Article (refereed) - postprint

Taylor, Christopher M.; de Jeu, Richard A. M.; Guichard, Francoise; Harris, Phil P.; Dorigo, Wouter A.. 2012 Afternoon rain more likely over drier soils. *Nature*, 489. 423-426. 10.1038/nature11377

Afternoon rain more likely over drier soils

Christopher M. Taylor[a], Richard A. M. de Jeu[b], Françoise Guichard[c], Phil P. Harris[a], Wouter A. Dorigo[d]

[a] Centre for Ecology and Hydrology, Wallingford, U.K.

[b] Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[c] CNRM (CNRS and Météo-France), Toulouse, France

[d] Vienna University of Technology, Vienna, Austria

This paper should be cited as…

Taylor, C. M., R. A. M. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo (2012), Afternoon rain more likely over drier soils, *Nature*, *489*(7416), 423-426 (10.1038/nature11377)

Land surface properties influence the partition of radiative energy between latent and sensible heat fluxes during daytime. A climatically important process during dry periods is the impact of soil water deficit on limiting evapotranspiration, leading to increased surface heating of the lower atmosphere[1-2]. Soil moisture can influence the development of convective storms through modification of low level atmospheric temperature and humidity[1,3], which in turn feeds back on soil moisture. There is considerable uncertainty in how soil moisture affects convective storms across the world, due to both a lack of observational evidence and uncertainty in large-scale models[4]. Here we present the first global-scale observational analysis of the coupling between soil moisture and precipitation. This shows that across all six continents studied, afternoon rain falls preferentially over soils which are relatively dry compared to the surrounding area. The signal emerges most clearly in the observations over semi-arid regions, where surface fluxes are sensitive to soil moisture and convective events are frequent. We find no evidence in our analysis of a positive feedback i.e. preference for rain over wetter soils, at the 50-100 km scale studied. In contrast, our analysis of six state-of-the-art global weather and climate models demonstrates the dominance of a positive feedback of soil moisture on simulated precipitation. These results suggest a fundamental weakness in the sensitivity of convection to the land surface in large-scale models, which may contribute to excessive simulated droughts.

Soil moisture influences precipitation on a range of time and space scales[5]. In drought-affected continental regions, weak evapo-transpiration leads to reduced atmospheric moisture content over a period of days, potentially suppressing subsequent precipitation[6]. When soil moisture anomalies are extensive, surface-induced perturbations to the atmospheric heat budget may modify synoptic scale circulations[2], in turn affecting moisture advection from the oceans[7]. On smaller scales, the development of convective clouds and precipitation can be influenced by local surface fluxes over the course of the day[1,3]. Theoretical considerations[8-9] suggest that, in an undisturbed atmosphere, the likelihood and sign of a surface feedback will be determined by the atmospheric profiles of temperature and humidity. Thus one might expect regional variations in the strength and sign of

convective sensitivity to soil moisture[10-11]. Mesoscale soil moisture variability can also influence the feedback via the development of daytime circulations[12], which provide additional convergence to trigger convection[13-14].

There have been a number of studies examining the impact of the land surface on observed rainfall in different regions of the world. Analyses in Illinois [15] and West Africa [16] have indicated positive correlations between antecedent soil moisture and precipitation consistent with a positive soil moisture feedback. A recent study[17] based on observationally-constrained reanalysis data showed an inreasing frequency of convective rainfall when evapotranspiration was higher across much of North America. On the other hand, examination of satellite cloud data tend to indicate locally enhanced afternoon precipitation frequency over surfaces with increased sensible heat fluxes, via mesoscale circulations due either to soil moisture[18] or vegetation cover[19-20].

At the regional scale, climate models tend to agree where feedbacks occur, these being constrained largely by where soil moisture limits evapotranspiration in the presence of convective activity[4]. The spread in simulated feedback strength is large however, highlighting both the uncertainty in surface flux sensitivity to soil moisture and the response of the Planetary Boundary Layer (PBL) and convection to surface fluxes[21-22]. Indeed, the feedback sign can change depending on model spatial resolution, with a strong influence of the convective parameterisation likely to be responsible [23].

This study addresses the lack of observations to evaluate feedbacks in large-scale models. We focus on the least well-understood aspect of soil moisture-precipitation feedbacks, namely the response of daytime moist convection to soil moisture anomalies. Global observational datasets of both surface soil moisture[24-25] and precipitation[26] have become available in the last decade at a resolution of $0.25° \times 0.25°$ on daily and 3-hourly time steps respectively. We use these to analyse the location of afternoon rain events relative to the underlying antecedent soil moisture. In particular we examine whether rain is more likely over soils that are wetter or drier than the surrounding area. The same methodology is then applied to six global models used in reanalyses or climate projections.

We focus on the development of precipitation events during the afternoon, when the sensitivity of convection to land conditions is expected to be maximised. An event is defined at a 0.25°×0.25° pixel location (L_{max}) with a maximum in precipitation, centred within a box of 1.25°×1.25° (see Methods and Supplementary Fig. S3). Each L_{max} is paired with one or more pixels within the box where afternoon rainfall is a minimum (L_{min}). We compute the difference in pre-rain event soil moisture, ΔS_e, between L_{max} and L_{min} having first subtracted a climatological mean soil moisture from both locations. We quantify the strength of the soil moisture signal on precipitation using a sample of events and assess how unexpected the observed sample mean value of ΔS_e is relative to a control sample, ΔS_c, from the same location pairs on non-event days. More precisely, we examine the difference in ΔS between the event and control samples, $\delta_e = \text{mean}(\Delta S_e) - \text{mean}(\Delta S_c)$, as a percentile of typical δ values (see Methods). Mountainous and coastal areas are excluded because of their effects on mesoscale precipitation, and we are unable to analyse the observations in tropical forest regions due to the limitations of soil moisture retrievals beneath dense vegetation.

The map in Figure 1 indicates regions of the world where afternoon precipitation is observed more frequently than expected over wet (blue) or dry (red) soils, based on analysis of δ_e at a scale of 5°. Globally, 28.9% of the grid cells analysed take percentile values P<10, as compared to an expected frequency (assuming no feedback) of 10%, and just 3.4% with percentiles P>90. Clusters of low percentiles are found in semi-arid and arid regions, most notably North Africa, but also in Eastern Australia, Central Asia and Southern Africa. These clusters indicate a clear preference for afternoon rain over drier soils in those regions, consistent with a previous study over the Western Sahel[18]. This signal is also evident when computing δ_e from all events across the world (Figure 1 insets). Further analysis (Supplementary Information and Tables S3 and S4) demonstrates that this signal is statistically significant at the 99% level over all continents and in all climate zones, with the exception of tropical forests, where accurate soil moisture retrievals are unavailable. We repeated the analysis having first degraded the spatial resolution from 0.25° to 1.0° . This produced only

about one tenth of the number of events compared to the 0.25° data, but a statistically robust preference for rain over drier soil was still found across the tropics, and in particular over parts of North Africa and Australia (Supplementary Fig. S10, Tables S3 and S4).

Using two alternative precipitation datasets, we found the same global preference for rain over drier soil and similar regions contributing to that signal (Supplementary Fig. S8, Table S3 and S4). Though all of the satellite-derived datasets are subject to errors at the event scale, analysing the data over many events should yield more accurate estimates of δ_e. Furthermore, our approach exploits an aspect of rainfall which is relatively well-captured by satellite, i.e. its spatial structure. Additional analysis (Supplementary Information Fig. S4) indicates a strong degree of mutual consistency in the spatial variability of soil moisture and rainfall in our independent datasets, providing further evidence to support our methodology.

We now consider whether the observed preference for rain over drier soil is consistent with land surface feedback. For a soil moisture feedback on precipitation, soil water deficit must limit evapotranspiration. This regime is found only in certain seasons and regions of the world[4] where water stress coincides with convective activity. Low percentiles in Figure 1 occur in areas that are relatively dry, and originate from seasons with convective storms (Supplementary Fig. S9). Using data from across the globe, the sensitivity of δ_e to the areal mean (1.25°x1.25°) soil moisture is explored in Figure 2a. The most negative values (rain over drier soil) are found for the driest mean conditions, and the signal loses significance at the 95% level above 0.20 m^3m^{-3}. This behaviour is consistent with soil moisture feedback as the sensitivity of sensible and latent heat fluxes to soil moisture increases as mean soil moisture decreases. Also, the use of surface soil moisture as a proxy for surface flux variability should be most effective for dry and sparsely-vegetated surfaces.

A land feedback requires a strong diurnal sensitivity in the observed signal. We repeated our analysis, this time detecting the onset of precipitation at varying lag times after a soil moisture observation at 0130LT. Values of δ_e (Figure 2b) illustrate a pronounced diurnal cycle, still evident 36

hours after the observation. The most negative values occur during daytime, in particular between 12 and 15LT. On the other hand, between 2100 and 0300LT the opposite signal emerges, that is events are more likely to be defined over wetter soils. The early afternoon minimum is consistent with a negative soil moisture feedback on convective initiation, when the impact of surface properties on the PBL, convective instability, and mesoscale flows are maximised. Mechanisms to explain the reverse signal in the hours around midnight may be more subtle. The impact of thermals and daytime surface-induced flows are likely to be relatively short-lived after dusk. On the other hand, nocturnal humidity anomalies may persist for longer, depending on the spatial scale of the surface features and wind conditions. From detailed examination of individual events, it appears that overnight, there is an increasing influence of pre-existing, fast-moving convective systems in our sample, particularly in the Sahel. Distinct mechanisms will be involved in the surface interaction with organised convective systems which may favour a positive feedback[16].

Finally, we repeat our analysis using 3-hourly diagnostics from six global models, ranging in horizontal resolution from 0.5 to 2.0°. Results (Figure 3) indicate a strong preference for rain over wet soils for large parts of the world, in contrast to observations. Only one model (Figure 3e) produces more than the expected 10% of grid cells with P<10, largely due to contributions at mid-latitudes. The cross-model signal favouring precipitation over wet soil, particularly across the tropics (Table S3), demonstrates a fundamental failing in the ability of convective parameterisations to represent land feedbacks on daytime precipitation. It is likely to be linked to the oft-reported phase lag in the diurnal cycle of precipitation; i.e. simulated rainfall tends to start several hours too early[27], and possibly amplified by a lack of boundary-layer clouds in some models. This weakness has been related to the crude criteria employed to trigger deep convection in large-scale models[28]. The onset of convective precipitation is overly sensitive to the daytime increase of moist convective instability, which is typically faster over wetter soils[3], favouring a positive feedback. Early initiation limits the effect of other daytime processes on triggering convection in the models. Indeed, our observational analysis points to the importance of *dry* boundary-layer dynamics for this phenomenon over land.

The observed preference for afternoon rain over locally drier soil on scales of 50 – 100 km is consistent with a number of regional studies based on remotely sensed data[18-20]. Our failure to find areas of positive feedback may indicate the importance of surface-induced mesoscale flows in triggering convection[18], though the coarse spatial resolution of our datasets prevents us from drawing firm conclusions on this issue. Equally,mixing processes in the growth stage of convective clouds prior to precipitation[23,29] may play an important role. Neither of these processes are captured in existing one-dimensional analyses[8]. Furthermore, our results raise questions about the ability of models reliant on convective parameterisations to adequately represent them. Whilst the coarser resolution models analysed here (HadGEM2, CNRM-CM5 and INMCM4) cannot resolve mesoscale soil moisture structures, nor their potential impacts on convective triggering [18], all the models have a strong tendency to rain over wetter soils for which we find no observational support. Our study does not however imply that the soil moisture feedback is negative at different time and space scales than those analysed here. The multi-day accumulation of moisture in the lower atmosphere from a freely-transpiring land surface may provide more favourable initial (dawn) conditions for daytime convection than the equivalent accumulation over a drought-affected region. Equally, the large-scale dynamical response to soil moisture may dominate in some regions. However, the erroneous sensitivity of convection schemes demonstrated here is likely to contribute to a tendency for large-scale models to "lock-in" dry conditions, extending droughts unrealistically, and potentially exaggerating the role of soil moisture feedbacks within the climate system[30].

Methods Summary

Surface soil moisture retrievals are used between 60°S and 60°N from the Advanced Microwave Scanning Radiometer for EOS (AMSR-E; June 2002 to October 2011)[24], and the MetOP Advanced Scatterometer (ASCAT; 2007-11)[25]. They have typically one overpass per pixel per day at either 0130 or 1330 (AMSR-E) local time (LT), and 0930 or 2130LT (ASCAT). Additional soil moisture quality

control procedures are described in Supplementary Information. The CMORPH[26] 3-hourly precipitation dataset is based on data from a combination of satellites.

Locations of afternoon events L_{max}, are defined within a box of 5x5 pixels by the maximum accumulated precipitation (1200-2100LT) which exceeds 3mm. Pixels with more than 1mm rain in the preceding hours are excluded, and an additional filter applied to remove cases close to active precipitation when using 1330LT soil moisture data. These steps ensure that the soil moisture measurement precedes the rainfall. Locations where topographic height variability exceeds 300m are excluded, along with regions containing water bodies or strong climatological soil moisture gradients.

The control sample ΔS_c is constructed from daily soil moisture differences between locations L_{max} and L_{min} using data for the same calendar month but from non-event years. This quantifies typical (non-event) soil moisture differences between the locations. Each value in samples ΔS_e and ΔS_c has an individual climatological mean ΔS subtracted that is calculated from ΔS values in the same calendar month in non-event years. For the models, soil moisture and rainfall accumulations are available every 3 hours (UTC). Because of the lower spatial resolution (0.5-2.0°), the event box is reduced to 3×3 pixels and the time slot between 0600 and 0859LT adopted to calculate ΔS. Convective rain is accumulated in the subsequent 9 hours, several hours in the day earlier to account for diurnal phase bias in model precipitation.

For further details see the Supplementary Information.

References

1 Betts, A. K. & Ball, J. H. *FIFE surface climate and site-average dataset 1987-89*, *J. Atmos. Sci.* **55**, 1091-1108 (1998).

2 Fischer, E. M. et al. *Soil moisture - Atmosphere interactions during the 2003 European summer heat wave*, *J. Climate* **20**, 5081-5099 (2007).

[3] Eltahir, E. A. B. *A soil moisture-rainfall feedback mechanism 1. Theory and observations*, *Water Resources Research.* **34**, 765-776 (1998).

[4] Koster, R. D. et al. *Regions of strong coupling between soil moisture and precipitation*, *Science.* **305**, 1138-1140 (2004).

[5] Goessling, H. F. & Reick, C. H. *What do moisture recycling estimates tell us? Exploring the extreme case of non-evaporating continents*, *Hydrol. Earth Syst. Sci.* **15**, 3217-3235 (2011).

[6] van der Ent, R. J., Savenije, H. H. G., Schaefli, B. & Steele-Dunne, S. C. *Origin and fate of atmospheric moisture over continents*, *Water Resour. Res.* **46**, W09525 (2010).

[7] Webster, P. J. *Mechanisms of Monsoon Low-Frequency Variability - Surface Hydrological Effects*, *J. Atmos. Sci.* **40**, 2110-2124 (1983).

[8] Findell, K. L. & Eltahir, E. A. B. *Atmospheric controls on soil moisture-boundary layer interactions. Part I: Framework development*, *J. Hydromet.* **4**, 552-569 (2003).

[9] Ek, M. B. & Holtslag, A. A. M. *Influence of soil moisture on boundary layer cloud development*, *J. Hydromet.* **5**, 86-99 (2004).

[10] Findell, K. L. & Eltahir, E. A. B. *Atmospheric controls on soil moisture-boundary layer interactions. Part II: Feedbacks within the continental United States*, *J. Hydromet.* **4**, 570-583 (2003).

[11] Ferguson, C. R. & Wood, E. F. *Observed Land–Atmosphere Coupling from Satellite Remote Sensing and Reanalysis*, *J. Hydromet.* **12**, 1221-1254 (2011).

[12] Ookouchi, Y., Segal, M., Kessler, R. C. & Pielke, R. A. *Evaluation of soil moisture effects on the generation and modification of mesoscale circulations*, *Mon. Weather Rev.* **112**, 2281-2292 (1984).

[13] Cheng, W. Y. Y. & Cotton, W. R. *Sensitivity of a cloud-resolving simulation of the genesis of a mesoscale convective system to horizontal heterogeneities in soil moisture initialization*, *J. Hydromet.* **5**, 934-958 (2004).

[14] Anthes, R. A. *Enhancement of convective precipitation by mesoscale variations in vegetative covering in semi-arid regions*, *J. Clim. Appl. Meteorol.* **23**, 541-554 (1984).

[15] Findell, K. L. & Eltahir, E. A. B. *An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois*, *Wat Resour Res.* **33**, 725-735 (1997).

[16] Taylor, C. M. & Lebel, T. *Observational evidence of persistent convective-scale rainfall patterns*, *Mon. Weather Rev.* **126**, 1597-1607 (1998).

[17] Findell, K. L., Gentine, P., Lintner, B. R. & Kerr, C. *Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation*, *Nature Geosci.* **4**, 434-439 (2011).

[18] Taylor, C. M. et al. *Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns*, *Nature Geosci.* **4**, 430-433 (2011).

[19] Wang, J. F. et al. *Impact of deforestation in the Amazon basin on cloud climatology*, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 3670-3674 (2009).

[20] Carleton, A. M. et al. *Synoptic circulation and land surface influences on convection in the Midwest US ''Corn belt'' during the summers of 1999 and 2000. Part II: Role of vegetation boundaries*, *J. Climate.* **21**, 3617-3641 (2008).

[21] Santanello, J. A., Peters-Lidard, C. D. & Kumar, S. V. *Diagnosing the Sensitivity of Local Land–Atmosphere Coupling via the Soil Moisture–Boundary Layer Interaction*, *J. Hydromet.* **12**, 766-786 (2011).

[22] Guo, Z. C. et al. *GLACE: The Global Land-Atmosphere Coupling Experiment. Part II: Analysis*, *J. Hydromet.* **7**, 611-625 (2006).

[23] Hohenegger, C., Brockhaus, P., Bretherton, C. S. & Schar, C. *The Soil Moisture-Precipitation Feedback in Simulations with Explicit and Parameterized Convection*, *J. Climate* **22**, 5003-5020 (2009).

[24] Owe, M., de Jeu, R. & Holmes, T. *Multisensor historical climatology of satellite-derived global land surface moisture*, *Journal of Geophysical Research-Earth Surface.* **113** (2008).

[25] Bartalis, Z. et al. *Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT)*, *Geophys. Res. Lett.* **34**, L20401 (2007).

[26] Joyce, R. J., Janowiak, J. E., Arkin, P. A. & Xie, P. *CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution*, *J. Hydromet.* **5**, 487-503 (2004).

[27] Dai, A. *Precipitation Characteristics in Eighteen Coupled Climate Models*, *J. Climate* **19**, 4605-4630 (2006).

[28] Guichard, F. et al. *Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models*, *Q. J. R. Meteorol. Soc.* **130**, 3139-3172 (2004).

[29] Zhang, Y. & Klein, S. A. *Mechanisms Affecting the Transition from Shallow to Deep Convection over Land: Inferences from Observations of the Diurnal Cycle Collected at the ARM Southern Great Plains Site*, *J. Atmos. Sci.* **67**, 2943-2959 (2010).

[30] McCrary, R. R. & Randall, D. A. *Great Plains Drought in Simulations of the Twentieth Century*, *J. Climate* **23**, 2178-2196 (2009).

Supplementary Information available online.

Acknowledgements This research was partly funded by the European Union (FP6) WATCH Integrated Project (contract 036946), the UK National Centre for Earth Observation, and the European Space Agency STSE Water Cycle Multi-mission Observation Strategy (WACMOS) project (ESRIN/Contract No. 22086/08/I-EC. We would like to thank Anton Beljaars, Sonia Seneviratne and Doug Parker amongst others for valuable discussions on this topic. We also thank the CMORPH, TRMM, PERSIANN and GPROF teams for the provision of their precipitation data.

Author Contributions CT and RJ conceived the study, CT performed the analysis and wrote the paper, RJ and WD provided expertise on soil moisture datasets, FG interpreted the convective

responses in models and observations and PH devised statistical tests. All authors discussed the results and edited the manuscript. The authors declare no competing financial interest.

Author Information Correspondence to cmt@ceh.ac.uk

Figure 1: **Preference for afternoon precipitation over soil moisture anomalies**. Percentiles of the observed variable $\delta_e = \text{mean}(\Delta S_e) - \text{mean}(\Delta S_c)$ for each 5°×5° box under a null assumption that no feedback exists. Null sampling distributions of δ values were estimated for each box by resampling without replacement from the combined set of event and non-event ΔS values. Low (high) percentiles indicate where rainfall maxima occur over locally dry (wet) soil more frequently than expected. Grey denotes 5°×5° cells containing fewer than 25 events. The map is based on a merging of two separate analyses using either ASCAT or AMSR-E soil moisture. For each 5°×5° cell, the relative quality of the two datasets is tested independently to determine which product is used (Supplementary Information Figs. S5 and S6). Insets: frequency histograms $F(\Delta S_c)$ of soil moisture difference in the global control sample (purple; %), and the difference $F(\Delta S_e) - F(\Delta S_c)$ between the histograms of the global event and global control samples (orange shading; %). The units of ΔS are m^3m^{-3} for AMSR-E and fractional saturation for ASCAT, and total number of events (n_e) noted.

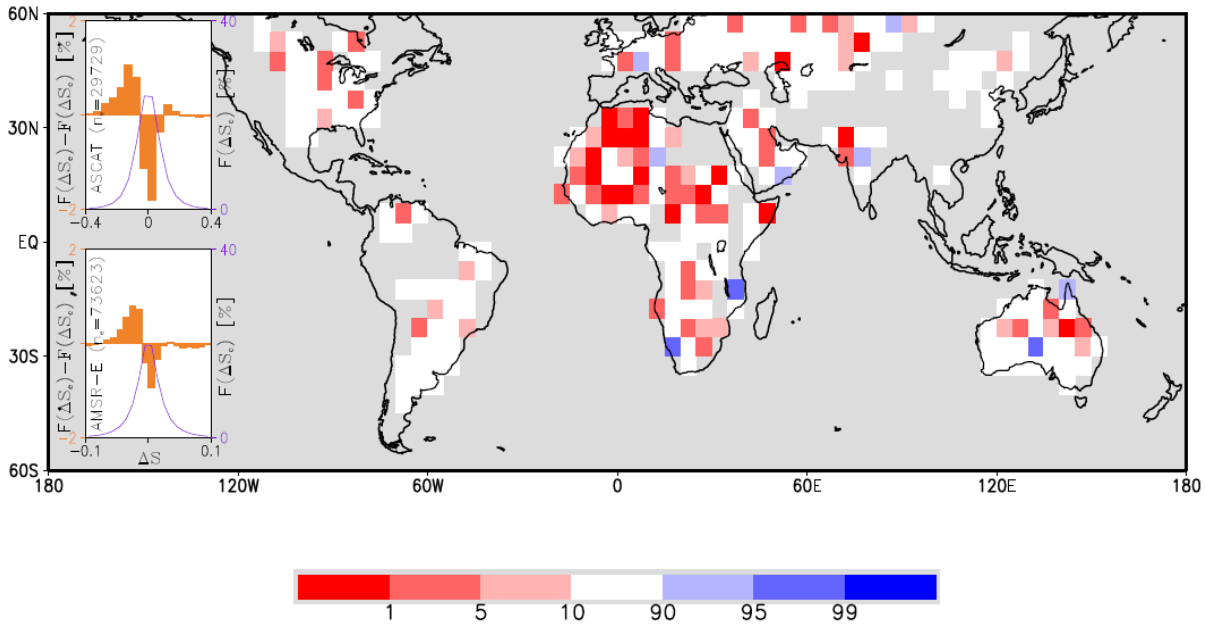

60N
30N
EQ
30S
60S
180
120W
60W
0
60E
120E
180
ASCAT (n=29729)
AMSR−E (n=73623)
F(ΔS_e)−F(ΔS_e) [%]
F(ΔS_e) [%]
ΔS
1
5
10
90
95
99

Figure 2: **Sensitivities to mean soil moisture and time of day**. The bars denote the anomalous pre-storm soil moisture difference δ_e averaged over every event globally, as a function of (a)pre-event soil moisture averaged over 1.25°×1.25° , and (b) time of first precipitation (at least 1mm over 3 hours) following a soil moisture measurement at 0130LT on day 1. A negative value of δ_e indicates a preference for precipitation over drier soil, and the error bars mark the 90% confidence limits. The number of events used is indicated by the triangles.

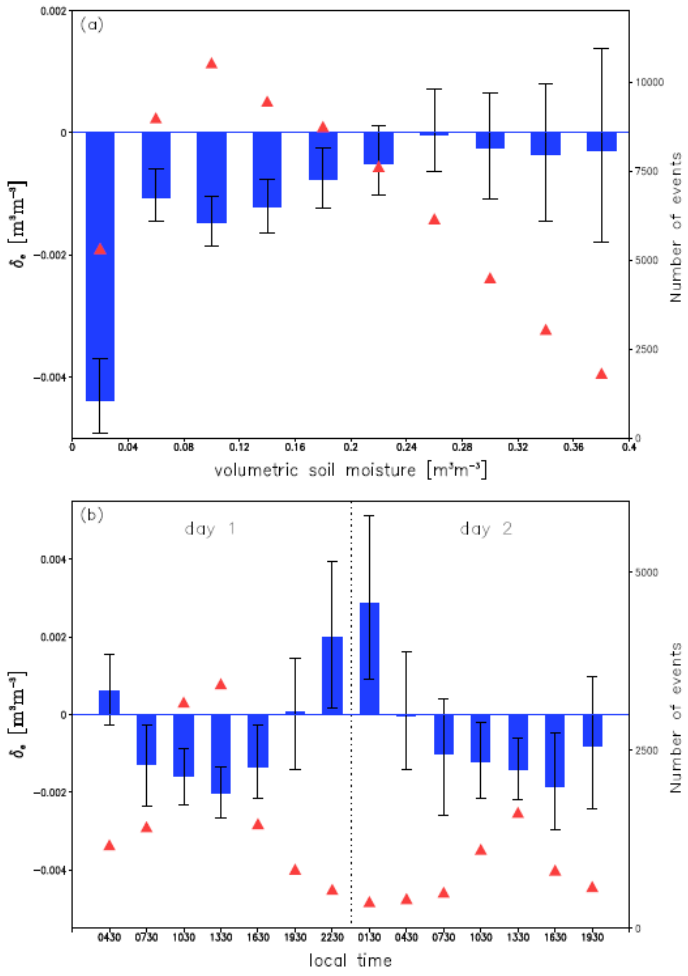
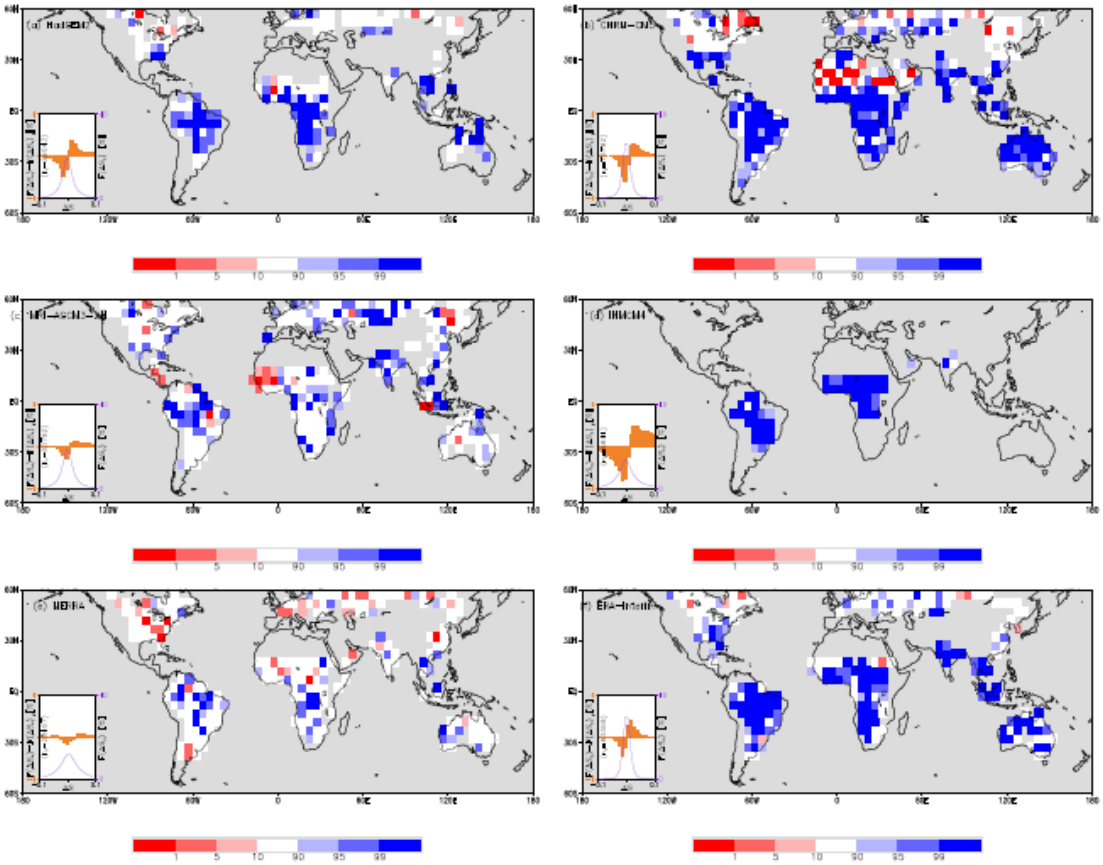


Figure 3: **Simulated preference for afternoon precipitation over soil moisture anomalies.** As for Figure 1 but using diagnostics from integrations by four climate models (a-d) and two atmospheric reanalysis models (e,f). Red (blue) shading indicates convective precipitation more likely over wetter (drier) soils. (a) HadGEM2, (b) CNRM-CM5, (c) MRI-AGCM3-2H, (d) INMCM4, (e) MERRA and (f) ERA-Interim. Maps of the number of events in each model are provided in Fig S11. Inset: as for Figure 1; ΔS in units of m^3m^{-3}. Further details on the models in Supplementary Information.



Afternoon rain more likely over drier soils: Supplementary Information

Christopher M. Taylor[a], Richard A. M. de Jeu[b], Françoise Guichard[c], Phil P. Harris[a], Wouter A. Dorigo[d]

[a] Centre for Ecology and Hydrology, Wallingford, U.K.

[b] Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[c] CNRM (CNRS and Météo-France), Toulouse, France

[d] Vienna University of Technology, Vienna, Austria

Observational Datasets

We use two different satellite soil moisture datasets, one derived from the Advanced Microwave Scanning Radiometer Earth Observing system (AMSR-E) on the Aqua satellite and another one from the Advanced Scatterometer (ASCAT) onboard the MetOp-A satellite. The AMSR-E soil moisture data set was derived from the Land Parameter Retrieval Model [*Owe et al.*, 2008]. This model was developed by VU University Amsterdam in collaboration with NASA and uses a radiative transfer model to convert the observed brightness temperatures into volumetric soil moisture values. The ASCAT soil moisture product was developed by Vienna University of Technology and is based on a change detection algorithm [*Wagner et al.*, 1999]. This product describes soil moisture in degrees of saturation. For both products, errors commonly remain below 4 volumetric percent (See Figure S2), although the ASCAT product is known to have difficulties over desert areas and both products become less accurate over dense vegetation [*Dorigo et al.*, 2010; *Parinussa et al.*, 2011].

The aim of the study was to examine the impact of spatial variability in soil moisture on spatial variability in afternoon rainfall. The approach was to build up statistics on the location of precipitation maxima relative to the underlying antecedent soil moisture. We defined the spatial extent of our analysis based on the need to minimise the impact of large-scale atmospheric variability on precipitation, and thus isolate the role of contrasting soil moisture conditions. This implies using as small a length scale as possible, depending on the resolution of the datasets. Whilst the AMSR-E soil moisture product is gridded at 0.25°, the footprint of the sensor is closer to 50km (~0.5° near the equator). Thus we chose a box size of 5x5 pixels, the minimum size for AMSR-E to properly resolve soil moisture contrasts between the centre of the box and its edge. Note that the same box size was employed for ASCAT, which has a higher nominal resolution of 25 km.

We employ 3 different precipitation datasets which use a combination of satellite data and, in some cases, surface raingauge data, to estimate precipitation at a resolution of 0.25° every 3 hours. Our primary precipitation dataset is the Climate Prediction Center Morphing Method (CMORPH) [*Joyce et al.*, 2004] which takes data from overpasses of satellites with passive microwave sensors onboard as a starting point, and are available from December 2002. These relatively high quality precipitation estimates are typically only available several times per day. The CMORPH product combines these estimates with thermal-infra-red data from geostationary satellite data, which are available on sub-hourly time scales over much of the world. The merging is done by using motion-vectors derived from geostationary data to propagate the microwave rainfall estimates. This provides a physical

basis for interpolating infrequent microwave estimates, though short-lived events may not be detected by CMORPH.

To test the sensitivity of our results to precipitation datasets we use 2 alternative sources. The first is the TRMM3B42 product from the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis [*Huffman et al.*, 2007]. This methodology uses the passive microwave data (mentioned above) when available for the 3-hourly estimates, and incorporating data from the active radar on the TRMM satellite. Infra-red data, calibrated at the monthly time scale with microwave overpasses, are used to fill in the gaps. Finally, the estimates are scaled to match monthly rain gauge observations, where they exist. At the time of analysis, TRMM3B42 data were available up to the end of June 2011. The third dataset we used, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [*Sorooshian et al.*, 2000] again takes the same input satellite datasets but uses a neural network approach to estimate precipitation. Note that unlike CMORPH and PERSIANN, the TRMM3B42 data are only available between 50°S and 50°N. The 3-hourly precipitation data were adjusted to local time based on linear interpolation according to longitude.

These precipitation datasets have been used for a wide variety of purposes, for example the examination of temporal variability of rainfall [*Ruane and Roads*, 2007]. The datasets are known to each have large uncertainities [*Tian and Peters-Lidard*, 2010], which tend to increase at higher latitudes, over complex terrain, coastlines and water bodies [*Tian and Peters-Lidard*, 2007]. However, in a study comparing the products with gauge-based analyses over the United States [*Tian et al.*, 2009], the contribution to total errors from false detection of events was relatively small, particular for CMORPH and TRMM3B42. This suggests that when we define a precipitation event (based on an afternoon accumulation exceeding 3mm), the probability that it really is raining is high, particularly having first excluded complex terrain, snow-covered surfaces and water bodies.

Model datasets

We used results from two meteorological reanalyses and four climate simulations (Table S1); all of them cover approximately the same period as the satellite observational products. ERA-Interim [*Dee et al.*, 2011] and MERRA [*Rienecker et al.*, 2011] are two recent reanalyses whose spatial resolution is finer than 1°. Their differences are mostly related to distinct assimilation procedures and sets of physical parametrizations. For ERA-Interim, we used both the 0000 UTC and 1200 UTC forecasts and excluded data in the first 9 hours of the forecast to minimize spin-up and drift issues. For MERRA, such considerations were not relevant as a correction term based on the analysis is continuously added to the forecast.

The climate simulations that we use were prepared by different research centres in the context of the Coupled Model Intercomparison Project Phase 5 (CMIP5; http://cmip-pcmdi.llnl.gov/cmip5/). We used outputs from AMIP-type simulations as they are forced by realistic oceanic boundary conditions for the contemporary period. Only climate models with a spatial grid finer than 2° x 2°, and with 3-hourly soil moisture and precipitation (convective and total) diagnostics available were analysed. Four models satisfied these criteria, listed in Table S1. When defining an event for the models, we used a much more restrictive threshold of 0.1 mm for the maximum total precipitation (convective and large-scale) allowed during the pre-event period, i.e. the 3 hour time slot between 0600 and 0859LT.

Model and modelling centre	Type of simulation	Spatial resolution zonal x meridional (°)	Period	Thickness of layer of soil moisture diagnosed [mm]
ERA-Interim, ECMWF	Reanalysis	0.703 x 0.703	2002-10	70
MERRA, NASA	Reanalysis	0.667 x 0.5	2002-10	20
HadGEM2-A, Met Office Hadley Centre	AMIP	1.875 x 1.25	1998-2009	100
CNRM-CM5, CNRM-CERFACS	AMIP	1.406 x 1.389	1999-2008	10
Inmcm4, Institute of Numerical Mathematics	AMIP	2.0 x 1.5	1989-2008	100
MRI-AGCM3-2H, MRI	AMIP	0.56 x 0.56	1999-2008	100

Table S1 Details of models used in analysis.

Additional quality control of soil moisture data

To limit the impact of noise in the soil moisture datasets, we applied a series of quality control tests. Firstly, we only used data for calendar months where at least 40 observations (20 for ASCAT) from a particular overpass (either am or pm) were available in the entire 9½ years (5 years for ASCAT) of the dataset. Those pixels containing more than 5% water were excluded. These were defined using the water body classification in the 1km GLC2000 dataset http://bioval.jrc.ec.europa.eu/products/glc2000/products.php. In addition, values of soil moisture exceeding 0.5 m^3m^{-3} were excluded from the AMSR-E dataset, which is known to be less accurate over wet soils [*Champagne et al.*, 2010]. Even when such high values are realistic, though rare, their removal should have little influence on this study, focussing as it does on the impact of soil water deficit on rainfall via evaporation. Retrievals from ASCAT and AMSR-E in regions of tropical forest and complex terrain were removed. With ASCAT, we imposed a maximum threshold in the associated error flag of 15%, and for AMSR-E we used a similar threshold, and also incorporated a vegetation density product [*Liu et al.*, 2011] to screen the data. A maximum vegetation optical depth threshold of 0.8 was used to mask the soil moisture values with high uncertainties [*Parinussa et al.*, 2011].

We developed a method to remove pixels of questionable quality from the soil moisture datasets, based on their consistency with the precipitation datasets. For each month of the year and each overpass (am or pm), we correlated the change in soil moisture between consecutive observations

with precipitation in the intervening period. In the case of the AMSR-E data, we first excluded cases where rainfall was present at overpass time from the Goddard Profiling Algorithm [GPROF; *Kummerow et al.*, 2001]. Only pixels with a positive linear correlation ($p<0.25$) were retained. This process removed previously unflagged pixels in months with frozen ground, wetlands and strong radio-frequency interference (RFI).

We wished to exclude regions where rainfall variability could arise from fixed features, notably topography. To this end, we removed pixels where the range of topographic height within a box of size 1.25° exceeded a threshold of 300m. We used the ETOPO1 [*Amante and Eakins*, 2009] topography dataset at 1 arc-minute (0.017°) resolution for this purpose. The same mask was used for the coarser resolution model outputs. We also excluded areas with strong climatological soil moisture gradients (greater than $0.1 m^3 m^{-3}$ per 0.25° for AMSR-E, or a saturation index greater than 0.2 per 0.25° for ASCAT) which could potentially influence the climatological rainfall field. The spatial analysis presented in Figures 1 and 3 was performed using all the events within a 5° box. This spatial scale was chosen as providing a compromise between maximising sample size, and retaining adequate spatial resolution of the output.

The January and July soil moisture climatologies for the daytime overpasses from AMSR-E and ASCAT are presented in Figure S1 (a-d). The Pearson regression coefficients determined from the correlation analysis described above are also presented (e-h) alongside the final masks used to identify precipitation events (i-l). Note that similar results were obtained with nocturnal overpasses.

Figure S1 Characteristics and filtering of soil moisture data (overleaf). Mean soil moisture (a-d), Pearson correlation coefficient (r; e-h) of soil moisture with precipitation, and pixel mask (i-l) are shown for January and July and for both AMSR-E soil moisture [$m^3 m^{-3}$] and ASCAT soil moisture index [ ] daytime overpasses. The means and correlation coefficients are only plotted where significance exceeds 75% according to a Student t-test. The pixel mask (i-l) indicates where pixels have been excluded for the month in question due to poor correlation with rainfall, topographic variability, water bodies or climatological soil moisture gradients. Only pixels shaded red are used in the subsequent analysis.

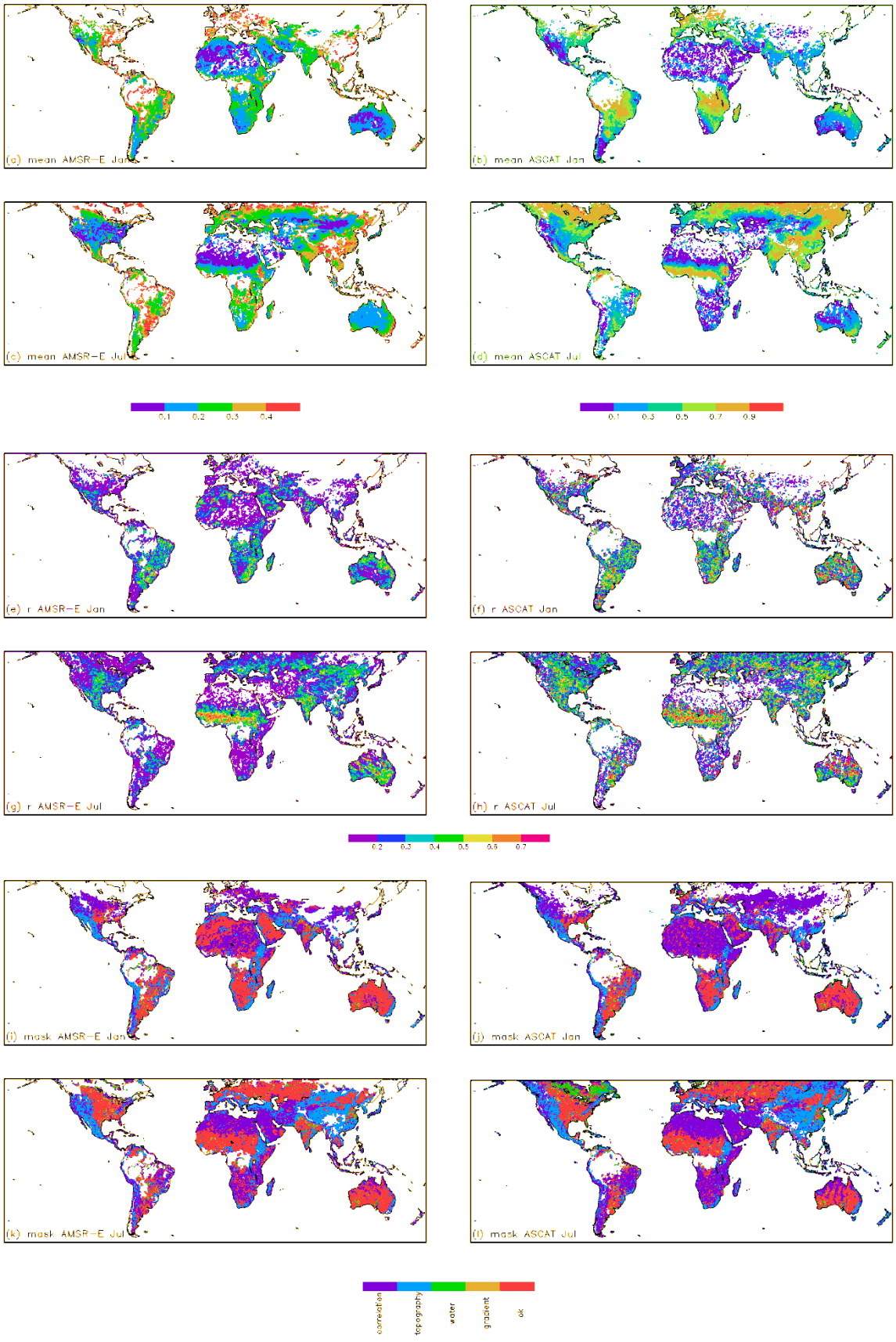

(a) mean AMSR−E Jan
(b) mean ASCAT Jan
(c) mean AMSR−E Jul
(d) mean ASCAT Jul
0.1 0.2 0.3 0.4
0.1 0.3 0.5 0.7 0.9
(e) r AMSR−E Jan
(f) r ASCAT Jan
(g) r AMSR−E Jul
(h) r ASCAT Jul
0.2 0.3 0.4 0.5 0.6 0.7
(i) mask AMSR−E Jan
(j) mask ASCAT Jan
(k) mask AMSR−E Jul
(l) mask ASCAT Jul
correlation
topography
water
gradient
ok

Figure S1

Soil Moisture Errors

Soil moisture errors exhibit strong spatial variability, mainly resulting from vegetation density and surface characteristics. For AMSR-E, in the regions of the world where a strong signal of rain on drier soil emerges at the 5° scale (e.g. the Sahel and Sahara, Central Australia, Mongolia; Figure S5a), random errors in soil moisture are relatively low, generally below 3% volumetric soil moisture. Similarly when using ASCAT data, regions with a strong signal on precipitation (Sudan, parts of western and southern Africa, central Asia; Fig S5b) exhibit relatively low errors, typically below 12% saturation. Note that these errors are representative of the annual cycle, whereas in the current study, we exclude months of the year containing poorer quality soil moisture data. Thus errors in soil moisture data used in our analysis will be lower than those depicted in Figure S2.

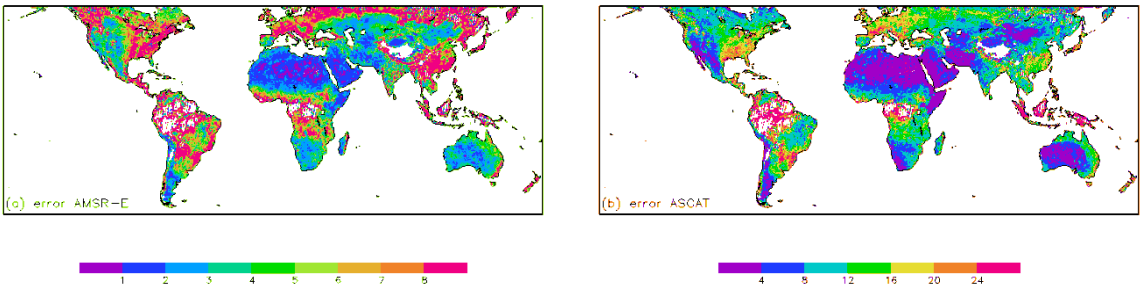
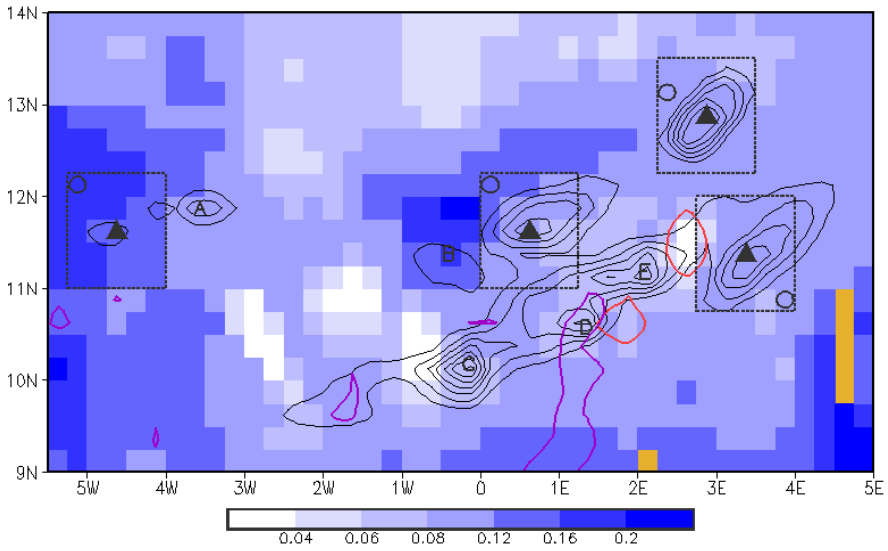


Figure S2 Geographical distribution of soil moisture errors. Mean random errors over the period 2007-2008, expressed as standard deviation of the absolute random deviations, based on Dorigo et al, [2010]. (a) AMSR-E [% volumetric soil moisture], (b) ASCAT [% saturation].

Example rain event

An example of the event detection method in action is provided in Figure S3, with explanation provided in the caption. It was chosen as it illustrates several aspects of the event detection technique in a single image. Note that there are often events with more than one L_{min}, when zero afternoon precipitation is detected over multiple pixels close to L_{max} , although none are shown in Figure S3. In such cases the mean soil moisture conditions across all L_{min} pixels are used. Also note that overlapping of event boxes on any day is not permitted, so we were able to interpret each event as effectively independent. The overall results were not affected in tests where events were excluded when they occurred within 3° of another event on the same day. Whilst not an issue in the image presented in Figure S3, events were excluded where less than 50% of the pixels within the surrounding 1.25°x1.25° box did not contain valid soil moisture data from the relevent overpass.

Figure S3 Detection of afternoon rain events on 28 June 2006 in West Africa. Black contours denote total precipitation (every 3mm, starting at 3mm) between 1200 and 2100 LT from CMORPH. Shading indicates soil moisture from AMSR-E [m^3 m^{-3}] detected from an overpass at 1330LT with missing pixels shown in yellow. Non-zero GPROF precipitation estimates at 0.25° resolution from the same overpass are shown by the red contour (0.1 mm hr^{-1}). Pixels containing topographic heights ranging by 300m or more are enclosed by the purple contour. In this image there are 4 events detected, which are then centred within 5x5 pixel boxes (dashed squares). The location of precipitation maxima (L_{max}) and minima (L_{min}) are shown by triangles and circles respectively. There are 5 other precipitation maxima in the image which are not defined as events, each denoted by a letter. Maximum “A” was excluded because of late morning precipitation which exceeded the threshold of 1 mm. No event was defined at “B” because its box overlapped another box enclosing a more intense precipitation maximum. The boxes surrounding maxima “C” and “D” both contain pixels where topographic variability exceeds the threshold of 300m whilst the boxes centred on maxima “D” and “E” include non-zero instantaneous precipitation rates from GPROF at the time of the soil moisture retrieval from the AQUA satellite.



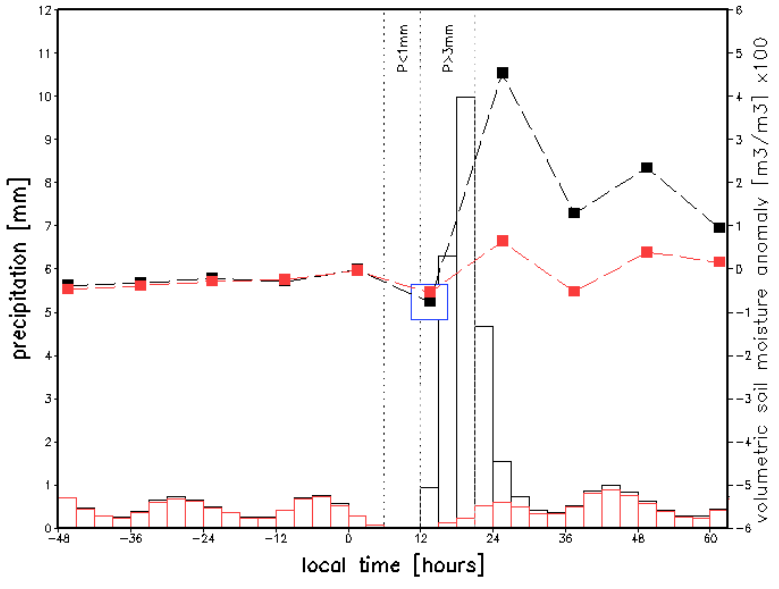
The precipitation datasets are provided in terms of total 3-hourly accumulations with no partition between stratiform and convective events. Our focus is on convective precipitation, and by focusing

on afternoon events without significant morning accumulations (within a region of 1.25°x1.25°), we aim to suppress the number of large-scale stratiform events sampled. Similarly we make no attempt to filter the events by season. However, the combination of our precipitation event detection and soil moisture quality control produces only a small number of events at mid-latitudes during the cold season (Figure S10).

Mean temporal evolution of soil moisture and precipitation about an event

To verify the mutual consistency of the rainfall and soil moisture datasets in our methodology, we computed composite time sequences around each of the 73623 rain events using CMORPH and AMSR-E data. Figure S4 illustrates the 3-hourly data which define an afternoon rain event (12-21 hours) at L_{max} and L_{min}, calculating ΔS_e from observations at 1330LT. By definition, precipitation is much higher during the afternoon over L_{max} compared to L_{min}. However, the subsequent evolution of soil moisture at the 2 locations is based on entirely independent observations. These show that soil moisture rises sharply at L_{max} but not L_{min}. The positive soil moisture anomaly created by the rain event is seen to decay in subsequent days, as expected. The composite also shows that for L_{max}, mean soil moisture has not risen relative to previous measurements, indicating that the values used to compute ΔS_e precede the rain itself. Similar results are obtained for events defined using the 0130LT soil moisture data. The composite sequences indicate that the soil moisture and precipitation datasets are fit for purpose, and that ΔS_e is a measure of pre-event conditions, uncontaminated by the event itself. Note also that whilst the definition of an event requires a minimum of only 3 mm precipitation between 12 and 21LT, on average, the total 12-hour accumulation at L_{max} exceeds 22mm. In a test with a doubled afternoon precipitation threshold (6 mm) qualitatively similar results to Figure 1 were produced, though based on fewer events.

Figure S4 Evolution of soil moisture and precipitation in the days either side of an event. Composite mean soil moisture (expressed as anomalies [m^3m^{-3}] from the pixel climatological mean; dashed lines) and precipitation [mm; bars] based on 34,946 events using soil moisture observations at 1330LT (blue box) to compute ΔS_e. The events are defined by precipitation at L_{max} exceeding 3mm between 12 and 21LT with less than 1 mm in the preceding 6 hours. Conditions at L_{max} are depicted in black, whilst those at L_{min} are shown in red.



Merging of results from AMSR-E and ASCAT

The 2 soil moisture datasets have different strengths and weaknesses [*Dorigo et al.*, 2010]. Broadly speaking, the passive microwave dataset performs best in sparsely-vegetated and desert areas, but can be sensitive to RFI in some regions, most notably the US. For intermediate vegetation cover, the ASCAT product is more reliable, though this too cannot provide information over dense forest cover. In Figure 1 we merged results from the 2 independent percentile maps, shown in Fig. S5.

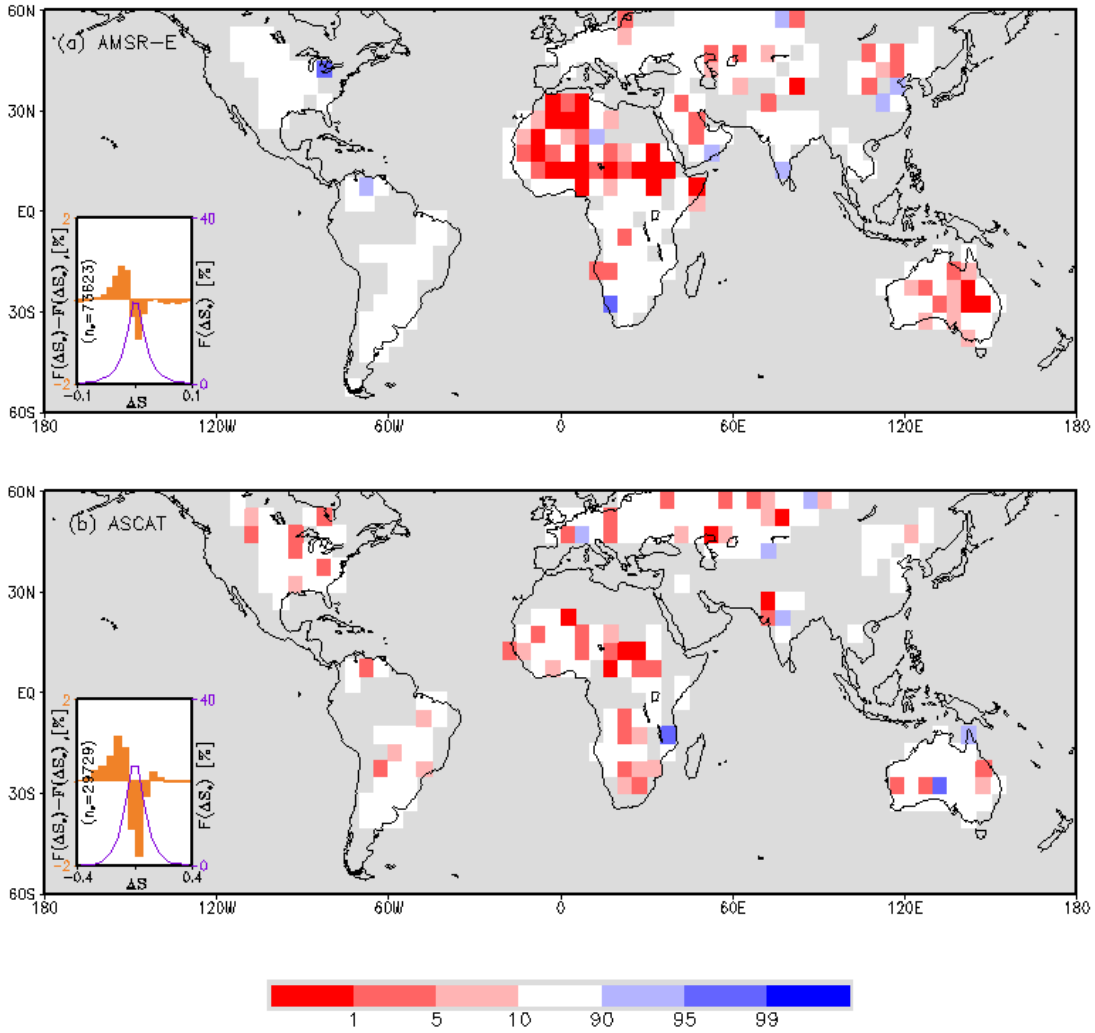


Figure S5 Preference for afternoon precipitation over soil moisture anomalies based on two soil moisture datasets. As for Figure 1 except that the percentile maps are based on single soil moisture datasets only (a) AMSR-E (b) ASCAT.

The merging for Figure 1 was based on a comparison of the Pearson correlation coeficients for the two datasets (e.g. Figure S1(e-h)) based on high frequency changes in soil moisture and precipitation. The comparison is depicted at the 5° scale in Figure S6. Blue areas indicate that the ASCAT data explains more of the variance in the precipitation than AMSR-E, and such areas predominate globally. As expected, AMSR-E outperforms ASCAT across semi-arid and arid zones. The overall pattern is remarkably consistent with a previous comparison of the 2 products based on an alternative methodology [*Dorigo et al.*, 2010]. The difference in quality between the datasets is consistent with the broad differences in locations of low percentiles in Figure S5. This implies that a

high quality soil moisture dataset is necessary to detect a signal in δ_e. For example across the US, where RFI compromises the quality of AMSR-E data, Figure S5a indicates no pixels with a percentile value below 10. The ASCAT data are unaffected by RFI and eight such pixels are evident in the US based on that dataset (Figure S5b). On the other hand over the Saharan, Australian, and Arabian deserts, AMSR-E outperforms ASCAT, and many more, low percentile pixels are found in these deserts in Figure S5a compared to Figure S5b.

We have used all available soil moisture data for the analysis in Figures 1 and S5. For AMSR-E, this means that for some of the events, soil moisture data is retrieved close in time to the event itself. We have used the GPROF dataset to flag out all rainfall maxima where non-zero rain is present at the time of the soil moisture retrieval (e.g. Figure S3). To further verify that our results are not sensitive to the use of daytime soil moisture data, we repeated the analysis using only 0130LT (AMSR-E) and 2130LT (ASCAT). This reduced the number of events substantially, but the results (not shown) indicate the same preference for events to occur over locally drier soil. This is consistent with the data presented in Figure 2b which show a similar diurnal signal on the second day after the the soil moisture measurement as the first.

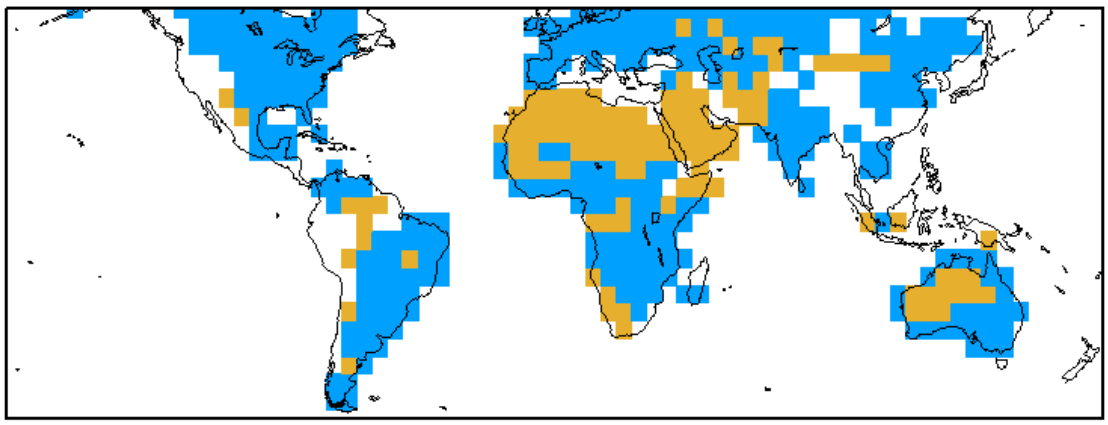


Figure S6 Map at the 5° scale comparing the annual sum of positive Pearson correlation coefficients between high frequency soil moisture fluctuations and precipitation using ASCAT and AMSR-E data. Monthly positive regressions not significant at the 75% level have been removed. Blue (gold) areas indicate where that ASCAT (AMSR-E) is superior to AMSR-E (ASCAT). This binary field determines whether percentiles from Figure S5a (AMSR-E) or S5b (ASCAT) were plotted in Figure 1. Note that the level of screening for dense vegetation between the two products differs, with a high threshold adopted for ASCAT. This mismatch results in larger apparent correlations in AMSR-E over tropical forests compared to (zero) correlation for the missing ASCAT data. This difference is evident in Figures S1 (i) and (k). In fact neither dataset provides accurate soil moisture retrievals beneath dense vegetation.

Comparison of results using different precipitation datasets

We repeated the correlation analysis between a single soil moisture dataset (AMSR-E) and the 3 alternative precipitation datasets. The differences in the annual means of all positive correlation coefficients are shown in Figure S7 and provide an indication of the quality of the precipitation datasets in different regions. Based on the correlations with high frequency soil moisture fluctuations, it appears that the CMORPH product (and to a lesser extent TRMM3B42), performs better across much of the mid-latitudes than PERSIANN. On the other hand, PERSIANN performs better by this criterion than TRMM3B42 over the humid tropics of South America and Africa. Similar results are obtained using ASCAT rather than AMSR-E.

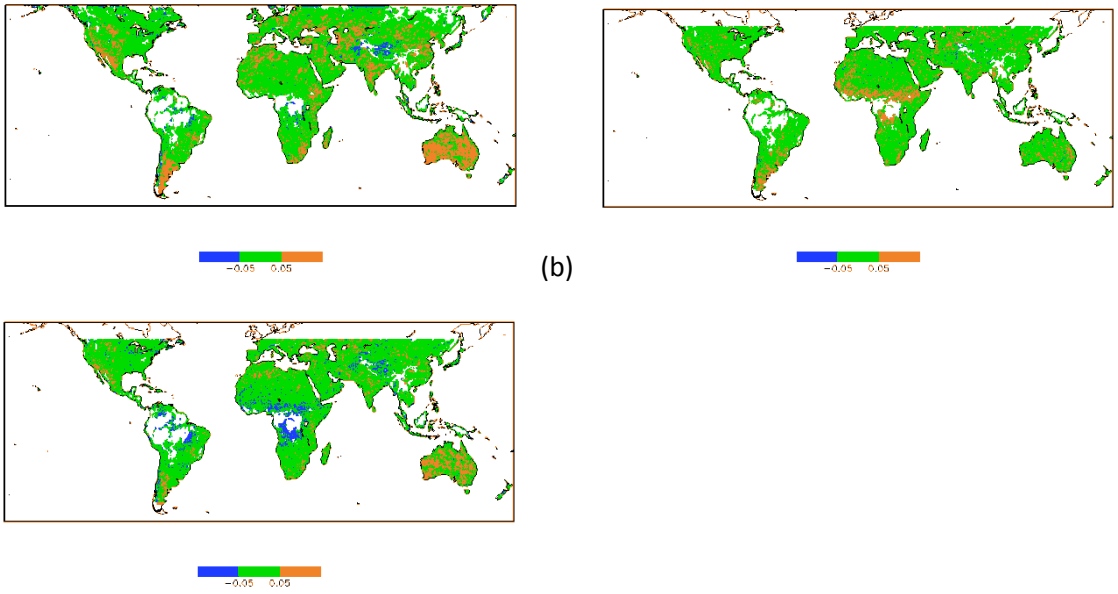
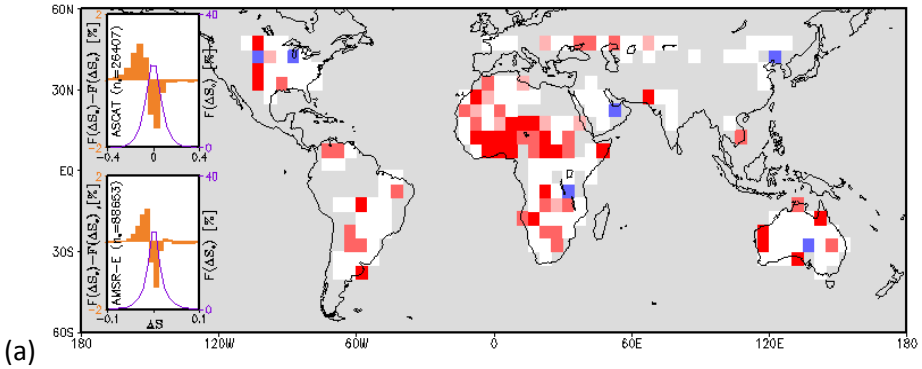


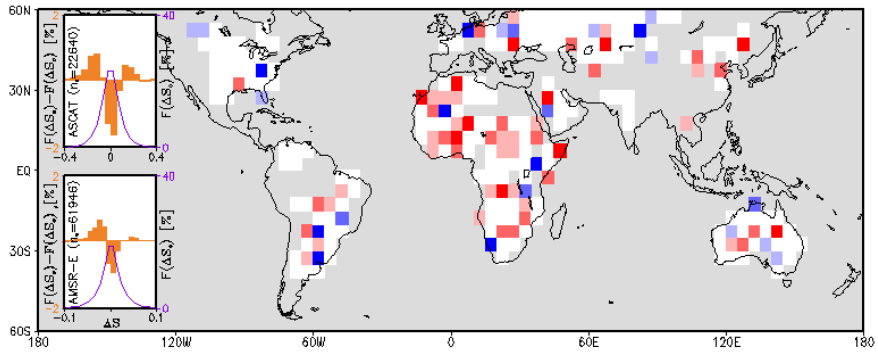
Figure S7 Differences in precipitation products in terms of their annual sum of positive Pearson correlation coefficients between high frequency soil moisture fluctuations and precipitation. (a) CMORPH-PERSIANN, (b) CMORPH-TRMM, and (c) TRMM-PERSIANN. The soil moisture data were taken from ascending (1330LT) overpasses of AMSR-E.

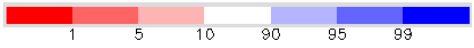
We repeated our comparison of pre-event soil moisture differences ΔS_e (Figure 1 main text), but using the 2 alternative precipitation datasets. The results (Figure S8) illustrate key features in common with the CMORPH analysis. Globally, precipitation is favoured over dry soils in all 3 datasets, with Africa and Australia providing the strongest signals. The percentage of non-grey 5x5° cells (i.e. with at least 25 events) with percentiles less than 10 is 31.8% for TRMM3B42 and 22.2% for PERSIANN. The percentage of cells with percentiles exceeding 90 is 2.9% for TRMM3B42 and 9.3% for PERSIANN. Further analysis of the sensitivity of the results to precipitation dataset is provided in Tables S3 and S4.

Figure S8 Preference for afternoon precipitation over soil moisture anomalies based on alternative precipitation datasets. As for Figure 1 in the main text but using (a) TRMM3B42 and (b) PERSIANN precipitation data instead of CMORPH.



(a)





(b)

Differences between Figure 1 and Figure S8 occur because each precipitation dataset, though using many common inputs, produce different numbers and locations of events, with different precipitation accumulations. Some summary statistics of the similarities between the 3 sets of events are provided in Table S2. The CMORPH and TRMM3B42 event datasets are more similar than either is to the PERSIANN dataset in terms of the number of events with the same location of the rainfall maximum L_{max}. However, only 1.3% of the events within the CMORPH and TRMM3B42 datasets have exactly the same location(s) of the rainfall minimum L_{min} as well as L_{max}. It is possible that the relatively weak signal emerging from the PERSIANN dataset in Figure S8 is due to weaknesses in the precipitation product, as quantified by the generally lower temporal correlations of that dataset with soil moisture fluctuations (Figure S7). Independent evidence to support this comes from the composite event analysis (as shown in Figure S4 for CMORPH) using events defined by PERSIANN. In this case (not shown), the difference in soil moisture between L_{max} and L_{min} at t=25.5 hours (that is 0130LT on the day following the event) is only 0.028 m^3m^{-3}, as compared to 0.038 m^3m^{-3} using CMORPH (Figure S4).

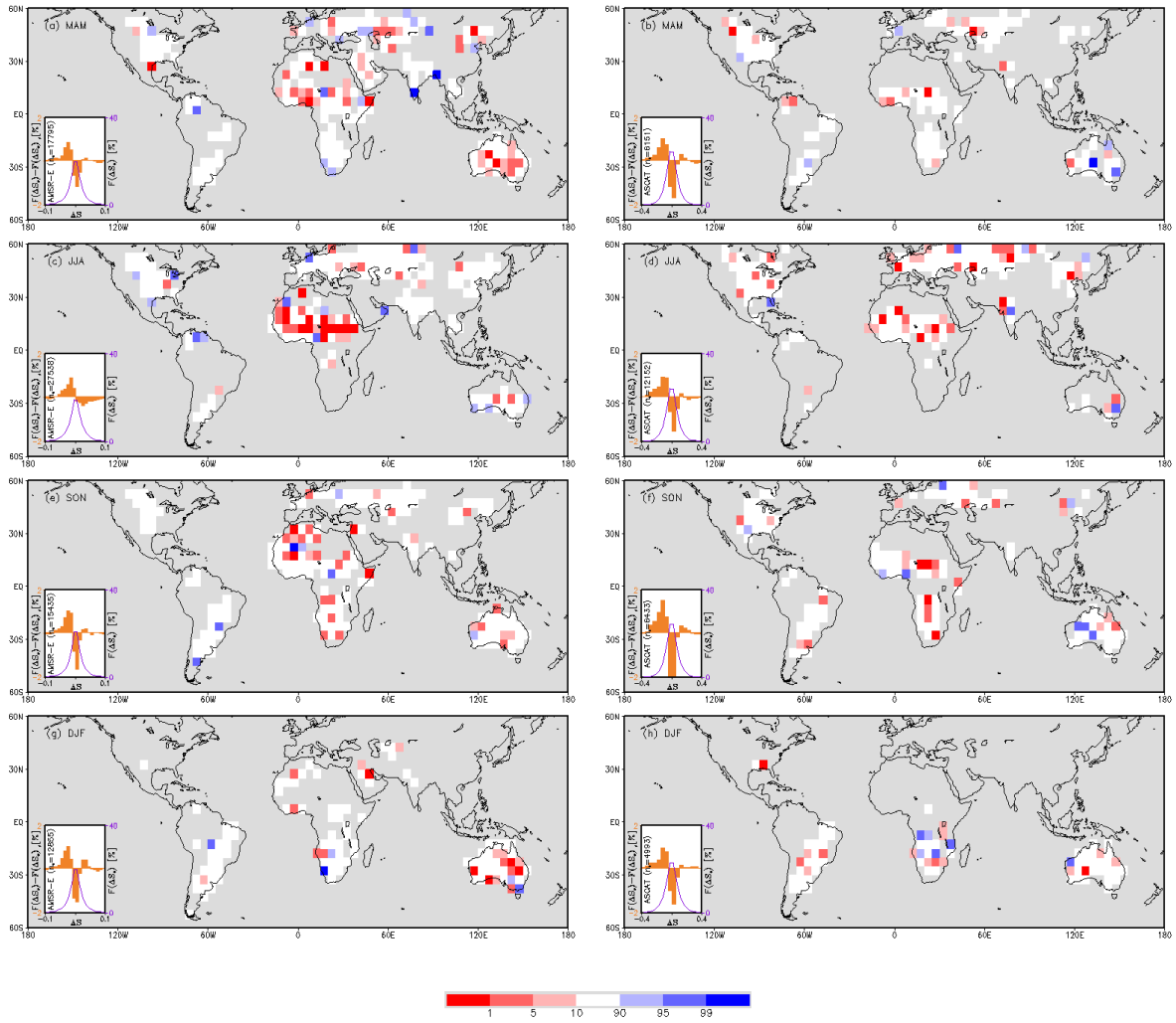
Table S2 Statistics of defined rain events in common between AMSR-E analysis using 3 alternative rainfall products. The comparison is only performed for events between 1st January 2003 and 4th October 2011 and between 50°S and 50°N.

Dataset	TRMM3B42				PERSIANN			
	Number of events	Number of events with common L_{max}	Number of events with common L_{max} and L_{min}	Number of common events within 5x5 box excluding those with common L_{max}	Number of events	Number of events with common L_{max}	Number of events with common L_{max} and L_{min}	Number of common events within 5x5 box excluding those with common L_{max}
CMORPH	64510	10077 (15.6%)	821 (1.3%)	15957 (24.7%)	64510	2020 (3.1%)	273 (0.4%)	13360 (20.7%)
PERSIANN	52055	2230 (4.3%)	174 (0.3%)	15610 (30.0%)	-	-	-	-

Seasonality of soil moisture feedback signal

The global analysis shown in Figures 1 and S5 is illustrated by season (rather than the full annual cycle) in Figure S9. As expected, the seasonal cycle of the signal tends to be linked to the seasonality of convective rainfall and soil moisture. When the surface is relatively dry and convective events are frequent, a preference for rain over drier soils tends to emerge.

Figure S9 Seasonality of preference for afternoon precipitation over soil moisture anomalies. As for Figure S5 but based on data from (a,b) March-May, (c,d) June-August, (e,f) September-October and (g,h) December-February. The analysis was performed with CMORPH precipitation and either AMSR-E (left-hand column) or ASCAT (right-hand column). Note that compared to Figure S5, a reduced threshold number of events of 15 is adopted here.



Sensitivity of results to spatial scale of input datasets

The impact of the spatial scale of the soil moisture and precipitation datasets on the analysis was assessed using either AMSR-E (Figure S10a) or ASCAT (Figure S10b) soil moisture in combination with CMORPH data. Soil moisture and precipitation datasets were degraded from 0.25 to 1.0° before processing. In addition, the size of the event-centred box was modified; rather than employ a box of 5x5 0.25° pixels, we used a box of 3x3 1.0° pixels, for better consistency with the analysis of the models in Figure 3. The precipitation thresholds used in the definition of an event were halved i.e. precipitation at L_{max} had to exceed 1.5mm and precipitation during the morning had to be less than 0.5mm. This analysis led to a substantial reduction in the total number of events; from 73,623 to 6,077 for AMSR-E, and from 29,729 to 3,411 for ASCAT. As a consequence, there is a fall in the number of 5x5° grid cells with at least 25 events from 295 to 78 for AMSR-E and from 220 to 53 for ASCAT. Nevertheless, the global preference for precipitation over dry soils remains (see histograms, and Tables S3 and S4), based on the relatively large number of events identified over North Africa and Australia.

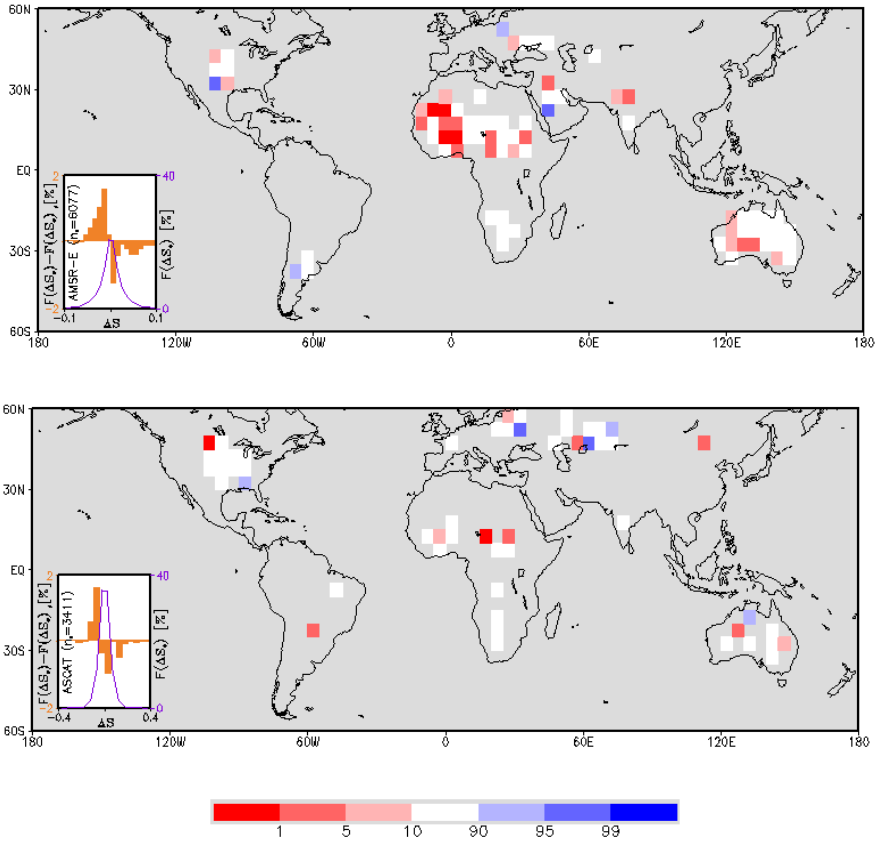


Figure S10 Preference for afternoon precipitation over soil moisture anomalies using coarse resolution data. As for Figure S5 except that the analysis was based on datasets with resolution of 1.0°.

Geographical distribution of events

There are large geographical differences in the number of defined events in each of the preceding analyses. These are presented in Figure S10. Areas of strong topographic variability are excluded from all the analyses (e.g. across the western US and Tibet). On the other hand, the criteria used to define precipitation events ensure that certain models in some regions (notably, INMCM4 and to a lesser extent, ERA-Interim) provide insufficient data to perform our statistical analysis. We adopted a minimum threshold of 25 events. Using a higher threshold reduces the number of 5° grid boxes available for analysis, but tends to increase the fraction of those grid boxes with extreme percentiles.

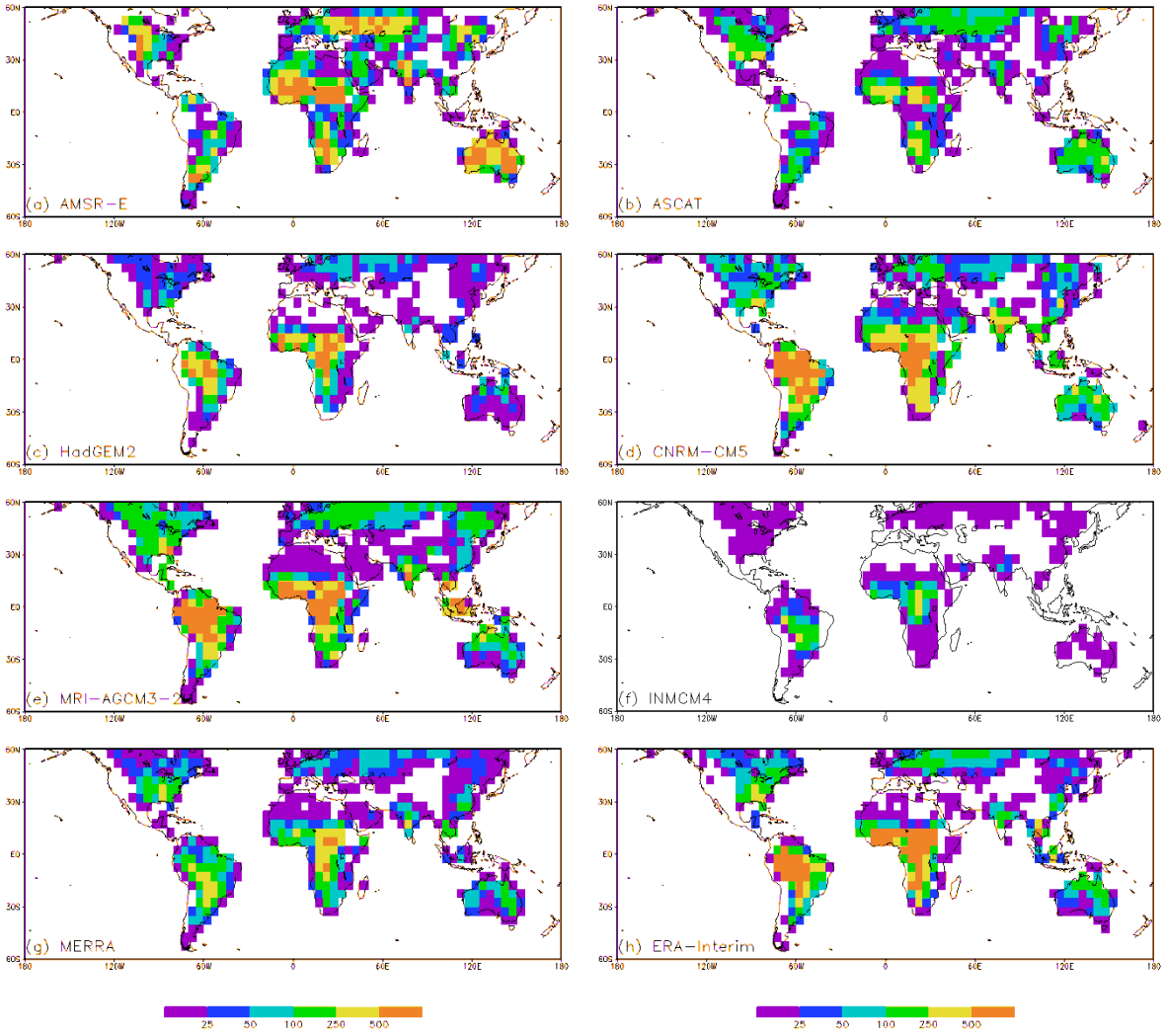


Figure S11 Number of afternoon rain events. Number of events for each 5x5° grid box used in calculations of percentiles in Figures 1 (observations; a and b) and 3 (models; c-h). Cells containing no events are shown in white.

Analysis of results by climatic region and continent

The calculation of percentiles of the observed variable δ_e = mean(ΔS_e) − mean(ΔS_c) under a null assumption that no feedback exists was repeated, but this time creating samples based on all events

within either climatic zones (Table S3) or continents (Table S4). This provided much larger sample sizes than were available for the calculation every 5° (see Figure S11). The results are presented for the different combinations of observed soil moisture and precipitation datasets at 0.25°, for the analysis using observations at 1° based on CMORPH data, and for the 6 atmospheric models. The climate zones were taken from the Köppen-Geiger classes, as computed by *Kottek et al.* [2006], and merged into 6 classes, depicted in Figure S12.

Table S3 Results of analysis (percentiles) by climatic zone. Extreme percentiles are shaded dark red (≤1), pale red (≤5), pale blue (≥95) and dark blue (≥99), and number of events in parentheses.

	Observations								Models					
	AMSR-E CMORPH	ASCAT CMORPH	AMSR-E TRMM	ASCAT TRMM	AMSR-E PERSIANN	ASCAT PERSIANN	AMSR-E CMORPH 1°	ASCAT CMORPH 1°	HadGEM2	CNRM-CM5	MRI-AGCM3_2H	INMCM4	MERRA	ERA-Interim
Moist tropics	89.5 (587)	19.8 (196)	93.5 (911)	18.1 (304)	27.2 (777)	18.1 (255)	-	-	>99 (7852)	>99 (20457)	>99 (28578)	>99 (1154)	>99 (3514)	>99 (25462)
Savanna	1.13 (12937)	<0.01 (6666)	<0.1 (17698)	<0.1 (8243)	<0.1 (12496)	1.1 (6248)	0.03 (632)	1.61 (451)	>99 (12021)	>99 (23777)	>99 (26422)	>99 (3256)	>99 (9612)	>99 (33595)
Semi-arid	<0.1 (23453)	<0.1 (7880)	<0.1 (33668)	<0.1 (8302)	1.1 (19539)	0.3 (5773)	0.39 (1872)	0.3 (821)	>99 (3438)	>99 (9517)	>99 (8284)	>99 (643)	86 (4467)	>99 (9761)
Arid	<0.1 (18810)	<0.1 (3752)	<0.1 (23450)	<0.1 (2926)	<0.1 (13952)	<0.1 (2216)	<0.01 (2613)	0.67 (620)	>99 (907)	<1 (6931)	84 (2957)	>99 (186)	97 (2401)	>99 (2496)
Temperate	58.1 (6224)	0.12 (4721)	13.1 (7917)	<0.1 (4858)	50.5 (5565)	54.5 (3844)	38.69 (474)	10.02 (455)	>99 (1844)	>99 (7151)	>99 (7174)	>99 (376)	1 (4811)	>99 (6092)
Continental	9.5 (11612)	<0.01 (6514)	78.4 (5009)	<0.1 (1774)	64.3 (9617)	49.0 (4304)	32.79 (485)	34.17 (485)	45 (2981)	61 (5949)	>99 (10377)	>99 (206)	<1 (3692)	>99 (5963)

The preference for rain over drier soil is a robust feature of all combinations of observational datasets in the savanna, semi-arid and arid zones (Table S3). The signal is generally weaker at mid-latitudes than over the tropics. At a resolution of 0.25° over mid-latitudes (Temperate and Continental zones), a signal for rain over drier soil only emerges above the noise when using ASCAT soil moisture data, in combination with either CMORPH or TRMM precipitation. The clearer mid-latitude signal using ASCAT rather than AMSR-E is consistent with both a better soil moisture sensitivity in the former over regions with intermediate vegetation, and the deleterious effects of RFI on the latter in certain regions. Furthermore, the lack of mid-latitude signal in Table S3 using PERSIANN precipitation is consistent with lower soil moisture correlations with antecedent precipitation at mid-latitudes for PERSIANN (Figure S7). The absence of a mid-latitude signal using data at 1° may be indicative of small sample sizes, in combination with an inherently weaker extra-tropical signal. The weaker signal in turn could be due to generally wetter soils (Figure 2a) and/or a higher contribution of frontal events, where rainfall is likely to be less sensitive to land surface

fluxes. On the other hand, the weakening of the mid-latitude signal in going from 0.25° to 1.0° data may indicate an important impact on convection of soil moisture variability on scales of 50 km or less.When analysing the data by climatic zone, the contrast between models and observations in the tropics is stark. All the models favour rain over wetter soil in the savanna regions whilst all combinations of observations indicate the opposite behaviour.

Figure S12 Merged Köppen-Geiger classes. Purple: Moist tropics (classes Af and Am), pale blue: savanna (As and Aw), yellow: semi-arid (BS), red: arid (BW), green: temperate (C), dark blue: continental (D).

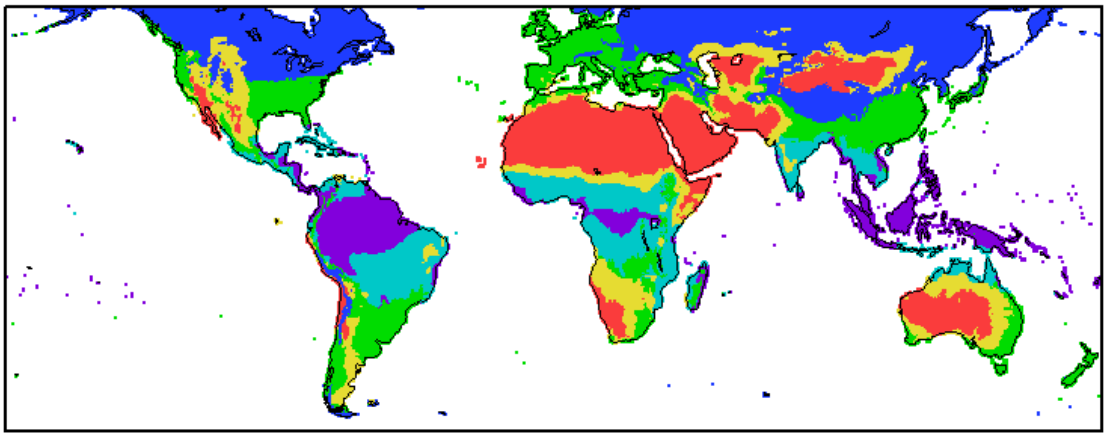


Table S4 Results of analysis (percentiles) by continent. Extreme percentiles are shaded dark red (≤1), pale red (≤5), pale blue (≥95) and dark blue (≥99), and number of events in parentheses.

	Observations								Models					
	AMSR-E CMORPH	ASCAT CMORPH	AMSR-E TRMM	ASCAT TRMM	AMSR-E PERSIANN	ASCAT PERSIANN	AMSR-E CMORPH 1°	ASCAT CMORPH 1°	HadGEM2	CNRM-CM5	MRI-AGCM3_2H	INMCM4	MERRA	ERA-Interim
South America	86.88 (4909)	0.26 (2956)	26.7 (7263)	<0.1 (3953)	43.0 (5304)	52.1 (2980)	3.74 (349)	0.32 (222)	>99 (8304)	>99 (25200)	>99 (28405)	>99 (2434)	>99 (7268)	>99 (261200)
North America	64.14 (5867)	0.24 (4535)	15.5 (6347)	0.1 (3595)	97.2 (5287)	88.8 (3455)	48.52 (1296)	4.26 (622)	88 (1623)	>99 (4721)	>99 (7461)	>99 (143)	<1 (3498)	>99 (4497)
Europe	8.83 (4374)	0.41 (2412)	33.8 (2451)	0.4 (736)	14.6 (3479)	26.2 (1536)	77.82 (474)	53.43 (336)	99 (1041)	>99 (1843)	>99 (2808)	53 (52)	<1 (937)	99 (1850)
Africa	<0.01 (29782)	<0.01 (10142)	<0.1 (43181)	<0.1)122806)	<0.1 (27506)	<0.1 (8575)	0.05 (1695)	0.04 (904)	>99 (14777)	>99 (29071)	>99 (29469)	>99 (3336)	>99 (10796)	>99 (426009)
Australasia	<0.01 (12837)	7.59 (4022)	<0.1 (17733)	0.1 (3510)	65.1 (7793)	3.5 (2309)	20.35 (860)	12.11 (485)	>99 (1075)	>99 (3788)	97 (2829)	>99 (87)	97 (1816)	>99 (2227)
Asia	0.2 (15854)	<0.01 (5662)	3.6 (11678)	8.8 (2327)	0.6 (12577)	17.9 (3785)	4.48 (1450)	27.7 (842)	>99 (2223)	>99 (9159)	>99 (12820)	>99 (343)	62 (4182)	>99 (6066)

Table S4 indicates that a preference for rain over drier soil is found in all continents, depending on the combination of observational datasets. The relative merits of the precipitation and soil moisture datasets discussed above provide some explanation of which combinations exhibit clearer signals. Of the models, only the MERRA dataset depicts a preference for rain over drier soil at the continental scale, and then only for Europe and North America.

Data Sources

The observational datasets used in this analysis are publically available and can be downloaded from the following:

AMSR-E soil moisture (version 2):
ftp://hydro1.sci.gsfc.nasa.gov/data/s4pa/WAOB/LPRM_AMSRE_SOILM2.002/

ASCAT soil moisture:

CMORPH 3-hourly precipitation: ftp://ftp.cpc.ncep.noaa.gov/precip/global_CMORPH/3-hourly_025deg/

TRMM3B42 3-hourly precipitation:
http://disc.sci.gsfc.nasa.gov/precipitation/documentation/TRMM_README/TRMM_3B42_readme.shtml

PERSIANN 3-hourly precipitation: http://chrs.web.uci.edu/persiann/data.html

GPROF instantaneous precipitation (version 4):
ftp://rain.atmos.colostate.edu/RAINMAP/data/amsre/

In addition we provide processed quality flags used for filtering soil moisture data for each month in the year (e.g. Figs S1(i)-(l)) as supplemental data.

References

Amante, C., and B. W. Eakins (2009), ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, 19 pp.
Champagne, C., A. Berg, J. Belanger, H. McNairn, and R. De Jeu (2010), Evaluation of soil moisture derived from passive microwave remote sensing over agricultural sites in Canada using ground-based soil moisture monitoring networks, *International Journal of Remote Sensing*, *31*(14), 3669-3690.
Dee, D. P., et al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553-597.
Dorigo, W. A., K. Scipal, R. M. Parinussa, Y. Y. Liu, W. Wagner, R. A. M. de Jeu, and V. Naeimi (2010), Error characterisation of global active and passive microwave soil moisture datasets, *Hydrol. Earth Syst. Sci.*, *14*(12), 2605-2616.
Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, *J. Hydromet.*, *8*(1), 38-55.
Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie (2004), CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution, *J. Hydromet.*, *5*(3), 487-503.

Kottek, M., J. Grieser, rgen, C. Beck, B. Rudolf, and F. Rubel (2006), World Map of the Koppen-Geiger climate classification updated, *Meteorologische Zeitschrift*, *15*(3), 259-263.
Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, D. B. Shin, and T. T. Wilheit (2001), The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors, *J. Appl. Meteorol.*, *40*(11), 1801-1820.
Liu, Y. Y., R. A. M. de Jeu, M. F. McCabe, J. P. Evans, and A. I. J. M. van Dijk (2011), Global long-term passive microwave satellite-based retrievals of vegetation optical depth, *Geophys. Res. Lett.*, *38*(18), L18402.
Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *Journal of Geophysical Research-Earth Surface*, *113*(F1).
Parinussa, R. M., A. G. C. A. Meesters, Y. Y. Liu, W. A. Dorigo, W. Wagner, and R. A. M. de Jeu (2011), Error Estimates for Near-Real-Time Satellite Soil Moisture as Derived From the Land Parameter Retrieval Model, *IEEE Transactions on Geoscience and Remote Sensing*, *8*, 779-783.
Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W. Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C. R. Redder, R. Reichle, F. R. Robertson, A. G. Ruddick, M. Sienkiewicz, and J. Woollen (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Climate 24*(14), 3624-3648.
Ruane, A. C., and J. O. Roads (2007), 6-hour to 1-year variance of five global precipitation sets, *Earth Interactions*, *11*.
Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite (2000), Evaluation of PERSIANN System Satellite–Based Estimates of Tropical Rainfall, *Bull. Am. Meteorol. Soc.*, *81*(9), 2035-2046.
Tian, Y., and C. D. Peters-Lidard (2007), Systematic anomalies over inland water bodies in satellite-based precipitation estimates, *Geophys. Res. Lett.*, *34*(14), L14403.
Tian, Y., C. D. Peters-Lidard, J. B. Eylander, R. J. Joyce, G. J. Huffman, R. F. Adler, K.-l. Hsu, F. J. Turk, M. Garcia, and J. Zeng (2009), Component analysis of errors in satellite-based precipitation estimates, *J. Geophys. Res.*, *114*(D24), D24101.
Tian, Y., and C. D. Peters-Lidard (2010), A global map of uncertainties in satellite-based precipitation measurements, *Geophys. Res. Lett.*, *37*(24), L24407.
Wagner, W., G. Lemoine, and H. Rott (1999), A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, *Remote Sens. Environ.*, *70*(2), 191-207.