

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Su, CH;Ryu, D;Dorigo, W;Zwieback, S;Gruber, A;Albergel, C;Reichle, RH;Wagner, W

Title:

Homogeneity of a global multisatellite soil moisture climate data record

Date:

2016-11-16

Citation:

Su, C. H., Ryu, D., Dorigo, W., Zwieback, S., Gruber, A., Albergel, C., Reichle, R. H. & Wagner, W. (2016). Homogeneity of a global multisatellite soil moisture climate data record. *Geophysical Research Letters*, 43 (21), pp.11-252. <https://doi.org/10.1002/2016GL070458>.

Persistent Link:

<https://hdl.handle.net/11343/291903>

**Homogeneity of a global multi-satellite soil moisture
climate data record**

Chun-Hsu Su,¹ Dongryeol Ryu,¹ Wouter Dorigo,² Simon Zwieback,³

Alexander Gruber,² Clement Albergel,⁴ Rolf H. Reichle,⁵ Wolfgang Wagner²

Author Manuscript

Corresponding author: Chun-Hsu Su, Department of Infrastructure Engineering, University of Melbourne, Victoria 3010, Australia (csu@unimelb.edu.au)

¹Department of Infrastructure

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/2016GL070458](https://doi.org/10.1002/2016GL070458)

D R A F T

October 16, 2016, 12:00am

D R A F T

Engineering, University of Melbourne,
Victoria, Australia

²Department of Geodesy and
Geoinformation, Vienna University of
Technology, Vienna, Austria

³Institute of Environmental Engineering,
ETH Zurich, Switzerland

⁴CNRM, UMR 3589 (Météo-France,
CNRS) Toulouse, France

⁵Global Modeling and Assimilation Office,
NASA Goddard Space Flight Center,
Greenbelt, MD 20771, USA

Author Manuscript

Key Points.

- A 36-year, multi-satellite soil moisture climate record is tested for temporal homogeneity for the first time using new homogeneity tests.
- Inhomogeneities due to changing observing systems modify the observed trends but an average error is less than $0.003 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$.
- Contiguous periods of data without inhomogeneities at satellite changeover times are generally longer than 10 years.

Climate Data Records (CDR) that blend multiple satellite products are invaluable for climate studies, trend analysis and risk assessments. Knowledge of any inhomogeneities in the CDR is therefore critical for making correct inferences. This work proposes a methodology to identify the spatiotemporal extent of the inhomogeneities in a 36-year, global multi-satellite soil moisture CDR as the result of changing observing systems. Inhomogeneities are detected at up to 24% of the tested pixels with spatial extent varying with satellite changeover times. Nevertheless, the contiguous periods without inhomogeneities at changeover times are

generally longer than 10 years. Although the inhomogeneities have measurable impact on the derived trends, these trends are similar to those observed in ground data and land-surface reanalysis, with an average error less than $0.003 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$. These results strengthen the basis of using the product for long-term studies and demonstrate the necessity of homogeneity testing of multi-satellite CDRs in general.

1. Introduction

Soil moisture (SM) plays many roles in the climate system by controlling feedback between temperature and precipitation, the return flow of water from land to the atmosphere via evaporation, and heterotrophic soil respiration [Seneviratne *et al.*, 2010]. Changes in SM supply have led to multi-year declines in continental evaporation [Miralles *et al.*, 2014], and SM anomalies can strengthen summer heat anomalies [Fischer *et al.*, 2007]. A long-term, homogeneous observational data set of soil moisture is therefore invaluable for improving understanding of climate variability by reducing uncertainties in modelling of the SM-precipitation coupling in climate models and the responses of vegetation and the carbon cycle to the changing climate. Such data can initialize, constrain and evaluate climate models' land surface schemes, and allow trend and anomaly detection for agro-environmental and hydrological changes and likelihood analysis of extremes [e.g. Albergel *et al.*, 2011a]. However, ground measurements are only available in few regions and mostly only for short periods [Dorigo *et al.*, 2011]. The ESA's Climate Change Initiative (CCI) SM project aims to close this gap by combining multiple satellite SM data sets to produce a 36-year (November 1978–December 2014) SM Climate Data Record (CDR) (referred to as ESA CCI SM henceforth) [Liu *et al.*, 2012; Wagner *et al.*, 2012]. This is particularly challenging for SM because different satellite products show different cumulative distribution functions (cdfs), owing to the fact that SM is highly variable across spatial scales, and different sensors have distinctive error characteristics and/or represent different soil volumes.

Several investigators have evaluated previous versions of the ESA CCI SM product. Trend analyses by *Dorigo et al.* [2012] and *Albergel et al.* [2013a, b] show similarities between ESA CCI SM and model re-analyses from ERA-Land, ERA-Interim and GLDAS-Noah, and/or precipitation records, while *Albergel et al.* [2013a] found significant correlations between ESA CCI and in situ SM 30-day moving averages and anomalies. These efforts are particularly important as satellite SM datasets, particularly ESA CCI SM, are increasingly being used for long-term studies, e.g., studying the impact of oceanic oscillations on surface water distributions [*Bauer-Marschallinger et al.*, 2013] and trends in changing global water fluxes [*Miralles et al.*, 2014]. However, *Dorigo et al.* [2015] evaluated ESA CCI SM against in situ SM measurements to find temporal changes in its quality over time, and *Loew et al.* [2013] observed discontinuities in zonal mean SM. Such abrupt changes, i.e., inhomogeneities, in multi-satellite timeseries data can result from non-climatic causes such as changes in instrumentation, retrieval algorithm, and instrumental drift and failure, and lead to spurious climatic behaviour. They must therefore be identified and addressed during or after blending.

To inform the use of this data for climate studies, this paper proposes homogeneity tests to characterise the possible inhomogeneities, and their extent, of the monthly ESA CCI SM timeseries, as a result of changing observing systems. The spatiotemporal extent of homogeneity in the data is quantified and the impacts of inhomogeneities on trend analysis are examined.

2. Homogeneity testing of ESA CCI SM

The algorithm behind the ESA CCI SM production consists of two stages of bias correction and merging of eight passive and active microwave satellite sensor types [Liu *et al.*, 2012; Wagner *et al.*, 2012]. In the first stage, passive SM retrievals from AMSR-E (Advanced Microwave Scanning Radiometer–Earth Observing System) serve as the reference to which SM estimates from SMMR (Scanning Multi-channel Microwave Radiometer), SSM/I (Special Sensor Microwave Imager), TMI (Tropical Rainfall Measuring Mission Microwave Imager) and WindSat were rescaled by matching their cdfs and merged on a per pixel basis. This merged data set was then extended with retrievals from the AMSR2 (Advanced Microwave Scanning Radiometer–2). In parallel, active retrievals from ERS-2 (European Remote Sensing) were rescaled to match ERS-1, which were subsequently rescaled to and merged with the SM from ASCAT (Advanced Scatterometer). In the second stage, the merged passive and active data sets were rescaled against GLDAS-Noah modeled SM, and merged using selection rules based on their relative sensitivity to vegetation density. Here we use the latest v2.2, daily 0.25° gridded product that is representative of the upper (up to 2 cm) soil layer.

Relative tests (RT), which involve using reference SM data, are applied to the monthly aggregates of the ESA CCI data to detect possible inhomogeneities at the changes of the instruments and/or retrieval methods, where ‘inhomogeneity’ is defined as either an artificial discontinuity in overall SM level or a change in variability. A form of absolute homogeneity test that does not require a reference series has also been trialled, but it is found not skillful (see *Supporting Information*). As shown in Fig. 1, the data can be decomposed into eight time periods T_i based on the observing sensors, such that

homogeneities could occur at the transition time $t_{i,j}$ between two successive periods T_i and T_j . Six out of seven possible breakpoints $t_{i,j}$ are tested for inhomogeneities; those at $t_{1,2}$ are not tested because there are only 10% valid data per month, on average, available to compute monthly averages during the operating period T_1 of SMMR.

RT is performed on a monthly difference (residual) series Q between the candidate series Y (i.e., ESA CCI SM) and a reference series X [Alexandersson and Moberg, 1997; Tuomenvirta and Alexandersson, 1995]. In circumstances where there are multiple (K) series X_i ($i = 1, 2, \dots, K$), e.g., from a network of in situ sensors, that are associated or coincident with Y_n , we define,

$$Q = Y - \frac{\sum_{i=1}^K V_i(\beta_i X_i - c_i)}{\sum_{i=1}^K V_i}, \quad (1)$$

where the second term refers to the representative reference series to which Y is compared. The normalization V_i weights the relative contribution of the i th source (X_i) to the construction of the reference series. Each weighting coefficient is taken as the square of positive (linear) correlation between X_i and Y . Distinguishing from the formalism of Alexandersson and Moberg [1997] who corrected additive or multiplicative biases for meteorological data, we use bivariate linear regression (with coefficients β_i and c_i) to correct for both kinds of biases between the candidate and the reference series, because a linear model is commonly used to describe the biases amongst in situ, modelled and satellite SM [Gruber et al., 2016].

Since each analysis period $T_i \cup T_j$ may contain inhomogeneities at a priori known transition time $t_{i,j}$, RT looks for significant statistical differences between samples of $\{Q_n; n \in T_i\}$ c.f. $\{Q_n; n \in T_j\}$. While any two periods T_i and T_j can be chosen arbitrar-

ily, our choice of $j = i + 1$ identifies the location of the inhomogeneities. The discontinuity is tested using the Wilcoxon rank-sum test for equality of median [Karl and Williams, 1987], which is a non-parametric test that is more robust against small sample sizes, outliers, and non-normal or tailed distributions, compared to a t -test for equality of means. The ranks of Q_n in each sample are summed and compared, and the p -value of the test is determined through a permutation test for small samples. Similarly, an artificial change in variance, which leads to unequal dispersion between the two samples of $\{Q_n\}$, is tested using the Fligner-Killeen test, which is a non-parametric, rank-sum based counterpart to the F -test [Conover et al., 1981]. Formally, RT at $t_{i,j}$ is a hypothesis test with the null hypothesis H_0^τ (i.e., homogeneous) and an alternative hypothesis H_1^τ (inhomogeneous) given by,

$$H_0^\tau : G(Q_n; n \in T_i) = G(Q_n; n \in T_j), \quad (2)$$

$$H_1^\tau : G(Q_n; n \in T_i) \neq G(Q_n; n \in T_j), \quad (3)$$

where $G(\cdot)$ computes the sum of ranks of Q_n in the Wilcoxon test ($\tau = W$) and the sum of ranks of the absolute deviations $|Q_n - \text{median}(Q)|$ in the Fligner-Killeen test ($\tau = FK$). Note that other forms of inhomogeneities such as inhomogeneous trends and non-climatic changes that do not coincide with instrumental changes [Alexandersson and Moberg, 1997] are not considered here, but such inhomogeneities could in principle lead to rejection of the null hypothesis.

RT is implemented with co-located ground SM measurements (referred to as RTG hereafter) at ≤ 10 cm soil depths from 44 operational and experimental monitoring networks worldwide with a variety of climatic regions, land cover types and soil textures, drawn

from ISMN (International Soil Moisture Network) [Dorigo *et al.*, 2011, 2013]. *Support-*
ing Information provides details on the processing of the data for testing. The ground
data are broadly accepted to provide the best available reference series, although they
are highly heterogeneous in spatial and temporal coverage. The discrepancies in spatial
scale need to be resolved, which is implicitly achieved via Eq. 1. It is of note that in
situ data can have large representativeness errors [Gruber *et al.*, 2013], but representivity
differences between station and satellite data are likely to diminish with temporal scales,
since monthly variations in various SM data generally show better agreement than their
sub-daily components [Su and Ryu, 2015]. Nevertheless it remains a working assumption
that the monthly variations and trends at local scales are similar to those at the coarse
scales.

We therefore also consider a second implementation of RT with modelled SM as reference
(referred to as RTM). Specifically, we use the top 2 cm SM from the MERRA (Modern
Era Retrospective-analysis for Research and Applications) Land, a land-only reanalysis
data product [Reichle *et al.*, 2011]. Similar to ESA CCI SM, reanalyses such as MERRA-
Land suffer from discontinuities in the observing systems [Robertson *et al.*, 2014], but the
impact of the evolving satellite system is somewhat mitigated by the fact that MERRA-
Land uses gauge-based precipitation forcing. Even so, the global precipitation gauge
network changes with time, and gauges are scarce in remote regions, which can also lead
to discontinuities. Thus, the use of MERRA-Land data in RTM needs to be evaluated
against RTG, where two tests are conducted independently with the satellite data being
sampled separately for each test.

3. Results and discussion

3.1. Testing for inhomogeneities with RTG

Tab. 1 shows percentages of test instances ($\%H_1^T$) where RTG detected inhomogeneities with different significance levels α . Across the periods, 6.8% of the tested instances at a significance level $\alpha = 0.01$ show discontinuities and 1.8% show changes in variance. *Supporting Information* shows timeseries plots of detected inhomogeneities and rainfall data [Adler *et al.*, 2003]. Limited ground monitoring before 2000 and less frequent spatio-temporal coverage of the satellites undermine the ability of RTG to test for inhomogeneities at $t_{2,3}$ and $t_{3,4}$. The occurrence of inhomogeneity in variance is consistently lower than that of discontinuities. This suggests that the cdf-matching approach of the CCI product may be more effective in reducing multiplicative biases between different satellite products than removing additive biases, or that the variances of monthly aggregates of SM are similar amongst the products.

The percentages of inhomogeneities vary between the analysis periods, with considerably higher values recorded for $t_{5,6}$ (at the advent of ASCAT) than for $t_{6,7}$ and $t_{7,8}$. This is partly due to the fact that RTG's statistical power varies with different test periods; the hypothesis testing has less statistical power during these latter two periods due to small data samples in period T_7 . However the remarkable differences in $\%H_1^T$ is likely to suggest that ESA CCI SM is relatively homogeneous over $T_6 - T_8$, owing to the instrumental similarities and retrieval algorithm consistency, amongst AMSR-E, WindSat and AMSR2. Their SM estimates are retrieved similarly from their C- and X-band observations via the Land Parameter Retrieval Model (LPRM) [Owe *et al.*, 2008], and these products have shown to have similar error structures in terms of biases (additive and multiplicative) and

signal-to-noise ratio [Su et al., 2016]. By contrast, the higher occurrence of discontinuities at $t_{5,6}$ could arise from different relative contributions of the active retrievals to the merged active-passive product before and after January 2007. While the ERS and ASCAT SM show similar error structures after inter-calibration [Reimer et al., 2013; Su et al., 2016], the ERS coverage became limited after 2003 and the error structures of active retrievals differ considerably from those of passive counterparts [e.g. Dorigo et al., 2015; Su et al., 2016]. The CCI's blending algorithm explicitly selects the active retrievals over passive counterparts in densely vegetated areas.

3.2. Global mapping of inhomogeneities in ESA CCI SM

We evaluate the skills of RTM by assessing its ability to match the test outcomes of RTG implemented with $\alpha = 0.01$, based on the odds ratio (OR) and the Hanssen-Kuipers discriminant (HK) (*Supporting Information*). OR measures the odds of an inhomogeneity being correctly detected relative to the odds of the detection being incorrect (with larger values being better, and 1 being not skillful), whereas HK indicates how well the candidate test separates the inhomogeneous events from homogeneous ones (HK = 0 indicates no skill, and the perfect score is 1). Based on 685–714 test instances across analysis periods and locations, the Wilcoxon (and Fligner-Killeen) RTM test with $\alpha = 0.01$ are found to be skillful, with high OR = 21 (93) and HK = 0.46 (0.57). These results suggest that RTM using MERRA-Land is skillful at detecting inhomogeneities, although our assessment is principally based on areas that tend to have raingauge measurements available for performing more accurate land surface analyses. RTM could yield false positives in other

187 areas with limited raingauges. Despite this limitation, the current results give cause to
 188 apply RTM globally.

189 Fig. 2 maps the locations of inhomogeneities detected by RTM ($\alpha = 0.01$) when applied
 190 to 27–57% of the land points. The occurrence of discontinuities is considerably higher for
 191 $t_{5,6}$ ($\sim 24\%$ of test locations) than for $t_{6,7}$ and $t_{7,8}$ ($\sim 7\%$), mirroring the results from
 192 RTG. For the earlier $t_{2,3}$, $t_{3,4}$ and $t_{4,5}$, there are also relatively high instances of inhomoge-
 193 neities (15–21%), corresponding to successive introductions of ERS, TMI, and AMSR-E
 194 (in addition to end of SSM/I and TMI), respectively. *Su et al.* [2016] find the seasonal-to-
 195 inter-annual variations in the three passive products (SSM/I, TMI, and AMSR-E) show
 196 different error structures, despite being derived using the same LPRM retrieval algorithm.
 197 TMI provided exclusively X-band retrievals compared to the preferred C-band retrievals
 198 from AMSR-E (which are used in all regions without C-band radio frequency interfer-
 199 ence), while the SSM/I retrievals are based on higher frequency (K_u -band) observations
 200 that are more sensitive to vegetation masking. The SSM/I SM is known to show erro-
 201 neous seasonality, such that the SSM/I SM seasonality during 1987–2002 was replaced
 202 with AMSR-E seasonality, which was derived from the 2002–2007 data [*Liu et al.*, 2012],
 203 before merging. This approach does not fully account for inter-annual variability in sea-
 204 sonal cycles, and the inability of LPRM-derived SSM/I SM to reproduce seasonality and
 205 possibly interannual variations in SM reflects deeper problems in K_u -band retrievals. The
 206 findings are similar to those from the RTM test for inequality in variances, although their
 207 occurrence is still lower than that of discontinuities, e.g., 8.8% for $t_{5,6}$ and 1.1% for $t_{6,7}$
 208 and $t_{7,8}$.

Using the bivariate categorical χ^2 -test and HK analyses to examine the predictability of the locations with detected inhomogeneities (details in *Supporting Information*), three measures, which characterise the per-pixel changes in sensors, sensing types (active c.f. passive), and sensing frequencies, are shown to be significantly and positively correlated (p -value < 0.01) with these locations. We find that the spatial extents of changes in sensing characteristics are considerably larger than those of the inhomogeneities (Fig. S7–S9 in *Supporting Information*). Changes in sensors are found to be associated with the inhomogeneities detected primarily at $t_{3,4}$ and $t_{7,8}$. Changes in sensing types are associated with detections at $t_{5,6}$ and $t_{6,7}$, and changes in sensing frequencies with those at $t_{3,4}$ and $t_{4,5}$. The latter two observations support the above assessments of the impacts of introducing ASCAT and the changes in the passive sensors.

Finally, Fig. 3 maps the length of the longest homogeneous periods over which inhomogeneities at the transition times are tested and not detected by RTM. It shows that about 50% of the pixels tested can provide monthly data that is homogeneous across some of these transition times over a period > 15 years, and about 35% of the pixels with > 20 -year periods of homogeneous data. There is considerable spatial variability in the homogeneous data lengths globally, with concentrations of shorter homogeneous data (≤ 15 years across transition times) observed in the Europe-Mediterranean region, northern Africa and the Middle East, and north-eastern Australia. These are mainly because there are insufficient data in the CCI product to apply RTM at $t_{2,3}$, $t_{3,4}$ and $t_{4,5}$ in the Europe-Mediterranean region and north-eastern Australia, and at $t_{6,7}$ and $t_{7,8}$ in northern

Africa and the Middle East (see Fig. S2 in *Supporting Information*), given that not all the satellite SM data had been used in creating the product.

3.3. Impact of inhomogeneities on trend analysis

Fig. 4 examines the possible influence of inhomogeneities on trend analyses by comparing the slopes of ESA CCI SM timeseries against those of the reference data. We follow *Dorigo et al.* [2012] to compute Sen median estimates of slopes in seasonal SM over periods (3–9 years) with inhomogeneities as detected by RTG or RTM. The biases between ESA CCI SM and the reference data are corrected via linear regression before slope estimation. Fig. 4(a) shows moderate agreement between the trends observed in the raw CCI SM data and that in the ground data at 35 sites, with linear correlation R of 0.64 ± 0.21 (approximated 95% confidence interval based on the Fisher’s z transformation) and root-mean-square-error (rmse) of $0.006 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$. The directions of the trends can differ and the largest difference in magnitude is $0.02 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$. Although the sample size is too small to provide a representative measure of the impact of inhomogeneities on the actual trends, Fig. 4(c) indicates that there may be better agreement between ESA CCI SM and the ground data when the former is regarded homogeneous, with $R = 0.84$ and $\text{rmse} = 0.005 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$. The similar contrasts, between inhomogeneous and homogeneous CCI data ($R = 0.40$ and $\text{rmse} = 0.004 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$ c.f. 0.61 and $0.002 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$), can also be observed in Fig. 4(b) and (d) when compared with trends in MERRA-Land. In other words, while trend differences between CCI and reference data could be partly attributed to the presence of inhomogeneities in ESA CCI SM, the trends observed in the inhomogeneous data can still be similar to the reference data; based on the differences

in rmse between the inhomogeneous and homogeneous samples, the errors in trends are within $\text{rmse} = 0.003 \text{ m}^3\text{m}^{-3}\text{y}^{-1}$ on average. This result shows that the stability of ESA CCI SM meets the current stability requirement target specified in *GCOS* [2011].

4. Conclusion

This work detected non-climatic changes in mean and variance in the ESA CCI SM product at times when the satellite instrumental configuration changes. By using these tests with model data as the reference (RTM), we presented the first global maps of inhomogeneities (Fig. 2). The RTM setup allows global-scale testing and is shown to be skillful when validated against RTG (where the ground data provides the reference); however, the potential of RTM for false detections due to inhomogeneities in the reanalysis data suggests that further refinements are needed, e.g., further quality assurance of the reference data and using independent observations of rainfall as a reference.

Inhomogeneities are detected at up to 24% of the tested pixels with this spatial extent varying with specific crossover times between the satellite sensors. In spite of this, the contiguous periods without inhomogeneities at some of six tested changeover times are generally longer than 10 years. (Fig. 3). The inhomogeneities are shown to have measurable impacts to modify the perceived trends in the data, but there is still broad agreement between the trends in ESA CCI SM and the reference data (Fig. 4). Both of these findings are important to give confidence in the quality of the product for long-term climate studies and trend detection. As new SM climate data produced with improved blending algorithms and derived from new remote sensing sensors are expected soon, homogeneity testing can be combined with standard validation practices to provide stringent quality

control of the data, and contribute to reproducing a more consistent SM climatology in observational data. More generally, the work demonstrates the necessity of homogeneity testing of multi-satellite CDRs, and the assessment of other CDRs is expected to benefit from the methods proposed in this study.

Acknowledgments. This research is supported by The University of Melbourne Early Career Researcher Grant Scheme, and the ESA CCI Programme “Phase 2 of the ESA CCI SM FCV” (contract no. 4000112226/14/I-NB). WD is supported by the personal grant “TU Wien Wissenschaftspreis 2015”. The ESA CCI SM and ISMN data were obtained freely from <http://www.esa-soilmoisture-cci.org> and <https://ismn.geo.tuwien.ac.at>, respectively, and the MERRA-Land data were from <http://disc.sci.gsfc.nasa.gov/mdisc>.

References

- Adler, R.F., C.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, et al. (2003). The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, 4, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.
- Albergel, C., et al. (2013a), Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing, *J. Hydrometeor.*, 14, 1259–1277, doi:10.1175/JHM-D-12-01611.
- Albergel, C., et al. (2013b), Monitoring multi-decadal satellite earth observation of soil moisture products through land surface reanalyses, *Remote Sens. Environ.*, 138, 77–89, doi:10.1016/j.rse.2013.07.009.

Alexandersson, H., and Moberg, A. (1997), Homogenization of Swedish temperature data: Part I: Homogeneity test for linear trends, *Int. J. Climatol.*, 17: 25–34, doi: 10.1002/(SICI)1097-0088(199701)17:1<25::AID-JOC103>3.0.CO;2-J.

Bauer-Marschallinger, B., Dorigo, W., Wagner, W., and Van Dijk, A. (2013), How oceanic oscillations drives soil moisture variations over mainland Australia: An analysis of 32 years of satellite observations, *J. Climate*, 26: 10159–10173, doi:10.1175/jcli-d-13-00149.1.

Conover, W. J., Johnson, M. E., and Johnson, M. M. (1981), A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, 23(4), 351–361.

Dorigo, W.A., et al. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements, *Hydrol. Earth Syst. Sci.*, 15, 1675–1698, doi:10.5194/hess-15-1675-2011.

Dorigo, W., et al. (2012), Evaluating global trends (1988–2010) in harmonized multi-satellite surface soil moisture, *Geophys. Res. Lett.*, 39, L18405, doi: 10.1029/2012GL052988.

Dorigo, W.A., et al. (2013). Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone J.*, 12(3), doi: 10.2136/vzj2012.0097.

Dorigo, W. A., et al. (2015), Evaluation of the ESA CCI soil moisture product using ground-based observations, *Rem. Sens. Environ.*, 162, 380–395, doi: 10.1016/j.rse.2014.07.023.

- 314 Ferguson, C. R., and Villarini, G. (2012), Detecting inhomogeneities in the Twentieth
315 Century Reanalysis over the central United States, *J. Geophys. Res.-Atmos.*, 117(D5),
316 D05123, doi:10.1029/2011JD016988.
- 317 Fischer, E. M., et al. (2007), Soil moisture-atmosphere interactions during the 2003 Eu-
318 ropean summer heat wave, *J. Climate*, 20, 5081, doi:10.1175/JCLI4288.1.
- 319 Gruber, A., Dorigo, W. A., Zwieback, S., Xaver, A., and Wagner, W. (2013). Charac-
320 terizing coarse-scale representativeness of in situ soil moisture measurements from the
321 International Soil Moisture Network. *Vadose Zone J.*, 12(2), doi:10.2136/vzj2012.0170
- 322 Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W. (2016). Re-
323 cent advances in (soil moisture) triple collocation analysis. *Int. J. Appl. Earth Obser.*
324 *Geoinform.*, 45, 200–211, doi:10.1016/j.jag.2015.09.002
- 325 Global Climate Observing System (GCOS) (2011), Systematic observation requirements
326 for satellite-based data products for climate: The Supplemental details to the satellite-
327 based component of the “Implementation Plan for the Global Observing System for
328 Climate in Support of the UNFCCC (2010 Update, GCOS-154). Accessed online,
329 <https://www.wmo.int/pages/prog/gcos/documents/SatelliteSupplement2011Update.pdf>
- 330 Hollmann, R., et al. (2013), The ESA Climate Change Initiative: Satellite data records for
331 essential climate variables, *Bull. Am. Met. Soc.*, 94(10), 1541–1552, doi:10.1175/BAMS-
332 D-11-00254.1.
- 333 Karl, T. R., and Williams Jr., C. N. (1987), An approach to adjusting climatological
334 time series for discontinuous inhomogeneities, *J. Clim. Appl. Meteorol.*, 26, 1744–1763,
335 doi:10.1175/1520-0450(1987)026<1744:AATACT>2.0.CO;2.

- Liu, Y. Y., et al. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals, *Rem. Sens. Environ.*, 123, 280–297, doi:10.1016/j.rse.2012.03.014.
- Loew, A., Stacke, T., Dorigo, W., De Jeu, R., and Hagemann, S. (2013), Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies, *Hydrol. Earth Syst. Sci.*, 17, 3523–3542, doi:10.5194/hess-17-3523-2013.
- Miralles, D. G., et al. (2014), El Niño-La Niña cycle and recent trends in continental evaporation, *Nat. Clim. Change*, 4, 122–126, doi:10.1038/nclimate2068.
- Owe, M., de Jeu, R., and Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture, *J. Geophys. Res.*, 113, F01002, doi:10.1029/2007JF000769.
- Peterson, T. C., et al. (1998), Homogeneity adjustments of in situ atmospheric climate data: A review, *Int. J. Climatol.*, 18, 1493–1517, doi:10.1002/(SICI)1097-0088(199811)18:13<1493::AID-JOC329>3.0.CO;2-T.
- Reichle, R. H., et al. (2011), Assessment and enhancement of MERRA land surface hydrology estimates, *J. Climate*, 24(24), 6322–6338, doi:10.1175/JCLI-D-10-05033.1.
- Reimer, C., Hohm, S., and Wagner, W. (2013), Inter-calibration of ERS AMI and MetOp ASCAT backscatter measurements, *Proc. ESA Living Planet Symposium 2013*, Edinburgh, UK, SP-722, 2–13, p. 46.
- Robertson, F. R., et al. (2014), Consistency of estimated global water cycle variations over the satellite era, *J. Clim.*, 27(16), 6135–6154, doi:10.1175/JCLI-D-13-00384.1.

Figure 1. Time periods of the eight passive and active sensor types used in the construction of the ESA CCI SM product. Dashed blue lines $t_{i,j}$ refer to the months in which possible inhomogeneities may occur due to changes in observing systems between two time periods T_i and T_j . Note that the ERS coverage became limited after June 2003 following the failure of the ERS-2 tape drive.

Seneviratne, S. I., et al. (2010), Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99, 125–161, doi: 10.1016/j.earscirev.2010.02.004.

Su, C.-H., and Ryu, D., (2015), Multi-scale analysis of bias correction of soil moisture. *Hydrol. Earth Syst. Sci.*, 19, 17–31, doi:10.5194/hess-19-17-2015.

Su, C.-H., et al., (2016), Error decomposition of nine passive and active microwave satellite soil moisture data sets over Australia. *Rem. Sens. Environ.*, 182, 128–140, doi: 10.1016/j.rse.2016.05.008.

Tuomenvirta, H. and H. Alexandersson (1995). Adjustment of apparent changes in variability of temperature time series. *Proceedings of the Sixth International Meeting on Statistical Climatology*, Galway, Ireland, 443–446.

Wagner, W., et al. (2012). Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture., *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vols. I-7, 316–321, doi:10.5194/isprsannals-I-7-315-2012.

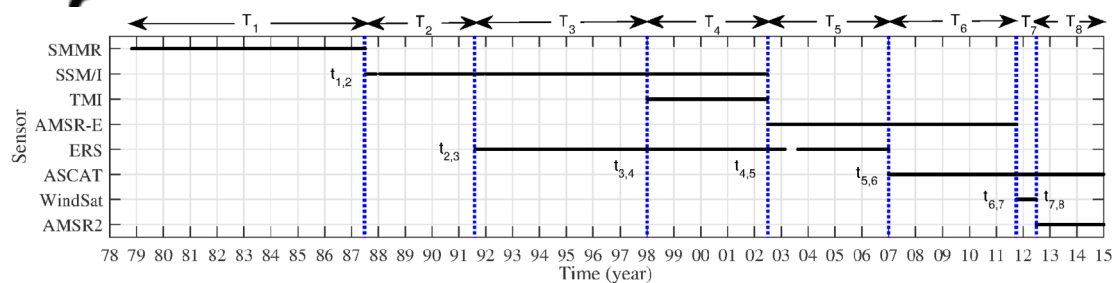
Figure 2. Spatial locations of inhomogeneities for each analysis period, identified by RTM with a significance level $\alpha = 0.01$. The colour coding indicates whether an inhomogeneity is detected (i.e., prefer H_1^τ) or not (prefer H_0^τ) for each test type $\tau = W$ (Wilcoxon) or FK (Fligner-Killeen). %cov denotes the coverage of the tests, and $\%H_1^W$ ($\%H_1^{FK}$) the percentages of locations tested that show discontinuities (change in variance).

Figure 3. Length of (longest) homogeneous periods, over which RTM is applied and no inhomogeneity is detected, in ESA CCI SM.

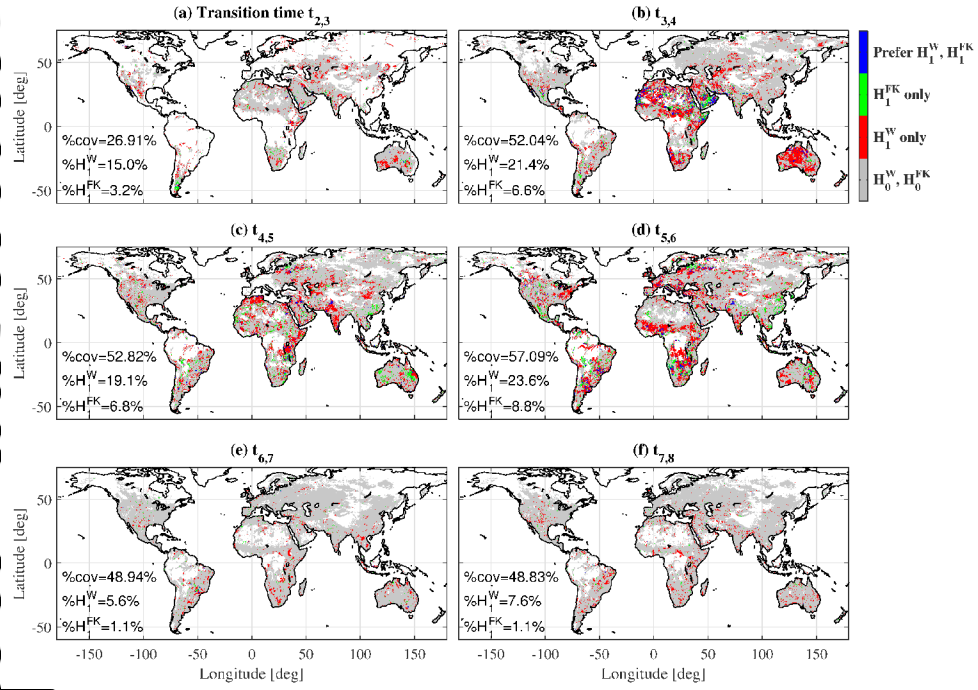
Figure 4. (a) Evaluation of trends in ESA CCI SM against ground measurements at locations where inhomogeneities are detected or not detected by RTG. (b) Same as (a), but for MERRA-Land with RTM. (c,d) Same as (a,b), respectively, but for locations where inhomogeneities are not detected. The range of values differs between the inhomogeneous and homogeneous samples (a,b c.f. c,d) because the latter is generally being estimated from longer timeseries.

Table 1. Percentages of Wilcoxon (Fligner-Killeen) RTG where the null hypotheses are rejected in favour of the alternative hypotheses H_1^W (H_1^{FK}), i.e., a discontinuity (a change in variability) exists, at different significance level α . N is the number of the tests conducted for a given transition time $t_{i,j}$; no test can be conducted for $t_{2,3}$.

	α	0.005		0.01		0.05	
$t_{i,j}$	N	$\%H_1^W$	$\%H_1^{FK}$	$\%H_1^W$	$\%H_1^{FK}$	$\%H_1^W$	$\%H_1^{FK}$
$t_{3,4}$	6	0.0	0.0	0.0	0.0	0.0	0.0
$t_{4,5}$	34	0.0	0.0	0.0	0.0	11.8	2.9
$t_{5,6}$	172	12.2	4.7	18.6	6.4	23.8	12.2
$t_{6,7}$	255	1.6	0.0	2.0	0.4	11.8	5.1
$t_{7,8}$	257	2.7	0.0	4.7	0.4	8.6	3.5

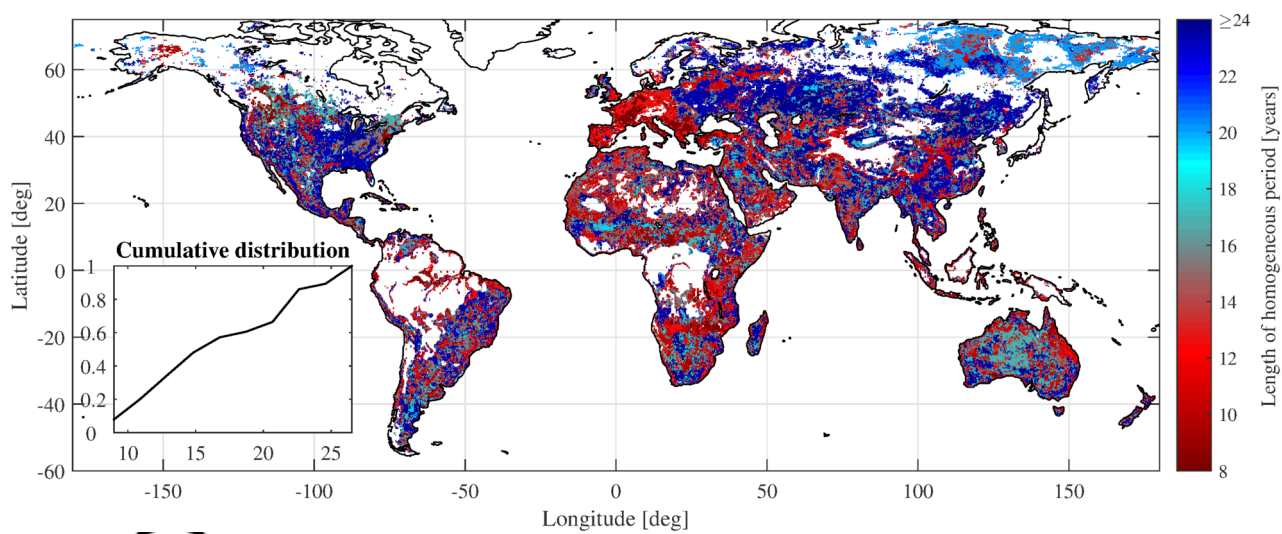


2016gl070458-f01-z-eps



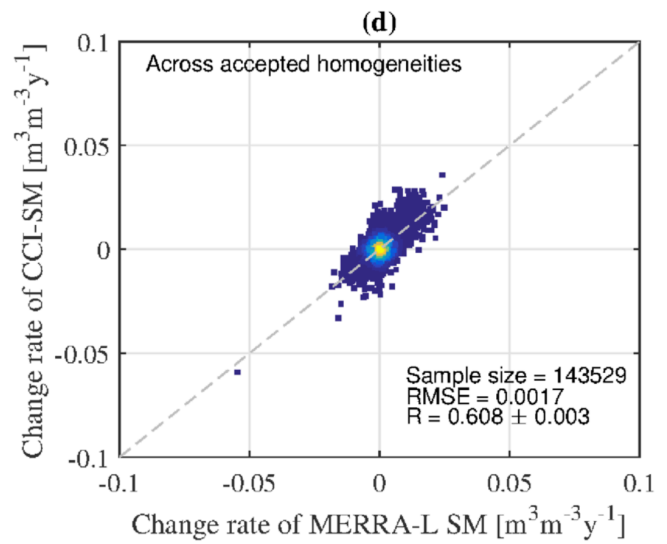
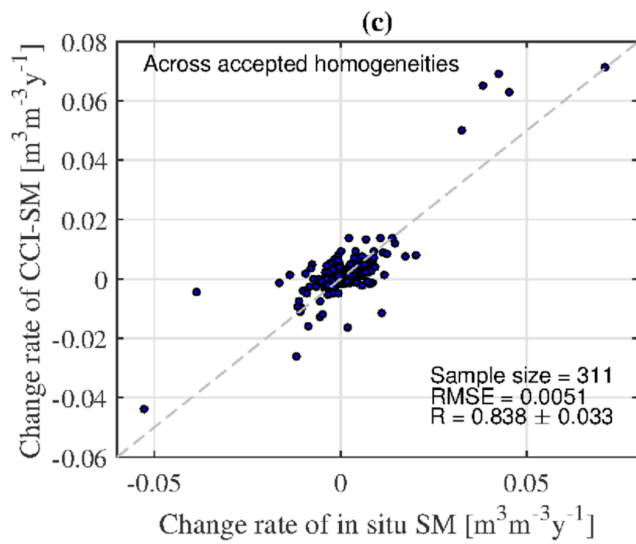
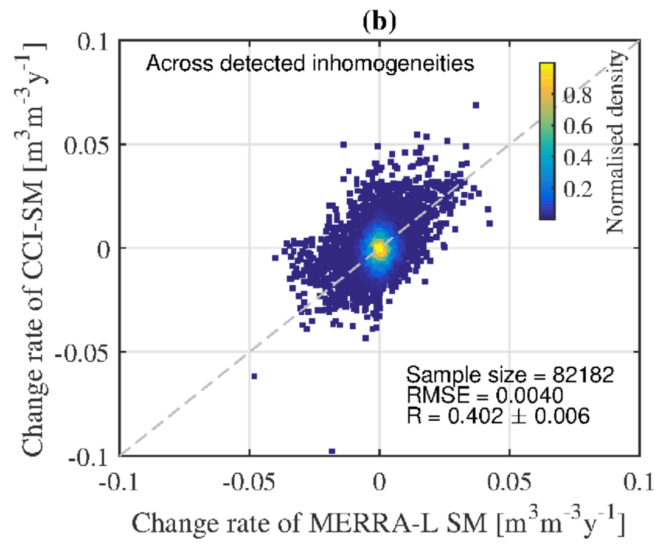
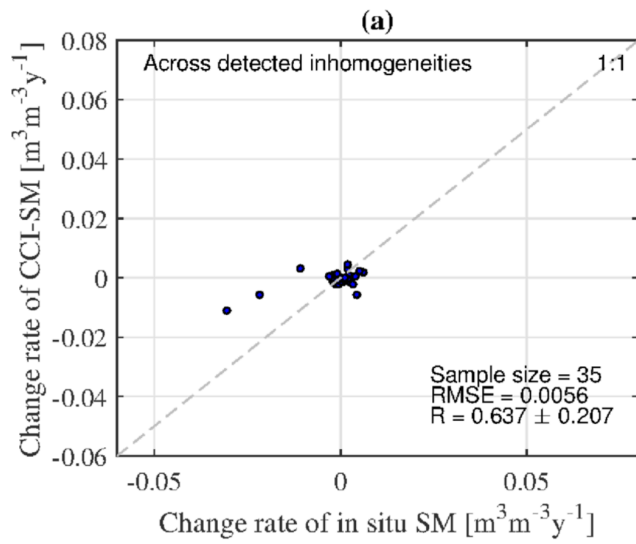
2016gl070458-f02-z-.eps

script



Authc

2016gl070458-f03-z-.eps



2016gl070458-f04-z-eps