can effectively describe reflectance behaviors in the spectral region of 1,350 to 1,410 nm of red cypress (*Chaemacyparis formosensis*) foliage and further show that our proposed approach is promising in restoration of field spectra measurements.

Introduction

Hyperspectral remote sensing is an emerging technique which takes advantage of high spectral resolution to explore subtle spectral variations of material substances and natural phenomena. It provides users with more accurate and detailed spectral signatures that can be used to build a reliable spectral library for data analysis. Since the Earth's atmosphere absorbs much of the Sun's ultraviolet radiation, and water vapor absorbs most of incoming shortwave infrared (SWIR) light around wavelengths of 1,400 nm, 1,900 nm, and those greater than 2,700 nm (Govender *et al.*, 2007), these factors reduce Signal-to-Noise Ratio (SNR) of spectroradiometer data and/or remote sensing images. As a result, data integrity and quality will be compromised. In addition, remote sensing data collection is also affected by atmospheric gases. On the other hand, the water absorption bands for liquid water are known to occur approximately at 970 nm, 1,200 nm, 1,450 nm, and 1,950 nm (Curran, 1989) with features at 1,450 nm and 1,950 nm being most pronounced. However, at about 1,400 nm and 1,900 nm broad absorption features usually occur due to the water vapor in the atmosphere. As a result, hardly any

radiation in those absorption regions reach the Earth's surface. Under such a circumstance, the bands at about 1,450 nm and 1,950 nm should not be used for data processing (Vikhamar *et al.*, 2004; Estep and Carter, 2005; Clevers *et al.*, 2007; Schaepman *et al.*, 2007).

Hyperspectral sensor developers have been trying to resolve this problem and offer some solutions. To reduce the atmospheric effect, Clark *et al.* (2003) proposed using a powerful artificial light source to measure a large spot in an enclosure and thus minimizing solar and atmospheric effects. However, the heavy batteries used to power the lights and the spectrometer are inconvenient to carry, especially over rough terrain. One solution to mitigate this problem is to sample surface materials and measure their spectra in the laboratory. The US Geological Survey has found that for some calibration sites, the laboratory spectra of field samples are usually adequate to measure the site's spectral data characteristics and remove artifacts from the field data. In this way, the laboratory spectral data provides a second check in establishing the quality of the *in situ* field data.

ASD (Analytical Spectral Devices) offers a software solution where a spectrum averaging method is used to reduce the effects of noise in the measured spectra because background noise is very natural. ASD (1999) suggested that each measurement of the field spectroradiometer could be done by averaging 10 to 25 readings at the same spot. Increasing the number of readings might improve the real measurement possibility. However, spectrum averaging takes more time to obtain individual spectra. One drawback of this procedure is that the resultant data can be compromised by introduction of low-frequency noise factors, such as varying cloud conditions. In addition, atmospheric conditions can also change quite frequently even in the case of a clear sky. To reduce the effects of low-frequency noise conditions such as those found outdoors, ASD (2002) recommended that multiple spectra should be taken with spectrum averaging set to 10 to 25, and then these spectra could be further averaged using a median filter after post-processing. Despite the fact that using averaged spectra or median-filtered spectra can improve SNR in dry atmospheric conditions for noise reduction, there is still no way to obtain noiseless spectra at normal and high moisture content atmospheric conditions (ASD, 1999).

Some environmental researchers do not use reflectance spectra from the atmospheric water absorption range (Ma and

Chinsu Lin and Khongor Tsogt are with the Department of Forestry and Natural Resources, National Chiayi University, 300 University Road, Chiayi (60004), Taiwan (chinsu@mail.ncyu.edu.tw).

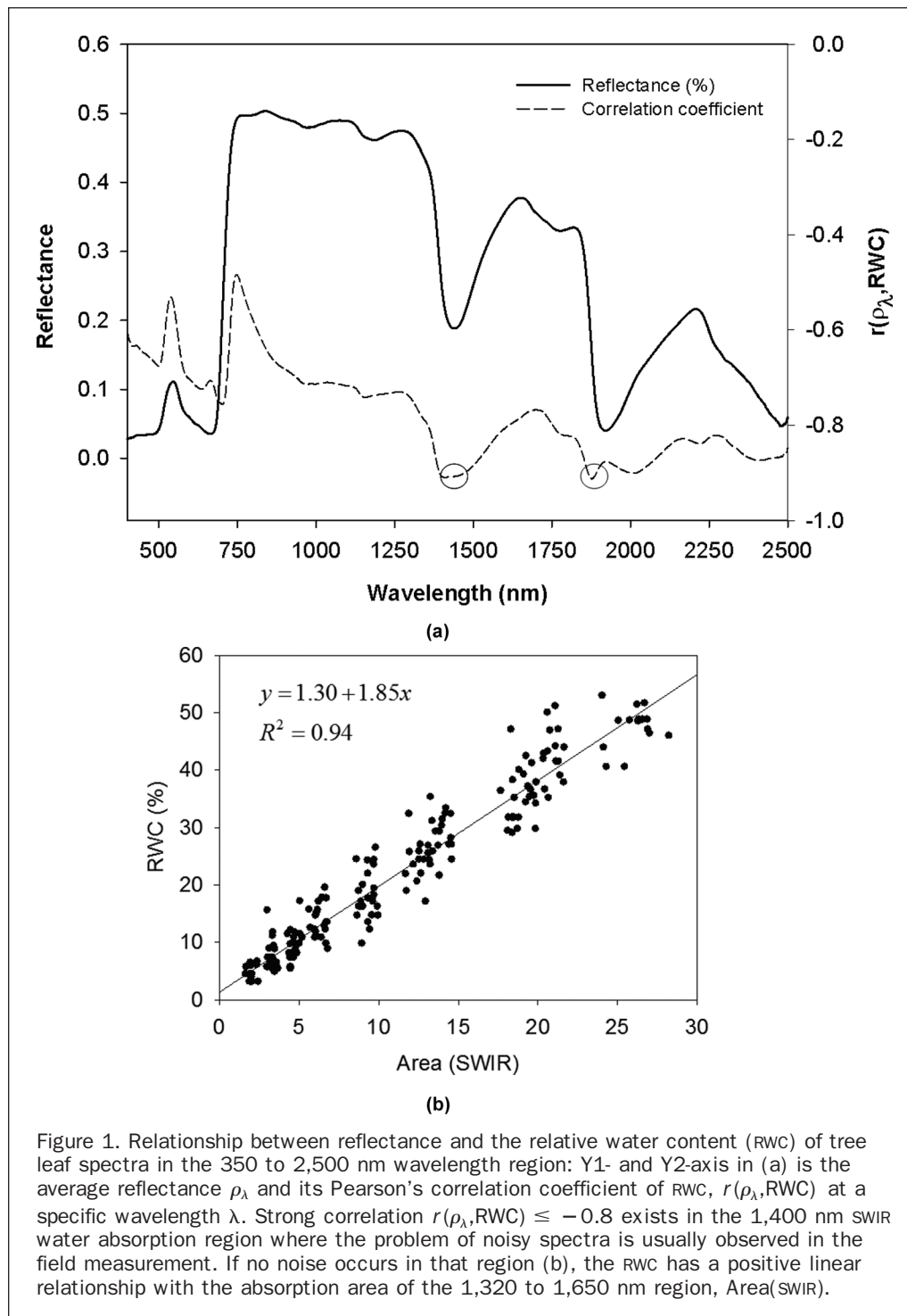
Chein-I Chang is with the Remote Sensing Signal and Image Processing Laboratory (RSSIPL), Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250.

Photogrammetric Engineering & Remote Sensing
Vol. 78, No. 2, February 2012, pp. 119–127.

0099-1112/12/7802-119/\$3.00/0
© 2012 American Society for Photogrammetry
and Remote Sensing

Chen, 2005; Psomas *et al.*, 2005; Clevers *et al.*, 2007, 2008, and 2010; Ge *et al.*, 2008), even though it is recognized that this spectral range can predict the water content. According to our extensive laboratory experiments (Lin 2005; Lin *et al.* 2006), the relative water content (RWC) of tree leaves can be predicted very accurately using spectral reflectance data in the SWIR water absorption regions. The RWC is calculated by (fresh weight - dry weight) / (turgid weight) in terms of a percentage. Figure 1a shows an example of the average spectral reflectance of *Cinnamomum camphora* leaves with varying relative water content (solid line) and the correlation

coefficient between the RWC and reflectance, $r(\rho_\lambda, RWC)$ in the region from 350 to 2,500 nm (dashed line). The highest negative correlation coefficient $|r(\rho_\lambda, RWC)| = 0.91$ indicated by the circle in Figure 1a is in the SWIR water absorption regions. Figure 1b shows the positive relationship between the absorption area of the 1,320 to 1,650 nm SWIR region and the RWC of leaves in percentage. The SWIR absorption area, determined by the continuum removal technique (Clark and Roush, 1984) is abbreviated as Area(SWIR). Clearly, there exists a strong linear relationship between the SWIR region and the RWC of leaves, which indicates that SWIR spectra are



strongly dependent on the water content of leaves. Thus, the SWIR is a useful spectral range for water stress detection. Consequently, the problem of signal/noise in the SWIR range needs to be addressed and methods of restoring the SWIR spectral signal become of central importance.

The objective of this study is to utilize empirical models currently available in the literature to develop a method that can remove noise and/or restore the SWIR signal *in situ* using ASD hyperspectral data. Linear, sigmoid, logistic, and Gaussian regression models were used to fit sample hyperspectral data during the laboratory stage of algorithm development. In the field, the proposed algorithm was applied to hyperspectral data obtained from mountainous forest and the obtained estimates were then used to evaluate the algorithm performance.

Empirical Regression Models

Four empirical regression models, (a) linear model, (b) four-parameter sigmoid model, (c) four-parameter logistic model, and (d) three-parameter Gaussian model used in this study are described below.

Linear Model

The linear model is the simplest linear function specified by a single regressor x that has a relationship with a response y in the form of a straight line. Such a function can be expressed as:

$$y = y_0 + ax \quad (1)$$

where y_0 and a are real constants and x is a real variable. The parameters a and y_0 are often called the slope and y -intercept, so the linear model is also referred to as the slope-intercept form. Changing a makes the line steeper or flatter, while changing y_0 moves the line up or down (Montgomery *et al.*, 2001).

Four-parameter Sigmoid Model

A sigmoid function is an “S” shape curve. It is real-valued and differentiable, and has either a non-negative or non-positive first derivative and also has exactly one inflection point. There are also two asymptotes, $x \rightarrow \pm\infty$. In Equation 2, the sigmoid model has four parameters, a , y_0 , b , x_0 , to describe the S-shape changes of the spectral behavior in SWIR where the constant a is often called the slope while y_0 is the y -intercept showing the point of intersection between the graph of the function and the y -axis. When the value of b is positive, the curve increases from left (shorter wavelength) to right (longer wavelength). Conversely, when b is negative, the curve decreases from left to right. The parameter x_0 sets the point of inflection of the curve (Mitchell, 1997):

$$y = y_0 + \frac{a}{1 + e^{-\left(\frac{x-x_0}{b}\right)}} \quad (2)$$

Four-parameter Logistic Model

A logistic curve is also an S-shaped curve used to describe growth data. In general, the initial stage of growth is approximately exponential; then, as saturation approaches, the growth slows and stops when it reaches maturity. The logistic function used in this paper has four parameters: \min , \max , EC_{50} , and $Hillslope$. In Equation 3, parameters \min and \max represent the upper and lower plateaus of the curve. The EC_{50} is an inflection point, which is the x value when y is half-way between the \min and \max plateaus and associated with the point of symmetry of the S-curve. $Hillslope$ is a curvature parameter and is related to the slope of the curve:

$$y = \min + \frac{(\max - \min)}{1 + \left(\frac{x}{EC_{50}}\right)^{Hillslope}} \quad (3)$$

Three-parameter Gaussian Model

Gaussian functions are widely used in statistics to describe a normal distribution curve (ReliaSoft Corporation, 2005). A Gaussian function is a function of the form specified by:

$$y = ae^{\left[-0.5\left(\frac{x-x_0}{b}\right)^2\right]} \quad (4)$$

where the three parameters, a , b , and x_0 are greater than zero, and e is the exponential function. The graph of a Gaussian curve is a “bell shape curve” centered at the mean and quickly falls off towards plus or minus infinity. The parameter a is the height of the curve’s peak and is determined by the reciprocal of square root of $2\pi\sigma^2$, x_0 is the position of the center of the peak, and b is the standard deviation, σ , which controls the width of the “bell.”

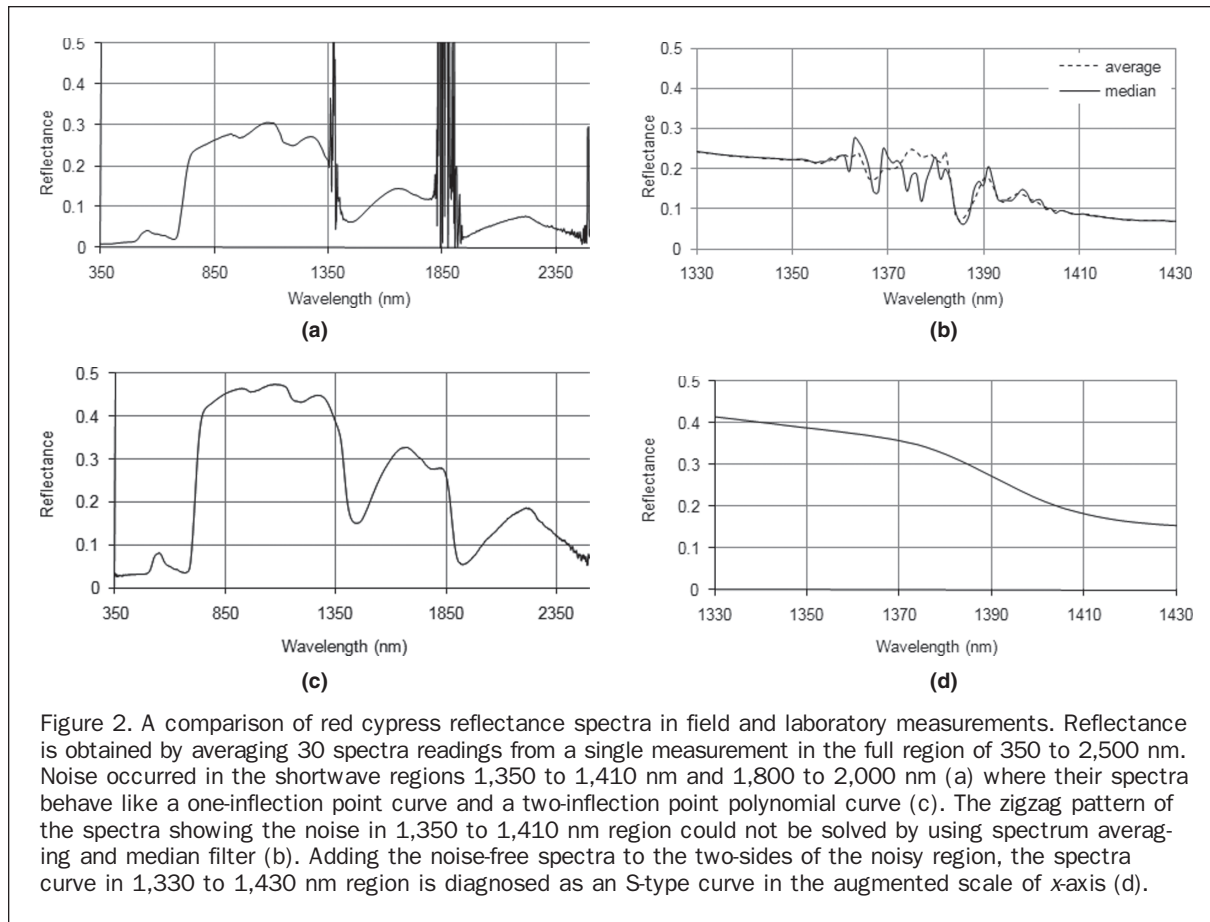
Materials and Methods

Study Site and Spectral Data Collection

The site chosen for study was a plantation of red cypress trees (*Chamaecyparis formosensis* Obtusa) located on the Alishan Mountain in the central Taiwan at the latitude of N 23°29’56”, E 120°49’08” with an average elevation of 2,267 m. The red cypress is a native species and is one of the most valuable wood resources in Taiwan. A field survey and study using an ASD FieldSpec Pro FR spectroradiometer was conducted on 09 October 2008. The experiments were performed using six trees centered around a 7-meter tall watchtower from which measurements were made. The spectroradiometer was deployed using a fiber optic cable with a 25° field of view. The height of made measurements was set to approximately 2.25 m above the crown. As a result, the field of view at the crown level was circular with a 1 m radius. The spectral data was collected under cloud-free conditions between 09:20 am and 10:00 am (local time) in the morning. Target measurements were collected immediately after the reference measurements were taken. The spectroradiometer measurement protocol was repeated five times per tree and each measurement was obtained by averaging 30 readings of the same tree crown. The sampling interval was 1 nm. The calibration was done by using a spectralon white reference panel every 20 minutes to account for the changing state of the visible near infrared detector, while removing the effects of dark current. The spectral reflectance of the trees was calculated by dividing the reflected target radiance by the irradiance of the white panel. A total of 30 spectral measurements were made. After the field survey, 30 samples of red cypress trees were taken to the laboratory for spectral measurements.

Diagnosing the SWIR Spectral Behavior of Red Cypress

The relative humidity of the air in the study site ranges from 81 percent in December to 93 percent in August with an annual average of 87 percent. It was expected that the SNR in the SWIR region would obviously be affected by the high humidity. Figure 2a shows the in-field spectra of the red cypress in the wavelength from visible-near infrared to shortwave infrared. It demonstrates that the noise always occur at the regions from 1,350 to 1,410 nm and 1,800 to 1,950 nm. Figure 2b shows the 1,350 to 1,410 nm region in an augmented scale along x-axis where the spectrum averaging and median filtering methods may not be sufficient to solve the noise problem. Figure 2c shows the in-laboratory spectrum



of the red cypress which was used to diagnose the normal behavior in the SWIR region of the red cypress spectrum where there is one inflection point in the 1,350 to 1,410 nm region and two inflection points in the 1,800 to 1,950 nm region. So, the spectral profile in these two wavelength ranges behaves like a polynomial curve. By adding noise-free signals to each side of the 1,350 to 1,410 nm noise region, the extended spectra from 1330 to 1430 nm in Figure 2d looks like a typical S-type curve. This offers a base for developing empirical model-based algorithms to restore the spectral signals in the SWIR water absorption region. According to the typical spectral reflectance, the inflection points are located at around 1,840 nm and 1,940 nm, while the in-field spectra have extremely high noise in the 1,800 to 1,950 nm region. It seems there are no noise-free signals that could be used to estimate the inflection of a polynomial curve. Therefore, this paper only focuses on the restoration of signals in the 1,350 to 1,410 nm region.

Algorithm Development Using the Behavior Modeling of Noise-free Signals

Given a wide range of spectral variations of natural material substances, an effective and better way to analyze such objects is to learn and fit their spectral behaviors using mathematical models. As mentioned in the previous section, the noise-free spectra of red cypress behaves like an S-type curve in the 1,350 to 1,410 nm region, while the spectra at 1,330 to 1,350 nm and 1,410 to 1,430 nm could be used as a base of spectral curve fitting. The S-type curve can be described by either a sigmoid model or a logistic model. The linear model is included only for comparative study.

Spectral behavior modeling is a three-step procedure including the development and the integration of the signal model and the bias model. In the initial step, each of linear,

sigmoid, and logistic models, abbreviated as LM, SIGM, and LOGM, respectively, is applied separately to individual spectra to derive the primary estimates of the signals (or major signals) in the noise region. The signal model produced by this step is called simplex signal model. Since the regression analysis assumes the residuals are normally distributed, the second step is therefore to apply a Gaussian model to fit the residuals of the corresponding signal models developed in the initial step. The resulting model is referred to as a bias model, denoted by GM, which can be used to derive the secondary estimates of the signals (or minor signals) in the noise region. Finally both signal model and bias model are integrated together using Equation 5 to accomplish the signal restoration of the noisy SWIR spectra. The model produced in the final step is called complex signal model:

$$\text{restored_}R_{\lambda} = \text{signalModel_estimate_}R_{\lambda} + \text{biasModel_estimate_}R_{\lambda}. \quad (5)$$

Three complex signal models resulting from such integration derived from Equation 5 are linear-Gaussian model, sigmoid-Gaussian model, and logistic-Gaussian model referred to as LGM (LM + GM_{linear}), SigGM (SigM + GM_{sigmoid}), and LogGM (LogM + GM_{logistic}), respectively. Let SME_λ and BME_λ be the primary signal estimate of signal model and the secondary signal estimate of bias model for a specific wavelength λ in the noisy SWIR region, the final reflectance estimate of restored noise-free spectrum (restored_R_λ) can be determined by adding up these two estimates. The restored spectrum for a specific wavelength in the SWIR region by this algorithm is described as follows:

- LGM restored_R_λ = LM_SME_λ + LM_BME_λ,

- SigGM restored $R_\lambda = \text{SigM_SME}_\lambda + \text{SigM_BME}_\lambda$, and
- LogGM restored $R_\lambda = \text{LogM_SME}_\lambda + \text{LogM_BME}_\lambda$.

Statistical Test of the Signal Models and Bias Models

An analysis of variance (ANOVA) coupled with F-test and Student's t-test was conducted to evaluate the significance of various regression models and their coefficients ($\hat{\beta}_i$). In addition, the variance inflation factor (VIF) is also used as an indicator of multicollinearity in the regression model since it measures how much the variance of a regression coefficient (square of the standard error) is increased as a result of multicollinearity. A large VIF value makes an empirical model unstable and gives poor prediction/estimation. A common rule of thumb is that if $\text{VIF}(\hat{\beta}_i) > 5$ then multicollinearity is high. So, a value of 10 has been proposed as a cut-off value (Longnecker and Ott, 2004). All of the signal models and the corresponding bias models were evaluated by using F-test, t-test, and VIF indicators.

Results and Discussion

Empirical Signal and Bias Models of Red Cypress Leaves

Table 1 shows the regression coefficients and VIFs of empirical signal models of LM, SIGM, and LOGM derived from the in-lab spectra of red cypress leaves. These signal models were statistically significant with $R^2 = 0.99$ indicating 99 percent of the reflectance variation in the 1,330 to 1,430 nm SWIR region could be well described by these signal models. Although all of the coefficients in these signal models were significant at the probability level $P < 0.0001$, the VIFs of the regression coefficients showed a very large variation among the LM, SIGM, and LOGM signal models. The LM signal model whose VIFs $> 1,000$ indicates the intercept and the slope of the fitted signal regression line has the largest prediction variance and will cause instability to the estimation. The regression coefficients of SIGM and LOGM work much better than LM does because the former has relatively lower VIFs while the latter has the smallest VIFs value. In particular, the slope coefficient (Hillslope) of LOGM has a VIF value of 7.52 less than 10 which is the cut-off value set for VIF.

Table 2 shows the regression coefficients and VIFs of the Gaussian bias models coupled with each of the corresponding signal models in Table 1 where the bias model of the LM signal has non-Gaussian distributed residuals in the 1,330 to 1,430 nm SWIR region. This is witnessed in Figure 3b

where the residual has an N-shape distribution along with the wavelength on the x-axis. In contrast to the LM bias model, the bias models of SIGM and LOGM signal models with $R^2 = 0.99$ have their spectra estimates with Gaussian or normal distribution residuals (Figure 3d and 3e). The regression coefficients of the SIGM bias model and the LOGM bias model have very low VIFs close to 1.0 which indicates that these two bias models worked very well in the prediction of the secondary signal estimates in the noisy SWIR region. The results of statistical tests in Table 1 and Table 2 suggest that the reflectance of red cypress leaves could be predicted quite well by SIGGM and LOGGM models which use the four-parameter sigmoid model or logistic model in conjunction with the three-parameter Gaussian model. However, from a statistical point of view the combination of the logistic signal model and the Gaussian bias model is more suitable for the signal restoration of the spectra in the 1,330 to 1,430 nm SWIR region because of the good fit of its regression coefficients.

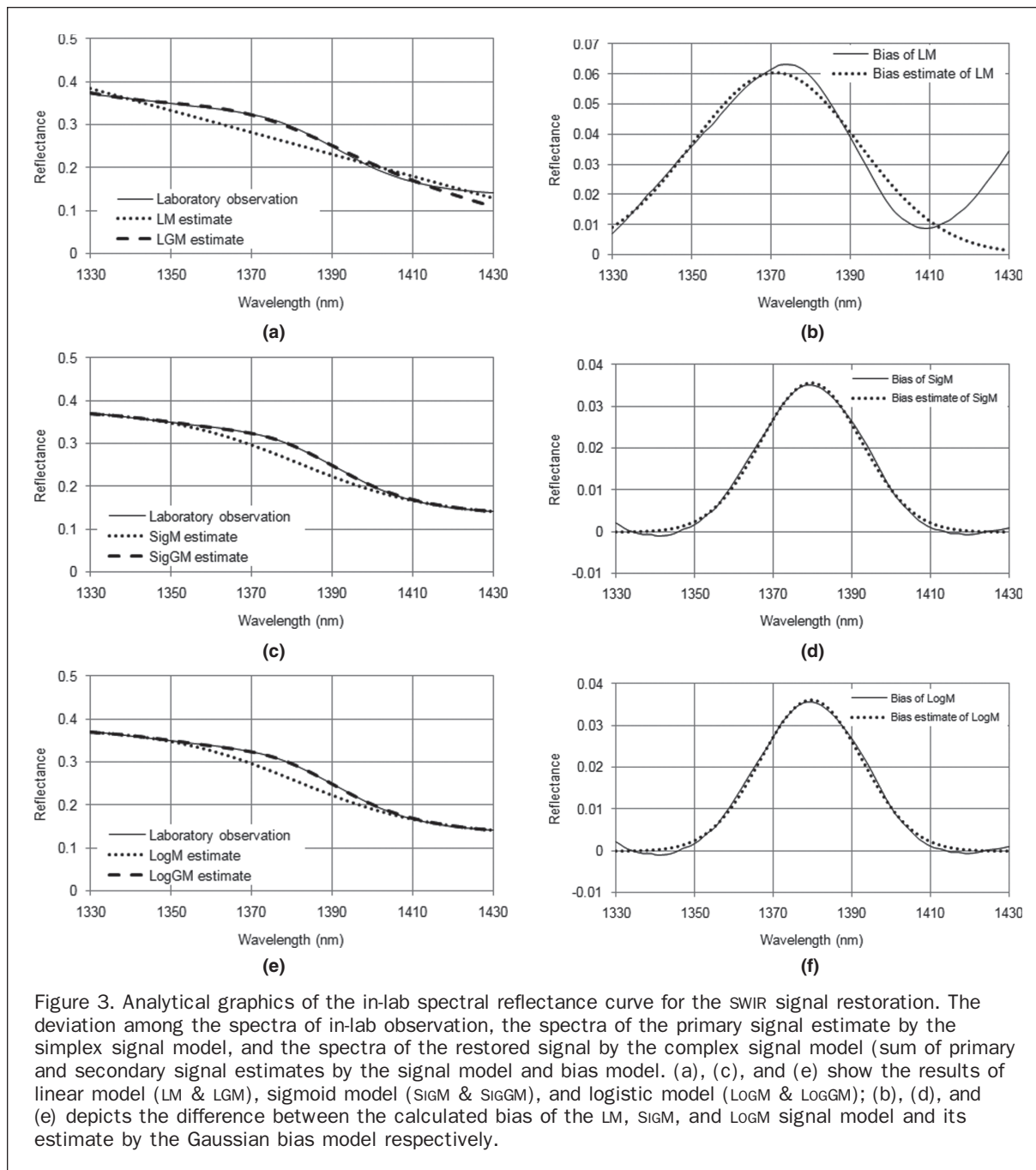
By examining the reflectance curve of the 1,330 to 1,430 nm SWIR region in Figure 3, the difference between the spectral estimates of the LM, SIGM, and LOGM signal models (shown with dotted line) and their counterparts of the LGM, SIGGM, and LOGGM (shown with long dashed line) are shown in Figure 3a, 3c, and 3e where the laboratory observation spectra (solid line) are also included for comparison. The reflectance predicted by the LM model is a line superimposed on top of the observed reflectance curve. It has three deviation portions at both ends and the center of the spectral region (Figure 3a). Such deviation could be partially corrected by the bias estimates (Figure 3b) predicted by the Gaussian bias model. The reflectance values predicted by the curvatures of SIGM and LOGM models show a small departure from the observed reflectance curve (see Figures 3c and 3e). In addition, their deviations from the observations, i.e., prediction biases, also show a uni-modal type of curves which could be well fitted in a Gaussian bias model. Figures 3d and 3f also demonstrate good fitness by small deviations between the bias of signal model (solid line) and the bias estimate (dashed line) for SIGM and LOGM. It is noted that both the curves of SIGGM estimate and LOGGM estimate fall on the laboratory observation curve. This overlap indicates that the sigmoid-Gaussian model and logistic-Gaussian model can work effectively and have very similar quality and effects in the modeling of 1,330 to 1,430 nm spectra.

TABLE 1. THE REGRESSION COEFFICIENTS AND THEIR VIFs OF THE IN-LAB SWIR SPECTRA FITTED SIGNAL MODELS

(a) LM			(b) SIGM			(c) LOGM		
Parameter	Coefficient	VIF	Parameter	Coefficient	VIF	Parameter	Coefficient	VIF
y_0	3.6092	1164.58	y_0	0.1259	48.70	min	0.1252	17.02
a	-0.0024	1164.58	x_0	1380.0480	13.16	EC_{50}	1380.0820	10.27
			a	0.2410	59.65	Hillslope	81.3998	7.52
			b	-16.9560	24.12	max	0.3663	1.27

TABLE 2. THE REGRESSION COEFFICIENTS AND THEIR VIFs OF THE BIAS MODELS OF THE IN-LAB SWIR SPECTRA FITTED SIGNAL MODELS

Source of bias (signal models)	(a) LM ($R^2 = 0.72$)		(b) SIGM ($R^2 = 0.99$)		(c) LOGM ($R^2 = 0.99$)	
	Coefficient	VIF	Coefficient	VIF	Coefficient	VIF
Parameters						
a	0.0539	1.5	0.0342	1.5	0.0348	1.5
b	22.510	1.5	12.8650	1.5	12.8810	1.5
x_0	1370.75	1.0	1379.4400	1.0	1379.4500	1.0



Application of the Signal Restoration Method

Experimental results obtained by the laboratory spectra modeling have shown the logic of integrating empirically nonlinear models and Gaussian model (SigGM and LogGM), and that it is a useful approach to use the noise-free spectra as samples to initiate the spectra in the 1,350 to 1,410 nm SWIR region of tree leaves. Meanwhile the residual of those empirical nonlinear models is a function of wavelength and can be well described by Gaussian model. It offers the shortage quantity of primary signal estimate of the simplex signal models that is LM, SigM, and LogM to completely derive the restored spectra in the 1,350 to 1,410 nm region. This quantity (named as secondary signal estimate in previous section) works like an offset for the reflectance restoration in satellite image processing where the gain and the offset for each band of the image will be used to transform digital number to radiance.

Let $FSME_{\lambda}$ be the primary signal estimate of a signal model for a specific wavelength λ in the noisy SWIR region for the in-field individual spectra of the red cypress leaves. Our proposed signal restoration method calculates the restored in-field signal (restored_FR $_{\lambda}$) in the noisy wavelength region by adding up the primary signal estimate $FSME_{\lambda}$ and the secondary signal estimate BME_{λ} produced by the bias model. This method can be described in the following equations:

- $LGM \text{ restored_FR}_{\lambda} = LM_FSME_{\lambda} + LM_BME_{\lambda}$,
- $SigGM \text{ restored_FR}_{\lambda} = SigM_FSME_{\lambda} + SigM_BME_{\lambda}$, and
- $LogGM \text{ restored_FR}_{\lambda} = LogM_FSME_{\lambda} + LogM_BME_{\lambda}$.

Table 3 lists the regression coefficients and VIFs of the LM, SigM, and LogM signal models of the field spectra based on the noise-free region of 1,310 to 1,330 nm and 1,410 to 1,430 nm of the tree crown of red cypress. These signal

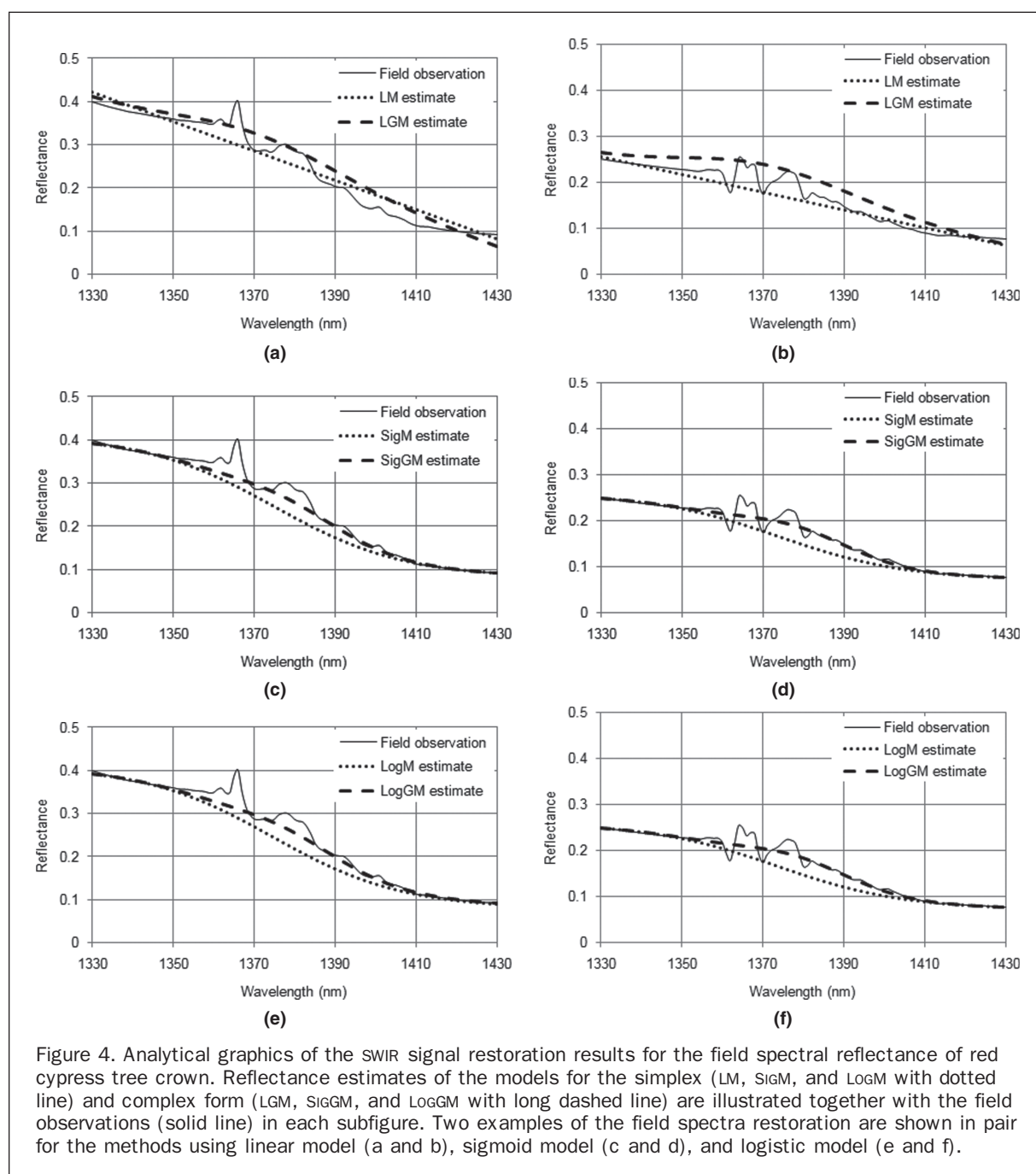
TABLE 3. THE REGRESSION COEFFICIENTS AND THEIR VIFs OF THE IN-FIELD SWIR SPECTRA FITTED SIGNAL MODELS

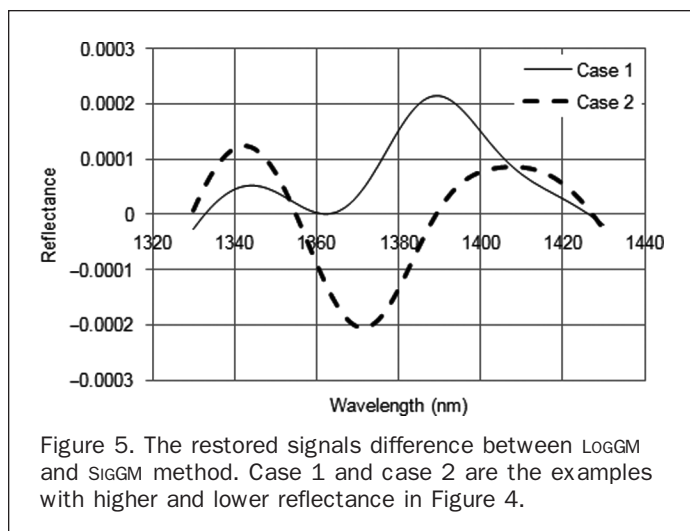
(a) LM			(b) SigM			(c) LogM		
Parameter	Coefficient	VIF	Parameter	Coefficient	VIF	Parameter	Coefficient	VIF
y_0	3.5644	1164.58	y_0	0.0674	35.98	min	0.0671	13.21
a	-0.0024	1164.58	x_0	1375.2510	12.18	EC_{50}	1375.2290	10.52
			a	0.2379	50.13	$Hillslope$	86.3627	9.30
			b	-15.9960	23.17	max	0.3053	1.26

models of the field spectra were statistically significant at the level $P < 0.0001$ and have a determination coefficient of $R^2 = 0.99$. In addition the coefficients and their VIFs for LM, SIGM, and LOGM signal models are very close to the ones for the laboratory spectra. Compared to the suitability of the indicator VIF, it is expected that the LogM signal model is

more favorable than SIGM, while the LM signal model is not suitable.

The combination of the signal model derived from the field spectra and bias model derived from the laboratory spectra has similar effects in the signal restoration of the noisy SWIR region. Figure 4 demonstrates a good example of an





application of our proposed restoration method; there are two examples shown in this figure. One of them is with the lowest reflectance and another one is with the highest reflectance in the field experimental cases. It can be seen that the restoration efficiency of SIGGM and LOGGM is nearly the same while the LGM has the problem of over-estimation. By comparing the results from laboratory reflectance modeling in the previous section, it was found that the sigmoid and the logistic models were visually similar to each other. The difference of restored signals or reflectance between LOGGM and SIGGM methods ($\text{LOGGM restored_FR}_\lambda - \text{SIGGM restored_FR}_\lambda$) for the in-field samples is less than ± 0.0003 (Figure 5). It therefore suggested that the integration of the four-parameter logistic model (LOGM) or the four-parameter sigmoid model (SIGM) with the three-parameter Gaussian model (GM) would be a good method for correcting noisy spectra in the SWIR region from 1,350 to 1,410 nm.

Conclusions

Hyperspectral remote sensing offers a wide variety of detailed and subtle spectral signatures of material substances for scientific research, product management, environmental management, and climate change, etc. Despite that spectroradiometer devices have been successfully used for the spectral measurement in field study, the noisy or low signal-to-noise ratio in the shortwave infrared (SWIR) wavelength range makes valuable signatures in this range unusable. For example, the water absorption bands in the vegetation reflectance curve such as 1,300 to 1,650 nm is an important feature for vegetation science. The higher the water content in leaves the stronger the absorption feature created in this SWIR region. This feature not only can be used to diagnose the water status of leaves, but also can be as an index for drought stress, moisture status of the soil surface and/or the soil profile pit, and the vegetation physiology. Moreover, the atmospheric correction might be another interesting issue. It is highly desirable to remove the noise in 1,350 to 1,410 nm region to make more effective use of the in-field SWIR spectra.

Based on the in-laboratory measurements of red cypress leaves, it was found that the spectra in this noisy region behave like an S-type curve and are very stable. According to our experiments a four-parameter sigmoid model (SIGM) or four-parameter logistic model (LOGM) can better dictate the trend of major spectral signals. Furthermore, a three-parameter Gaussian model ($\text{GM}_{\text{sigmoid}}$ or $\text{GM}_{\text{logistic}}$) can be also used to simulate the minor signals or the residuals of these two signal models. Combining the major signal model and

the minor signal model, two new models ($\text{SigGM} = \text{SigM} + \text{GM}_{\text{sigmoid}}$ and $\text{LogGM} = \text{LogM} + \text{GM}_{\text{logistic}}$) can be also derived to precisely describe the spectral behavior of tree leaves. Our proposed models were evaluated by in-field spectral measurements of red cypress and shown to be very useful in removing noise as well as in the restoration of the signal in the noisy SWIR region. Besides, our proposed signal restoration method is also shown to be able to overcome high humidity and/or low solar intensity induced noise problem. It can further improve the accuracy of spectroradiometer field data and provides valuable SWIR water absorption features to be used for research of forestry, crop and soil, and radiometric calibration. This restoration method can be directly applied to individual spectra or mean spectra of the in-field reflectance measurements. However, it should be noted that our method is not suitable for the noise in 1,800 to 2,000 nm, which is another water absorption region in SWIR. This is because there is no noise-free spectrum available to be used for reference in the 1,800 to 2,000 nm region.

Acknowledgments

This paper was finished with a budget partially from the National Science Committee (NSC) and the Taiwan Forestry Bureau (TFB). The first author therefore would like to thank both governmental organizations due to their financial supports for 2005 (NSC: 94-2313-B-415-017) and 2006 to 2007 (TFB95-B2-122 and TFB96-B2-109) which were sponsored by NSC and TFB, respectively.

References

- ASD, 1999. *ASD Technical Guide*, Third edition, Analytical Spectral Devices, Inc., Boulder, Colorado, 140 p.
- ASD, 2002. *FieldSpec Pro User's Guide*, Analytical Spectral Devices, Inc., Boulder, Colorado, 136 p.
- Clark, R.N., and T.L. Roush, 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing application, *Journal of Geophysical Research*, 89:6329–6340.
- Clark, R.N., G.A. Swayze, K.E. Livo, R.F. Kokaly, T.V.V. King, J.B. Dalton, J.S. Vance, B.W. Rockwell, T. Hoefen, and R.R. McDougal, 2003. Surface reflectance calibration of terrestrial imaging spectroscopy data: A tutorial using AVIRIS, *Proceedings of the 11th JPL Airborne Earth Science Workshop* (R.O. Green, editor), 05-08 March 2002, Pasadena, California (JPL Publication 03-4), pp. 43–63.
- Clevers, J.G.P.W., G.W.A.M. van der Heijden, S. Verzakov, and M.E. Schaepman, 2007. Estimating grassland biomass using SVM band shaving of hyperspectral data, *Photogrammetric Engineering & Remote Sensing*, 73(10):1141–1148.
- Clevers, J.G.P.W., L. Kooistra, and M.E. Schaepman, 2008. Using spectral information from the NIR water absorption features for the retrieval of canopy water content, *International Journal of Applied Earth Observation and Geoinformation*, 10:388–397.
- Clevers, J.G.P.W., L. Kooistra, and M.E. Schaepman, 2010. Estimating canopy water content using hyperspectral remote sensing data, *International Journal of Applied Earth Observation and Geoinformation*, 12:119–125.
- Curran, P.J., 1989. Remote sensing of foliar chemistry, *Remote Sensing of Environment*, 30(3):271–278.
- Estep, L., and G.A. Carter, 2005. Derivative analysis of AVIRIS data for crop stress detection, *Photogrammetric Engineering & Remote Sensing*, 71(12):1417–1421.
- Ge, S., R.I. Carruthers, D.F. Spencer, and Q. Yu., 2008. Canopy assessment of biochemical features by ground-based hyperspectral data for an invasive species, giant reed (*Arundo donax*), *Environment Monitoring Assessment*, 147:271–278.
- Govender, M., K. Chetty, and H. Bulcock, 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies, *Water SA*, 33(2):145–152.

- Longnecker, M.T., and R.L. Ott, 2004. *A First Course in Statistical Methods*, Thomson Brooks/Cole, Belmont, California, 615 p.
- Lin, C. 2005. *Exploration on Biochemical Properties of Forest Vegetation Using Hyperspectral Remote Sensing Techniques*, Project Report No. NSC94-2313-B-415-017, National Science Council, Taipei, Taiwan, 9 p.
- Lin, C., S.C. Huang, H.L. Wen, H.C. Cheng, and M.H. Yang, 2006. *Hyperspectral Signatures Measurement and Database Construction for Calocedrus formosana and Cinnamomum camphora*, Project Report No: 95AS-12.2.3-e4, Council of Agriculture, Taipei, Taiwan, 28 p.
- Ma, J., and X. Chen, 2006. Adjacency effect estimation by ground spectra measurement and satellite optical sensor synchronous observation data, *Chinese Optics Letters*, 4:546–549.
- Mitchell, T.M., 1997. *Machine Learning*, WCB-McGraw-Hill, Boston, Washington, 417 p.
- Montgomery, D.C., E.A. Peck, and G.G. Vining, 2001. *Introduction to Linear Regression Analysis*, Third edition, John Wiley and Sons, New York, 641 p.
- Psomas, A., M. Kneubuehler, S. Huber, K.I. Itten, and N.E. Zimmermann, 2005. Seasonal variability in spectral reflectance for discriminating grasslands along a dry-mesic gradient in Switzerland, *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, 27–29 April, Warsaw University, Krakowskie Przedmiescie 30, Warsaw, Poland, pp. 709–722.
- ReliaSoft Corporation, 2005. *Weibull++: Life Data Analysis Reference*, version 7, ReliaSoft Publishing, Tucson, Arizona, 580 p.
- Vikhamar, D., R. Solberg, and K. Seidel, 2004. Reflectance modeling of snow-covered forests in hilly terrain, *Photogrammetric Engineering & Remote Sensing*, 70(9):1069–1079.
- Schaepman, M.E., G.W.W. Wamelink, H.F. van Dobben, M. Gloor, G. Schaepman-Strub, L. Kooistra, J.G.P.W. Clevers, A. Schmidt, and F. Berendse, 2007. River floodplain vegetation scenario development using imaging spectroscopy derived products as input variables in a dynamic vegetation model, *Photogrammetric Engineering & Remote Sensing*, 73(10):1179–1188.

(Received 12 August 2010; accepted 22 December 2010; final version 08 August 2011)