The International Soil Moisture Network - An observational network for soil moisture product validations

W. Dorigo[1], A. Gruber[1], P. Van Oevelen[2], W. Wagner[1], M. Drusch[3], S. Mecklenburg[4], A. Robock[5], T. Jackson[6]

[1] Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Vienna, Austria
[2] International GEWEX Project Office, Silver Spring, Maryland, USA
[3] European Space Agency, ESTEC, Noordwijk, the Netherlands
[4] European Space Agency, ESRIN, Frascati, Italy
[5] Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA
[6] Hydrology and Remote Sensing Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, USA

Abstract – In situ soil moisture measurements are invaluable for calibrating and validating land surface models and satellite-based soil moisture retrievals. On a worldwide basis the number of meteorological networks measuring soil moisture is still limited and their data lack standardization of technique and protocol. In response to this, the International Soil Moisture Network (ISMN) was initiated to serve as a centralized data hosting facility where globally available in situ soil moisture measurements from operational networks and validation campaigns are collected, harmonized, and made available to users through a web interface, while downloads are provided according to common standards for data and metadata. Currently the network contains data of 16 networks with more than 400 stations and is rapidly expanding. As the ISMN is fed and utilized by the scientific community it will become a valuable resource for validating and improving satellite-derived soil moisture products and studying climate related trends.

Keywords: Soil moisture, global climate system, validation, data hosting center, quality

1. INTRODUCTION

Across many landscapes, soil moisture and its freeze/thaw state control evapotranspiration, thus providing the link between terrestrial and atmospheric water, energy and carbon cycles (Robock et al., 2000). Additionally it determines the partitioning of precipitation into infiltration and runoff. The availability of better spatial estimates of surface soil moisture conditions can therefore help to improve forecasting of precipitation, droughts and floods as well as climate projections and predictions (Dirmeyer et al., 2006). The importance of soil moisture in the global climate system has recently been underlined by the Global Climate Observing System (GCOS) by endorsing soil moisture as an Essential Climate Variable. Obviously there is a high need for calibration and validation of satellite missions for global soil moisture observations such as AMSR-E, ASCAT, SMOS and SMAP. In situ observations are crucial for this and need to be available for large number of sites worldwide representing a wide range of climatological conditions and stand characteristics. Although several local and regional meteorological and hydrological networks routinely measure soil moisture, globally the number of long-term ground based monitoring networks is still small and largely restricted to mid-latitude regions. Moreover, the lack of standard measurement techniques and protocols complicates the use of network data. As a result, there are many differences between the measurements such as measurement depth, units of soil moisture, sampling interval and precision. Also the fact that the various data sets are managed by a large number of different organizations means that global studies incorporating ground-based soil moisture are tedious to perform. To overcome these issues, the Global Energy and Water Cycle Experiment (GEWEX) of the Word Climate Research Group (WCRP) with continued support of the Group on Earth Observations (GEO) and the Committee on Earth Observation Satellites (CEOS) was instrumental in developing an integrative platform for ground-based soil moisture measurements. The launch of the Soil Moisture and Ocean Salinity (SMOS) mission by the European Space Agency (ESA) provided the decisive impetus as, in support of calibration and validation of SMOS products, an integrative system was needed to host quality controlled and harmonized soil moisture measurements emerging from various ground validation campaigns and operational networks. Consequently, ESA provided the financial support for the development and first phase of operation of the ISMN which is implemented at the Vienna University of Technology (Dorigo et al., 2011).

2. THE INTERNATIONAL SOIL MOISTURE NETWORK

In short, the ISMN is a centralized data hosting centre where globally available ground-based soil moisture measurements are collected, quality controlled, harmonized, and made available to users through a web interface. In support to quality control and to address fundamental science questions related to soil moisture and its role in the integrated water cycle, the ISMN also stores soil moisture measurements of the deeper layers and relevant auxiliary variables such as precipitation, temperature of air and soil, soil porosity and soil texture.

2.1 The Database

The database can be considered the core of the International Soil Moisture Network data hosting facility. Its design is very critical since inadequate choices may soon slow down operation when the database gets filled or be incapable of including new networks with a deviating design of measurement set up. Hence, the design of the database has been guided by its potential content. The entries have been established after consultation of data providers, possible users, and standards for data and metadata (Coordinated Energy and water cycle Observations Project (CEOP[1]), ISO 19115, and Infrastructure for Spatial Information in the European Community (INSPIRE[2])).

[1]http://www.eol.ucar.edu/projects/ceop/dm/documents/refdata_report/; renamed GEWEX Hydroclimatology Projects in September, 2010
[2] http://inspire.jrc.ec.europa.eu/

2.1 Harmonization

The data provided by the networks highly differ in temporal resolution, the quantities measured, the units used, and the configuration of the measurements. Besides, there are differences in the structure of stored data and metadata and the dissemination medium (e.g. FTP or HTTP). To achieve a fully automated data transmission between the network and the ISMN and also to enable the possibility of near real time (NRT) data updating (for fully automated networks such as the american Soil Climate Analysis Network (SCAN) or the finnish FMI that share their NRT data sets with the ISMN), it was indispensable to implement several steps to harmonize the data sets. User consultation revealed that an hourly (60 minutes) sampling interval is sufficient for all envisaged applications, including numerical weather prediction, climate studies and run-off prediction. Resampling data with higher temporal resolution to one hour has the advantage that the data amount is significantly reduced, thus leading to a better performance of the database. The measurements with sub-hourly temporal resolutions are resampled to an hourly resolution. This is done by selecting the measurements occurring closest to the full hour (UTC) reference time steps. No averaging of data sets takes place. All soil moisture measurements in the ISMN are expressed as volumetric soil moisture [m^3 m^{-3}]. If data is provided in another unit (e.g. kg m^{-3} or Plant Available Water), these are converted. The bottle neck in this conversion is the availability of accurate metadata. Nevertheless, the more recent data are usually provided as volumetric soil moisture content. No harmonization in vertical direction is taking place, as there is no agreement on the optimum depth of in situ soil moisture measurements for satellite and LSM validation because this strongly depends on the band width of the studied satellite sensor or the vertical resolution of models, respectively.

2.2 Quality Checking

Due to the inhomogeneous data structure of the different networks and as, up to date, most contributing networks do not provide quality indicators with their data, it was decided to integrate the allocation of quality indicators in the data hosting facility itself, which is based on the CEOP-Data Flag Definitions. This rather conservative quality indicator scheme was implemented, so that suspect observations can be detected more easily. It was decided to provide only quality indicators that can be checked in an objective way by the ISMN. This leads to the subset of CEOP quality flags presented in Table 1.

Table 1. CEOP quality flags adopted in International Soil Moisture Network

Flag value	Definition
C	Reported value exceeds output format field size OR was negative precipitation.
M	Parameter value missing OR derived parameter can not be computed.
D	Questionable/dubious
U	Unchecked

Table 2 shows the possible ranges for the auxiliary variables most relevant to the ISMN. If a measured data value exceeds this range (at either side of the range) the measurement receives the flag value “C”. If a measurement is missing (a NaN entry) for a data set value, its quality flag is set to “M” (Parameter value missing). All other data set values have been set to “U” for unchecked.

Table 2. Plausible variable ranges for most important meteorological data.

Variable name	Variable range
Soil moisture	0 – 60 %
Soil temperature	-60 – 60 °C
Air temperature	-60 – 60 °C
Precipitation	0 – 100 mm h^{-1}
Soil suction[4]	0 – 2500 kPa

For soil moisture also the quality flag “Questionable/Dubious” is applied. This quality flag is adopted when a soil moisture measurement in combination with another variable leads to a suspicious result (Table 3).

Table 3. Parameter combinations leading to the quality flag “Questionable/Dubious” for soil moisture

1	Valid soil moisture measurement in combination with a negative soil temperature (measured at same depth)
2	Valid surface soil moisture measurement in combination with a negative air temperature
3	A decreasing or stable surface soil moisture content (with respect to the previous time step) while precipitation is measured

Even though the CEOP flagging methods are able to detect extreme outliers, they are not able to detect gradual shifts (e.g. due to sensor drifting) of jumps that are within plausible limits but cannot be attributed to geophysical phenomena. To overcome these issues, one future vision of the ISMN is the improvement of a quality control system in two ways: (1) improve online detection of random errors, spikes, biases, and jumps and (2) characterize the quality and stability of single stations and networks. Therefore adjacent stations, which generally show a high spatio-temporal correlation, will be used to detect outliers, jumps, and temporary biases through so-called “buddy checking”, which is already a standard quality control method in several other geophysical disciplines (Rayner et al., 2006;Ingleby and Huddleston, 2007). Systematic differences between sites or different networks, caused by varying observation techniques and/or inconsistent sensor calibrations can be characterized by comparing them to a common baseline that is stable in space and time, e.g. ECMWF’s ERA-Interim reanalysis data sets. The random uncertainty of an individual station can be established e.g. by using the triple collocation technique (Dorigo et al., 2010;Miralles et al., 2010), which is a powerful tool to estimate the root mean square error while simultaneously solving for systematic differences in the climatologies of a set of three time series with independent error characteristics.

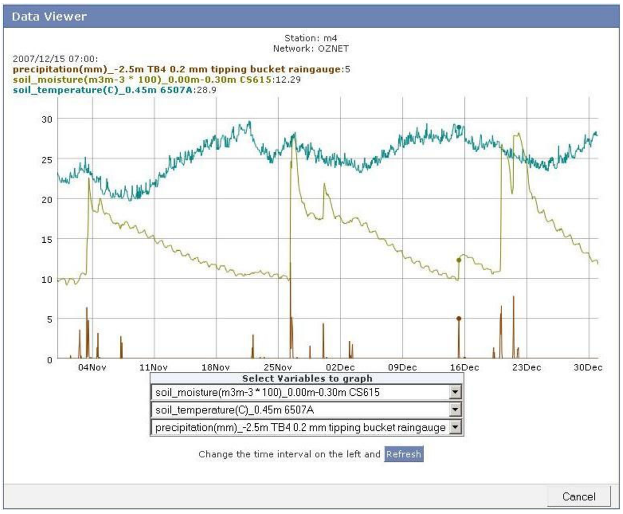
2.3 Web Interface

The ISMN can be accessed at http://www.ipf.tuwien.ac.at/insitu and consists of two major parts: (1) A project website providing details about networks, partners and the project itself and (2) the data hosting centre itself, where users can query, view and download the data contained in the database. Figure 1 shows a screenshot of the initial view of the data hosting centre. It embeds a Google Maps application programming interface (API), which offers the user a map-based selection tool to view and download the available data. In addition, through the satellite data and orthophotos in Google Maps the user is able to obtain additional information about the land use/land cover and topography in the vicinity of the measurement stations.

Figure 1. The data hosting centre of the ISMN. Blue dots indicate the locations of soil moisture stations.

By clicking on the markers, the user obtains a summary of important metadata about the networks or stations, respectively, depending on the zoom level within Google Maps. For a selected station the available data can be displayed in a data viewer to get a first impression of the availability and quality of the data. Figure 2 illustrates this analysis with data from the Australian OZNET Network. You can see soil temperature measured at 0.45 m depth (blue), soil moisture measured from 0.00 m to 0.30 m (yellow) and precipitation (brown) from November 1st until December 31st 2007. This shows clearly the correlation between the precipitation and the soil moisture.

Figure 2. Visualization of data sets



The map-based graphical data selection can also be refined by selecting continent, network, latitude/longitude and time period. After selecting the required data sets for download, the output is prepared according to the data and metadata standards of the CEOP. CEOP's main goal is to understand and predict continental to local-scale hydroclimates for hydrologic applications and coordination of the regional hydroclimate projects. CEOP has accelerated the adoption of standards for various types of observations, including those of soil moisture. These standards do not only specify the formats of the data but also provide prescriptions on metadata formats and file naming conventions. Metadata descriptions compliant with ISO 19115 and INSPIRE are directly retrieved from the database and saved in an XML file that is appended to the data download.

2.4 Data sets

As shown in Table 4, currently the ISMN contains 16 stations with more than 400 stations located in North America, Europe, Asia, and Australia. Besides several currently operational networks, the ISMN now also contains all data sets from the historical Global Soil Moisture Data Bank (Robock et al., 2000). Hence, the time period spanned by the entire database runs from 1952 until the present, although most data sets have originated during the last decade. The database is rapidly expanding, which means that both the number of stations and the time period covered by the existing networks are still growing. Every potential network is invited to enrich the collection by sharing its in situ soil moisture data with the ISMN.

Table 4. Networks currently available within the ISMN

Network	Country	Stations	Data availability
CAMPANIA	Italy	2	2000/11/26 – 2008/12/31
CHINA	China	40	1981/01/08 – 1999/12/28
CNR-IRPI	Italy	4	2002/10/09 – 2008/05/31
ICN	USA	19	1983/01/03 – 2008/07/15
IOWA	USA	6	1972/04/04 – 1994/11/15
MOL-RAO	Germany	2	2003/01/01 – 2008/12/31
MONGOLIA	Mongolia	44	1964/04/08 – 2002/10/28
OZNET	Australia	64	2001/01/01 – 2010/08/31
REMEDHUS	Spain	18	2005/01/01 – 2008/12/31
RUSWET-AGRO	Former Soviet Union	78	1986/12/28 – 1988/12/28
RUSWET-GRASSE	Former Soviet Union	122	1952/01/08 – 1985/12/28
RUSWET-VALDAI	Former Soviet Union	3	1960/01/15 – 1990/12/15
SMOSMANIA	France	12	2007/01/01 – 2010/01/01
SWEX-POLAND	Poland	2	2006/08/03 – 2010/01/03
UDC_SMOS	Germany	11	2007/11/08 – 2010/07/25
UMSUOL	Italy	1	2007/07/01 – 2010/09/30

3. VALIDATION WITH ISMN DATA

Evaluating satellite- and model-derived soil moisture retrievals with in situ soil moisture measurements is commonly based on the root-mean-square metric (Jackson et al., 2010). However, to use in situ soil moisture measurements from the ISMN in satellite and land surface model validation and calibration, the user should be aware of the systematic differences that may exist between in situ measurements and soil moisture estimates from models and satellite observations (Entekhabi et al., 2010) Even though the ISMN provides soil moisture measurements in the same volumetric unit that is returned by most satellite products and models, biases and differences in the dynamic range may exist between the datasets, e.g., by assumptions and generalizations made within the retrieval concept, scaling issues, or due to the different soil layers or soil depths considered. Thus, other metrics, such as the Pearson or Spearman correlation coefficient, often provide valuable and complementary information on the performance (Entekhabi et al., 2010). To combine in situ soil moisture measurements with satellite retrievals and a first guess predicted by a land surface model, e.g., in the framework of data assimilation, it is often necessary to minimize systematic differences between the individual data sets (Drusch, 2007). These correction methods include standard rescaling techniques, e.g. based on simple statistic descriptors of both datasets such as minimum, maximum, mean, and variance (Dorigo et al., 2010; Miralles et al., 2010), linear regression (Scipal et al., 2008) and others.

4. CONCLUSIONS

In addition to the 16 already contributing networks, several others, including SCAN and FLUXNET (Baldocchi et al., 2001), have evinced their interest in participating to the ISMN. SCAN consists of more than 150 automated remote sites throughout the USA which collect soil moisture and soil temperature data along with precipitation, wind, and solar radiation data. Due to the fully automation of the sites, data sets from SCAN are processed, quality checked and made available

for users in near real time. As the ISMN was also designed for processing incoming data on a fully automated basis, these NRT data sets within the ISMN will become very important for time critical applications, ranging from several hours to several days. FLUXNET coordinates regional and global analysis of observations from approximately 500 micrometeorological tower sites, also in areas with few potentially contributing networks such as South America and Africa.
This directly leads into the direction of realizing a supersite program with wide spread (approximately satellite footprint size), high density measurements also in high-latitude areas and in the southern hemisphere, which are required for satellite sensor evaluation and calibration as well as to develop soil wetness algorithms for satellite measurements and to evaluate climate model outputs. However, the scope of the ISMN is to go beyond the role of a satellite validation resource and also serve other communities, such as hydrologists, meteorologists, climate modelers and water managers.
The positive contribution of international organizations such as WCRP, GEWEX and GEO, the support of Space Agencies and the voluntary efforts of many individual scientists clearly underlined the consciousness and willingness to realize such an integrated soil moisture observing system.
Despite the great achievements reached since the initial implementation of the ISMN, the success of these efforts, especially on long term basis, highly depends upon the commitment of financial support and the cooperation of data providers. As it was demonstrated by the success of the Global Soil Moisture Data Bank the benefit of sharing soil moisture data free of cost with the scientific community is valuable not only for data users but also for the networks. They may become embedded as key networks in international calibration and validation activities or climate monitoring programs, e.g. like happened to the flux tower sites participating to FLUXNET or the stations participating to the Baseline Surface Radiation Network (Ohmura et al., 1998), what in turn may lead to vast international scientific recognition and pave the way for access to extended funding resources. Additionally we want to emphasize that the ISMN is a growing entity animated by the scientific community itself and solicit users to download, use, and give feedback on the data sets currently contained in the database.

REFERENCES

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem–Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, Bulletin of the American Meteorological Society, 82, 2415-2434, 2001.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006), GSWP-2: Multimodel analysis and implications for our perception of the land surface, Bull. Amer. Meteorol. Soc., 87, 1381-1397.

Dorigo, W.A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture data sets, Hydrology and Earth System Sciences, 14, 2605-2616, 10.5194/hess-14-2605-2010, 2010.

Dorigo, W.A, Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A. and Jackson, T.: The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci. Discuss, accepted, 2011.

Drusch, M.: Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's integrated forecast system and the TMI soil moisture data set, Journal of Geophysical Research D: Atmospheres, 112, D03102 10.1029/2006JD007478, 2007.

Entekhabi, D., Reichle, R. H., Koster, R. D., and Crow, W. T.: Performance Metrics for Soil Moisture Retrievals and Application Requirements, Journal of Hydrometeorology, 11, 832-840, doi:10.1175/2010JHM1223.1, 2010.

Ingleby, B., and Huddleston, M.: Quality control of ocean temperature and salinity profiles - Historical and real-time data, Journal of Marine Systems, 65, 158-175, 2007.

Jackson, T., Entekhabi, D., Van Oevelen, P. J., and Kerr, Y.: Towards integrated global soil moisture observations, GEWEX News, 15, 8-9, 2005.

Miralles, D. G., Crow, W. T., and Cosh, M. H.: Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations, Journal of Hydrometeorology, 0, doi:10.1175/2010JHM1285.1, 2010.

Ohmura, A., Dutton, E. G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., Heimo, A., König-Langlo, G., McArthur, B., Müller, G., Philipona, R., Pinker, R., Whitlock, C. H., Dehne, K., and Wild, M.: Baseline Surface Radiation Network (BSRN/WCRP): New Precision Radiometry for Climate Research, Bulletin of the American Meteorological Society, 79, 2115-2136, 1998.

Rayner, N. A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell, T. J., and Tett, S. F. B.: Improved analyses of changes and uncertainties in sea surface temperature measured in Situ since the mid-nineteenth century: The HadSST2 dataset, Journal of Climate, 19, 446-469, 2006.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S., and Namkhai, A.: The Global Soil Moisture Data Bank, Bulletin of the American Meteorological Society, 81, 1281-1299, 2000.

Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., and Wagner, W.: A possible solution for the problem of estimating the error structure of global soil moisture data sets, Geophysical Research Letters, 35, -, Artn L24403, doi 0.1029/2008gl035599, 2008.

ACKNOWLEDGEMENTS

We greatly thank all networks that have shared their data with the ISMN. Without their valuable contributions this initiative would not have been possible in the first place. We also thank Angelika Xaver for helping with the processing of the data sets. The ISMN has been founded through the SMOS Soil Moisture Network Study (ESA ESTEC Contract No.22954/09/NLCT). Alan Robock is supported by NASA grant NNX09AJ99G.