

APPROACHES TO ESTABLISHING A METADATA STANDARD FOR FIELD SPECTROSCOPY DATASETS

Barbara A. Rasaiah^{a,*}, Tim J. Malthus^b, Chris Bellman^c, Laurie Chisholm^d, John Gamon^e,
Andreas Hueni^f, Alfredo Huete^g, Simon D. Jones^h, , Cindy Ongⁱ, Stuart Phinn^j, Chris
Roelfsema^k, Lola Suarez^l, Philip Townsend^m, Rebecca Trevithickⁿ, Mat Wyatt^o

^{a,c,h,l}RMIT University Melbourne, Australia

^bCSIRO Land and Water, Australia

^dUniversity of Wollongong, Australia

^eUniversity of Alberta, Canada

^fRemote Sensing Laboratories, University of Zurich, Switzerland

^gUniversity of Technology, Sydney, Australia

ⁱCSIRO Exploration and Mining, Australia

^{j,k}University of Queensland, Australia

^mUniversity of Wisconsin-Madison, USA

ⁿQueensland Department of Environment and Resource Management, Australia

^oiVEC, Australian National Data Service, Australia

ABSTRACT

There is an urgent need within the international remote sensing community to establish a metadata standard for field spectroscopy that ensures high quality, interoperable metadata sets that can be archived and shared efficiently within Earth observation data sharing systems. Careful examination of all stages of metadata collection and analysis can inform a robust standard that is applicable to a range of field campaigns. This paper presents approaches towards a standard that encompasses *in situ* metadata collection and initiatives towards sharing metadata within intelligent archiving systems.

Keyword (Index Terms) -- *Databases, Metadata, Field Spectroscopy, Validation, Calibration, Data Quality*

1. INTRODUCTION

With the ever-growing volume of *in situ* spectroscopy datasets from a broad variety of campaigns and instrumentation, there is an urgent need within the international remote sensing community to record, store, and share field spectroscopy data and metadata in consistent formats within intelligent archiving systems. Maintaining optimal integrity of the data is a key priority to ensure effective re-use of the data, enabling more efficient and higher impact remote sensing research.

National initiatives are providing the infrastructure, procedures and people that allow data collection, sharing and exchange such as NEON, EuroSpec, SpecNet and TERN, and are prompting a closer examination of requirements for consistency and interoperability of datasets, including those from field spectrometry.

Metadata are an important component in the cataloguing and analysis of *in situ* spectroscopy datasets because of their central role in identifying and quantifying the quality and reliability of spectral data and the products derived from them. The time invested in metadata collection is outweighed by its benefits in enabling detailed quality assessment of the data, and its multiple reuses [1]. Although standard protocols exist for some environments there is currently no internationally standardized methodology for collecting *in situ* spectroscopy data or metadata protocols for remote sensing applications across terrestrial, marine and atmospheric environments. Such protocols are likely to remain elusive, due to the variety of instruments and purposes of study. This makes rich and flexible metadata capabilities even more important, but also emphasizes the need for emerging standards tied to particular protocols.

2. BUILDING A METADATA STANDARD

Every stage of data collection and analysis (e.g. processing methods) also warrant good metadata. A correctly implemented “end-to-end”

metadata approach can help with provenance and in many cases, good metadata on calibration and instrument characterization is critical to the proper interpretation of the final product.

The criteria for building a metadata standard must address requirements for a wide range of terrestrial, marine, and atmospheric remote sensing campaigns suitable for use with airborne and satellite image data. These requirements were identified in an international experiment comprising a web-based survey [2] and an expert panel evaluation [3,4] that investigated protocols for recording metadata in field spectroscopy (see Figure 1).

The survey showed consensus among expert groups using a range of instrumentation allowing for a prioritization of critical metadata elements applicable to all campaigns, irrespective of the target being measured – these include metadata about the instrument, reference standards, and illumination conditions, among others. A secondary set of optional metadata elements standards was also identified, for the purpose of increasing the robustness, flexibility and discoverability of the metadataset. These include specific target properties and environmental conditions that could be influencing factors on the measurements. Consensus on the metadata elements

that are critical in a metadata standard was considerably larger among respondents from the same expert group (ex: vegetation/marine/mineralogy). Panel discussions [3,4] recommend that a metadatasets’ fitness-for-purpose and best practice in the field must be a central consideration when building a metadata standard that ensures robustness, flexibility, quality, and discoverability. This must also be balanced with the diversity of instrumentation used in the field and the need for strictness in standards [3]. Achieving consistency with emerging national standards is an additional factor in designing a field spectroscopy metadata standard that ensures interoperability among metadatasets.

3. SHARING METADATA

Earth observation data sharing systems such as ESDIS (Earth Science Data and Information System) and GEOSS (Global Earth Observation System of Systems), currently lack any explicit metadata specifications for *in situ* spectroscopy campaigns, hence the need for this study. These systems can be defined as ‘networks of networks’ that provide intelligent storage of Earth observation data from distributed data centres and delivery

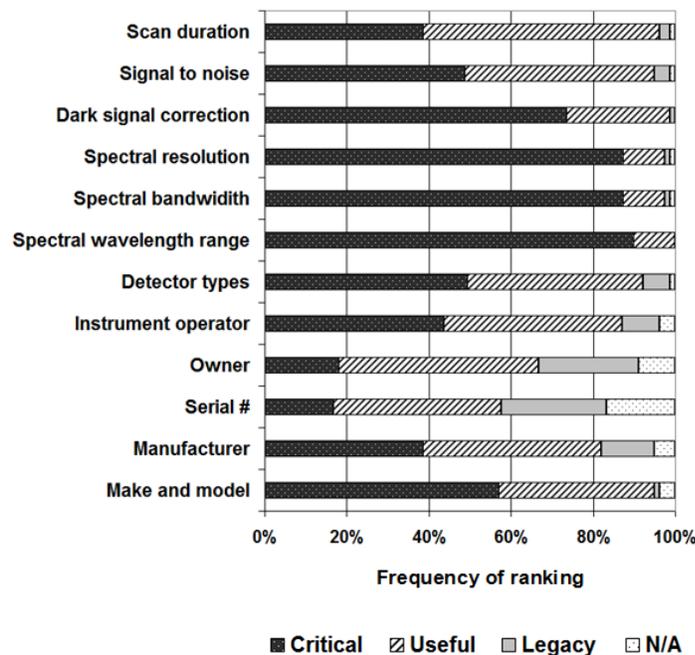


Figure 1 Results from the metadata survey showing frequency of criticality ranking where respondents ($n=79$) were asked to rate each field within the ‘instrument’ metadata category by one of four rankings: critical, useful, legacy potential, N/A

information through a web portal. They aggregate data from multiple databases and other sources in varying formats to a single point of access for a large population of users. Future integrations of spectral information systems within systems such as GEOSS would greatly benefit from adherence to international metadata standards, and with this in mind, according to standardisations should be pursued with alacrity; existing systems should be upgraded accordingly, once such standards have been defined. Adapting these systems to incorporate existing spectroscopy databases (SPECCHIO [5], Hyperspectral.info, DLR Spectral Archive) and spectral libraries (USGS Digital Spectral Library) requires applying the requirements for robust and high quality metadata unique to field spectroscopy (see Figure 2).

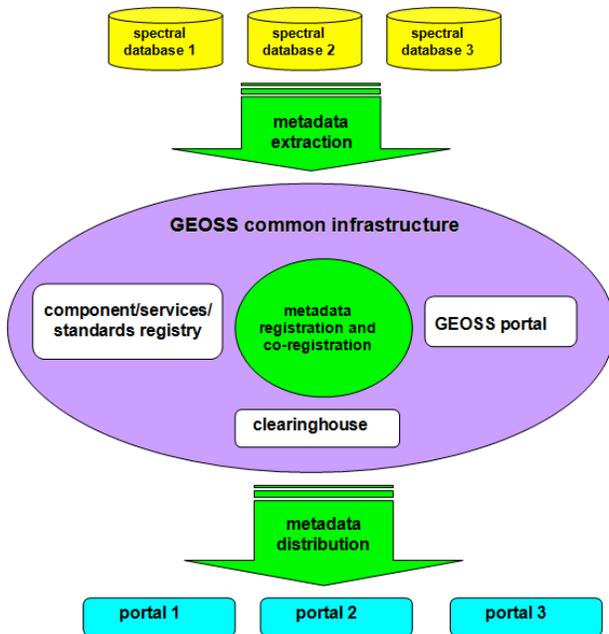


Figure 2 An adapted GEOSS data model illustrating the flow of field spectroscopy metadata through the GEOSS data infrastructure integrated with current

4. CONCLUSIONS

Longevity, interoperability, and quality of field spectroscopy datasets are dependent on establishing and adopting strict metadata standards on local and international scales. Within the context of archiving, sharing, and distributing metadata within Earth observation data systems, the need for such standards becomes imperative with the growing volume of *in situ* spectroscopy datasets and the increasingly

advanced analytics and reporting capabilities within intelligent archiving systems.

Existing challenges to efficient sharing and intercomparison of field spectroscopy metadata – including lack of quality control process, inconsistency in metadata collection, and *ad hoc* archiving systems – can be overcome with a recognition of the need to standardize and a cooperative spirit within the international remote sensing community.

REFERENCES

- [1] K. Pfitzner, A. Bollhöfer, and G. Carr, 2006. “A Standard Design for Collecting Vegetation Reference Spectra: Implementation and Implications for Data Sharing” *Spatial Science*, 52:2, pp. 79-92.
- [2] B. Rasaiah, T. J. Malthus, S. D. Jones, C. Bellman, A. Schembri, 2012. “Defining Critical Metadata for Field Spectroscopy Campaigns”, *International Journal of Remote Sensing* (under review).
- [3] T. J. Malthus, 2012. "Bio-optical datasets: best practice and legacy datasets", report on ACEAS Bio-Optical Data Workshop, Brisbane, 18-22 June 2012.
- [4] B. Rasaiah, 2012. “GSR 2 Spectral Libraries Workshop” report on GSR 2 Spectral Libraries Workshop, Melbourne, 12 December 2012.
- [5] A. Hueni, M. Kneubühler, 2010. “The Spectral Database SPECCHIO In Support of Cal/Val Activities”, ESA WORKSHOP.