On the reproducibility of reflectance factors: implications for EO science

Karen Anderson
School of Geography,
University of Exeter, UK
karen.anderson@exeter.ac.uk

E. J. Milton and **Vincent Odongo**
School of Geography,
University of Southampton, UK
e.j.milton@soton.ac.uk

Jennifer L. Dungan
NASA Ames Research Center,
Moffett Field, CA, USA
Jennifer.L.Dungan@nasa.gov

Abstract— **Measurements of reflectance quantities collected in natural radiation conditions underpin quantitative Earth observation (EO) science through calibration, validation and atmospheric correction techniques. Despite their importance to the longevity of EO data, only a few studies have commented on the reliability of such measurements. This paper will report on results from three experiments. Each was designed to explore a different facet of measurement uncertainty in field measurements of hemispherical conical reflectance factors (HCRF), and to consider the broader implications of this for EO. From an end-user's standpoint we will describe a simple methodology for standard uncertainty characterisation for any reflectance-measurement scenario. The work provides a broad basis for considering standardised approaches to uncertainty characterisation and reporting, which is a necessary step towards improving the reproducibility and traceability of EO data and associated products.**

Keywords: reflectance, uncertainty, spectroradiometer, HCRF

I. INTRODUCTION

Measurements of reflectance quantities collected in natural environmental conditions are of fundamental importance in Earth observation (EO) science, because they underpin a variety of quantitative pre-processing techniques including: vicarious calibration (Moran et al. 2001; Teillet et al. 2001; Thorne 2001); atmospheric correction (Smith and Milton, 1999, Liang et al. 2002), and product validation (Salomon et al. 2006). Additionally, close-range spectroradiometric studies are useful for measuring changes in surface properties temporally and spatially (Gamon et al. 2006). Despite these successes, one of the biggest challenges continues to be the accurate, reproducible characterisation of natural surface reflectance properties measured in the solar radiation environment (Milton et al. 2009). An emerging question in the community is: how reliable are those measurements? Inconsistencies in the description of reflectance quantities, coupled with user-to-user variations in data collection methods, and limited uncertainty reporting have all resulted in non-reproducible measurements of "reflectance". This paper will report on an experiment to study how the uncertainty in HCRF depends upon measurement conditions (lab vs field) and surface type (artificial vs quasi-natural) and will show how this varies over timescales from minutes to months.

II. APPROACH

A. Instrument Uncertainty

Our first research question focused on determining the value of laboratory-determined instrument characterizations (i.e. noise equivalent delta radiance, NEΔL) in defining operational measurement uncertainty. NEΔL provides a measure of inherent system uncertainty which is usually considered more useful as a measure of system performance than commonly-reported signal-to-noise ratios because it is not dependent on the magnitude of the signal. A combined laboratory and field experiment allowed comparison of NEΔL characterized using a stable radiance source in lab conditions with field-derived standard uncertainties. The latter was the standard uncertainties in the HCRF characterization of a grey Spectralon panel (calibrated HCRF=0.75 at 700nm)[†] measured in the field under optimal atmospheric conditions[‡]. Laboratory-characterized NEΔL was propagated to a noise equivalent delta reflectance (NEΔρ) according to equation 1 for comparison with the field-derived uncertainty measure.

$$NE\Delta\rho = \rho\sqrt{\left(\frac{NE\Delta L_{tar}}{L}\right)^2 + \left(\frac{NE\Delta L_{ref}}{E}\right)^2} \quad [1]^{\S}$$

B. Influence of Surface Type

An automated tramway-mounted multiband radiometer was used to measure the uncertainty in the HCRF of three different surfaces relative to a calibrated white ceramic tile. This experiment addressed the question – *how does HCRF uncertainty vary with surface type?* Samples comprised a flat panel of brown ceramic tiles (artificial) and trays of air-dry sand and air-dry angular white gravel (quasi-natural). These were placed at intervals adjacent to an eight metre long track, along which a calibrated radiometer travelled every 10 minutes between 11:00 and 13:00 UTC. Radiance was measured in eight spectral bands (FWHM 10 nm) and stored locally in a data logger. The diffuse to global irradiance ratio

---

[†] Expressed as the standard deviation (s.d.) in 10 HCRF replicates of panel serial # SRT75-180 10051B collected in a 5 minute time period.

[‡] Diffuse to global irradiance ratio (400-700 nm) = 0.129

[§] Where: $NE\Delta L_{tar}$ = laboratory NEΔL for target sensor (s/n #2002); $NE\Delta L_{ref}$ = laboratory NEΔL for reference sensor (s/n #2003); L = mean radiance of target surface (n=10); E = mean irradiance measured by reference sensor (n=10); and ρ = mean reflectance factor of target surface (n=10).

(400-700 nm) was 0.21 and did not change during the period of measurements which covered solar zenith angles from 57.1° to 59.7°.

C. *Temporal Stability of Calibration Surface HCRF*

Many studies make a central assumption that calibration surfaces maintain invariant reflectance characteristics over a range of timescales. The final part of the project addressed the question: *how stable is calibration surface HCRF over monthly timescales?* A carefully designed experiment was used that allowed good spatial positioning precision of a spectroradiometer over a concrete calibration surface, at a UK test site on different dates. Full details of the experiment are given in (Anderson and Milton 2006a). The experiment allowed precise measurement of calibration surface HCRFs (useable range 400-1050 nm) over two years. An intercalibrated dual-beam GER1500 spectroradiometer was used (Anderson and Milton 2006b). The reproducibility of the spectral measurements was determined empirically in field conditions. The data showed that for a calibration surface where HCRF = 0.27 (700 nm), the measurement reproducibility was HCRF = 0.27± 0.0026 (±1 s.d.).

III. RESULTS

A. *Instrument Uncertainty*

The results of the laboratory instrument characterization showed a "typical" (Markham et al. 1995) response where NEΔL increased with wavelength and was much higher in the NIR – results for a single instrument (serial #2002) are shown in Figure 1. Once this was propagated to NEΔρ and compared with standard uncertainties measured in the field (Figure 2), the discrepancy between the laboratory-derived and field-derived uncertainties become more apparent. Although both show a similar pattern, there is a considerable difference in the magnitude of the uncertainty, with field measured- NEΔρ showing several orders of magnitude difference in the visible region. Beyond 1000 nm, laboratory-derived measures of instrument uncertainty explain most of the field-derived variability. This illustrates that NEΔL should not be considered a complete description of measurement uncertainty – in the field scenario, a range of other factors play a role, including solar angle effects, skylight variability and methodological-induced uncertainty (i.e. positioning precision over spatially non-uniform targets).

B. *Influence of Surface Type*

Table 1 shows the results of the tramway experiment, with the uncertainty in HCRF expressed both as NEΔρ and as the coefficient of variation (CV=s.d./mean), so that the different surface types may be compared. The values of NEΔρ for the inert tile are comparable with those obtained in experiment 1, although the HCRF of the tile shows lower uncertainty at most wavelengths compared with the field-measured NEΔρ, most likely due to the wider bandwidth of the multiband radiometer providing a better signal-to-noise ratio. The reflectance factors measured from the artificial tile surface showed least variability, with an average CV over all eight bands of 0.14%. The uncertainty of the natural surfaces was an order of magnitude greater, with averages of 1.58% for the gravel and 2.07% for the sand surface.

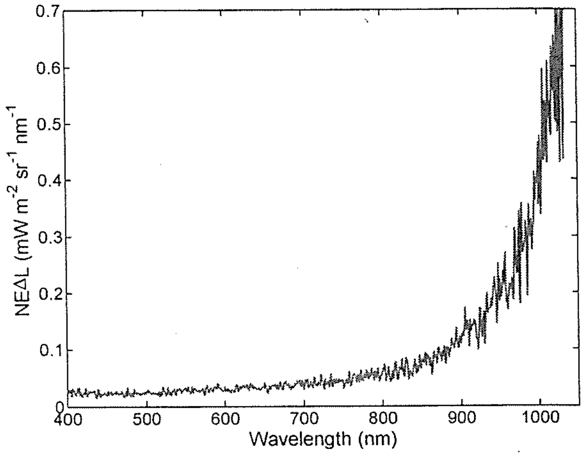

Figure 1. NEΔL for GER1500 serial #2002

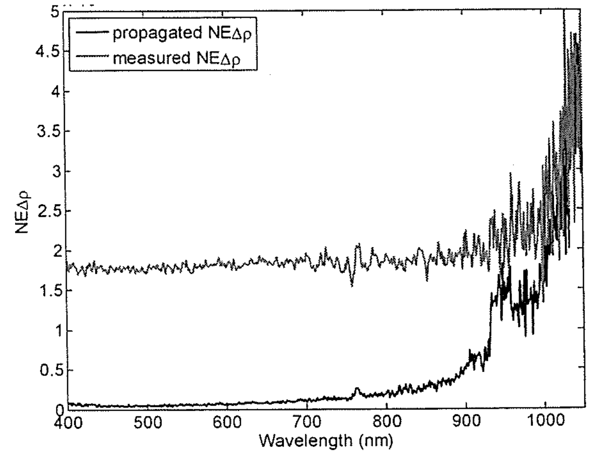

Figure 2. Comparison of laboratory-propagated NEΔρ and field-measured NEΔρ

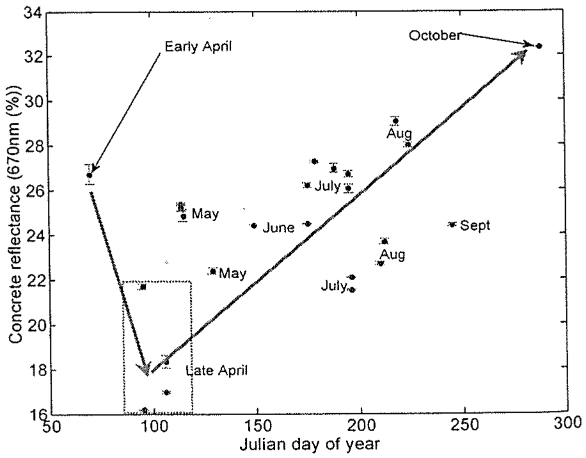

Figure 3. Trend in HCRF of a concrete calibration surface over a 2 year period (670 nm)

TABLE I. UNCERTAINTY IN HCRF FROM TRAMWAY EXPERIMENT

	Tile NE$\Delta\rho$ (x 10^{-3})	Tile CV(%)	Gravel CV(%)	Sand CV(%)
430nm	0.42	0.10	2.54	1.90
500nm	0.50	0.11	1.89	2.40
670nm	0.67	0.10	1.19	2.26
781nm	0.67	0.10	1.13	2.13
820nm	0.63	0.09	1.28	2.10
831nm	1.10	0.15	1.38	1.98
882nm	1.90	0.28	1.58	1.95
949nm	1.10	0.18	1.68	1.82

This additional uncertainty most likely arose from interaction between the surface and the illumination/atmospheric conditions during the measurements. Small changes in solar geometry and atmospheric turbidity were sufficient to affect the HCRF of the natural surfaces due to their irregular microrelief, but had little effect on the HCRF of the tiled surface. HCRFs from the sand surface were generally less precise than those from the gravel surface, probably because the absolute values were much lower (e.g. mean HCRF at 500nm: sand=0.06, gravel=0.52). These results suggest that the intrinsic uncertainty of HCRF measured in field conditions from sand and gravel surfaces with this particular radiometer is around ±2% relative to the mean.

C. *Temporal Variability in Calibration Surface HCRF*

The results from a two-year field experiment set up to test the temporal variability of a calibration surface HCRF showed variability over a range of timescales. The most dramatic change was the brightening of the calibration surface over seasonal timescales (Figure 3; Anderson and Milton 2006a). This was caused by seasonal growth of a biological material, which caused the reflectance factor to vary by a factor of two during the year (range = 0.164 at 670 nm). The spectral effect of this was most noticeable in field spectra collected in April. The same patterns in HCRF repeated over the-two year period, indicating the predictable nature of the biological signature. The research also suggested that the biological material was measurably affected on a daily basis by changes in relative humidity. This research highlights the dynamic nature of calibration target HCRF, and shows that the assumption of "invariance" can be invalid over a range of timescales. This is important for those using field sites for vicarious calibration or atmospheric correction purposes.

IV. DISCUSSION AND FORWARD DIRECTION

The results have demonstrated the range of uncertainties which should be considered when collecting field measurements of HCRF for supporting EO science. Key messages from this work are:

1. Measures such as NEΔL provide a useful baseline from which to assess instrument performance but are not helpful as a stand-alone indicator of spectral measurement uncertainty. This is because NEΔL does not take into account variability imposed by the methodology used, and measurement complexities from the field environment Wettle et al. 2004). We suggest that field-measured NE$\Delta\rho$ (red line, Fig. 2) is more representative of measurement uncertainty, because it captures NEΔL, coupled with environmental and methodological variability. Users of field spectroradiometers can generate this measure by calculating the s.d. of a set of measurements collected over the same target, in identical conditions to those used for data capture. Reviewers of research articles using such data should request all results be accompanied by such measures.
2. In EO studies the assumption of invariance in calibration surface HCRF is common. Our results have shown that seemingly inert surfaces such as sand, gravel and concrete exhibit temporal variability in HCRF over timescales ranging from hours to months. The dynamic nature of such targets should not be underestimated. Users of empirical atmospheric correction methods should know that the accuracy of any correction will degrade as the time delay between field measurements and EO data capture increases.

To conclude, we suggest that refinements in methodology are needed if field spectroscopy is to be successful in establishing its credentials as a reliable method of environmental measurement, underpinning quantitative EO science. Simple adjustments to measurement protocol will go a long way towards improving the reliability of data.

ACKNOWLEDGEMENTS

The Winston Churchill Memorial Trust provided funding for KA's research at NASA Ames and VO was supported by the EU Erasmus Mundus programme. The NERC FSF loaned a a variety of instruments to KA and EJM for all experiments. Ames' Airborne Science and Technology Lab and Dr. Liane Guild (NASA Ames) are thanked for their assistance in experiment 1. Dr Ted Hildum, Bradley Lobitz and Dr Vern Vanderbilt (NASA Ames) and Dr Nick Hamm (ITC) are thanked for their assistance and technical insight during the field and laboratory experiments. Mr Bill Damon's engineering expertise was fundamental to the success of experiments 2 and 3.

REFERENCES

Anderson, K. and E.J. Milton. (2006a). On the temporal stability of ground calibration targets: implications for the reproducibility of remote sensing methodologies. *International Journal of Remote Sensing*. 27, p. 3365-3374.

Anderson, K. and E.J. Milton. (2006b). Calibration of dual-beam spectroradiometric data. *International Journal of Remote Sensing*. 27 (5), 975-986.

Gamon, J.A., et al., (2006). A mobile tram system for systematic sampling ecosystem optical properties. *Remote Sensing of Environment*. 103, 246-254.

Liang, S., et al., (2002). Atmospheric correction of Landsat ETM+ land surface imagery: II. Validation and applications. *IEEE transactions on Geoscience and Remote Sensing*. 40 (12), 2736-2746.

Markham, B.L., et al. (1995). Radiometric characterisation of diode array field spectroradiometers. *Remote Sensing of Environment*. 41, 317-330.

Milton, E.J., et al., (2009). Progress in field spectroscopy. *Remote Sensing of Environment*. 113(SI), S92-S109

Moran, M.S., et al., (2001). A refined empirical line approach for reflectance factor retrieval from Landsat-5 TM and Landsat-7 ETM+. *Remote Sensing of Environment*. 78, 71-82.

Salomon, J.G., et al., (2006). Validation of the MODIS bidirectional reflectance distribution function and albedo retrievals using combined observations from the Aqua and Terra platforms. *IEEE transactions on Geoscience and Remote Sensing*. 44 (6), 1555-1565.

Smith, G.M. and E.J. Milton, (1999). The use of the Empirical Line Method to calibrate remotely sensed data to reflectance. *International Journal of Remote Sensing*. 20 (13), 2653-2662.

Teillet, P.M., et al., (2001). A generalised approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data. *Remote Sensing of Environment*. 77, 304-327.

Thome, K.J. (2001). Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method. *Remote Sensing of Environment*. 78, 27-38.

Wettle, M., V.E. Brando, and A.G. Dekker, (2004). A methodology for retrieval of environmental noise equivalent spectra applied to four Hyperion scenes of the same tropical coral reef. *Remote Sensing of Environment*. 93, 188-197.