RESEARCH DEPARTMENT

MEMORANDUM

To:	RD, MetOps, OD Division and Section Division heads
Copy:	
From:	Clément Albergel
Date:	10 April 2013
Subject:	**A new tool and dataset to evaluate soil moisture at ECMWF** File: RD13-132

Brief summary: Soil moisture measurements from nearly 800 stations from 13 networks across the world were collected and arranged in a similar format. A script to use them to evaluate soil moisture from ECMWF different experiments has been developed. The main outputs are the scores (correlation, significance of the correlation, 95% confidence interval, bias, root mean square differences and standard deviations), plots of time series and Taylor diagrams. Evaluation can be done at daily, decadal or monthly time steps.

Soil_moisture_tool_v0.tar.gz available under: ec:/da5/

1. ecp ec:/da5/ Soil_moisture_tool_v0.tar.gz
2. tar xvfz Soil_moisture_tool_v0.tar.gz
3. cd Soil_moisture_tool_v0
4. ./get_in_situ.sc (copy and extract in situ measurements of soil moisture from ECFS to $SCRATCH, ~900Mb)
5. ./Comparisons_main_scripts.sc (start the evaluation as described below in section 'Case study')

Introduction

The importance of soil moisture in the global climate system has recently been underlined by the Global Climate Observing System (GCOS) Programme endorsing soil moisture as an Essential Climate Variable (ECV). It is a crucial variable for numerical weather prediction (NWP) and climate projections because it plays a key role in hydrological processes. A good representation of soil moisture conditions can therefore help improve the forecasting of precipitation, droughts and floods. For many applications global or continental scale soil moisture maps are needed.
Among the first soil moisture analysis systems used for operational NWP was the system implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) in 1994 to prevent the Land Surface Model (LSM) drifting to dry conditions in summer. Since then, major

upgrades have been implemented in the land surface modelling and analysis systems of the high-resolution component of the Integrated Forecasting System (IFS) used operationally at ECMWF.
One very important aspect of the environmental variables retrieval such as soil moisture is the evaluation of their performance. The typical validation approach for model based data products is to compare them to in situ observations; in situ measurements of soil moisture are a highly valuable source of information for assessing the quality of model moisture products. While in the 1980s and 1990s records of in situ soil moisture measurements were available for only a few regions and often for only very short periods, huge efforts have been made in the last decade to install long-term observations networks in contrasting biomes and climate conditions.
In this context soil moisture measurements from nearly 800 stations from 13 networks across the world were collected and arranged in a similar format. A script allowing the evaluation of soil moisture from our different experiments (e.g. analyses, re-analysis, and offline simulations) has been developed.

In situ measurements

Most of the in situ soil moisture measurements were obtained through the International Soil Moisture Network (ISMN, http://www.ipf.tuwien.ac.at/insitu/), a data hosting centre where globally-available ground-based soil moisture measurements are collected, harmonized and made available to users. Data from 13 networks are considered so far over 2007-2012: NRCS-SCAN (Natural Resources Conservation Service - Soil Climate Analysis Network), SNOTEL (SNOwpack TELemetry) and USCRN (U.S. Climate Reference Network) over the United States (177, 348 and 114 stations), SMOSMANIA (Soil Moisture Observing System-Meteorological Automatic Network Integrated Application) and SMOSMANIA-E in France (12 and 9 stations), REMEDHUS (REd de MEDición de la HUmedad del Suelo) and VAS (Valencia Anchor Stations) in Spain (20 and 2 stations), MAQU in China (20 stations), OZNET in Australia (38 stations), AMMA (African Monsoon Multidisciplinary Analyses) in western Africa (7 stations), UMBRIA in Italy (4 stations), HOBE (Hydrological Observatory) in Denmark (30 stations) and GTK in Finland (4 stations). Figure 1 gives an overview of the location of the stations.

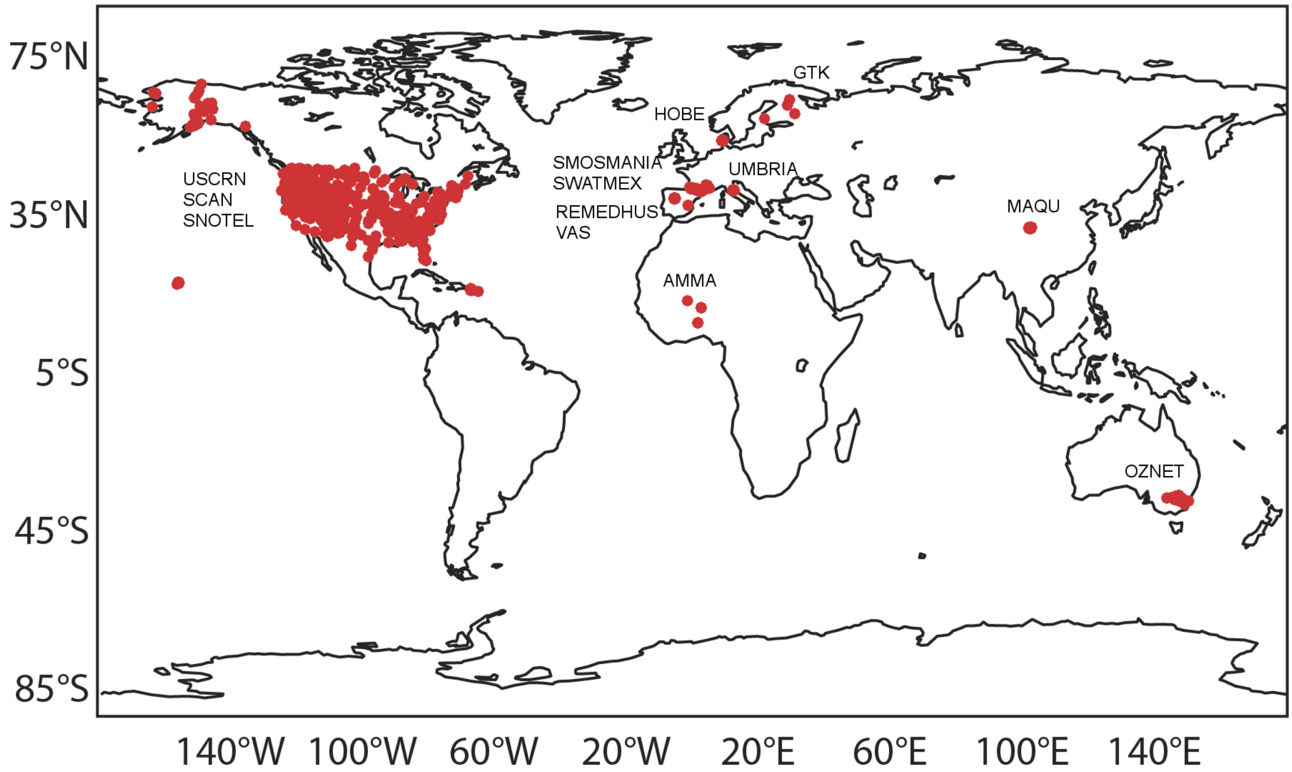


Figure 1: location of the stations with in situ measurements available.

While more networks might be available, it was decided to focus on those that also measure soil temperature; it permits to remove observations potentially affected by frozen condition. The different stations cover a wide variety of climatological conditions. Figure 2 illustrates measurements at 6 stations, 2 of the SCAN network in USA, 2 from the OZNET networks in south-eastern Australia and 2 from the AMMA network in western Africa.

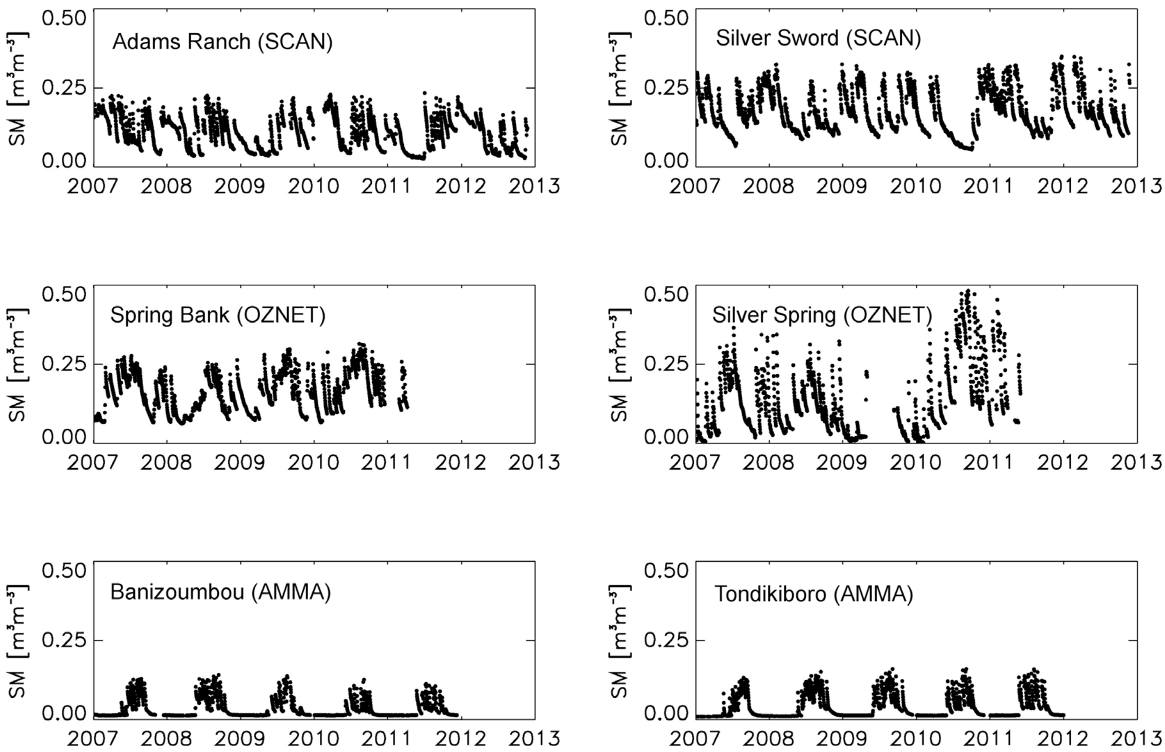


Figure 2: in situ measurements of soil moisture (5 cm depth) for; 2 stations of the SCAN network over the USA (top), 2 stations of the OZNET network in south-eastern Australia (middle) and 2 stations of the AMMA network in Western Africa (bottom).

Metrics used for the evaluation

The choice of performance metrics to be used is of crucial interest; it is governed by the nature of the variable itself and is influenced by the purpose of the investigation and its sensitivity to the considered variables. No single metric or statistic can capture all the attributes of environmental variables. Some are robust in respect to some attributes while insensitive to others. The most commonly used metrics to evaluate the accuracy of soil moisture retrievals are the correlation coefficient (R, Eq.1), the root mean square difference (RMSD, Eq.2) and the bias (Bias, Eq.3).

$$R = \frac{\frac{1}{N}\sum_{n=1}^{N}(SM_n - \overline{SM})(insitu_n - \overline{insitu})}{\sigma_{SM}\sigma_{insitu}} \qquad \text{Eq.1}$$

$$RMSD = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(insitu_n - SM_n)^2} \qquad \text{Eq.2}$$

$$Bias = \frac{1}{N}\sum_{n=1}^{N}(insitu_n - SM_n) \qquad \text{Eq.3}$$

The p-value (a measure of the correlation significance) is calculated; this indicates the significance of the test. Also, the normalised standard deviation (SDV, Eq.4) and the centred unbiased RMSD (E, Eq.5) between model and in situ observations are used. SDV is the ratio between standard deviations of the soil moisture product and in situ measurements; it gives the relative amplitude whilst E quantifies errors in the pattern variations. E does not include any information on biases because means of the fields are subtracted before computing second order errors.

$$SDV = \sigma_{SM}/\sigma_{insitu} \qquad \text{Eq.4}$$

$$E^2 = (RMSD^2 - Bias^2)/\sigma_{insitu}^2 \qquad \text{Eq.5}$$

The main reason for computing the two last scores is that R, E and SDV can be displayed on a single two-dimensional diagram (Taylor diagram) and this helps with the interpretation of the results. These three parameters are complementary but not independent as they are related by Eq.6.

$$E^2 = SDV^2 + 1 - 2.SDV.R \qquad \text{Eq.6}$$

In a Taylor diagram the SDV is displayed as a radial distance and R as an angle in the polar plot. In situ measurements are represented by a point located on the x axis at R=1 and SDV=1; the distance to this point represents E.

For each R estimate a 95% Confidence Interval (CI) was calculated using a Fisher Z transform. Whereas p-value gives an indication on the significance of the correlation, the 95% CI permits to identify periods or areas that are significantly better than others.

Main script for the evaluation

The main script to use is called 'Comparisons_main_scripts.sc', there is a header (see figure 3 below) where you have to indicate;

- the period you want to consider for the evaluation (a full year is so far the longest period to be evaluated),
- the experience Id,
- the frequency you want (evaluation based on daily, decadal or monthly value),
- the class of your experience Id (rd, od, ei…),
- the layer of soil you want to consider (e.g. 39 for the first layer of soil [0-7 cm]) and
- the network you want to use for the evaluation (a full list of networks is given in Annex I).

```
#################
year_start=2010
month_start=01
month_stop=12
day_start=01
day_stop=31
expid=fiv7
# daily - decade - monthly
freq=daily
class=rd
paramid=39
network=smosmania
################
```

Figure 3: Header of the main script for evaluating ECMWFs' soil moisture

The script (./Comparisons_main_scripts.sc) will create a folder /EXPERIMENT and then sub-folders /EXPERIMENT/expId/year/ where the results will be gathered in two folders;

- /plots with all the figures (postscript and png format),
- /tables with 2 files with results; expId_year_python.dat and expId_year_taylor.dat.

Also in the main folder 'Soil_moisture_tool_v0' a file named 'final_score.dat' is created indicated the number of stations with significant level of correlations, the mean, minimum and maximum values of correlations, RMSD and bias.

Finally a last file is created; data_taylor.dat, it used as an input of code_taylorDiagram.py (e.g.: python code_taylorDiagram.py) to build a Taylor diagram. Next section illustrates the evaluation of one RD experiment (called fiv7) using the twelve stations of the SMOSMANIA network in southwestern France for 2010.

Case study

The following command; "./Comparisons_main_scripts.sc" will start the evaluation in the configuration proposed by figure 3. In this configuration, the fiv7 experiment will be evaluated over the stations of the SMOSMANIA network over the full 2010. Main statistical scores will be in the folder EXPERIMENT/fiv7/2010/tables/fiv7_2010_python.dat, they are; the Pearson correlation associated with its p-value (significance of the test), the 95% lower and upper confidence interval, root mean square difference, bias and sample size. It is illustrated by figure 4.

```
fiv7 smosmania.cdm 2010 pearsonR 0.872186 pvalue 0.000000 lower 0.844792 upper 0.895021 rmse 0.036528 bias -0.012416 N 355.000000
fiv7 smosmania.crd 2010 pearsonR 0.824385 pvalue 0.000000 lower 0.787906 upper 0.855099 rmse 0.195894 bias -0.192720 N 355.000000
fiv7 smosmania.lhs 2010 pearsonR 0.871883 pvalue 0.000000 lower 0.844037 upper 0.895040 rmse 0.046298 bias 0.000536 N 346.000000
fiv7 smosmania.lzc 2010 pearsonR 0.741226 pvalue 0.000000 lower 0.640562 upper 0.816846 rmse 0.195196 bias -0.190318 N 105.000000
fiv7 smosmania.mnt 2010 pearsonR 0.844162 pvalue 0.000000 lower 0.811161 upper 0.871804 rmse 0.057480 bias 0.023148 N 351.000000
fiv7 smosmania.mtm 2010 pearsonR 0.855813 pvalue 0.000000 lower 0.825236 upper 0.881387 rmse 0.086924 bias -0.075593 N 355.000000
fiv7 smosmania.nbn 2010 pearsonR 0.841870 pvalue 0.000000 lower 0.808539 upper 0.869817 rmse 0.090845 bias -0.080956 N 353.000000
fiv7 smosmania.prg 2010 pearsonR 0.844475 pvalue 0.000000 lower 0.811378 upper 0.872174 rmse 0.096784 bias -0.089376 N 348.000000
fiv7 smosmania.sbr 2010 pearsonR 0.863679 pvalue 0.000000 lower 0.834067 upper 0.888329 rmse 0.086044 bias -0.078734 N 343.000000
fiv7 smosmania.sfl 2010 pearsonR 0.851299 pvalue 0.000000 lower 0.819809 upper 0.877655 rmse 0.084694 bias -0.075698 N 354.000000
fiv7 smosmania.svn 2010 pearsonR 0.817323 pvalue 0.000000 lower 0.779553 upper 0.849169 rmse 0.100191 bias -0.074263 N 355.000000
fiv7 smosmania.urg 2010 pearsonR 0.865778 pvalue 0.000000 lower 0.837129 upper 0.889690 rmse 0.097425 bias 0.048019 N 355.000000
```

Figure 4: main results of the evaluation, here for expId fiv7 over 2010 and the 12 stations of the SMOSMANIA network in France: Pearson correlation, pvalue, 95% lower and upper confidence interval, root mean square difference, bias and sample size.

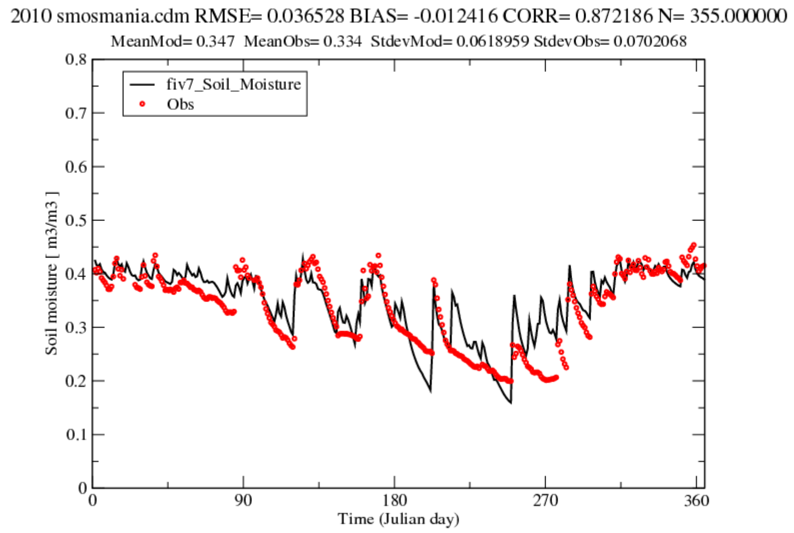
Files EXPERIMENT/fiv7/2010/tables/fiv7_2010_taylor.dat contains the standard deviation of fiv7 soil moisture for the location of the stations as well as that of the corresponding observations (additionally to the bias, RMSD and Correlation).

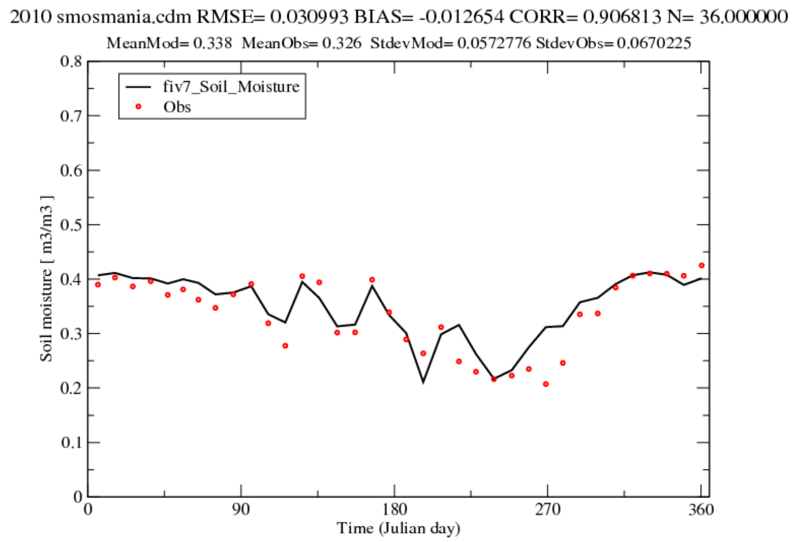
File 'final_score.dat' (figure 5) summarizes the main statistics.

```
N stations with significant R = 12
Mean_R / Min_R / Max_R
0.841173 0.741226 0.872186
Mean_RMSD / Min_RMSD / Max_RMSD
0.0978586 0.036528 0.195894
Mean_Bias / Min_Bias / Max_Bias
-0.0665309 -0.192720 0.048019
```

Figure 5: Summary of the main statistics

Figure 6 illustrates 3 graphics from EXPERIMENT/fiv7/2010/plots/.





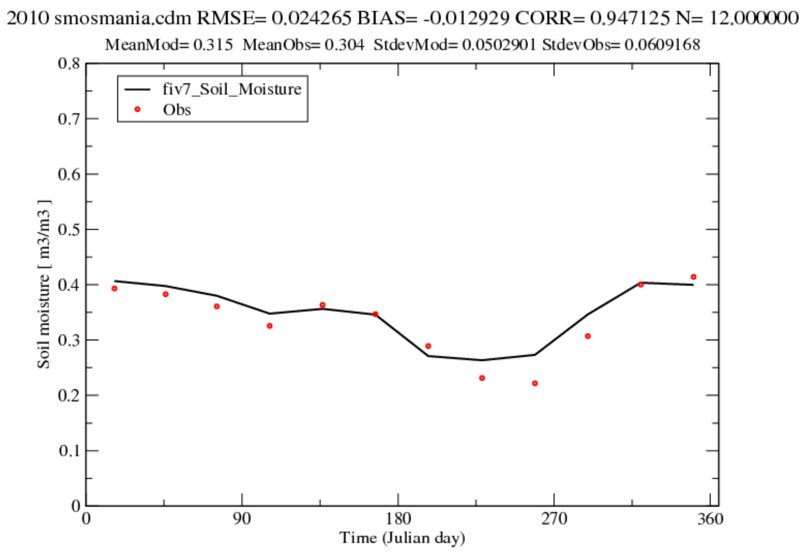


Figure 6: time series of observed and modelled soil moisture for one station on the SMOSMANIA network in south western France over 2012. From top to bottom the evaluation uses daily, decadal and monthly values. Main statistical scores are reported on the graphics.

In this configuration the code_taylorDiagram.py (e.g.: 'python code_taylorDiagram.py', or other version of python according to your Desktop) will create the following figure (figure 7).

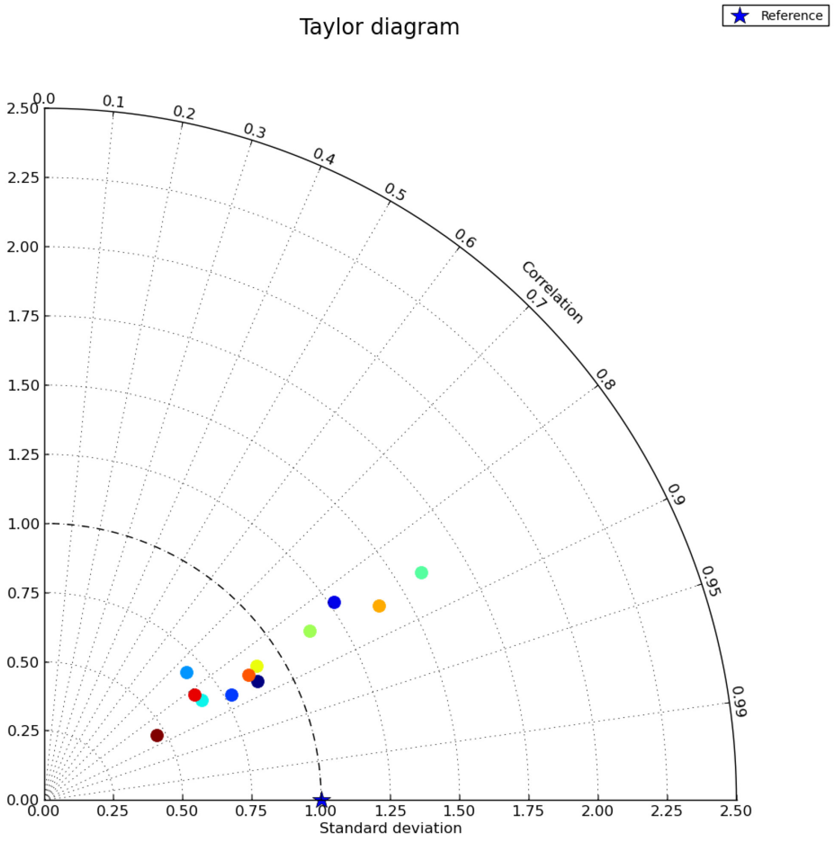


Figure 7: Taylor diagram illustrating the statistics of the comparison between soil moisture from fiv7 and the twelve stations of the SMOSMANIA network in south western France for 2012.

Last tips: If you wish to re-start an evaluation of the same experiment, the same year but without the same starting/ending days, first delete the following files; dat_6hr and dat_hr in EXPERIMENT/ExpId/YEAR/.

Acknowledgments

I'd like to thanks Gianpaolo Balsamo and Souhail Boussetta for providing the bases of development of this script, Tomas Kral for helping me handling python codes. Patricia de Rosnay and Lars Isaksen are thanked for their support to this initiative as well as Joaquin Muñoz-Sabater for being the first to test the script. Robert Hine is thanked for is help improving figure 1. This RD memorandum is documenting an on-going effort and I will be grateful for comments and suggestions aiming at improving it.

Further reading

More information on soil moisture evaluation (metrics, in situ measurements, quality of ECMWF's soil moisture products) can be found at:

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y. & Wagner, W. (2012a). Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. Remote Sensing of Environment, 118, 215-226.

(Also Tech. Mem. #652, http://www.ecmwf.int/publications/library/do/references/show?id=90275)

Albergel, C., De Rosnay, P., Balsamo, G., Isaksen, L. & Munoz-Sabater, J. (2012b). Soil Moisture Analyses at ECMWF: Evaluation Using Global Ground-Based In Situ Observations. Journal of Hydrometeorology, 13, 1442-1460

(Also Tech. Mem. #651, http://www.ecmwf.int/publications/library/do/references/show?id=90273)

Albergel, C., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J. & Boussetta, S. (2012c). A bare ground evaporation revision in the ECMWF land-surface scheme: evaluation of its impact using ground soil moisture and satellite microwave data. Hydrology and Earth System Sciences, 16, 3607-3620

(Also Tech. Mem. #685, http://www.ecmwf.int/publications/library/do/references/show?id=90545)

Albergel, C., Dorigo, W., Reichle, R., Balsamo, G., de Rosnay, P.., Muñoz-Sabater, J., Isaksen, L.,Reichle, R., Jeu, R.d., & Wagner, W. (2013). Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. Journal of Hydrometeorology, in press, doi:10.1175/JHM-D-12-0161.1.

(Also Tech. Mem. #695, http://www.ecmwf.int/publications/library/do/references/show?id=90708)

Annexe 1: Soil moisture networks currently available at ECMWF

	Number of stations	**Periods**	**Considered depth (cm)**
SCAN (USA) Schaefer and Paetzold, 2010	177	2007-2012	5, 10, 20, 50
SNOTEL (USA) Schaefer and Paetzold, 2010	348	2012	5
USCRN (USA) http://www.ncdc.noaa.gov/crn/	114	2007-2012	5, 10, 20, 50
SMOSMANIA (France) Albergel et al., 2008, Calvet et al., 2007	12	2007-2012	5, 10, 20, 30
SWATMEX (France) Parrens et al., 2011	9	2009-2012	5, 10, 20, 30
AMMA (Western Africa) Pellarin et al., 2009	7	2007-2011	5
OZNET (Australia) Smith et al., 2012	38	2007-2011	0-30, 30-60, 60-90
REMEDHUS (spain) Ceballos et al., 2005	20	2007-2011	5
MAQU (China) Su et al., 2011	20	2008-2010	10
GTK (Finland) Geological Survey of Finland	4	2007-2011	10, 30, 50, 70
VAS (Spain) http://nimbus.uv.es/	2	2010-2011	5

HOBE (Denmark) Birsher et al., 2011	30	2009-2011	5
UMBRIA (Italy) Brocca et al., 2011	4	2007-2012	5, 15, 35

Albergel, C., C. Rüdiger, T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet & E. Martin (2008). From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in situ observations and model simulations, Hydrol. Earth Syst. Sci., 12, 1323–1337, doi:10.5194/hess-12-1323-2008.

Bircher, S., Skou, N., Jensen, K.H., Walker, J.P., and Rasmussen, L. (2011): A soil moisture and temperature network for SMOS validation in Western Denmark. Hydrology and Earth System Sciences Discussions, 8, 9961-10006, doi:10.5194/hessd-8-9961

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C., Bittelli, M. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe. Remote Sensing of Environment, 115, 3390-3408, doi:10.1016/j.rse.2011.08.003.

Calvet, J.-C., N. Fritz, F. Froissard, D. Suquia, A. Petitpa & B. Piguet (2007) in situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, International Geoscience and Remote Sensing Symposium, IGARSS, Barcelona, Spain, doi:10.1109/IGARSS.2007.4423019.

Ceballos, A., Scipal, K., Wagner, W & Martinez-Fernandez, J., (2005). Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero-Basin, Spain. Hydrol. Process., vol. 19, no.8, pp.1549-1566, May 2005.

Parrens, M, E. Zakharova, S. Lafont, J.-C. Calvet, Y. Kerr, W. Wagner & J.-P. Wigneron (2011). Comparing soil moisture retrievals from SMOS and ASCAT over France, Hydrol. Earth Syst. Sci. Discuss., 8, 8565-8607, doi:10.5194/hessd-8-8565-2011.

Schaefer, G.L. & R.F. Paetzold (2000). SNOTEL (SNOwpack TELemetry) and SCAN (Soil Climate Analysis Network) Presented at the Automated Weather Station (AWS) workshop. March 6-10. Lincoln, NE.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S. & Richter, H. (2012). The Murrumbidgee soil moisture monitoring network data set. Water Resources Research, 48, W07701, doi:10.1029/2012WR011976.

Su, Z., Wen, J., Dente, L., van der Velde, R., Wang, L., Ma, Y., Yang, K., and Hu, Z. 2011, The Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products, Hydrol. Earth Syst. Sci., 15, 2303–2316, 2011