

Response to Editor's Comments  
**Global Downscaling of Remotely-Sensed Soil Moisture using Neural Networks**  
**(Manuscript # hess-2017-680)**

Thank you for providing a detailed and pointed response to the reviewers' comments. Overall I am satisfied with the response however I wonder if more can be done to address reviewer #1's comment #1. E.g. can you not generate a 2.25 Km version directly from the 36 Km version and compare it with the version of 2.25 Km that is generated using 9km? I think the reviewer has raised an important point about the uncertainty in the downscaling relationship and it should be satisfactorily addressed.

***Response/Action:** We appreciate Editor's suggestion, and acknowledge the importance of the scaling assumption. However, as we mentioned in the response to referee #1's comment there is no way to generate a downscaled soil moisture estimate at 2.25km directly from the 36km product. The NN approach needs to be trained on a similar scaled product (in this case 36km to 9km), and then be used with the 9km product as input to generate the 2.25 km product. If we want to generate the 2.25km product directly from 36km product, we need a training datasets with 1/8 of scale between the spatial resolution of input and output which is not available.*

Response to Referee #1 Comments  
**Global Downscaling of Remotely-Sensed Soil Moisture using Neural Networks**  
**(Manuscript # hess-2017-680)**

*We thank Referee #1 for their positive comments. We respond to the general comments and specific points in the following.*

**General Comments:**

- The manuscript presents a new global soil moisture product provided at an unprecedented spatial resolution of 2.25 km. It is built from a neural network (NN) and data comprised of SMAP 36 km resolution level 3 soil moisture, an enhanced soil moisture product derived from 36 km SMAP observations and posted on a 9 km resolution grid (Chan et al. 2017) and MODIS NDVI data aggregated at various resolutions between 2.25 and 45 km. The authors have also tested the inclusion of a topographic index at the target downscaling resolution. The approach is evaluated by analyzing global soil moisture maps and by comparing downscaled soil moisture estimates against in situ data from the international soil moisture network (ISMN).

A global soil moisture product at 2.25 km resolution is of high interest to the hydrological and Earth system science community. I also find that the comparison of the NN method with simpler methods (linear interpolation, and the null-hypothesis i.e. no disaggregation) is quite positive as well. In fact, my comments mainly concern the underlying assumptions of the approach (comments #1 and 2) and the evaluation of the downscaled data set (#3).

- 1) On the use of the 9 km resolution soil moisture product. The basis for the proposed approach is to calibrate a relationship between 36 km (SPL3SMP) and 9 km (SPL3SMP\_E) resolution soil moisture products, and then to apply it at 9 km resolution to derive the 2.25 km soil moisture. The point is that the actual spatial resolution of SPM3SMP\_E (the so-called “9 km resolution product”) is 33 km while it is resampled at 9 km resolution (Chan et al. 2017). The 33 km resolution is so close to the original 36 km resolution SMAP level 3 data that one may wonder how a relationship derived from 36 and 33 km resolution data can be valid between 9 km and 2.25 km resolutions. At the very least, I recommend a sensitive analysis to assess the impact on the results of a coarser spatial resolution (33 km instead of 9 km) for training.

**Response/Action:** *We agree with the difference between native resolution and grid spacing of the SPM3SMP\_E product. Indeed, we had included this in the product description (Section 2.1) of the original submission. However, the methodology to generate gridded brightness temperature (TB) (and subsequently soil moisture) from original observations of the SMAP antenna is different in the 36km and 9km product. The 9km product applies the Backus-Gilbert (BG) optimal interpolation technique to the oversampled radiometer measurements to estimate TB at each point, which is the center of each 9km pixel. Meanwhile, the native*

*resolution of the data which is the spatial extent projected on earth surface by the 3-dB beamwidth of the radiometer is very close to the 36km product.*

*Contrary to this approach, the SMAP 36km product uses the antenna pattern information to estimate TB from the original radiometer measurements, and does not take into account the oversampling effect. This results in a coarser grid resolution data.*

*Unfortunately, we cannot conduct a sensitivity analysis since the main assumption is that the downscaling has the same scaling ratio during training and estimation (from 36km to 9km, and from 9km to 2.25km). Moreover, the enhanced soil moisture product is already on a 9km grid. If we want to upscale that to 33km it adds some uncertainty to the sensitivity analysis (there is no soil moisture product from SMAP on 33km grid).*

*In this study, we made the assumption that the scaling relationship from the 36km product to the 9km product can be used to estimate soil moisture at 2.25km. We have independently validated the soil moisture values at the 2.25km scale; therefore, we believe this is a reasonable assumption.*

*Unfortunately, we cannot test such relationship between the different scales (we would need indeed 2.25, 9 and 36km data and even more spatial dependence to test the scaling relationship across scales). This is currently impossible nonetheless but we hope that future high resolution mission could allow such test of the scaling relationship to inform such retrieval which could be adapted in the future to include some scale dependence.*

- 2) The NN is trained and run using NDVI data as auxiliary information about the sub-pixel soil moisture variability. Some limitations related to the soil moisture-NDVI relationship are mentioned in the conclusion (such as presence of vegetation, saturation effects). However, I think that the discussion should be deepened. It is true that NDVI and topography are variables available at global scale, but they are not the only factors explaining the soil moisture variability. In addition, the soil moisture-NDVI relationship established at the monthly time scale (phenological time scale) may not be valid at the daily time scale, at which SMAP observes the Earth and the observed surface soil moisture evolves. For clarity, the assumptions underlying the implementation of the NN using NDVI data should be better highlighted in the manuscript.

**Response/Action:** *We agree with the referee on the limitations of NDVI – SM relationship. The assumption we make in using NDVI as an auxiliary information to predict variability of soil moisture is that within a neighborhood of a specific pixel NDVI of that pixel ‘relative’ to the NDVI of the neighboring pixels is an indicator of the wetness or dryness of that pixel with respect to its neighboring pixels. This is different than using the ‘absolute’ value of NDVI for soil moisture prediction.*

*However, to clarify this point, we added the following paragraph in the conclusion section (P. 16, L. 3-7):*

*“In this study, we use the relative value of NDVI (in a given pixel with respect to the neighboring pixels) as an auxiliary information to predict spatial variability of*

*soil moisture in each coarse-scale pixel. While the relationship between soil moisture and NDVI at phenological time scales may not be valid at the temporal scale of SMAP observations (couple of days), our assumption builds on the relative value of NDVI within a small region and not the absolute value. Therefore, it is reasonable to use NDVI as a predictor in this case”*

- 3) Evaluation of the NN output: Line 22 page 9: "NN is appropriately explaining the spatial variability of soil moisture using NDVI as ancillary data". Line 14 page 13: "our evaluation shows that the downscaling algorithm has high accuracy in terms of temporal correlation, anomaly correlation and ubRMSE when compared to in situ soil moisture estimates from ISMN". It is difficult to assess the quality of the downscaled soil moisture at fine scale using global maps. Global maps convey the message that the high-resolution product is global, but some fine scale assessment is missing. Evaluation of the results over focus (perhaps instrumented) areas would be very useful. Regarding the temporal aspect, validation using 2-year averages does not allow for assessing the relevance of soil moisture-NDVI relationships at the temporal scale of SMAP observations/surface soil moisture dynamics. In addition, I do not think that the comparison with in situ measurements shows "that the downscaling algorithm has high accuracy". I would soften this point of view as results are very similar for all products (from 2.25 km to 36 km, see Figure 8). Even though the downscaling method does not degrade low resolution information, the improvement is hard to detect. The authors mention that "accuracy is better than or equal to the SMAP 9 km soil moisture estimates". I take them at their word, but from Figure 8 it seems that the original 36 km product be more accurate at several stations. More explanations are needed to clarify the improvement provided at 2.25 km and at which temporal scale.

***Response/Action:*** *We agree with referee’s comment that conducting evaluations over focus regions would be informative. We added this to the revised manuscript. The goal of the comparison with in-situ observations is to assess temporal accuracy of the downscaled product. Since the input data to the soil moisture estimates at 2.25km is the 9km product, we did not expect to get significantly higher temporal accuracy with respect to the 9km product. The value of the 2.25km product is an enhanced spatial resolution, while having the same temporal accuracy. We revised the statements in the results section highlighted by the referee to clarify this.*

### **Specific points:**

- a) I may have missed something but the two statements re-written below look contradictory: - Page 9, line 11: "Moreover, the latitudinal average plots (on the right side of each panel of Figure 5) show that at higher spatial resolutions there is more spatial heterogeneity. The latitudinal average for the 36 km product is much

smoother than the 2.25 km one." - page 9, line 18 : "For comparison, we also calculate CV for the 9km soil moisture estimates from SMAP at the 36km grid (Figure 6 bottom panel). The two panels in Figure 6 have different range of CV which is expected given the difference in their spatial scales." Since aggregation tends to reduce variabilities, one would expect an increase in the spatial variability at higher spatial resolution. However, the CV is divided by about 5 at 9 km resolution compared to the CV at 36 km resolution. Could the authors comment on their seemingly opposite findings ?

**Response/Action:** *Indeed, these two statements are not contradictory. The first one referring to Figure 5 is based on the fact that at higher spatial resolution we are seeing more variability. Mainly, the latitudinal plots capture the smaller changes in soil moisture that were not observable with the coarse resolution product.*

*On the other hand, the CV plots in Figure 6 show the variability of soil moisture within a 9km pixel and within a 36km pixel. Within a 9km pixel the variability of soil moisture is smaller than that of a 36km pixel, since the surface is less heterogeneous (this is true for a 9km pixel which is spatially within the 9km pixel). It is true that the 2.25km product should show more heterogeneity compared to the 9km product but this would be true if both 9km and 2.25km product variabilities are compared at the 36km scale. Currently, we are comparing variability of the 2.25km product within the 9km pixel, and variability of the 9km pixel within the 36km pixel. We will add the CV plot of the 2.25km product within each 36km pixel in the revised manuscript to better explain the spatial variability of soil moisture.*

- b) Line 13 page 2: "Some of them use linear relationships (i.e. projection) to define the impact of spatial heterogeneity using ancillary data, typically in combination with a radiative-transfer model to relate surface temperature and soil moisture (Colliander et al., 2017a; Merlin et al. 2005, 2008a, 2008b, 2008c)." I noted two errors in this sentence: 1) physical models that relate surface temperature and soil moisture are energy balance models (not radiative transfer models) and 2) the projection technique used in Merlin et al. 2005, 2008a does not implement linear relationships, but a non- linear energy balance model.

**Response/Action:** *We agree with the referee's comment, and revised this section to correctly summarize the work done in the literature.*

- c) Line 15 page 2: "A major issue is that surface temperature at finer spatial scales from satellites cannot be estimated under cloudy conditions". Agree and I would add another essential limitation that the surface temperature cannot be used as a signature of soil moisture in energy-limited conditions.

**Response/Action:** *We agree with this limitation, and added it to the revised manuscript.*

d) Line 21 page 4: "we assume that the scaling relationship between 36 and 9 km soil moisture estimates is the same as the scaling relationship between 9 km and 2.25 km resolution. To the best of our knowledge, this is the first time that the assumption of similar scaling relationship is used to downscale soil moisture". I would like to mention that the same scaling relationship has already been used to downscale soil moisture from 40 km to 1 km and from 1 km and 100 m in Merlin et al. 2009 and Merlin et al. 2013.

Merlin, O., Al Bitar, A., Walker, J. P., & Kerr, Y. (2009). A sequential model for disaggregating near-surface soil moisture observations using multi-resolution thermal sensors. *Remote Sensing of Environment*, 113(10), 2275-2284.

Merlin, O., Escorihuela, M. J., Mayoral, M. A., Hagolle, O., Al Bitar, A., & Kerr, Y. (2013). Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain. *Remote sensing of environment*, 130, 25-38.

***Response/Action:*** *We acknowledge new references provided by the referee. However, neither of the studies make the same assumption as we did between different scales of soil moisture. Both of the studies developed a sequential disaggregation in which the first step uses MODIS data to disaggregate soil moisture from SMOS resolution to MODIS resolution, next the MODIS-disaggregated soil moisture is further downscaled using ASTER and/or Landsat data. Unlike these two studies, we did not develop a sequential downscaling approach. The 9km product used in our study is estimated by different technique from direct brightness observations of the satellite. We find a relationship between the soil moisture estimates at 36km and 9km using some ancillary data, and then apply this relationship to the 9km product to estimate soil moisture at 2.25km. Those references have in any case been added to the list of references.*

Response to Referee #2 Comments  
**Global Downscaling of Remotely-Sensed Soil Moisture using Neural Networks**  
**(Manuscript # hess-2017-680)**

*We thank Referee #2 for their comments. Here, we respond to the general comments and specific points:*

**General Comments:**

The authors have downscaled global satellite-based soil moisture observations to 2.25km spatial resolution using neural networks. The manuscript is well written, and the methodology is sound. The high-resolution products will be useful for global hydro- logical and climate studies. However, there are a few issues in this manuscript, hence I suggest some major revisions.

My major concerns are:

1. The training and retrieval in the NN algorithm are based on the hypothesis that it shares the same "relationship" between inputs and outputs to downscale soil moisture from 36km to 9km and to downscale soil moisture from 9km to 2.25km. The author should provide more information about how the SPL3SMP\_E product is enhanced from the SPL3SMP product. As the author mentioned, the native resolution of SPL3SMP\_E is actually about 33km. So what is the relationship among 36km, 33km, and 9km?

***Response/Action:*** *In this study, we made the assumption that the scaling relationship from the 36km product to the 9km product can be used to estimate soil moisture at 2.25km. We have independently validated the soil moisture values at the 2.25km scale; therefore, we believe this is a valid assumption.*

*We agree with the difference between native resolution and grid spacing of the SPM3SMP\_E product. The 9km product applies the Backus-Gilbert (BG) optimal interpolation technique to the oversampled radiometer measurements to estimate brightness temperature (TB) at each point, which is the center of each 9km pixel. Meanwhile, the native resolution of the data which is the spatial extent projected on earth surface by the 3-dB beamwidth of the radiometer is very close to the 36km product.*

*Contrary to this approach, the SMAP 36km product uses the antenna pattern information to estimate TB from the original radiometer measurements, and does not take into account the oversampling effect. This results in a coarser grid resolution data.*

*Unfortunately, we cannot test such relationship between the different scales (we would need indeed 2.25, 9 and 36km data and even more spatial dependence to test the scaling relationship across scales). This is currently impossible nonetheless but we hope that future high resolution mission could allow such test of the scaling relationship to inform such retrieval which could be adapted in the future to include some scale dependence.*

2. It is a little confusing to add standard deviation of NDVI in the downscale scheme. All the schemes have included NDVI at the higher resolution (9 km for training and 2.25km for retrieval). What additional improvement will the standard deviation of NDVI provide? As the author stated, "This estimate provides a proxy of the heterogeneity within the coarse scale grid". Hasn't the higher-resolution NDVI already provided information about "sub-pixel heterogeneity"? This makes the conclusions somehow self-contradictory. The key of the downscaling algorithm is the higher-resolution NDVI "as an ancillary data to quantify sub-pixel heterogeneity of soil moisture". However, the similar performance of the four schemes suggests that "variability of NDVI within the coarse scale pixel does not provide additional useful information on the spatial heterogeneity of soil moisture for the downscaling".

**Response/Action:** *We would like to highlight that the final algorithm (scheme R1) which is used to generate soil moisture at the downscaled 2.25km scale does not include standard deviation of NDVI (NDVI\_std). The reason we included NDVI\_std in scheme R2 and R4 was to check if an explicit input of the variability of the NDVI within each coarse scale pixel (36km during training and 9km during estimation) would increase the accuracy of the soil moisture estimation at the finer scale. NN algorithms apply a combination of linear operations to the inputs, but not necessarily power operations that might inform the network of the variance/std of the inputs. So we made an assumption, added this variability as an input, tested the results and this hypothesis, and reported that it does not improve the accuracy of estimation. Therefore, to reduce complexity of the model we did not include NDVI\_std (or TI for the same reason) in the final algorithm. We believe reporting this analysis is informative for the readers who might be interested in expanding our approach and building new downscaling algorithms.*

### Specific points:

- P2, L29: "Our final product is soil moisture estimates at 2.25km spatial resolution with full global coverage every 2-3 days ... " Is it possible to explain why the final product is downscaled at 2.25km spatial resolution here?  
**Response/Action:** *We added the following sentence in this paragraph to explain the reason for choice of 2.25km:  
"The 2.25km spatial resolution is chosen since we use SMAP 36km and 9km (1/4 of 36km) soil moisture products for training, and developing the scaling relationship. This relationship is used to estimate soil moisture at the finer 2.25km resolution which is 1/4 of 9km."*
- P4, L8: What are the temporal resolution of the in situ soil moisture observations?

**Response/Action:** They are hourly measurements, but we have only used the measurement at 6am local time which is concurrent with SMAP measurement.

- P4, L13: 20 measurements at one station?

**Response/Action:** We apologize for the confusion, we meant 20 measurements in time in each station. We have excluded any station that has less than 20 measurements during the study period because temporal statistics won't be meaningful. We changed the sentence in the revised manuscript.

- P4, L14: What is the resolution of the "pixel" here? Does one pixel only contain one station?

**Response/Action:** We apologize for the mistake. We mean stations not pixels. The numbers for each network refer to the number of stations used from each. We corrected this in the revised manuscript.

- P7, Table 1: Are they inputs only for training? It would be better to include the information about retrieval.

**Response/Action:** The information for training and retrieval are the same but their spatial resolution is different. We changed the table in the revised manuscript to reflect the specifications of the input data for retrieval.

- P9, L8: The maps look pretty similar, can you tell any difference between them, or which one is better?

**Response/Action:** We expect them to be similar at this scale. However, there are differences in regions such as the Sahara and regions with high soil moisture gradient (near Congo basin, and the Amazon). Other differences are in the latitudinal averages and pixels that do not have soil moisture estimate (such as major rivers). Due to the suggestion of referee 1, we included soil moisture maps from focus regions in the revised manuscript that depict differences between the products. (Figure 6)

- P9, L12-13: ". . . at higher spatial resolutions there is more spatial heterogeneity. The latitudinal average for the 36 km product is much smoother than the 2.25 km one." What is the purpose of this? Does more spatial heterogeneity indicate better quality?

**Response/Action:** It is a measure of quality relative to the spatial resolution. Soil moisture at higher spatial resolution is more informative for applications at small scale. This comparison shows that the soil moisture estimate at 2.25km has indeed more variability, and it not an interpolated version of the coarser resolution product.

- P11, Figure 7: Any explanation for the dry bias over the arid area?

**Response/Action:** We had noted this in the conclusion section of the original manuscript (P13, L19-22):

"Use of NDVI as an ancillary measurement to disaggregate soil moisture builds on the assumption that there is a moderate vegetation cover in the pixel of interest. Therefore, this lowers the quality of the downscaling algorithm in bare soil or

*sparsely vegetated regions. Moreover, NDVI estimates tend to saturate in highly dense vegetated regions such as the tropical forests which results in limited ability of the downscaling algorithm to resolve the heterogeneity of the soil moisture in those regions.”*

*We will revised this paragraph as following:*

*“Use of NDVI as an ancillary measurement to disaggregate soil moisture builds on the assumption that there is a moderate vegetation cover in the pixel of interest. Therefore, this lowers the quality of the downscaling algorithm in bare soil or sparsely vegetated regions. Figure 7 shows a dry bias in arid regions which is an indication of lack of NDVI – soil moisture relationship. Moreover, NDVI estimates tend to saturate in highly dense vegetated regions such as the tropical forests which results in limited ability of the downscaling algorithm to resolve the heterogeneity of the soil moisture in those regions.”*

- P12, L2: How are the metrics calculated? calculate correlation temporally, then average among the pixels/stations? About the error bars on Figure 8, are they standard deviation among pixels or anything else?

**Response/Action:** *For each pixel that has in situ observations one set of metrics is calculated. If more than one station from a network fall within a pixel, their observations are averaged before calculating the metrics. Finally, metrics from each station are summarized in Figure 8 with mean and standard deviation among pixels.*

- P12, L4: For Figure S6, It would be better to show the 10 networks with different colors or symbols on the map... Also, this map should have been mentioned in section 2.4.

**Response/Action:** *We changed the figure as requested and referenced it in Section 2.4.*

- P12, Figure 8: It is really hard to tell from the figure if there is any better performance of the downscaled products than the 9km SMAP product.

**Response/Action:** *This figure summarizes the performance of different products with respect to temporal dynamics metrics. Since the downscaled product is derived from the 9km product, we do not expect to see significant improvement in temporal dynamics of soil moisture estimates. Rather we check to verify that the increased spatial resolution in the downscaled product does not decrease temporal performance. We emphasized this in the revised manuscript.*

- P12, Figure 8: Any idea why there is poor performance over some network (e.g., SNOTEL) but good performance over some other networks (e.g. SOILSCAPE). Why does SMAP 36km have the best agreement with iRON?

**Response/Action:** *While this assessment is beyond the scope of this manuscript, we would like to note that while iRON has the best agreement with SMAP data*

*with respect to ubRMSE it does not have the best correlation. SOILSCAPE has the best correlation, but it has an average ubRMSE metric. More detailed assessment of SMAP retrievals across different ground truth stations is provided in:*

*Chan, S. K. et al. (2016) 'Assessment of the SMAP Passive Soil Moisture Product', IEEE Transactions on Geoscience and Remote Sensing, 54(8), pp. 4994–5007. doi: 10.1109/TGRS.2016.2561938.*

- P13, L7-9: This conclusion might only be true among the different NN downscaling schemes.

**Response/Action:** *We agree with referee, and changed this statement as following in the revised manuscript:*

*“Our investigation shows that topographic index and variability of NDVI within the coarse scale pixel do not provide additional useful information on the spatial heterogeneity of soil moisture for the downscaling using our proposed NN technique and are thus omitted in the final product.”*

- P13, L20: "... this lowers the quality of the downscaling algorithm in bare soil or sparsely vegetated." Do any of the results support this statement? According to Figure 4, there is very high correlation coefficient over those regions.

**Response/Action:** *Lower quality is with respect to both correlation and bias. Figure 4 shows that those regions have high correlation, but Figure 7 shows the same regions have a negative bias in soil moisture estimates. We revised the sentence as following as also explained in response to an earlier comment:*

*“Therefore, this lowers the quality of the downscaling algorithm in bare soil or sparsely vegetated regions. Figure 7 shows a dry bias in arid regions which is an indication of lack of NDVI – soil moisture relationship..”*

# Global Downscaling of Remotely-Sensed Soil Moisture using Neural Networks

Seyed Hamed Alemohammad<sup>1,2</sup>, Jana Kolassa<sup>3,4</sup>, Catherine Prigent<sup>1,2,5</sup>, Filipe Aires<sup>1,2,5</sup>, Pierre Gentine<sup>1,2,6</sup>

<sup>1</sup>Department of Earth and Environmental Engineering, Columbia University

<sup>2</sup>Columbia Water Center, Columbia University

<sup>3</sup>Universities Space Research Association, Columbia, MD

<sup>4</sup>Global Modelling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD

<sup>5</sup>Observatoire de Paris

<sup>6</sup>Earth Institute, Columbia University

Correspondence to: Pierre Gentine (pg2328@columbia.edu)

**Abstract.** Characterizing soil moisture at spatio-temporal scales relevant to land surface processes (i.e. of the order of a kilometer) is necessary in order to quantify its role in regional feedbacks between land surface and the atmospheric boundary layer. Moreover, several applications such as agricultural management can benefit from soil moisture information at fine spatial scales. Soil moisture estimates from current satellite missions have a reasonably good temporal revisit over the globe (2-3 days repeat time); however, their finest spatial resolution is 9km. NASA's Soil Moisture Active Passive (SMAP) satellite estimates soil moisture at two different spatial scales of 36km and 9km since April 2015. In this study, we develop a neural networks-based downscaling algorithm using SMAP observations and disaggregate soil moisture to 2.25km spatial resolution. Our approach uses mean monthly Normalized Differenced Vegetation Index (NDVI) as an ancillary data to quantify sub-pixel heterogeneity of soil moisture. Evaluation of the downscaled soil moisture estimates against in situ observations shows that their accuracy is better than or equal to the SMAP 9km soil moisture estimates.

## 1 Introduction

Soil moisture is a key variable constraining the fluxes between the land surface and atmosphere boundary, and therefore it plays a key role in regulating the feedbacks between the terrestrial water, carbon and energy cycles (Berg et al., 2014; McColl et al., 2017; Seneviratne et al., 2010). Soil moisture partitions the surface energy between latent heat and sensible heat fluxes (Entekhabi et al., 1996; Gentine et al., 2007, 2011; Koster et al., 2010). Moreover, plant photosynthesis is regulated by the water available to them through their roots along with atmospheric conditions (Rodríguez-Iturbe and Porporato, 2007; Seneviratne et al., 2010; Volk et al., 2000). Finally, soil moisture regulates the surface water fluxes including infiltration and surface runoff generation (Salvucci, 1993; Salvucci and Entekhabi, 1994; Sun et al., 2011). Therefore, there is a need to characterize soil moisture at spatial scales relevant to the representation of land-surface and mesoscale processes in the atmosphere. This can potentially improve representation of evapotranspiration, runoff and precipitation in hydrologic and weather prediction models and result in improved predictive skills (Gedney and Cox, 2003). In addition, knowledge of soil moisture at fine spatial scales is necessary to improve farming practices and optimizing irrigation scheduling.

Soil moisture spatial variability is regulated by several factors including but not limited to precipitation, soil texture, surface vegetation and topography. A combination of these factors results in high spatial heterogeneity for soil moisture (Blöschl and Sivapalan, 1995; Famiglietti et al., 2008; Manfreda et al., 2007; Peng et al., 2017; Western and Blöschl, 1999).

Current global soil moisture estimates are available at coarse spatial scales (between 9 and 40 km) which limits their suitability for applications such as evapotranspiration modeling, particularly in regions with high atmospheric convection as well as farm management. Therefore, several approaches have been introduced to downscale soil moisture retrievals at finer resolution (~1km) (Hatfield, 2001; Mascaro et al., 2011; Merlin et al., 2006, 2008c; Peng et al., 2017; Piles et al., 2011; Srivastava et al., 2013). Here, our objective is to generate a global downscaling product using Soil Moisture Active Passive (SMAP) measurements (Chan et al., 2016; Colliander et al., 2017b; Entekhabi et al., 2010).

Soil moisture downscaling can be conducted through different strategies. The first one is to use the synergy of active and passive observations to take advantage of highly accurate but coarse resolution passive observations and active measurements available at higher spatial resolution but not as directly related to soil moisture as passive microwave measurements (Das et al., 2011; Jagdhuber et al., 2015, 2016; Leroux et al., 2016; Montzka et al., 2016; Njoku et al., 2002; Piles et al., 2009; Wu et al., 2017). However, due to the failure of SMAP active instrument in July 2015 these approaches are not directly applicable to SMAP.

Another approach to downscaling is to use ancillary data in combination with coarse scale soil moisture estimates to describe the spatial heterogeneity of soil moisture. Several studies have used this approach to develop soil moisture at fine spatial scales (Chakrabarti et al., 2016, 2017; Mascaro et al., 2011; Piles et al., 2011, 2014; Srivastava et al., 2013; Verhoest et al., 2015).

However, there are limitations in these techniques. Some of them use a linear relationship or non-linear relationship (i.e. a projection) to define the impact of spatial heterogeneity using ancillary data, typically in combination with a radiative transfer energy balance model to relate surface temperature and soil moisture (Colliander et al., 2017a; Merlin et al., 2005, 2008a, 2008b, 2008c). A major issue is that surface temperature at finer spatial scales from satellites cannot be estimated under cloudy conditions. (Colliander et al., 2017a; Merlin et al., 2005, 2008c, 2008a, 2008b). A major issue is that surface temperature at finer spatial scales from satellites cannot be estimated under cloudy conditions. Moreover, surface temperature cannot be used as a signature of soil moisture in energy-limited conditions. The other group of downscaling methods use complex and computationally intensive disaggregation algorithms that are typically unsuitable for global applications. Another group of methods use physical land surface models at fine spatial scales as an information useful for downscaling which adds significant uncertainty to the downscaling due to the errors of the model parameter estimates (Ines et al., 2013; Merlin et al., 2006; Roerink et al., 2000; Shin and Mohanty, 2013). Merlin et al., 2009 and Merlin et al., 2013 also developed a sequential technique to use satellite observations at different scales and sequentially downscale soil moisture observations from a 40km scale to 1km scale and then 100m. Peng et al. (2017) provide a more detailed review of the current status of soil moisture downscaling algorithms and their advantages and disadvantages.

Our approach in this study is to use Neural Networks (NN) to develop a downscaling algorithm for SMAP soil moisture estimates at the global scale. The capability of NNs in learning complex relationships between inputs and target data as well as their quick run-time after training are among the reasons for their popularity in Earth science and remote sensing problems. In recent years, NN has been used to develop soil moisture retrieval algorithms from either passive or active instruments or combination of both (Aires et al., 2012; Alemohammad et al., 2017a; Jiménez et al., 2013; Kolassa et al., 2013, 2016, 2017a, 2017b; Rodríguez-Fernández et al., 2017; Rodríguez-Fernández et al., 2015) (Aires et al., 2012; Alemohammad et al., 2017a; Jiménez et al., 2013; Kolassa et al., 2013, 2016, 2017, 2018; Rodríguez-Fernández et al., 2017; Rodríguez-Fernández et al., 2015). They have also been used to retrieve surface turbulent fluxes from remote sensing observations (Alemohammad et al., 2017b; Jiménez et al., 2009).

Our final product is soil moisture estimates at 2.25km spatial resolution with full global coverage every 2-3 days as dictated by the SMAP orbit from April 2015 until end of March 2017 (the first two years of SMAP mission). The 2.25km spatial resolution is chosen since we use SMAP 36km and 9km (1/4 of 36km) soil moisture products for training, and developing the scaling relationship.

Field Code Changed

This relationship is used to estimate soil moisture at the finer 2.25km resolution which is  $\frac{1}{4}$  of 9km. We explicitly emphasize that we assume that the scaling between 36km and 9km is similar to between 9km and 2.25km. Unfortunately, there is no current way to systematically test a potential scale-dependence of such scaling but the systematic comparisons with observations are used as an indirect validation of this assumption. Moreover, we use ancillary data that are available at global scale at all times (monthly NDVI and topographic index) to mitigate issues with current downscaling algorithms.

## 2 Datasets

In order to perform the downscaling algorithm, we use three remote sensing based observations: (1) SMAP estimated coarse-resolution soil moisture (Section 2.1), (2) Normalized Difference Vegetation Index (NDVI) as a vegetation index that is strongly correlated with soil moisture spatial patterns (Section 2.2) and (3) a topographic index to account for the location of water in the landscape (Section 2.3).

These data are provided on different spatial grids and are re-projected to a common grid for our analysis (Section 3.1). Moreover, in situ soil moisture data that are used for evaluation are introduced in Section 2.4.

### 2.1 SMAP Soil Moisture

SMAP satellite estimates surface soil moisture (within the top 5cm of the soil) across the globe with a 2-3 days repeat time (Entekhabi et al., 2010). SMAP uses a passive microwave radiometer in L-band (1.4 GHz) which is known to be sensitive to surface soil moisture and transparent to clouds and atmospheric moisture. Brightness temperature observations from SMAP are used together with ancillary data on vegetation condition and surface temperature to estimate soil dielectric constant from a zeroth-order radiative transfer model, commonly referred to as  $\tau$ - $\omega$  model (Jackson, 1993; Jackson and Schmugge, 1991; Kurum et al., 2011). Finally, soil dielectric constant estimates are converted to volumetric soil moisture estimates using soil texture data (Chan et al., 2016; O'Neill et al., 2015). Measurements from SMAP are available from April 2015 to present.

In this study, we use two SMAP Level-3 (L3) passive soil moisture products. The first is the version 4.0 SPL3SMP product, which estimates soil moisture on a 36km the Equal-Area Scalable Earth Grid (EASE-grid 2.0) (O'Neill et al., 2016b). This product provides soil moisture estimates for both the morning (6 AM) and evening (6 PM) overpasses of the satellite; however, we only use the morning overpasses in this study since it minimize observation errors due to Faraday rotation and soil and canopy temperatures are in equilibrium in the morning. The second is the version 1.0 SPL3SMP\_E product, which estimates soil moisture with an enhanced spatial resolution posted on 9km EASE-grid 2.0 (O'Neill et al., 2016a). Both of the products use the same brightness temperature observations from SMAP radiometer; however, in the enhanced product spatial resolution is enhanced using the Backus Gilbert interpolation technique (Chan et al., 2017; Chaubell et al., 2016). While the SPL3SMP\_E product is posted on a 9 km EASE 2.0 grid, its native resolution is coarser ( $\sim 33$  km). Chan et al., 2017, provides a detailed explanation of how the native resolution and grid spacing of this product are different. Both products are downloaded from National Snow and Ice Data Center (NSIDC) at <https://nsidc.org/data/smap/smap-data.html>.

To be consistent with the SMAP grid, we re-project our ancillary data to EASE-grid 2.0 at the respective spatial resolutions. Moreover, our final downscaled soil moisture estimates are at 2.25km spatial resolution on the EASE-grid 2.0 which is nested within the 9km and 36km grids of SMAP L3 soil moisture estimates (Section 3.1).

## 2.2 MODIS NDVI

Plant's photosynthesis and subsequently vegetation greenness are regulated by the moisture available to plants through their roots, as well as atmospheric conditions (Béziat et al., 2013; Kool et al., 2014; Wang et al., 2014). In turn, vegetation modifies its soil moisture environment through changes in the partitioning of evapotranspiration and through precipitation interception (Coenders-Gerrits et al., 2013; Markewitz et al., 2010). Therefore, surface vegetation cover can be a proxy for the spatial heterogeneity of soil moisture. Vegetation greenness indices such as NDVI are estimated by observations at red and near-infrared frequencies and have relatively high spatial resolution. In this study, we use monthly mean values of NDVI from the MODerate resolution Imaging Spectroradiometer (MODIS) instrument on board Terra satellite. We use version 006 of MYD13A3 product that provides vegetation indices on a 1km spatial resolution at global scale (Didan, 2015).

MODIS estimates are provided on a sinusoidal grid, and due to its high resolution, data are divided into  $10^\circ \times 10^\circ$  tiles. In order to re-project these data to SMAP EASE-grid 2.0, we used two open source libraries in Python. First, the pyModis library was used to mosaic all the tiles together and generate one global map for each measurement. Next, GDAL library was used to reproject MODIS estimates from its original 1km sinusoidal grid to the EASE-grid 2.0s.

## 2.3 Topographic Index

Topographic index (TI) or topographic wetness index is a measure of the soil's saturation tendency given the upstream drainage area and the slope of the local outflow (Marthews et al., 2015). This index is a good proxy for the heterogeneity of soil moisture at the watershed scale and thus could be expected to provide useful information on landscape moisture organization. Here, we use TI at 15 arcsec spatial resolution developed by Marthews et al., 2015. We upscaled the TI values to the EASE-grid 2.0 at 36km, 9km and 2.25km to be used as an ancillary data.

## 2.4 International Soil Moisture Network (ISMN)

In situ soil moisture observations from a set of local networks are provided through the ISMN network and are used to independently evaluate the performance of our downscaling algorithm. ISMN collects and standardizes soil moisture observations from several networks around the globe (Dorigo et al., 2011, 2013). The spatial density of probes and their measurement depth varies across different networks. In this study, we only use data from networks that have measurement at a depth of 5cm below surface (the nominal penetration depth of SMAP) during the first two years of the SMAP mission. Moreover, we only keep [pixelsstations](#) that have at least 20 measurements collocated [coincidentally](#) in time with SMAP observations which are 6am local time. As a result, we only use data from the following 10 networks: FMI (8 [pixelsstations](#)), iRON (4 [pixelsstations](#)), REMEDHUS (20 [pixelsstations](#)), RSMN (19 [pixelsstations](#)), SCAN (159 [pixelsstations](#)), SMOSMANIA (17 [pixelsstations](#)), SNOTEL (344 [pixelsstations](#)), SOILSCAPE (10 [pixelsstations](#)), TERENO (5 [pixelsstations](#)), and USCRN (98 [pixelsstations](#)) (Albergel et al., 2008; Bell et al., 2013; Calvet et al., 2007; Moghaddam et al., 2010, 2016; Zacharias et al., 2011). [Figure S6 shows the spatial distribution of these stations.](#)

## 3 Methodology

Our downscaling algorithm is based on an NN approach that uses the two soil moisture estimates from SMAP at 36km and 9km for training, and then retrieves soil moisture at 2.25km spatial resolution using the 9km estimates from SMAP. The NN relates the coarse scale soil moisture and NDVI estimates as well as the fine scale NDVI estimates as input to the fine scale soil moisture

estimates as output. To do so, we assume that the scaling relationship between 36km and 9km soil moisture estimates is the same as the scaling relationship between 9km and 2.25km estimates. To the best of our knowledge, this is the first time that the assumption of similar scaling relationship is used to downscale soil moisture, while it has been shown that soil moisture has a fractal scaling property across scales (Famiglietti et al., 2008). In the following the spatial grid setup, data preprocessing and NN training are explained.

### 3.1 Spatial Grids

SMAP observations are provided on the EASE-grid 2.0 at different spatial resolutions. The advantage of this grid is the possibility of having nested grids at different spatial resolutions. In this case, SMAP's original radiometer-only soil moisture estimates are at 36km resolution and the enhanced estimates are at 9km resolution which is nested within the 36km one. Since we use the relationship between the 36 and 9km grids ( $\frac{1}{4}$  of the original scale) to train our NN algorithm, our final downscaled estimates also have a spatial resolution of  $\frac{1}{4}$  of the 9km product which is used as input in the retrieval step. This results in a spatial resolution of 2.25km on the EASE-grid 2.0. [Figure 1](#) shows the setup of grids used in the training and retrieval steps.

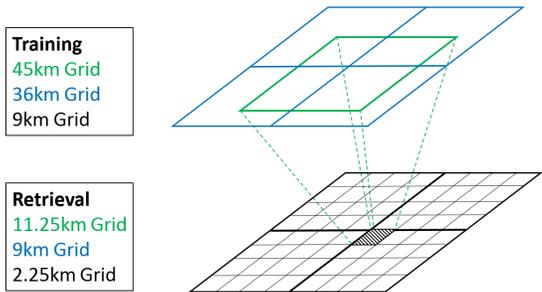


Figure 1- Two levels of spatial grids used for training and retrieval steps in the NN algorithm. While both steps have similar grid structures, the spatial resolutions are different as listed here.

To the best of our knowledge, EASE-grid 2.0 at 2.25km has not been used in any other product before. Therefore, we defined the grid using the same parameters as for 36km and 9km, listed in (Brodzik et al., 2012, 2014), other than the grid spacing which is set to 2252.0138025365 m (equals to  $\frac{1}{4}$  of the grid spacing for the 9km grid).

Our NN algorithm estimates soil moisture at a given pixel of the finer resolution grid (the hatched pixel on the black grid in [Figure 1](#)). To generate the coarse scale soil moisture estimates that are used as input to the NN algorithm, we use a moving window averaging at the coarser resolution grid (the green pixel in [Figure 1](#)) that is centered on the target pixel at the finer grid. In the training step, this moving window is 45 km (applied over the 36 km product) and in the retrieval step, the moving window is 11.25km (applied over the 9 km product). All the inputs to the algorithm are averaged at the scale of the moving window from the coarse resolution grid using an area-weighted averaging. The only disadvantage of the moving window technique is that we cannot retrieve soil moisture at the finer scale near coastal regions or places where SMAP does not provide estimates of soil moisture (such as urban areas, ice covered regions and frozen soil conditions). While this is a limitation of our approach, its impact on the spatial coverage of our downscaled soil moisture estimates is very limited geographically. This limitation can potentially be improved by incorporating an extrapolation algorithm that uses NDVI to estimate soil moisture in coastal pixels (Aires et al., 2017). Moreover, we used the static water bodies' mask from SMAP 9km soil moisture product and masked out all pixels that

Formatted: Font: Not Bold  
Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic

Formatted: Font: Not Bold  
Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic  
Formatted: Font: Not Bold  
Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic

have a water fraction of more than 10%. The reason is that high water fraction impacts brightness temperature observations and results in a biased estimate of soil moisture. We also observed higher discrepancies between the SMAP 36km and 9km soil moisture estimates in regions with high water fraction (not shown here).

### 3.2 Neural Networks Setup

5 Our retrieval algorithm is a statistical approach based on NN. NN uses a set of training data to learn the relationship between inputs and outputs without explicitly modelling the physical relationship between them. This is a powerful statistical approach and has been shown to be able to approximate any continuous function (Cybenko, 1989). NN consists of a set of layers and neurons in each layer that mimic the neural system in humans' brain. Each neuron has a weight and bias corresponding to the neurons from its previous layer. The output of a neuron is a linear summation of the weighted inputs plus the bias that goes through an activation  
10 function. In each Multi-Layered Perceptron (MLP) NN (Rumelhart et al., 1985), there is an input layer, an output layer and one or more hidden layers in between. Number of neurons in the input layer is equal to the number of inputs provided, and number of neurons in the output layer is equal to the number of outputs given to the NN during training.

The right choice of the number of neurons in the hidden layer and number of hidden layers depend on the complexity of the relationship between inputs and outputs. While having more hidden layers and neurons can potentially increase estimation accuracy, it might result in over-fitting to training data. This means that the NN has lower accuracy when applied to other  
15 independent data. To avoid over-fitting, we use the simplest network (i.e., minimum of hidden layers and neurons) that has an acceptable estimation accuracy. Moreover, we use an extensive and representative set of data for training that covers the full range of soil moisture dynamics across different climates and surface conditions. Furthermore, we test the generalization ability of the NN in an independent dataset.

20 We use the first two years of data from SMAP beginning on April 1<sup>st</sup>, 2015 until March 31<sup>st</sup>, 2017. For training data, we sample SMAP soil moisture estimates every 10 days at global scale. This results in 46.3 million data points after removing pixels with high percentage of water bodies. For validation, we sample the SMAP data every 5 days with one day lag with respect to the training data at global scale; therefore, training and validation samples are mutually exclusive. This results in 92 million data points after removing pixels with high percentage of water bodies.

25 During training the weights and biases of each neuron are estimated iteratively by minimizing a mean squared error cost function using a gradient-descent algorithm with back propagation (Hagan and Menhaj, 1994; Rumelhart et al., 1985). For this purpose, training data are divided to three categories of training (60%), validation (20%) and testing (20%). After each iteration on the training category, the validation data are used to check for over-fitting and test data are used to check convergence. When changes in the cost function are smaller than a threshold, training stops, and weights and biases that resulted in the best performance are  
30 selected as the parameters of the network. The validation data used here should not be mistaken with the independent validation dataset sampled mutually exclusive to the training dataset. Throughout the paper, we will use "validation data" terminology to refer to the dataset sampled mutually exclusive to the training data (explained in the previous paragraph) and will be used to evaluate the performance of the NN training in terms of over-fitting.

We trained a set of networks with number of hidden layers between 1 and 5 and number of neurons in each hidden layer between  
35 1 and 15. After evaluating their performance, we selected the network with one hidden layer and 5 neurons, which has a good accuracy, while adding more hidden layers or neurons did not change the performance. We use the hyperbolic tangent sigmoid function as the activation function for the hidden layer, a standard practice, and a linear function as the activation function for output layer.

### 3.3 Inputs and Downscaling Schemes

We develop four different schemes with increasing number of ancillary inputs for downscaling SMAP soil moisture estimates. In this section, we introduce each of them, and we compare their performance in Section 4. At the end, only one scheme was selected to provide the final downscaled soil moisture estimates.

5 Scheme R1 has the three following inputs: soil moisture estimates from SMAP on the moving window grid at the coarse scale resolution (45km for training and 11.25km for retrieval), mean NDVI at the coarse scale moving window, and NDVI in the target pixel at the fine scale grid (9km for training and 2.25km for retrieval).

Scheme R2 has all the inputs in scheme R1 plus standard deviation of NDVI at the fine scale pixels within the moving window. For example, in the training step there are 25 9km pixels within the 45km moving window. We estimate the standard deviation of those, and input that to the network. This estimate provides a proxy of the heterogeneity within the coarse scale grid.

10 Scheme R3 has all the inputs in scheme R1 plus TI at the moving window of the coarse scale and TI in the target pixel at the fine scale grid.

Scheme R4 has all the inputs from schemes R1-R3. [Table 1](#) lists all the four schemes and their respective inputs. For each scheme, the 9km SMAP soil moisture estimates are used as the NN target data.

**Table 1- Inputs used in each of the downscaling schemes. The output is always soil moisture estimates at 2.25km resolution.**

#	Usage	Input 1	Input 2	Input 3	Input 4	Input 5	Input 6
R1	Training	SM at 45 km	NDVI at 45 km	NDVI at 9 km for target pixel	--	--	--
	Retrieval	SM at 11.25 km	NDVI at 11.25 km	NDVI at 2.25 km for target pixel	--	--	--
R2	Training	SM at 45 km	NDVI at 45 km	NDVI at 9 km for target pixel	$\sigma_{NDVI}$ at 45 km for all pixels at 9 km	--	--
	Retrieval	SM at 11.25 km	NDVI at 11.25 km	NDVI at 2.25 km for target pixel	$\sigma_{NDVI}$ at 11.25 km for all pixels at 2.25 km	--	--
R3	Training	SM at 45 km	NDVI at 45 km	NDVI at 9 km for target pixel	TI at 45 km	TI at 9 km for target pixel	--
	Retrieval	SM at 11.25 km	NDVI at 11.25 km	NDVI at 2.25 km for target pixel	TI at 11.25 km	TI at 2.25 km for target pixel	--
R4	Training	SM at 45 km	NDVI at 45 km	NDVI at 9 km for target pixel	TI at 45 km	TI at 9 km for target pixel	$\sigma_{NDVI}$ at 45 km for all pixels at 9 km
	Retrieval	SM at 11.25 km	NDVI at 11.25 km	NDVI at 2.25 km for target pixel	TI at 11.25 km	TI at 2.25 km for target pixel	$\sigma_{NDVI}$ at 11.25 km for all pixels at 2.25 km

### 4 Results

In this section, we first present the results of NN training and evaluation of the downscaling from 36km to 9km. We only present the results from Scheme R1, since all four schemes have similar performances. Next, we apply the downscaling scheme to the entire 2-year SMAP data record and generate soil moisture at 2.25km spatial resolution. Finally, we evaluate the accuracy of downscaled soil moisture from all the four schemes using in situ soil moisture estimates from the ISMN dataset.

20

Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic

Formatted: Font: 9 pt

Formatted: Line spacing: single

Inserted Cells

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Line spacing: single

Merged Cells

Formatted: Line spacing: single

Formatted: Font: 9 pt

Inserted Cells

Formatted: Font: 9 pt

Formatted: Indent: Before: -0.04", After: -0.08", Line spacing: single

Formatted: Font: 9 pt

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Font: 9 pt

Formatted: Font: 9 pt

Formatted: Indent: Before: -0.04", After: -0.08", Line spacing: single

Formatted: Font: 9 pt

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Font: 9 pt

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Font: 9 pt

Formatted: Line spacing: single

Formatted: Indent: Before: -0.04", After: -0.08", Line spacing: single

Formatted: Line spacing: single

Formatted: Line spacing: single

Inserted Cells

Formatted: Font: 9 pt

Formatted: Font: 9 pt

Formatted: Indent: Before: -0.04", After: -0.08", Line spacing: single

#### 4.1 Evaluating NN Training

To evaluate the success of the NN training, we compute the correlation coefficient ( $R^2$ ) and the unbiased Root Mean Square Difference (ubRMSD) between the NN estimates and the target data for the training and validation data. We compare the metrics aggregated across all the data and spatially at each pixel.

Figure 2 shows the density scatterplot of 9km NN estimates versus target 9km soil moisture data in the training and validation datasets. Both datasets have similar ubRMSD and  $R^2$  which shows that the NN is able to generalize beyond the training data.  $R^2$  of ~0.98 is an almost perfect correlation between the target and estimates, and indicates that the NN setup with the inputs provided is capable of learning the relationship between coarse and fine scale soil moisture.

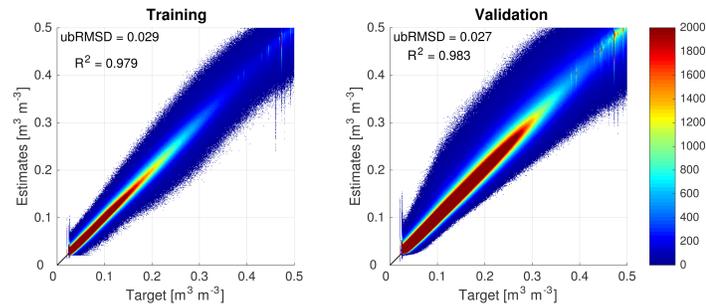
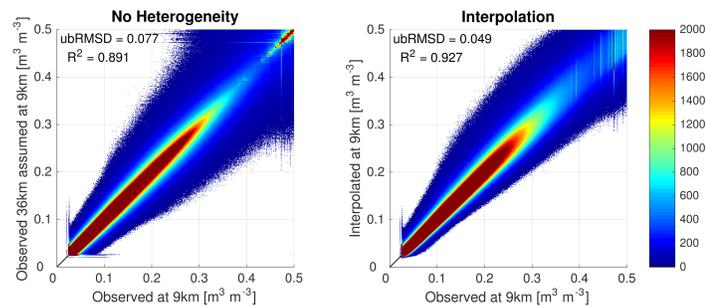


Figure 2- Density scatterplot of NN estimates vs target data for training (left) and validation (right) datasets. Color shades show number of samples in each part of the scatterplot.

In order to compare the performance of NN with other techniques, we present similar results for two other approaches. The first approach assumes a uniform distribution of soil moisture (i.e. no heterogeneity) within each 36km grid pixel. This means that all the 9km pixels that fall within a 36km grid pixel are attributed the same soil moisture value as the 36km one (i.e. no downscaling). The second approach is to use a linear interpolation (i.e. a crude downscaling). This is used as a reference to a simpler approach than the NN. In this approach, soil moisture values on the 36km grid are assumed to be at the center point of the pixel. Then using a linear interpolation, soil moisture is estimated at the center point of each 9km grid pixel. We implement the interpolation at the global scale for each SMAP observation. Figure 3 shows the density scatterplots for these two approaches. As the scatterplots and performance metrics show, the NN algorithm has higher correlation and lower ubRMSD compared to both no heterogeneity and interpolation approaches. As expected, the homogeneous algorithm has the worst performance.



Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic

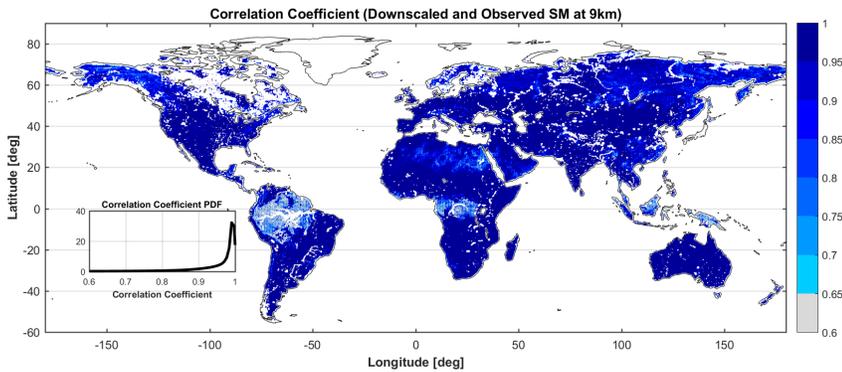
Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic, Check spelling and grammar

**Figure 3- Density scatterplots of uniform 9km soil moisture estimate (left) and interpolated soil moisture at 9km (right) vs observed 9km soil moisture. Color shades show number of samples in each part of the scatterplot.**

To further evaluate the performance of the NN downscaling algorithm, we evaluate the correlation between target soil moisture and NN estimates at 9km in each pixel (Figure 4). These estimates are based on the validation dataset, and on average 50 data points are used to calculate the correlations (in some snowy regions like Himalaya Mountains as few as 20 data points were used, the minimum number of samples set to estimate the correlation in this evaluation). Correlations range between a minimum of 0.6 in dense tropical forests (where NDVI saturates and limited spatial inhomogeneity can be found) to ~1 in most other parts of the world, except in very northern regions where the correlation is closer to 0.9. The smaller value of correlation in the tropics is a result of the low seasonality and saturation effect of NDVI in those regions, so that NDVI does not provide useful information for downscaling (Morton et al., 2014). These regions have relatively high soil moisture and relatively low variability in NDVI throughout the year. In addition, NDVI and EVI tend to saturate at high vegetation cover so that there is only limited variability in those regions. As a result, disaggregating soil moisture using NDVI has a lower accuracy compared to other regions. Nevertheless, NDVI and EVI are the only global fine scale vegetation indices that can be used as ancillary data to describe the sub-pixel heterogeneity of soil moisture, and both lack seasonality in the dense tropical forests. Moreover, Figure S1 shows the percentage difference between the NN retrieval and SMAP estimates at 9km. Largest differences are observed in mountainous and coastal regions where soil moisture variations are influenced by factors such as the terrain's slope. However, inclusion of other ancillary data such as slope in other downscaling schemes did not improve the bias and schemes R2-R4 have similar performance both in terms of aggregated R<sup>2</sup> and ubRMSD and spatial patterns of correlation between NN estimates and target soil moisture. Moreover, Figures S2-S5 show the spatial correlation maps and percentage difference between SMAP estimates at 9km and the estimates from Interpolation and No Heterogeneity approaches. These figures show that both approaches have significantly higher bias with respect to the SMAP estimates compared to the NN retrieval.

Formatted: Font: Not Bold  
Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic



**Figure 4- Correlation coefficient (R<sup>2</sup>) between SMAP observed soil moisture at 9km and NN downscaled soil moisture at 9km. White regions indicate no data.**

#### 4.2 Downscaled Soil Moisture

Using the NN that was trained and evaluated in the previous section, we then estimate soil moisture at 2.25km EASE-grid 2.0 resolution. For this, we use the SMAP 9km observed soil moisture as input together with ancillary data for each scheme as described in Section 3.3.

Figure 5 shows the average global soil moisture at three different spatial resolutions (36km and 9km from SMAP observations, and 2.25km downsampled from SMAP observations). Overall, the spatial patterns are as expected, with some heterogeneity being added as the resolution increases. These maps show that our downscaling algorithm preserves the large-scale spatial patterns of soil moisture while disaggregating it at local scales. Moreover, the latitudinal average plots (on the right side of each panel of Figure 5) show that at higher spatial resolutions there is more spatial heterogeneity. The latitudinal average for the 36 km product is much smoother than the 2.25 km one.

Figure 6 shows a snapshot of soil moisture estimates from SMAP 36km product, SMAP 9km product and the downsampled product at 2.25km across western Africa from Oct. 18<sup>th</sup>, 2015, in order to further demonstrate the added information of the downscaling algorithm. From left to right in Figure 6 with increasing spatial resolution, finer spatial patterns are revealed, with finer structure, while respecting the overall 9km structures. We note that there is no random spatial noise in the higher spatial resolution data emphasizing that the additional NN input (NDVI) is useful to preserve the spatial connectivity of different high-resolution pixels.

Formatted: Font: Not Bold, Complex Script Font: 10 pt

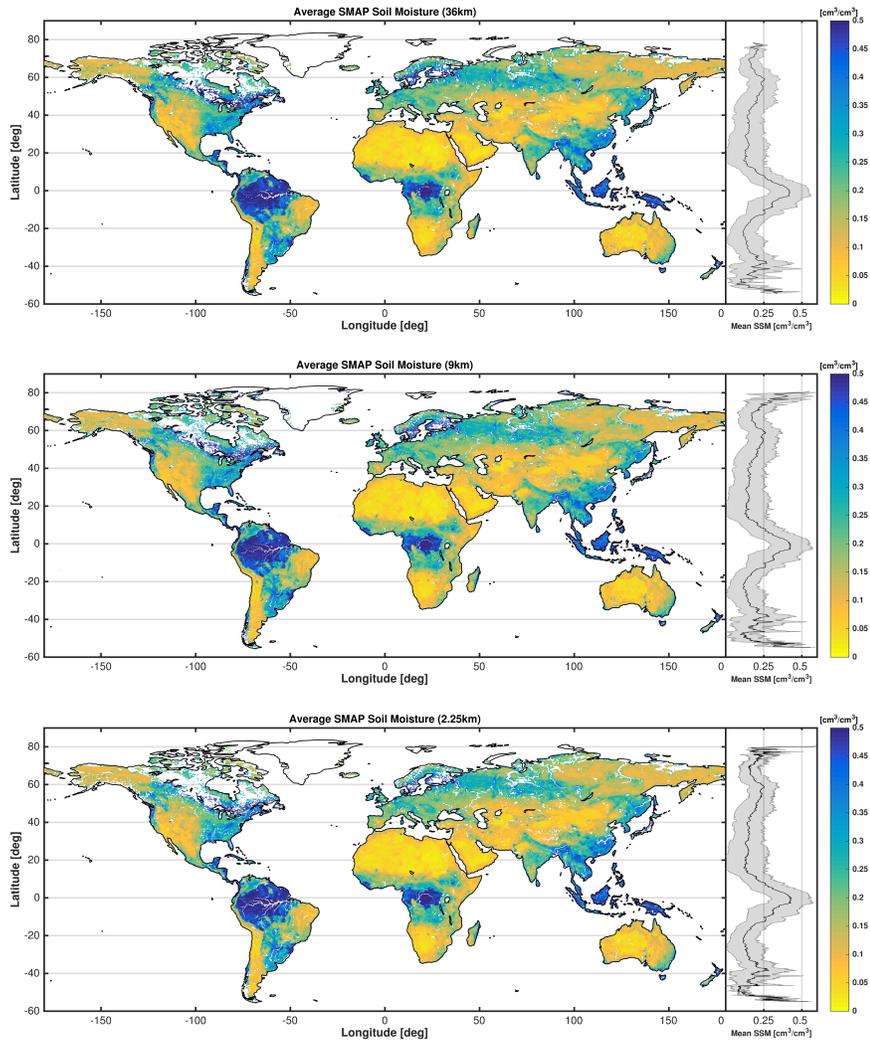
Formatted: Font: Not Bold, Complex Script Font: 10 pt

Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic, Check spelling and grammar

Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic, Check spelling and grammar



**Figure 55- Average soil moisture at 36km (top), 9km (middle) and 2.25km (bottom) between April 2015 and March 2017.** Shaded plot on the right side of each panel shows the mean and one standard deviation of soil moisture at each latitude.

- Formatted: Normal
- Formatted: Font: 9 pt, Italic
- Formatted: Font: Not Bold, Italic, Font color: Auto

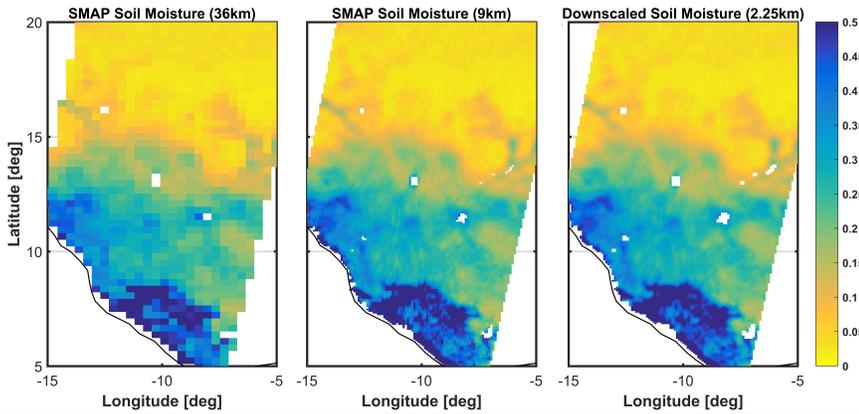
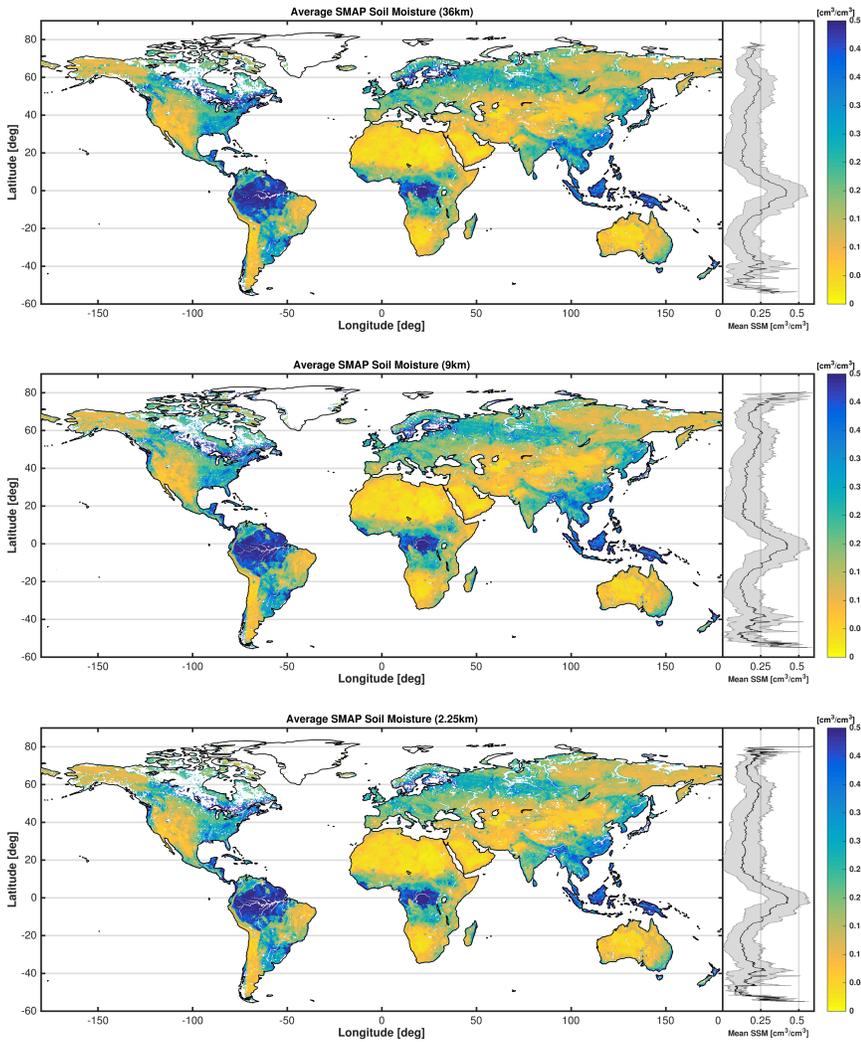


Figure 6- Example of a snapshot of soil moisture estimates from SMAP at 36km (left) and 9km (middle), and the downscaled estimates at 2.25km (right) in western Africa.

#### 4.3 Large Scale Evaluation of Downscaled Soil Moisture

5 In this section, we analyze large-scale patterns of the downscaled soil moisture estimates at 2.25km spatial resolution. Our first analysis is the spatial heterogeneity of the downscaled soil moisture. We calculate coefficient of variation (CV), which is the spatial standard deviation divided by the mean, of the 16 2.25km pixels within each 9km pixel at each time. [Figure 7](#) [Figure 6 top middle](#) panel shows the mean CV at each pixel for the downscaled soil moisture. For comparison, we also calculate CV for the 9km soil moisture estimates from SMAP at the 36km grid ([Figure 7](#) [Figure 6 top panel](#)), and CV for the 2.25km downscaled estimates at the [36km grid](#) ([Figure 7](#) bottom panel). [The two panels](#) in [Figure 7](#) [Figure 6](#) have different range of CV which is expected given the difference in their spatial scales. ~~However, they, but inset PDFs have similar scales for better comparison.~~ [These plots](#) reveal similar spatial patterns across different scales, yet with finer and contiguous structures at higher resolution. This means that in regions that have high spatial heterogeneity in the 9km SMAP estimates, we also [see](#) high heterogeneity in the 2.25km downscaled soil moisture estimates. [Meanwhile, the 2.25km estimates have higher variation \(i.e. heterogeneity\) at the 36km grid scale compared to the 9km product.](#) This indicates that the 2.25km product is not just a smoothed version of the 9km product but can add spatial variability not captured by the coarser 9km pixels. Therefore, our NN algorithm is appropriately explaining the spatial variability of soil moisture using NDVI as ancillary data. The other downscaling schemes (R2-R4) exhibit similar spatial patterns for CV. This is an indication that the other inputs to these schemes do not provide any extra information on the heterogeneity of soil moisture at the finer spatial scale.

- Formatted: Font: Not Bold
- Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic
- Formatted: Font: Not Bold
- Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic
- Formatted: Font: Not Bold, Complex Script  
Font: 10 pt
- Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: 10 pt, Not Italic
- Formatted: Font color: Text 1
- Formatted: Font: Not Bold
- Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic



a  
**Figure 5** Average soil moisture at 36km (top), 9km (middle) and 2.25km (bottom) between April 2015 and March 2017. Shaded plot on the right side of each panel shows the mean and one standard deviation of soil moisture at each latitude.

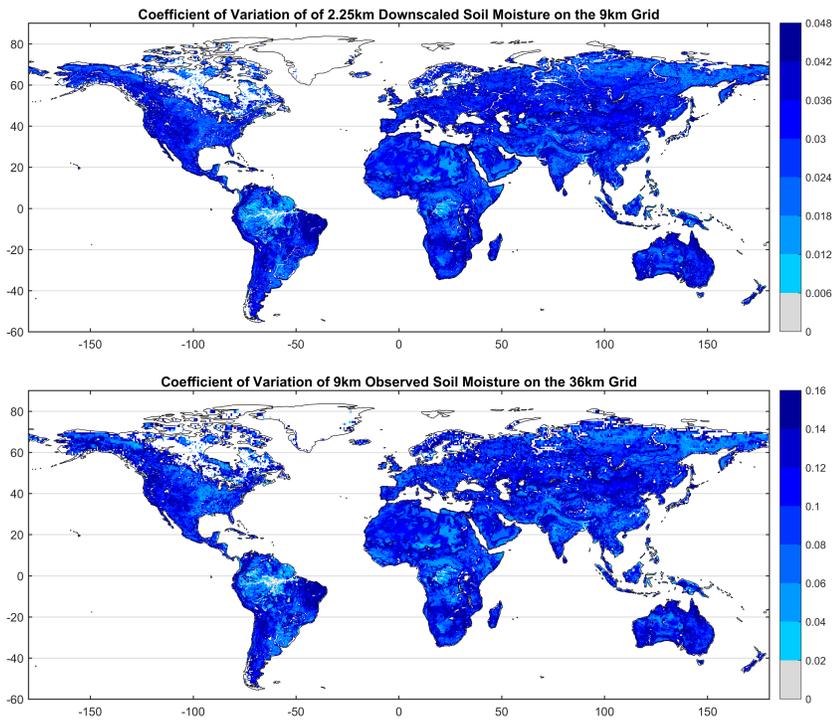
5 Our second analysis evaluates the conservation of the water balance within each coarse scale pixel. Since we are disaggregating soil moisture from 9km grid to 2.25km grid, the downscaling algorithm may not preserve the mean soil moisture within each coarse scale pixel (water balance). [Figure 8](#) [Figure 7](#) shows the spatial map of the percentage average difference (with respect to their

Formatted: Font: 9 pt, Italic  
 Formatted: Normal  
 Formatted: Font: Not Bold, Italic, Font color: Auto

Formatted: Font: Not Bold  
 Formatted: Font: Not Bold, Not Italic, Complex Script  
 Font: Not Italic

mean) between the mean soil moisture from the 16 2.25km pixels within each 9km pixel and soil moisture estimates at 9km from SMAP. The inset PDF also shows the probability distribution of these differences. On average, there is a positive bias of ~1.5% in the differences, and the downscaled soil moisture estimates at 2.25km are slightly positively biased, i.e. more humid. However, the differences are small (mostly less than 3% in absolute value) and within the SMAP retrieval mission accuracy ( $<0.04 \text{ m}^3 \text{ m}^{-3}$ ).

5 This analysis also reveals that our downscaling algorithm is robust in preserving the water balance within each coarse scale pixel and still within the original requirements of the SMAP mission.



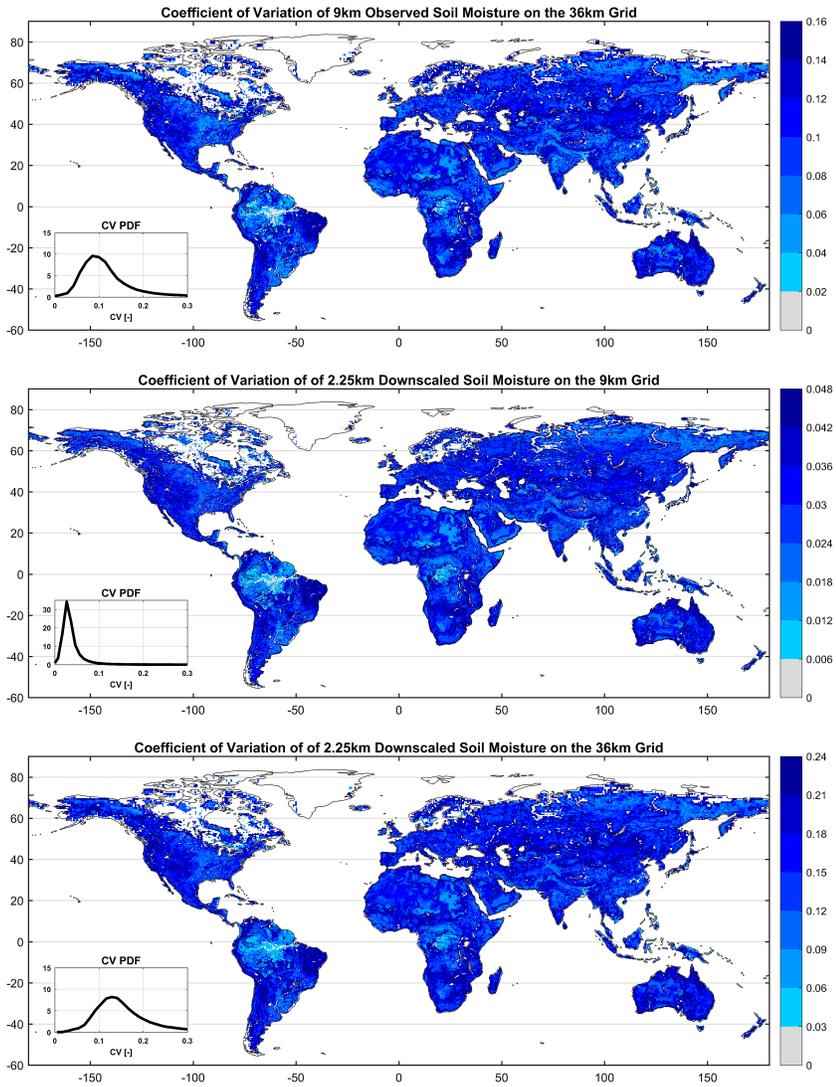


Figure 76- Coefficient of variations of (top) 9km observed soil moisture within each 36km grid pixel, (middle) 2.25km downscaled soil moisture within each 9km grid pixel (top), and of the 9km observed soil moisture within each 36km grid pixel (bottom); 2.25km downscaled soil moisture within each 36km grid pixel .

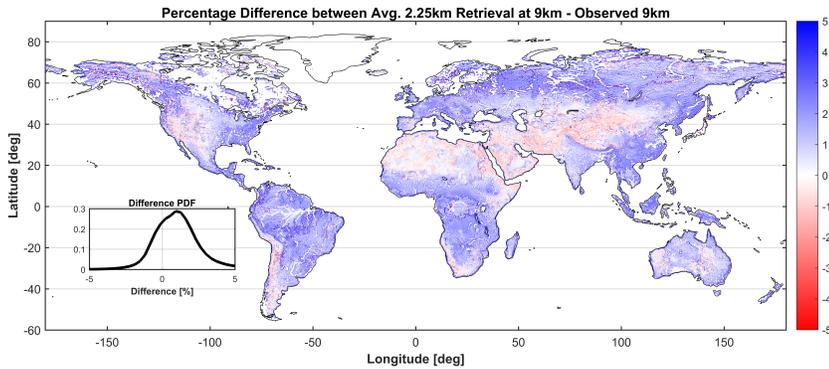


Figure 87- Percentage difference with respect to mean of average 2.25km soil moisture estimates at 9km and observed 9km soil moisture.

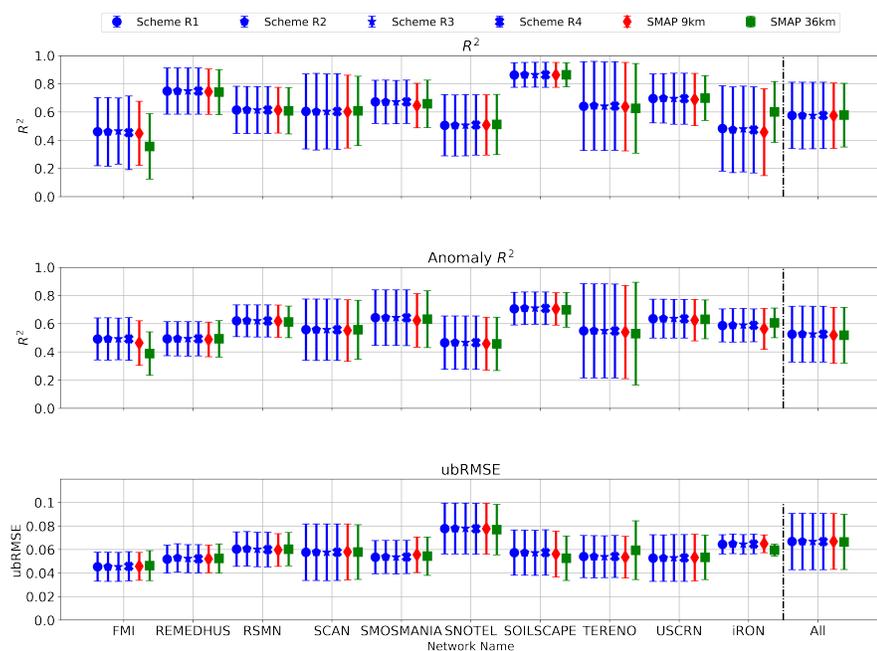
#### 4.4 Comparison against ISMN Data

In this section, we evaluate downscaled soil moisture estimates against ISMN dataset. We calculate  $R^2$ , anomaly  $R^2$  and ubRMSE for each of the four downscaling schemes as well as SMAP estimates at 36km and 9km. Anomalies are calculated based on a 30 day moving average window. We use stations from 10 networks within the ISMN dataset, and Figure S6 shows the spatial distribution of these stations. For networks that have more than one station within each SMAP pixel, we average the stations within the 2.25km pixels and then calculate the metrics for each of the 36km, 9km and 2.25km products. Figure 9 shows the average of each metric across stations for different networks and the average among all of the networks.

In general, all four schemes have quite similar performance within each network and across all networks. During the NN training, all four schemes had similar performance. In some cases, scheme R1 had slightly better performance (i.e. ~2% higher  $R^2$ ). For these reasons, and to reduce complexity of the downscaling algorithm, our final downscaling algorithm is scheme R1.

Our downscaling scheme always has equal or better performance compared to the 9km product across individual networks. Since the 9km product is the input to our downscaling, this shows that in terms of temporal correlation and ubRMSE our downscaling algorithm either improves the 9km product or is similar in terms of accuracy. In comparison to the 36km product, our downscaling algorithm's performance follows the 9km product performance. At stations where the 9km product has better performance compared to the 36km product (e.g. ubRMSE in iRON), the 2.25km product also has better performance compared to the 36km product, and vice versa. The main added value of the 2.25km product is its enhanced spatial resolution that provides a measure of soil moisture heterogeneity at finer spatial resolution using satellite observations while having an acceptable temporal accuracy.

Formatted: Font: Not Bold  
Formatted: Font: Not Bold, Not Italic, Complex Script  
Font: Not Italic



**Figure 98-**  $R^2$  (top), anomaly  $R^2$  (middle) and ubRMSE (bottom) of each of the downscaling schemes and SMAP 9km and 36km products across 10 different networks of ISMN. Error bars show one standard deviation of each metric around the mean value. Average values across all sites is presented on the right of each panel.

## 5 Conclusions

In this study, we have developed a statistical disaggregation algorithm using neural networks to downscale soil moisture observations from SMAP to a spatial resolution of 2.25km. We use the two level-3 soil moisture estimates from SMAP at 36km and 9km spatial resolutions to train our downscaling algorithm. Assuming the relationship is consistent between scales, we apply the downscaling algorithm to 9km soil moisture estimates from SMAP and disaggregate those to 2.25km. Our downscaling algorithm uses only mean monthly NDVI estimates as ancillary data at both the coarse and fine spatial scales to estimate the heterogeneity of soil moisture at sub-pixel level. Our investigation shows that topographic index and variability of NDVI within the coarse scale pixel do not provide additional useful information on the spatial heterogeneity of soil moisture for the downscaling

using our proposed NN technique and are thus omitted in the final product. This is the first study to estimate soil moisture at global scale at very fine spatial resolution (~2km). This new product can be used in a range of other studies and applications including land surface – atmosphere interaction modeling, evapotranspiration modeling and agricultural management. However, in some parts of the world that farm sizes are smaller than the resolution of this product, further disaggregation of soil moisture estimates is needed. Moreover, we will extend these estimates to near real time as long as SMAP soil moisture estimates are available.

Our evaluation shows that the [downscaling algorithm has high accuracy performance of the downscaled soil moisture estimates is equal or better than that of the SMAP 9km estimates](#) in terms of temporal correlation, anomaly correlation and ubRMSE when

compared to *in situ* soil moisture estimates from ISMN. ~~The performance of the downscaled soil moisture estimates is equal or better than that of the SMAP 9km estimates.~~ Moreover, averaging the estimates at 2.25km within each 9km pixel shows that our downscaling algorithm has high accuracy in preserving the water balance (less than 5% error, within the SMAP mission requirements).

Formatted: Font: Italic

~~In this study, we use the relative value of NDVI (in a given pixel with respect to the neighboring pixels) as an auxiliary information to predict spatial variability of soil moisture in each coarse-scale pixel. While the relationship between soil moisture and NDVI at phenological time scales may not be valid at the temporal scale of SMAP observations (couple of days), our assumption builds on the relative value of NDVI within a small region and not the absolute value. Therefore, it is reasonable to use NDVI as a predictor in this case.~~ Use of NDVI as an ancillary measurement to disaggregate soil moisture builds on the assumption that there is a moderate vegetation cover in the pixel of interest. Therefore, this lowers the quality of the downscaling algorithm in bare soil or sparsely vegetated regions. ~~Figure 8 shows a dry bias in arid regions which is an indication of lack of NDVI – soil moisture relationship.~~ Moreover, NDVI estimates tend to saturate in highly dense vegetated regions such as ~~the~~in tropical forests, which results in limited ability of the downscaling algorithm to resolve the heterogeneity of ~~the~~ soil moisture in those regions.

Formatted: Font: Not Bold, Complex Script Font: 10 pt

Formatted: Font: Not Bold, Not Italic, Complex Script Font: 10 pt, Not Italic

This study shows the potential of using neural networks with a large number of training data to develop a downscaling algorithm for soil moisture. Data generated using this algorithm can be used to provide an improved understanding of the dynamics of soil moisture and land surface atmosphere interactions at global scale.

#### Acknowledgements

The authors would like to acknowledge funding from NASA grant NNX15AB30G. The MYD13A3 data product was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, [https://lpdaac.usgs.gov/data\\_access/data\\_pool](https://lpdaac.usgs.gov/data_access/data_pool).

#### Data Availability

Downscaled soil moisture estimates from this study are publicly available, and we intend to extend the temporal coverage of the data periodically. Please contact the corresponding author to access the data: Pierre Gentine ([pg2328@columbia.edu](mailto:pg2328@columbia.edu))

#### References

- Aires, F., Aznay, O., Prigent, C., Paul, M. and Bernardo, F.: Synergistic multi-wavelength remote sensing versus a posteriori combination of retrieved products: Application for the retrieval of atmospheric profiles using MetOp-A, *J. Geophys. Res. Atmos.*, 117(D18), doi:10.1029/2011JD017188, 2012.
- Aires, F., Miolane, L., Prigent, C., Pham, B., Fluet-Chouinard, E., Lehner, B. and Papa, F.: A Global Dynamic Long-Term Inundation Extent Dataset at High Spatial Resolution Derived through Downscaling of Satellite Observations, *J. Hydrometeorol.*, 18(5), 1305–1325, doi:10.1175/JHM-D-16-0155.1, 2017.
- Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Pignat, B. and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrol. Earth Syst. Sci.*, 12(6), 1323–1337, doi:10.5194/hess-12-1323-2008, 2008.

- Alemohammad, S. H., Kolassa, J., Prigent, C., Aires, F. and Gentine, P.: Statistical retrieval of surface and root zone soil moisture using synergy of multi-frequency remotely-sensed observations, in 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 4943–4946, IEEE., 2017a.
- Alemohammad, S. H., Fang, B., Konings, A. G., Aires, F., Green, J. K., Kolassa, J., Miralles, D., Prigent, C. and Gentine, P.:  
5 Water, Energy, and Carbon with Artificial Neural Networks (WECANN): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence, *Biogeosciences*, 14(18), 4101–4124, doi:10.5194/bg-14-4101-2017, 2017b.
- Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., Hall, M. E., Kochendorfer, J., Meyers, T. P., Wilson, T. and Diamond, H. J.: U.S. Climate Reference Network Soil Moisture and Temperature Observations, *J.*  
10 *Hydrometeorol.*, 14(3), 977–988, doi:10.1175/JHM-D-12-0146.1, 2013.
- Berg, A., Lintner, B. R., Findell, K. L., Malyshev, S., Loikith, P. C. and Gentine, P.: Impact of Soil Moisture–Atmosphere Interactions on Surface Temperature Distribution, *J. Clim.*, 27(21), 7976–7993, doi:10.1175/JCLI-D-13-00591.1, 2014.
- Béziat, P., Rivalland, V., Tallec, T., Jarosz, N., Boulet, G., Gentine, P. and Ceschia, E.: Evaluation of a simple approach for crop evapotranspiration partitioning and analysis of the water budget distribution for several crop species, *Agric. For. Meteorol.*, 177,  
15 46–56, doi:10.1016/j.agrformet.2013.03.013, 2013.
- Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review, *Hydrol. Process.*, 9(3–4), 251–290, doi:10.1002/hyp.3360090305, 1995.
- Brodzik, M. J., Billingsley, B., Haran, T., Raup, B. and Savoie, M. H.: EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets, *ISPRS Int. J. Geo-Information*, 1(1), 32–45, doi:10.3390/ijgi1010032, 2012.
- 20 Brodzik, M. J., Billingsley, B., Haran, T., Raup, B. and Savoie, M. H.: Correction: Brodzik, M.J., et al. EASE-Grid 2.0: Incremental but Significant Improvements for Earth-Gridded Data Sets. *ISPRS International Journal of Geo-Information* 2012, 1, 32–45, *ISPRS Int. J. Geo-Information*, 3(3), 1154–1156, doi:10.3390/ijgi3031154, 2014.
- Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A. and Piguet, B.: In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, in 2007 IEEE International Geoscience and Remote Sensing Symposium, pp. 1196–1199,  
25 IEEE., 2007.
- Chakrabarti, S., Judge, J., Bongiovanni, T., Rangarajan, A. and Ranka, S.: Disaggregation of Remotely Sensed Soil Moisture in Heterogeneous Landscapes Using Holistic Structure-Based Models, *IEEE Trans. Geosci. Remote Sens.*, 54(8), 4629–4641, doi:10.1109/TGRS.2016.2547389, 2016.
- Chakrabarti, S., Judge, J., Rangarajan, A. and Ranka, S.: Utilizing Self-Regularized Regressive Models to Downscale Microwave  
30 Brightness Temperatures for Agricultural Land Covers in the SMAPVEX-12 Region, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 10(2), 478–488, doi:10.1109/JSTARS.2016.2637927, 2017.
- Chan, S. K., Bindlish, R., O’Neill, P., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M. H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., González-Zamora, A., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J.,  
35 Palecki, M., Small, E. E., Zreda, M., Calvet, J.-C., Crow, W. T. and Kerr, Y.: Assessment of the SMAP Passive Soil Moisture Product, *IEEE Trans. Geosci. Remote Sens.*, 54(8), 4994–5007, doi:10.1109/TGRS.2016.2561938, 2016.
- Chan, S. K., Bindlish, R., O’Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M. H., Caldwell, T., Walker, J., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Uldall, F., Seyfried, M., Bosch, D., Starks, P., Holifield Collins, C., Prueger, J., van der Velde, R., Asanuma, J., Palecki, M., Small,

- E. E., Zreda, M., Calvet, J., Crow, W. T. and Kerr, Y.: Development and assessment of the SMAP enhanced passive soil moisture product, *Remote Sens. Environ.*, (August), doi:10.1016/j.rse.2017.08.025, 2017.
- Chaubell, J., Yueh, S., Entekhabi, D. and Peng, J.: Resolution enhancement of SMAP radiometer data using the Backus Gilbert optimum interpolation technique, in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 284–287, IEEE., 2016.
- Coenders-Gerrits, A. M. J., Hopp, L., Savenije, H. H. G. and Pfister, L.: The effect of spatial throughfall patterns on soil moisture patterns at the hillslope scale, *Hydrol. Earth Syst. Sci.*, 17(5), 1749–1763, doi:10.5194/hess-17-1749-2013, 2013.
- Colliander, A., Fisher, J. B., Halverson, G., Merlin, O., Misra, S., Bindlish, R., Jackson, T. J. and Yueh, S.: Spatial Downscaling of SMAP Soil Moisture Using MODIS Land Surface Temperature and NDVI During SMAPVEX15, *IEEE Geosci. Remote Sens. Lett.*, 14(11), 2107–2111, doi:10.1109/LGRS.2017.2753203, 2017a.
- Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., Cosh, M. H., Dunbar, R. S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E. G. and Yueh, S.: Validation of SMAP surface soil moisture products with core validation sites, *Remote Sens. Environ.*, 191, 215–231, doi:10.1016/j.rse.2017.01.021, 2017b.
- Cybenko, G.: Approximation by superpositions of a sigmoidal function, *Math. Control. Signals, Syst.*, 2(4), 303–314, doi:10.1007/BF02551274, 1989.
- Das, N. N., Entekhabi, D. and Njoku, E. G.: An Algorithm for Merging SMAP Radiometer and Radar Data for High-Resolution Soil-Moisture Retrieval, *IEEE Trans. Geosci. Remote Sens.*, 49(5), 1504–1512, doi:10.1109/TGRS.2010.2089526, 2011.
- Didan, K.: MYD13A3 MODIS/Aqua Vegetation Indices Monthly L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC, , doi:10.5067/modis/myd13a3.006, 2015.
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A. and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrol. Earth Syst. Sci.*, 15(5), 1675–1698, doi:10.5194/hess-15-1675-2011, 2011.
- Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W. and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, *Vadose Zo. J.*, 12(3), doi:10.2136/vzj2012.0097, 2013.
- Entekhabi, D., Rodriguez-Iturbe, I. and Castelli, F.: Mutual interaction of soil moisture state and atmospheric processes, *J. Hydrol.*, 184(1–2), 3–17, doi:10.1016/0022-1694(95)02965-6, 1996.
- Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L. and Van Zyl, J.: The Soil Moisture Active Passive (SMAP) Mission, *Proc. IEEE*, 98(5), 704–716, doi:10.1109/JPROC.2010.2043918, 2010.
- Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M. and Jackson, T. J.: Field observations of soil moisture variability across scales, *Water Resour. Res.*, 44(1), doi:10.1029/2006WR005804, 2008.
- Gedney, N. and Cox, P. M.: The Sensitivity of Global Climate Model Simulations to the Representation of Soil Moisture Heterogeneity, *J. Hydrometeorol.*, 4(6), 1265–1275, doi:10.1175/1525-7541(2003)004<1265:TSOGCM>2.0.CO;2, 2003.

- Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G. and Duchemin, B.: Analysis of evaporative fraction diurnal behaviour, *Agric. For. Meteorol.*, 143(1–2), 13–29, doi:10.1016/j.agrformet.2006.11.002, 2007.
- Gentine, P., Entekhabi, D. and Polcher, J.: The Diurnal Behavior of Evaporative Fraction in the Soil–Vegetation–Atmospheric Boundary Layer Continuum, *J. Hydrometeorol.*, 12(6), 1530–1546, doi:10.1175/2011JHM1261.1, 2011.
- 5 Hagan, M. T. and Menhaj, M. B.: Training feedforward networks with the Marquardt algorithm, *IEEE Trans. Neural Networks*, 5(6), 989–993, doi:10.1109/72.329697, 1994.
- Hatfield, J. L.: Upscaling and Downscaling Methods for Environmental Research, *J. Environ. Qual.*, 30(3), 1100, doi:10.2134/jeq2001.3031100x, 2001.
- Ines, A. V. M., Mohanty, B. P. and Shin, Y.: An unmixing algorithm for remotely sensed soil moisture, *Water Resour. Res.*, 49(1), 408–425, doi:10.1029/2012WR012379, 2013.
- 10 Jackson, T. J.: III. Measuring surface soil moisture using passive microwave remote sensing, *Hydrol. Process.*, 7(2), 139–152, doi:10.1002/hyp.3360070205, 1993.
- Jackson, T. J. and Schmugge, T. J.: Vegetation effects on the microwave emission of soils, *Remote Sens. Environ.*, 36(3), 203–212, doi:10.1016/0034-4257(91)90057-D, 1991.
- 15 Jagdhuber, T., Entekhabi, D., Hajnsek, I., Konings, A. G., McColl, K. A., Alemohammad, S. H., Das, N. N. and Montzka, C.: Physically-based active-passive modelling and retrieval for SMAP soil moisture inversion algorithm, in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1300–1303, IEEE., 2015.
- Jagdhuber, T., Entekhabi, D., Konings, A. G., McColl, K. A., Alemohammad, S. H., Das, N. N., Montzka, C. and Piles, M.: Physically-based retrieval of SMAP active-passive measurements covariation and vegetation structure parameters, in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3078–3081, IEEE., 2016.
- 20 Jiménez, C., Prigent, C. and Aires, F.: Toward an estimation of global land surface heat fluxes from multisatellite observations, *J. Geophys. Res.*, 114(D6), D06305, doi:10.1029/2008JD011392, 2009.
- Jiménez, C., Clark, D. B., Kolassa, J., Aires, F. and Prigent, C.: A joint analysis of modeled soil moisture fields and satellite observations, *J. Geophys. Res. Atmos.*, 118(12), 6771–6782, doi:10.1002/jgrd.50430, 2013.
- 25 Kolassa, J., Aires, F., Polcher, J., Prigent, C., Jimenez, C. and Pereira, J. M.: Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology, *J. Geophys. Res. Atmos.*, 118(10), 4847–4859, doi:10.1029/2012JD018150, 2013.
- Kolassa, J., Gentine, P., Prigent, C. and Aires, F.: Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis, *Remote Sens. Environ.*, 173, 1–14, doi:10.1016/j.rse.2015.11.011, 2016.
- 30 Kolassa, J., [Reichle, R. H.](#), [Liu, Q.](#), [Gentine, P.](#), [Prigent, C.](#), [Aires, F.](#), Alemohammad, S. H., [Gentine, P.](#), [Aida, K.](#), [Asanuma, J.](#), [Bircher, S.](#), [Caldwell, T.](#), [Colliander, A.](#), [Cosh, M.](#), [Holfield-Collins, C.](#), [Jackson, T. J.](#), [Martínez-Fernández, J.](#), [McNairn, H.](#), [Pacheco, A.](#), [Thibeault, M.-H.](#), [Prigent, C.](#), [Aires, F.](#), [Prigent, C.](#) and [Walker, J. P.](#): Estimating surface soil moisture retrieval from SMAP observations using a Neural Network technique, *AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation*, *Remote Sens. Environ.*, 195, 202–217, doi:10.1016/j.rse.2017.10.045, 2017a, 2017.
- 35 Kolassa, J., [Gentine, P.](#), [Prigent, C.](#), [Aires, F.](#), [Reichle, R. H.](#), [Liu, Q.](#), Alemohammad, S. H., [Prigent, C.](#), [Aires, F.](#), [Prigent, C.](#), [Gentine, P.](#), [Aida, K.](#), [Asanuma, J.](#), [Bircher, S.](#), [Caldwell, T.](#), [Colliander, A.](#), [Cosh, M.](#), [Holfield-Collins, C.](#), [Jackson, T. J.](#), [Martínez-Fernández, J.](#), [McNairn, H.](#), [Pacheco, A.](#), [Thibeault, M.](#) and [Walker, J. P.](#): Estimating surface soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation, *SMAP observations using a Neural Network technique*, *Remote Sens. Environ.*, 195, 202–217, doi:10.1016/j.rse.2017.10.045, 2017b, 2018.

- Kool, D., Agam, N., Lazarovitch, N., Heitman, J. L., Sauer, T. J. and Ben-Gal, A.: A review of approaches for evapotranspiration partitioning, *Agric. For. Meteorol.*, 184, 56–70, doi:10.1016/j.agrformet.2013.09.003, 2014.
- Koster, R. D., Mahanama, S. P. P., Yamada, T. J., Balsamo, G., Berg, A. A., Boissarie, M., Dirmeyer, P. A., Doblas-Reyes, F. J., Drewitt, G., Gordon, C. T., Guo, Z., Jeong, J.-H., Lawrence, D. M., Lee, W.-S., Li, Z., Luo, L., Malyshev, S., Merryfield, W. J., Seneviratne, S. I., Stanelle, T., van den Hurk, B. J. J. M., Vitart, F. and Wood, E. F.: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment, *Geophys. Res. Lett.*, 37(2), n/a-n/a, doi:10.1029/2009GL041677, 2010.
- Kurum, M., Lang, R. H., O'Neill, P. E., Joseph, A. T., Jackson, T. J. and Cosh, M. H.: A First-Order Radiative Transfer Model for Microwave Radiometry of Forest Canopies at L-Band, *IEEE Trans. Geosci. Remote Sens.*, 49(9), 3167–3179, doi:10.1109/TGRS.2010.2091139, 2011.
- Leroux, D. J., Das, N. N., Entekhabi, D., Colliander, A., Njoku, E., Jackson, T. J. and Yueh, S.: Active–Passive Soil Moisture Retrievals During the SMAP Validation Experiment 2012, *IEEE Geosci. Remote Sens. Lett.*, PP(99), 1–5, doi:10.1109/LGRS.2015.2491643, 2016.
- Manfreda, S., McCabe, M. F., Fiorentino, M., Rodríguez-Iturbe, I. and Wood, E. F.: Scaling characteristics of spatial patterns of soil moisture from distributed modelling, *Adv. Water Resour.*, 30(10), 2145–2150, doi:10.1016/j.advwatres.2006.07.009, 2007.
- Markewitz, D., Devine, S., Davidson, E. A., Brando, P. and Nepstad, D. C.: Soil moisture depletion under simulated drought in the Amazon: impacts on deep root uptake, *New Phytol.*, 187(3), 592–607, doi:10.1111/j.1469-8137.2010.03391.x, 2010.
- Marthews, T. R., Dadson, S. J., Lehner, B., Abele, S. and Gedney, N.: High-resolution global topographic index values for use in large-scale hydrological modelling, *Hydrol. Earth Syst. Sci.*, 19(1), 91–104, doi:10.5194/hess-19-91-2015, 2015.
- Mascaro, G., Vivoni, E. R. and Deidda, R.: Soil moisture downscaling across climate regions and its emergent properties, *J. Geophys. Res. Atmos.*, 116(D22), n/a-n/a, doi:10.1029/2011JD016231, 2011.
- McCull, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S. and Entekhabi, D.: The global distribution and dynamics of surface soil moisture, *Nat. Geosci.*, 10(2), 100–104, doi:10.1038/ngeo2868, 2017.
- Merlin, O., Chehbouni, A. G., Kerr, Y. H., Njoku, E. G. and Entekhabi, D.: A combined modeling and multispectral/multiresolution remote sensing approach for disaggregation of surface soil moisture: application to SMOS configuration, *IEEE Trans. Geosci. Remote Sens.*, 43(9), 2036–2050, doi:10.1109/TGRS.2005.853192, 2005.
- Merlin, O., Chehbouni, A., Kerr, Y. H. and Goodrich, D. C.: A downscaling method for distributing surface soil moisture within a microwave pixel: Application to the Monsoon '90 data, *Remote Sens. Environ.*, 101(3), 379–389, doi:10.1016/j.rse.2006.01.004, 2006.
- Merlin, O., Chehbouni, A., Walker, J. P., Panciera, R. and Kerr, Y. H.: A Simple Method to Disaggregate Passive Microwave-Based Soil Moisture, *IEEE Trans. Geosci. Remote Sens.*, 46(3), 786–796, doi:10.1109/TGRS.2007.914807, 2008a.
- Merlin, O., Walker, J. P., Kalma, J. D., Kim, E. J., Hacker, J., Panciera, R., Young, R., Summerell, G., Hornbuckle, J., Hafeez, M. and Jackson, T.: The NAFE'06 data set: Towards soil moisture retrieval at intermediate resolution, *Adv. Water Resour.*, 31(11), 1444–1455, doi:10.1016/j.advwatres.2008.01.018, 2008b.
- Merlin, O., Walker, J. P., Chehbouni, A. and Kerr, Y. H.: Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency, *Remote Sens. Environ.*, 112(10), 3935–3946, doi:10.1016/j.rse.2008.06.012, 2008c.
- [Merlin, O., Al Bitar, A., Walker, J. P. and Kerr, Y.: A sequential model for disaggregating near-surface soil moisture observations using multi-resolution thermal sensors, \*Remote Sens. Environ.\*, 113\(10\), 2275–2284, doi:10.1016/j.rse.2009.06.012, 2009.](#)
- [Merlin, O., Escorihuela, M. J., Mayoral, M. A., Hagolle, O., Al Bitar, A. and Kerr, Y.: Self-calibrated evaporation-based](#)

- [disaggregation of SMOS soil moisture: An evaluation study at 3km and 100m resolution in Catalunya, Spain, Remote Sens. Environ., 130\(2013\), 25–38, doi:10.1016/j.rse.2012.11.008, 2013.](#)
- Moghaddam, M., Entekhabi, D., Goykhman, Y., Li, K., Liu, M., Mahajan, A., Nayyar, A., Shuman, D. and Teneketzis, D.: A Wireless Soil Moisture Smart Sensor Web Using Physics-Based Optimal Control: Concept and Initial Demonstrations, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 3(4), 522–535, doi:10.1109/JSTARS.2010.2052918, 2010.
- Moghaddam, M., Silva, A. R., Clewley, D., Akbar, R., Hussaini, S. A., Whitcomb, J. B., Devarakonda, R., Shrestha, R., Cook, R. B., Prakash, G., Santhana Vannan, S. k. and Boyer, A. G.: Soil Moisture Profiles and Temperature Data from SoilSCAPE Sites, USA. ORNL DAAC, Oak Ridge, Tennessee, USA, , doi:10.3334/ORNLDAAC/1339, 2016.
- Montzka, C., Jagdhuber, T., Horn, R., Bogen, H. R., Hajsek, I., Reigber, A. and Vereecken, H.: Investigation of SMAP Fusion Algorithms With Airborne Active and Passive L-Band Microwave Remote Sensing, *IEEE Trans. Geosci. Remote Sens.*, 54(7), 3878–3889, doi:10.1109/TGRS.2016.2529659, 2016.
- Morton, D. C., Nagol, J., Carabajal, C. C., Rosette, J., Palace, M., Cook, B. D., Vermote, E. F., Harding, D. J. and North, P. R. J.: Amazon forests maintain consistent canopy structure and greenness during the dry season, *Nature*, 506(7487), 221–224, doi:10.1038/nature13006, 2014.
- Njoku, E. G., Wilson, W. J., Yueh, S. H., Dinardo, S. J., Li, F. K., Jackson, T. J., Lakshmi, V. and Bolten, J.: Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99, *IEEE Trans. Geosci. Remote Sens.*, 40(12), 2659–2673, doi:10.1109/TGRS.2002.807008, 2002.
- O'Neill, P. E., Njoku, E. G., Jackson, T. J., Chan, S. and Bindlish, R.: SMAP algorithm theoretical basis document: Level 2 & 3 soil moisture (passive) data products. Jet Propulsion Lab., California Inst. Technol., Pasadena, CA, USA, JPL D-66480., 2015.
- O'Neill, P. E., Chan, S. K., Njoku, E. G., Jackson, T. J. and Bindlish, R.: SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center [Date Accessed: 30 May 2017], , doi:10.5067/ZRO7EXJ8O3XI, 2016a.
- O'Neill, P. E., Chan, S. K., Njoku, E. G., Jackson, T. and Bindlish, R.: SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center [Date Accessed: 30 May 2017], , doi:10.5067/OBBHQ5W22HME, 2016b.
- Peng, J., Loew, A., Merlin, O. and Verhoest, N. E. C.: A review of spatial downscaling of satellite remotely sensed soil moisture, *Rev. Geophys.*, 1–26, doi:10.1002/2016RG000543, 2017.
- Piles, M., Entekhabi, D. and Camps, A.: A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations, *IEEE Trans. Geosci. Remote Sens.*, 47(12), 4125–4131, doi:10.1109/TGRS.2009.2022088, 2009.
- Piles, M., Camps, A., Vall-llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y. H. and Walker, J.: Downscaling SMOS-Derived Soil Moisture Using MODIS Visible/Infrared Data, *IEEE Trans. Geosci. Remote Sens.*, 49(9), 3156–3166, doi:10.1109/TGRS.2011.2120615, 2011.
- Piles, M., Sanchez, N., Vall-llossera, M., Camps, A., Martinez-Fernandez, J., Martinez, J. and Gonzalez-Gambau, V.: A Downscaling Approach for SMOS Land Observations: Evaluation of High-Resolution Soil Moisture Maps Over the Iberian Peninsula, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 7(9), 3845–3857, doi:10.1109/JSTARS.2014.2325398, 2014.
- Rodríguez-Fernández, N., Muñoz Sabater, J., Richaume, P., de Rosnay, P., Kerr, Y., Albergel, C., Drusch, M. and Mecklenburg, S.: SMOS near real time soil moisture product: processor overview and first validation results, *Hydrol. Earth Syst. Sci. Discuss.*, (May), 1–23, doi:10.5194/hess-2017-211, 2017.

- Rodríguez-Fernández, N. J., Aires, F., Richaume, P., Kerr, Y. H., Prigent, C., Kolassa, J., Cabot, F., Jimenez, C., Mahmoodi, A. and Drusch, M.: Soil Moisture Retrieval Using Neural Networks: Application to SMOS, *IEEE Trans. Geosci. Remote Sens.*, 53(11), 5991–6007, doi:10.1109/TGRS.2015.2430845, 2015.
- Rodríguez-Iturbe, I. and Porporato, A.: *Ecohydrology of water-controlled ecosystems: soil moisture and plant dynamics*, Cambridge University Press., 2007.
- Roerink, G. ., Su, Z. and Menenti, M.: S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance, *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.*, 25(2), 147–157, doi:10.1016/S1464-1909(99)00128-8, 2000.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J.: Learning internal representations by error propagation., 1985.
- Salvucci, G. D.: An approximate solution for steady vertical flux of moisture through an unsaturated homogeneous soil, *Water Resour. Res.*, 29(11), 3749–3753, doi:10.1029/93WR02068, 1993.
- Salvucci, G. D. and Entekhabi, D.: Equivalent steady soil moisture profile and the time compression approximation in water balance modeling, *Water Resour. Res.*, 30(10), 2737–2749, doi:10.1029/94WR00948, 1994.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B. and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, *Earth-Science Rev.*, 99(3–4), 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.
- Shin, Y. and Mohanty, B. P.: Development of a deterministic downscaling algorithm for remote sensing soil moisture footprint using soil and vegetation classifications, *Water Resour. Res.*, 49(10), 6208–6228, doi:10.1002/wrcr.20495, 2013.
- Srivastava, P. K., Han, D., Ramirez, M. R. and Islam, T.: Machine Learning Techniques for Downscaling SMOS Satellite Soil Moisture Using MODIS Land Surface Temperature for Hydrological Application, *Water Resour. Manag.*, 27(8), 3127–3144, doi:10.1007/s11269-013-0337-9, 2013.
- Sun, J., Salvucci, G. D., Entekhabi, D. and Farhadi, L.: Parameter estimation of coupled water and energy balance models based on stationary constraints of surface states, *Water Resour. Res.*, 47(2), doi:10.1029/2010WR009293, 2011.
- Verhoest, N. E. C., van den Berg, M. J., Martens, B., Lievens, H., Wood, E. F., Pan, M., Kerr, Y. H., Al Bitar, A., Tomer, S. K., Drusch, M., Vernieuwe, H., De Baets, B., Walker, J. P., Dumedah, G. and Pauwels, V. R. N.: Copula-Based Downscaling of Coarse-Scale Soil Moisture Observations With Implicit Bias Correction, *IEEE Trans. Geosci. Remote Sens.*, 53(6), 3507–3521, doi:10.1109/TGRS.2014.2378913, 2015.
- Volk, M., Niklaus, P. A. and Körner, C.: Soil moisture effects determine CO<sub>2</sub> responses of grassland species, *Oecologia*, 125(3), 380–388, doi:10.1007/s004420000454, 2000.
- Wang, L., Good, S. P. and Caylor, K. K.: Global synthesis of vegetation control on evapotranspiration partitioning, *Geophys. Res. Lett.*, 41(19), 6753–6757, doi:10.1002/2014GL061439, 2014.
- Western, A. W. and Blöschl, G.: On the spatial scaling of soil moisture, *J. Hydrol.*, 217(3–4), 203–224, doi:10.1016/S0022-1694(98)00232-7, 1999.
- Wu, X., Walker, J. P., Rudiger, C., Panciera, R. and Gao, Y.: Intercomparison of Alternate Soil Moisture Downscaling Algorithms Using Active–Passive Microwave Observations, *IEEE Geosci. Remote Sens. Lett.*, 14(2), 179–183, doi:10.1109/LGRS.2016.2633521, 2017.
- Zacharias, S., Bogaena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G. and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, *Vadose Zo. J.*, 10(3), 955, doi:10.2136/vzj2010.0139, 2011.

**Formatted:** No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers