

# DEVELOPMENT OF GLOBAL SOIL INFORMATION FACILITIES

*N H Batjes<sup>1\*</sup>, H I Reuter, P Tempel, T Hengl, J G B Leenaars, and P S Bindraban*

*<sup>1</sup>ISRIC - World Soil Information, P.O. Box 353, 6700 AJ Wageningen, The Netherlands*

*Email: niels.batjes@wur.nl*

## ABSTRACT

*ISRIC - World Soil Information has a mandate to serve the international community as custodian of global soil information and to increase awareness and understanding of the role soils in major global issues. To adapt to the current demand for soil information, ISRIC is updating its enterprise data management system, including procedures for registering acquired data such as lineage, versioning, quality assessment and control. Data can be submitted, queried and analysed using a growing range of web-based services—ultimately aiming at full and open exchange of data, metadata and products—through the ICSU-accredited World Data Centre for Soils.*

**Keywords:** Soils, Data curation, Analysis services, Web services, ICSU World Data System

## 1 INTRODUCTION

Human activities are a major driver of current changes in the Earth's atmosphere, oceans and landscapes (Watson *et al.*, 2000; UNEP, 2007) and increasingly affecting key environmental, social and economic functions of soil (Blum, 2005). Such changes need to be measured, monitored and modelled to allow governments, civil society and other sectors to take informed decisions about climate change mitigation and adaptation, world-food security, production of biofuels, combatting environmental degradation, and other global issues. In this context, ISRIC – World Soil Information is strengthening its capabilities in collecting, storing, processing and disseminating global soil and terrain information for research and development of sustainable and productive land use (Bindraban *et al.*, 2010).

ISRIC has a mandate to serve the international community as custodian of global soil information and to increase awareness and understanding of the role of soils in major global issues. It was created in 1966, following a request by the 1964 UNESCO General Council, and obtained the status of *ICSU World Data Centre* in 1989, initially named *WDC for Soil Geography and Classification*, later *WDC for Soils*. In August 2011, ISRIC was accredited as regular member of the new *ICSU World Data System (WDS)*. In that capacity, it deals directly with data curation and data analysis services.

This paper discusses recent developments at ISRIC, focusing on the on-going development and implementation of distributed global soil information facilities to ultimately provide quality-assured soil information for better science and development.

## 2 DATA CURATION

These days, new soil data are collected less and less while the same old (legacy) data and derived information are repeatedly used. Hence, it is vital to secure, maintain and expand data that support current information, including soil reference collections, soil maps and reports, and this ideally for public, free access (see 3.1).

Various collections are being maintained and regularly expanded by ISRIC. Since its establishment, a primary responsibility has been to create and maintain a world soil reference collection. The collection comprises over 1000 soil monoliths sampled to be representative of the soil units of the *FAO-UNESCO Soil Map of the World (FAO, 1971-1981; FAO et al., 2009)*, forming a unique educational, cultural and scientific resource. A selection of the monoliths is on display in the World Soil Museum in Wageningen, and the whole collection (e.g., monoliths, reference samples, soil morphological and analytical data, thin sections, imagery) has been documented in a relational database, queryable on-line. This information is supported by a large collection of, often unique, country reports, books and maps that originate largely from third world countries. The map collection contains mainly small-scale (1:250,000 or smaller) maps, part of which have been digitised in the framework of the *EuDASM* project (Panagos *et al.*, 2011). So far, some 25% out of some 15,400 of the historic books and country reports has been scanned, while this is some 72% (out of ~ 7600) for the map collection

(<http://www.isric.org/services/world-soil-library-and-maps>). This wealth of soil geographic and attribute information has been used to develop a range of broad scale, GIS databases (e.g., Batjes, 2009; FAO et al., 2009; Nachtergaele *et al.*, 2011) that were generally distributed on-line as stand-alone, compressed files both through the ISRIC website and Global Change Master Directory (GCMD). The information is being used for a wide range of assessments, including agro-ecological zoning, global crop production, soil vulnerability to pollution, soil carbon stocks and changes, and soil gaseous emission potentials (e.g., Milne *et al.*, 2007; Gibbs *et al.*, 2008; Fischer *et al.*, 2010; Mekonnen & Hoekstra, 2011; Wang *et al.*, 2011).

### 3 WEB-BASED INFORMATION FACILITIES

#### 3.1 Aims

To better manage the current demand for soil information, ISRIC is in the process of implementing an enterprise data management system, including procedures for registering acquired data using versioning, quality assessment and control. The centralized enterprise database will contain quality-controlled and authorized data with defined and registered accuracy and quality (e.g., documented data lineage; detailed metadata). Subsequent to international peer-review and adjustments, the spatially-enabled RDMS will go into production and be accessible through different facilities (e.g., Data Entry, MetaDataService, WMS, and WPS).

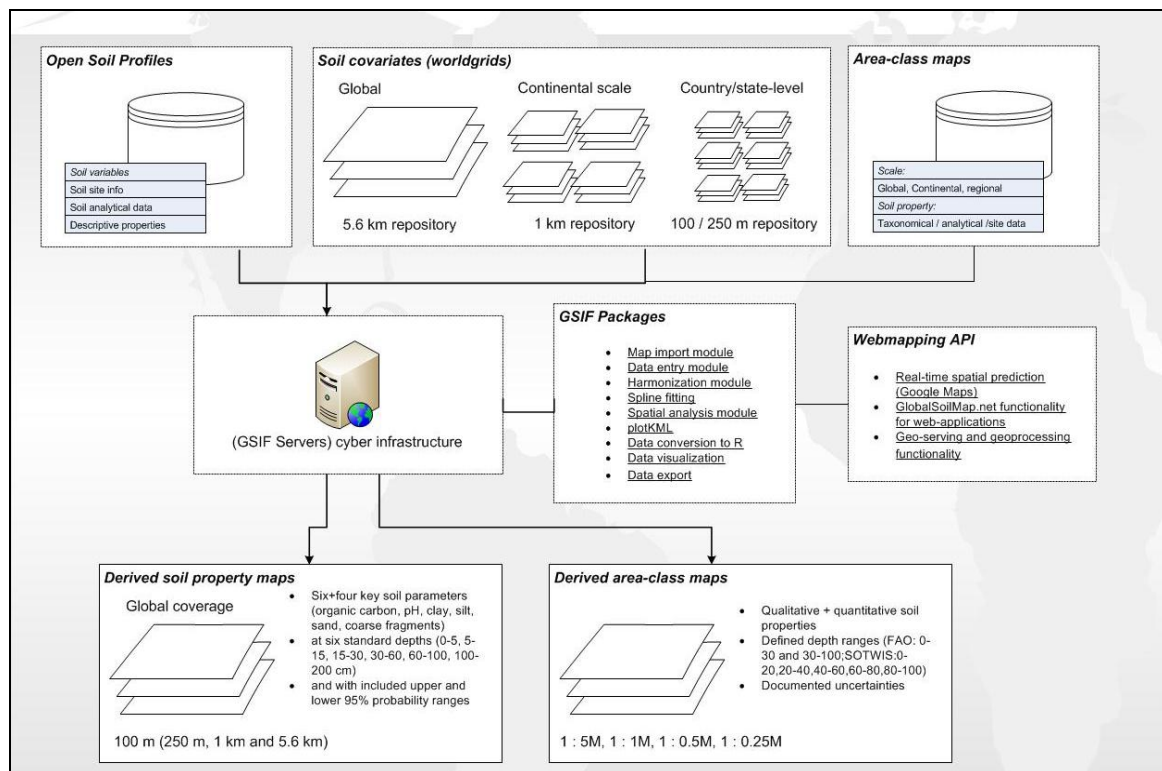
ISRIC's data policy follows the same principles for data sharing as those of ICSU-WDS, aiming at "full and open exchange of data, metadata and products shared within WDS, recognizing relevant international instruments and national policies and legislation" (see Uhlir *et al.*, 2009). When registering, the "contributor should present any existing restriction that is applicable to the exchange of the data, metadata and products submitted" (GEO, 2010).

#### 3.2 Main components

Since June 2011, ISRIC's website provides access to a range of new facilities that are in various stages of development:

- ISRIC enterprise database: All soil-related data managed or maintained by ISRIC will be made available online from one central database environment in one uniform format (Tempel & van Kraalingen, 2011). The spatially-enabled relational database management system (RDBMS) will be used to hold and manipulate a variety of data types, scales and sources. Version 1.0 is being populated using a range of disparate, stand-alone soil databases developed by ISRIC and its partners (e.g., Batjes, 2009; Van Engelen, 2011); for the future, other data(sets) will be added using both on-line and off-line facilities for data entry (<http://www.isric.org/data/wosis>). In conjunction with this, a range of web-based tools is being developed to extract user-defined data from the system using agreed data-exchange formats (e.g., SoilML, GeoSciML). The system is implemented using open standards (e.g., OGC, PostGresQL).
- Vocabulary service: This multilingual, soil-environmental thesaurus is being developed as a reference indexing and retrieval tool. It will provide ontologies of interest in the Soil Science domain and was created at the request of the Working Group on Soil Information Standardisation of the International Union of Soil Sciences (IUSS). Currently, the service is based on selected databases (FAO - Agrovoc; EEA - Gemet). Over the next 2-4 year, the service will be developed further in collaboration with colleagues from Europe (GS-Soil project), Australia (CSIRO SEEGrid) and similar projects around the world. Currently, it uses SEEGrid Vocabulary Service implementation as a frontend while editing is based on FAO's VocBench (<http://aims.fao.org/tools/vocbench-2>), which is open source.
- Metadata service (operational): Implemented as a multilingual catalogue application to enter, search and retrieve metadata records and manage spatially referenced resources. Currently, the service harvests soil-related metadata from various sources worldwide, including ISRIC's Library, FAO, and the Consultative Group on International Agricultural Research (CGIAR). The service is based on *GeoNetwork open source* (<http://geonetwork-opensource.org/>) and follows the principles of Free and Open Source Software (FOSS) and International and Open Standards for services and protocols (e.g., ISO/TC211 and OGC, Open Geospatial Consortium); see <http://www.isric.org/data/metadata-service>.
- OneSoil Map Viewer: Provides visualisations of qualitative and quantitative soil properties worldwide. Maps are accessed from various data sources and providers in a distributed way around the world. The viewer will create a dynamic soil map for different soil properties with standardized legends and data exchange formats (<http://www.isric.org/data/mapviewer>).
- Web mapping service (WMS): This service provides OGC-conform maps for different layers and allows users to visualize datasets hosted at the WDC for Soils (<http://www.isric.org/data/web-map-service>).

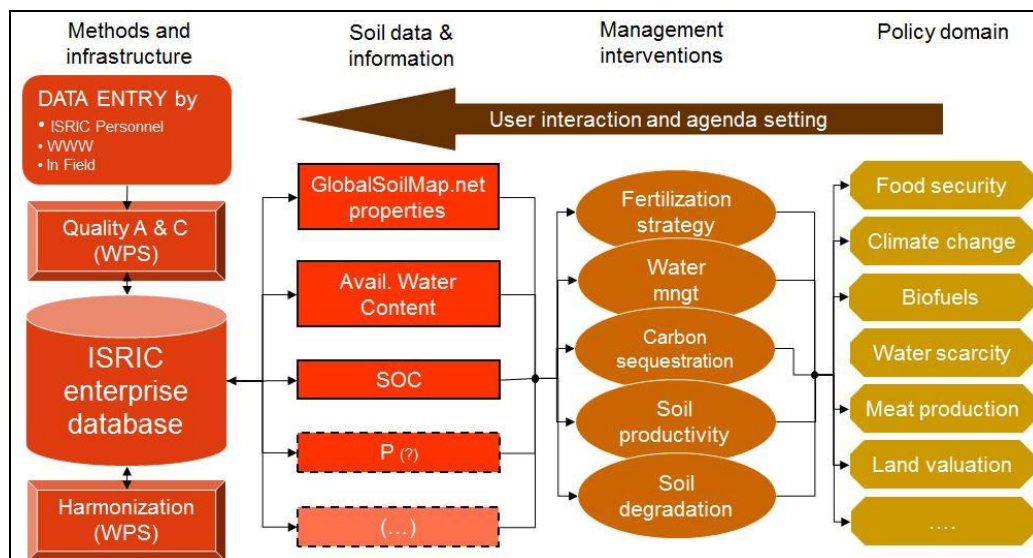
- Web processing service (WPS): Capability to compute soil functions or parameters of interest in the Soil Science domain. Four different services are currently available: a) interpolation over depth of quantitative soil profile properties using spline function; b) extraction of point information of soil information or auxiliary data sources for a given location on the planet, and soil information or auxiliary data sources for a given area on the earth specified by the corner coordinates; c) overlay of two raster data sources to provide statistical information for a predefined zone-grid (e.g., watershed, countries) for soil information or auxiliary data sources. The WPS delivers OGC conforming XML files (<http://www.isric.org/data/wprocessservice>).
- Worldgrids: Provides a repository of gridded predictors with global or at least partial global coverage. As WDC- Soils, ISRIC facilitates the collation and use of multi-thematic gridded repositories of Digital Soil Mapping covariates. Worldgrids already contains over 100 grid maps at a 5.6 km resolution and several 1 km resolution maps. Functionalities include querying, extraction, and creation of an overlay of a user-specified zone grid and a defined covariate map (see: <http://www.isric.org/data/worldgrids>).
- Global Soil Information Facility: GSIF will provide an overarching methodological framework for the production of consistent global soil information products of various types and scales/resolutions (Hengl *et al.*, 2011). Key principles for the design of GSIF, for production of open soil information, are outlined elsewhere (<http://www.isric.org/projects/global-soil-information-facilities-gsif>). Major components, as schematised in Figure 1, are expected to be fully operational by mid-2012. Where appropriate, algorithms and model states will be documented and preserved – recommended procedures are being developed and tested to align with the frameworks of on-going, large international projects such as *GlobalSoilMap.net* (Sanchez *et al.*, 2009) and *e-SOTER* (Van Engelen, 2011).



**Figure 1.** Proposed components of ISRIC's Global Soil Information Facilities and their relations

### 3.3 Applications

Upon their completion and testing, the ISRIC data and web facilities may be used freely —subject to any existing restrictions, see 3.1— to increase awareness and understanding of soils in major global issues; this is illustrated in Figure 2 under “policy domain”. Area-class and soil-property maps derived from the evolving web-processing services will also serve the global soil observing system as part of the *Global Earth Observing System of Systems (GEOSS)* and support collaborative activities with partner organisations such as the Food and Agricultural Organization (UN-FAO) and the Joint Research Centre (EU-JRC) in the broader framework of the *Global Soil Partnership* ([http://www.fao.org/nr/water/landandwater\\_gsp.html](http://www.fao.org/nr/water/landandwater_gsp.html)).



**Figure 2.** Management and use of soil information in support of research and informed decision making

## 4 CONCLUSIONS

The importance of judicious soil use and management in addressing major global issues, such as climate change, food security, energy, water conservation, and preserving biodiversity is increasingly recognised. As a result, there is a growing demand for quality-assessed soil data and information, of various types and scales, by other scientific disciplines, practitioners, policy and society. ISRIC is developing, testing and implementing a new enterprise data management system and a range of WMS and WPS services using international standards. Typically, new procedures and approaches are developed within the framework of on-going large international projects. Within the new ICSU World Data System, ISRIC will continue to ensure the long-term stewardship and provision of quality-assessed data and supporting web facilities to the international science community and other stakeholders.

## 5 REFERENCES

- Batjes, N.H. (2009) Harmonized soil profile data for applications at global and continental scales: updates to the WISE database. *Soil Use and Management* 25, 124-127
- Bindraban, P.S., Batjes, N.H., Leenaars, J.G.B. & Bai, Z. (2010) Relevance of soil and terrain information in studies of major global issues. In: *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World*. eds R. J. Gilkes & N. Prakongkep). International Union of Soil Sciences, 1-6 August, Brisbane, Australia, pp. 38-41.
- Blum, W. (2005) Functions of soil for society and the environment. *Reviews in Environmental Science and Biotechnology* 4, 75-79.
- FAO (1971-1981) *FAO-Unesco Soil Map of the World, 1:5,000,000 (Vol. 1 to 10)*. United Nations Educational, Scientific, and Cultural Organization, Paris.
- FAO, IIASA, ISRIC, ISSCAS & JRC (2009) Harmonized World Soil Database (version 1.1). Food and Agriculture Organization of the United Nations (FAO), International Institute for Applied Systems Analysis (IIASA), ISRIC - World Soil Information, Institute of Soil Science - Chinese Academy of Sciences (ISSCAS), Joint Research Centre of the European Commission (JRC), Laxenburg, Austria
- Fischer, G., Prieler, S., van Velthuizen, H., Lensink, S.M., Londo, M. & de Wit, M. (2010) Biofuel production potentials in Europe: Sustainable use of cultivated land and pastures. Part I: Land productivity potentials. *Biomass and Bioenergy* 34, 159-172.

GEO (2010) GEOSS data sharing action plan. Group on Earth Observations, Geneva, pp. 8 ([http://www.earthobservations.org/documents/geo\\_vi/07\\_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf](http://www.earthobservations.org/documents/geo_vi/07_Implementation%20Guidelines%20for%20the%20GEOSS%20Data%20Sharing%20Principles%20Rev2.pdf))

Gibbs, H.K., Johnston, M., Foley, J.A., Holloway, T., Monfreda, C., Ramankutty, N. & Zaks, D. (2008) Carbon payback times for crop-based biofuel expansion in the tropics: the effects of changing yield and technology. *Environ. Res. Lett.* 3, doi: 10.1088/1748-9326/1083/1083/034001.

Hengl, T., MacMillan, R.A., Walsh, M.G. & Reuter, H.I. (2011) Global Soil Information Facilities: A methodological framework for open soil information. ISRIC - World Soil Information, Wageningen, pp. 231 (*in the press*)

Mekonnen, M.M. & Hoekstra, A.Y. (2011) The green, blue and grey water footprint of crops and derived crop products. *Hydrology and Earth System Sciences* 15, 1577-1600.

Milne, E., Adamat, R.A., Batjes, N.H., Bernoux, M., Bhattacharyya, T., Cerri, C.C., Cerri, C.E.P., Coleman, K., Easter, M., Falloon, P., Feller, C., Gicheru, P., Kamoni, P., Killian, K., Pal, D.K., Paustian, K., Powlson, D.S., Rawajfih, Z., Sessay, M., Williams, S. & Wokabi, S. (2007) National and sub-national assessments of soil organic carbon stocks and changes: The GEFSOC modelling system. *Agriculture, Ecosystems & Environment* 112, 3-12.

Nachtergaele, F.O., Van Engelen, V.W.P. & Batjes, N.H. (2011) Qualitative and quantitative aspects of world and regional soil databases and maps. In: *Handbook of Soil Sciences (2nd ed.)*. eds Pan Ming Huang, Yuncong Li & M. E. Sumner). Taylor and Francis Group, (*in the press*).

Panagos, P., Jones, A., Bosco, C. & Kumar, P.S.S. (2011) European digital archive on soil maps (EuDASM): preserving important soil data for public free access. *International Journal of Digital Earth* 4, 434-443.

Sanchez, P.A., Ahamed, S., Carre, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M.d.L., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vagen, T.-G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A. & Zhang, G.-L. (2009) Digital Soil Map of the World. *Science* 325, 680-681.

Tempel, P. & van Kraalingen, D. (2011) Global Soil Information Facility: Database component. ISRIC - World Soil Information, ISRIC Report 2011/03. Wageningen, pp. 260

Uhlir, P.F., Chen, R.S., Gabrynowicz, J.I. & Janssen, K. (2009) Toward implementation of the Global Earth Observation System of Systems data sharing principles. *Data Science Journal* 8, GEO1-GEO91.

UNEP (2007) *Global Environmental Outlook: Environment for development (GEO-4)*. United Nations Environment Programme, Nairobi.

Van Engelen, V.W.P. (2011) Standardizing soil data (*e-SOTER* regional pilot platform as EU contribution to a Global Soil Information System). *International Innovation* June, 48-49.

Wang, M., Li, Y., Ye, W., Bornman, J.F. & Yan, X. (2011) Effects of climate change on maize production, and potential adaptation measures: A case study in Jilin province, China. *Climate Research* 46, 223-242.

Watson, R.T., Noble, I.R., Bolin, B., Ravindramath, N.H., Verardo, D.J. & Dokken, D.J. (2000) *Land Use, Land-Use Change, and Forestry (a special Report of the IPCC)*. Cambridge University Press, Cambridge.