INFLUENCE OF SIMULATED SIGNAL DISTURBANCES ON PREDICTION ACCURACY OF CHEMICAL SOIL PROPERTIES

T. Jarmer [a, *], T. Udelhoven [b], J. Hill [c]

[a] Faculty of Civil & Environmental Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel, jarmer@technion.ac.il
[b] CRP-Gabriel Lippmann, D`partement 'Environnement et Agro-biotechnologies', 4422 Belvaux, Luxembourg, udelhove@lippmann.lu
[c] Remote Sensing Department, University of Trier, Behringstrasse 108, 54286 Trier, Germany, hillj@uni-trier.de

KEY WORDS: prediction accuracy, signal-to-noise, inorganic carbon, PLSR, simulation

ABSTRACT:

In this simulation study reflectance spectra of soil samples have been measured in the lab in the wavelength range 380 - 2,500 nm and resampled to HyMap spectral resolution (128 spectral bands). Reflectance values have been converted into top of atmosphere radiance considering different atmospheric conditions and sun elevations. Sensor-specific noise depending on a given Noise Equivalent Δ Radiance (NEΔR) was randomly simulated and added to the radiances. The resulting 'noisy' radiance spectra have been recalculated into reflectance spectra at surface assuming identical radiometric conditions.
Soil chemical properties from reflectance spectra were estimated using partial-least-square regression. Predictions accuracy of inorganic carbon concentrations were used to investigate the influence of radiometric disturbances as well as randomly distributed noise and spectral resolution. Cross-validated regression coefficient and root mean square error were used to compare results.
Results indicated that prediction accuracy of inorganic carbon depends on sensor-specific and radiometric parameters. The main influence was attributed to the Noise Equivalent Δ Radiance of the simulated sensor whereas a higher NEΔR to a limited extent could be compensated by the number of spectral bands used for prediction. In case of atmospheric influences it turned out that lower sun elevations reduced prediction accuracy and water vapour showed substantial influence on prediction accuracy.

1. INTRODUCTION

Hyperspectral remotely sensed data are more and more frequently used for quantitative assessment of chemical or physical properties under varying climatic conditions in fields like precision agriculture or land degradation (e.g., Ben-Dor et al., 2002; Shrestha et al. 2005, Seelige et al., 2006; Jarmer et al., 2007) However, the full potential of remote sensing for such applications has yet to be fully realized. To exploit minor differences in soil properties from hyperspectral signals, the quality of imaging spectrometer data plays an important role. Noise level in hyperspectral data is high as their narrow bandwidth can only capture very little energy which may be overcome by the self-generated noise inside

* Corresponding author

the sensors (Vaiphasa, 2006). Additionally, physical disturbances such as varying atmospheric or illumination conditions may make the situation worse as the disturbances decrease the precision of spectral signals recorded by the sensor.

Smoothing techniques are commonly used for noise removal from hyperspectral data. However, smoothing should be carefully applied to preserve original data properties and alternatively investigation on how much noise is acceptable to guarantee precision of results on original data without smoothing is required. Du et al. (2008) have investigated the impacts of noise on the accuracy of hyperspectral image classification by support vector machine while Gong and Zhang (1999) analysed the effect of noise on linear spectral unmixing. Aiazzi et al. (2006) modelled and estimated noise from hyperspectral data.

The aim of this study was to investigate the influence of different noise levels on the prediction accuracy of soil properties using hyperspectral data. Additionally, the impact of atmospheric conditions and sun elevations were simulated and studied. The influence of radiometric disturbances and normal distributed artificial noise on the prediction accuracy was analysed for inorganic carbon which is an important parameter in the assessment of soil development and soil condition of carbonatic soils.

2. DATA AND METHODS

The study site is located near Trier in the Eifel region (Rhineland-Palatinate, SW-Germany). For the investigation, a plot ("Dietrichskreuz", Helenenberg) with a size of approximately 14 hectares was selected. The soil type of this plot was an eutric cambisol derived from airblown silt over limestone. At dry weather conditions in August 1999 114 topsoil samples were taken using a sampling raster of 60x60 m to cover the spatial variability of soil properties.

Soil samples were air-dried in the laboratory, gently crushed in order to pass a 2 mm-sieve and carefully homogenised. The total amount of carbon in the soil samples was analysed by dry combustion at 1100°C with a Leco CHN 1000 analyser. The soil organic carbon (C_{org}) was measured at a temperature of 600°C with the Leco-RC 412 analyser.

Bi-directional reflectance measurements of the soil samples were carried out in the laboratory with an ASD FieldSpec-II spectroradiometer. Spectral readings were taken in 1 nm steps between 350 nm and 2500 nm using a reflectance standard of known reflectivity (Spectralon©). The optical head of the spectroradiometer was mounted on a tripod in nadir position with a distance of 10 cm to the sample. For illumination, a 1000 W quartz-halogen lamp set in a distance of approximately 30 cm and with a zenith angle of 30° was used.

Absolute bi-directional reflectance spectra were obtained by multiplying the raw reflectance spectra by the certified reflectivity of the Spectralon panel. Only the spectral range from 0.4 to 2.4 µm was used for the further study to exclude the noisy parts of the spectra. Further on, spectra have been resampled to 128 bands using the position, band width and filter characteristics of the airborne HyMap sensor. Vector-normalization after centring each spectrum on its average (Otto, 1999) was applied to reduce albedo differences between spectra while spectral signatures are kept.

Soil inorganic carbon concentrations (C_{inorg}) were used to investigate varying noise effects on reflectance spectra. Estimation of C_{inorg} from reflectance spectra was obtained by PLSR („partial-least-square regression"). PLSR results were cross-validated according to the 'leave-one-out-method', which means that each sample is estimated by an empirical-statistical model that was calibrated using the remaining (n-1) samples. Coefficient of determination (r^2), and root mean squared error (RMSE) were calculated to assess prediction accuracy.

3. SIMULATION

Reflectance spectra have been converted into at-sensor radiance for the altitude 'space' considering different atmospheric conditions and sun elevations. Radiometric conditions were calculated for the geographic location of the city of Trier (49°45' / 6°38'). Artificial normal distributed sensor-specific noise against Noise Equivalent Δ Radiance was generated and added to at-sensor radiance. This noise considered the following wavelength-dependent amount of energy:

- 400 - 1000nm: 0.010 mWatt/cm²/srµm
- 1000 - 1900nm: 0.006 mWatt/cm²/srµm
- 1900 - 2400nm: 0.004 mWatt/cm²/srµm

The real 'signal-to-noise' ratio of the HyMap sensor however was disregarded. Resulting 'noisy' radiance spectra have then converted back to reflectance spectra at ground reassuming identical radiometric condition as used upwards. Sensitivity assessment of PLSR models against an error term was simplified carried out by adding artificial normal distributed noise spectra. The error terms were generated for different noise levels by using the above mentioned values as maximum and multiply them by different factors.

4. RESULTS AND DISCUSSION

The investigated soil samples vary between 0.54 g kg^{-1} C_{inorg} in the minimum and 46.27 g kg^{-1} C_{inorg} in the maximum with a mean of 12.4 C_{inorg} and a standard deviation of 11.2 C_{inorg}. The data were found normal distributed at 99 percent level.
In a first step, a PLSR model for C_{inorg} prediction was developed on original laboratory reflections spectra (no radiometric manipulation, no noise) at HyMap resolution. Estimation of C_{inorg} with this model using seven latent variables resulted in a cross-validated r^2 of 0.953 (cross-validated RMSE = 2.63). The radiometric manipulated and noise influenced spectra were used as a test data set for this model and compared concerning their retrieved prediction accuracy and RMSE. Different scenarios were included in the simulation to evaluate the varying influencing parameters, such as noise levels, atmospheric conditions, aerosol variations, sun elevation, water vapour concentrations and spectral resolution.
Hazy atmospheric conditions with low sun elevation (30th March) and normal atmospheric conditions with relatively high sun elevation (15th June) were considered to analyse the influence of different radiometric conditions. Different noise levels (1x noise, 2x noise, 3x noise, 4x noise, 5x noise, 10x noise) have been applied to radiance spectra for these two simulated radiometric conditions. While for both radiometric conditions adding single noise at least a r^2 higher than 0.8 (RMSE: 5 - 6 g kg^{-1} C_{inorg}) was achieved, the accuracy of the PLSR model already substantially decreased below 0.7 for doubled noise (Tab. 1). Further increase of noise level spiralled downward model accuracy. Poorer radiometric conditions in relation to good conditions produced 40 percent worse RMSE in average at the same noise levels.

Noise level	30th March r^2	30th March RMSE	15th June r^2	15th June RMSE
1x	0.803	6.22	0.875	4.66
2x	0.553	12.11	0.687	8.74
3x	0.384	18.15	0.521	13.03
4x	0.282	24.22	0.401	17.35
5x	0.220	30.25	0.319	21.69
10x	0.102	60.35	0.145	43.39

Original data (no noise added): r^2_{cv}: 0.953; $RMSE_{cv}$: 2.63.

Table 1. Influence of noise at different radiometric conditions (30th March: hazy atmospheric conditions, low sun elevation; 15th June: normal atmospheric conditions, high sun elevation)

Three different aerosol functions (Tab. 2) were used to analyse effects of aerosols' concentration and size distribution with respect to different noise levels. The variation of aerosol concentration and size distribution showed no noteworthy influence on simulations. The

Ångstrom relation	single noise		double noise	
	r^2	RMSE	r^2	RMSE
α = -1.1; β = 0.1	0.910	3.82	0.778	6.76
α = -1.0; β = 0.2	0.910	3.82	0.775	6.82
α = -1.3; β = 0.05	0.911	3.79	0.778	6.74

Original data (no noise added): r^2_{cv}: 0.953; $RMSE_{cv}$: 2.63.

Table 2. Influence of noise at different aerosol functions (9.00 h GMT)

model accuracy was only affected by the added noise.

Water vapour concentration in the atmosphere was expected to influence model accuracy. Therefore, the influence of three different water vapour concentrations (1.46 g cm^{-3}; 2.92 g cm^{-3}; 4.38 g cm^{-3}) on radiance spectra were considered while all other parameters were left constant. Model accuracy strongly decreased with higher water vapour concentration (Tab. 3). While a water vapour concentration of 1.46 g cm^{-3} for both radiance spectra with single and double noise added permitted model accuracy beyond 0.8, this could only be achieved for radiance spectra with single noise added at higher water vapour concentrations. Radiance spectra with single noise added even still provided better results with a simulated water vapour concentration of 4.38 g/cm^3 than radiance spectra with double noise added did with a simulated water vapour concentration of 1.46 g/cm^3.

Water vapour concentration	single noise		double noise	
	r^2	RMSE	r^2	RMSE
1.46 g cm-3	0.92	3.5	0.81	6.1
2.92 g cm-3	0.88	4.7	0.69	8.7
4.38 g cm-3	0.82	6.0	0.57	11.5

Original data (no noise added): r^2_{cv}: 0.953; $RMSE_{cv}$: 2.63

Table 3. Influence of noise at different water vapour concentrations

Four different sun zenith angles (63°, 52°, 43°, 40°) were considered to the influence of sun altitude while all other conditions were left constant. Radiance spectra with single noise added provided an r^2 higher than 0.75 with an RMSE below 8 g kg^{-1} C_{inorg} (Tab. 4). Model accuracy for sun elevation higher than 45° increased to an r^2 higher than 0.85 (RMSE < 5 g kg^{-1} C_{inorg}). The radiance spectra with double noise added only allowed a model accuracy of 0.7 and an RMSE of 9 g kg^{-1} C_{inorg} at sun zenith angles less than 40°.

Sun zenith angle	single noise		double noise	
	r^2	RMSE	r^2	RMSE
63°	0.755	7.25	0.486	14.15
52°	0.838	5.47	0.613	10.51
43°	0.868	4.81	0.673	9.06
40°	0.875	4.66	0.687	8.74

Original data (no noise added): r^2_{cv}: 0.953; $RMSE_{cv}$: 2.63.

Table 4. Influence of noise at different sun zenith angles

Hyperspectral data provide much more spectral information than often needed. Consequently, the relationship between sensor noise and the number of spectral bands was finally investigated in this simulation study. Four different spectral resolutions were simulated reducing the original radiance spectra stepwise at half of the number of spectral bands and further excluding bands within water absorption features at 1.4 µm and 1.9 µm (Fig. 1). For further analysis spectral bands influenced by water vapour absorption were excluded, resulting in 121, 60, 30 and 14 spectral bands included in simulations respectively. Results indicated that noise had more impact when less spectral bands were involved (Fig. 2). As expected, the prediction error rose with increasing noise component. Similar results were reported by Udelhoven et al. (Udelhoven et al., 2001) who compared prediction accuracy for simulated data of different hyperspectral sensors.

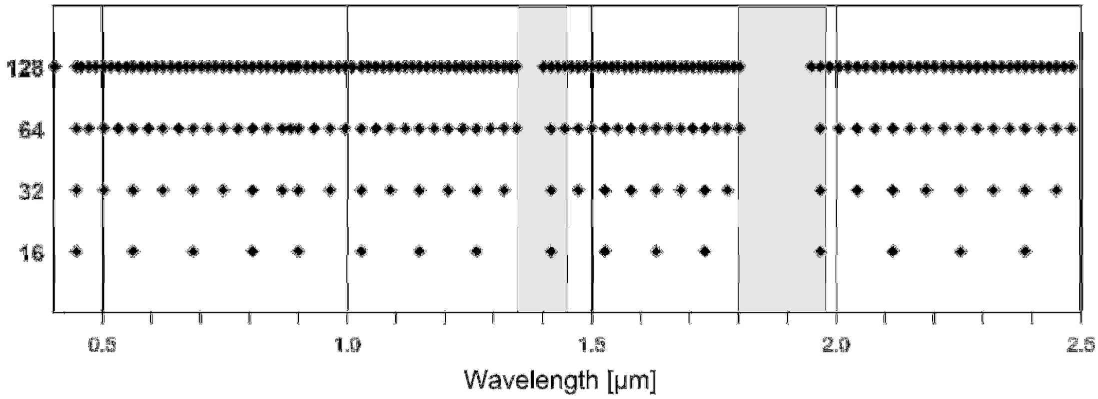

Figure 1. Wavelength position of spectral bands used in simulations on effects of different numbers of spectral bands (in grey: bands in water vapour absorptions at 1.4 µm and 1.9 µm excluded from analysis)

Number of spectral bands	no noise added r^2	RMSE	single noise r^2	RMSE	double noise r^2	RMSE
14	0.810	5.28	0.568	8.83	0.265	15.61
30	0.944	2.86	0.729	7.19	0.420	12.90
60	0.955	2.56	0.724	6.91	0.415	13.68
121	0.953	2.63	0.875	4.66	0.687	8.74

Original data (no noise added): r^2_{cv}: 0.953; $RMSE_{cv}$: 2.63.

Table 5. Influence of numbers of spectral bands at different noise levels noise

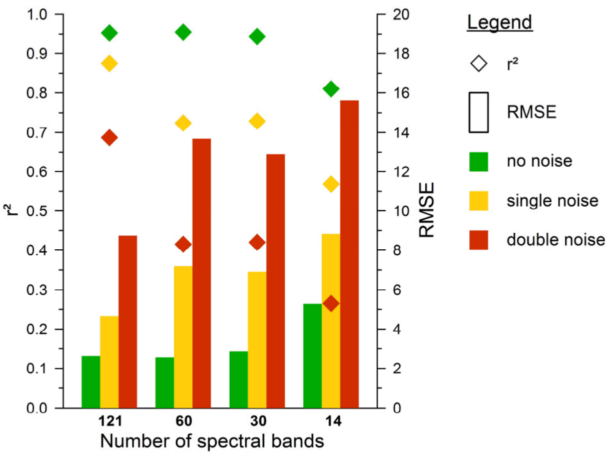

Figure 2. Influence of numbers of spectral bands at different noise levels

Radiance spectra without artificial noise added still permitted a model accuracy of 0.8 with only 14 spectral bands used while radiance spectra manipulated with single noise allowed an r^2 of 0.7 with 30 spectral bands. To achieve a model accuracy of 0.7 for radiance spectra with double noise added the whole spectral information (121 bands) were required. Reducing the spectral resolution to 60 spectral bands led to a model accuracy of 0.4 and an RMSE higher than 13 g kg^{-1} (Tab. 5).

An explanation for the rise of the RMSE is given by decrease of collinearity degree within spectra of considered spectral bands. PLS uses redundancy of x-matrix to stabilise prediction results against noise influence (Beebe et al. 1998). This redundancy was highest with 121 bands due to the relatively closely adjacent spectral sampling points while for only 14 bands it was least. Consequently, adding noise to this data reduced collinearity degree in particular strong.

5. CONCLUSIONS

This simulation study showed that the quality of quantitative assessment of soil properties is depending on sensor-specific and atmospheric parameters. The most important disturbance variable was found the Noise Equivalent Δ Radiance of the sensor. A higher NEΔR could be compensated to a certain extent by a high number of spectral bands. With respect to atmospheric influences it has to be considered that lower sun elevation decreased prediction accuracy and water vapour concentration was found to have a strong influence on prediction accuracy in general. Consequently, future remote sensing sensors should offer a high spectral resolution within atmospheric windows whereas a higher number of spectral bands stabilise prediction results against decreasing signal-to-noise.

ACKNOWLEDGEMENTS

This study was financially supported by the German Federal Ministry of Education and Research and by the German Research Foundation which is gratefully acknowledged.

REFERENCES

Aiazzi, B., Alparone, L., Barducci, A., Baronti, S., Marcoionni, P., Pippi, I., Selva, M., 2006. Noise modelling and estimation of hyperspectral data from airborne imaging spectrometers. *Annals of Geophysics*, 49(1), 9 pp.

Beebe, K.R., Pell, R.J., Seasholtz, M.B., 1998. *Chemometrics: A practical guide*. Wiley & Sons, New York, 348 pp.

Ben-Dor, E., Patkin, K., Banin, A., Karnielli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data - a case study over clayey soils in Israel. *International Journal of Remote Sensing*, 23(6), pp. 1043-1062.

Du, P., Wang, X., Tan, K., Foody, G.F., 2008. Impacts of noise on the accuracy of hyperspectral image classification by SVM. In: *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, Shanghai, P. R. China, June 25-27, 2008, pp. 138-144.

Gong, P., Zhang, A., 1999. Noise effect on linear spectral unmixing. *Geographic Information Sciences*, 5, pp. 52–57.

Jarmer, T., Hill, J., Mader, S., 2007. The use of hyperspectral remote sensing data for the assessment of chemical properties of dryland soils in SE-Spain. In: *Proceedings of the 5th EARSeL Workshop on Imaging Spectroscopy*, Brugge, April 23-25th 2007.

Otto, M., 1999. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*. Wiley-VCH, 330 pp.

Selige, T., Böhner, J., Schmidhalter, U., 2006. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma*, 136(1-2), pp. 235-244.

Shrestha, D.P., Margate, D.E., van der Meer, F., Anh, H.V., 2005. Analysis and classification of hyperspectral data for mapping land degradation: an application in southern Spain. *International Journal of Applied Earth Observation and Geoinformation*, 7, pp. 85–96.

Udelhoven, T., Jarmer, T., Hill, J., 2001. Die Nutzung von Hyperspektraldaten zur Erfassung von Bodeneigenschaften - Grundlagen und Anwendungsperspektiven. In: *Photogrammetrie und Fernerkundung - Geoinformation: Geodaten schaffen Verbindungen*, Proc. 21. Jahrestagung der DGPF und 18. Nutzerseminar des Deutschen Fernerkundungsdatenzentrums, 588 pp.

Vaiphasa, C., 2006. Consideration of smoothing techniques for hyperspectral remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2), pp. 91-99.