Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-191-SC1, 2019

ESSDD

Interactive comment

Interactive comment on "Gap-Free Global Annual Soil Moisture: 15km Grids for 1991–2016" *by* Mario Guevara et al.

Mario Antonio Guevara Santamaría

mguevara@udel.edu

Received and published: 31 December 2019

REVIEWER COMMENT: This paper presents a 15 km annual-average soil moisture product that is generated by machine-learning the relation between 0.25 degree ESA CCI soil moisture estimates and topographic indices derived from a higher-resolution DEM. I have several major concerns regarding the hypothesis/assumptions on which the methodology is based as well as the employed validation methodology, and consequently also the conclusions drawn from the presented analysis:

AUTHORS RESPONSE: We appreciate the reviewer comments as they provide valuable feedback to increase the impact of our work. We revised our work and will provide analyzes (including new datasets and analyzes of variable importance) to improve the

Printer-friendly version

Discussion paper

validation methodology. We have already started to update our work with new datasets (e.g., eco-climatic and soil type classes) and updated our results with the recently released version of the ESA-CCI soil moisture product (4.5, up to 2018). These results will be uploaded in a fully revised version.

To increase reproducibility of our prediction framework we have started compiling an R program (that will be available in a revised version) able to predict soil moisture based on remote sensing and machine learning, from the global to the country specific scales. This new code is able to perform a bootstrapping approach given a user defined sample size (if not specified, it will use 1/3 of available data/pixels for each year) to analyze the variance of model predictions as a function of variations in training data. Thus in a revised version we will report a spatial explicit metric of model-based uncertainty.

REVIEWER COMMENT: The methodology is based on the hypothesis that topography is a main driving factor for soil moisture patterns. However, the reference used to support this claim (Mason et al., 2016) presents only a very local analysis of differences between soil moisture values at low-slope and high-slope areas over grasslands only, and only in a small region over the UK. The observed relation is relatively low (RËĘ2 = 0.21) and the authors conclude "[...] a topographic signal can be seen in high resolution remotely sensed surface soil moisture data [...]. Unfortunately this signal is relatively weak." Moreover, Mason et al. (2016) uses 1 km SAR data for which topographic corrections are applied in the pre-processing, which likely induces a sensitivity of the measurements to topography parameters. These topography corrections are usually not applied to coarse resolution measurements such as the sensors used within the ESA CCI SM, because topographic effects average out at these scales.

AUTHORS RESPONSE: We agree in that more references could improve the main hypothesis driving this effort, so we will include a more detailed background of the main principles driving this approach in a revised version of our work. We argue that new alternative approaches are needed to provide different downscaling soil moisture outputs to compare data-model agreements and advance science.

We also argue that satellite-derived soil moisture data is retrieved by a direct measurement of the dielectric constant of soils representing specific vegetation types and climate conditions (within each pixel) that are also influenced by topographic patterns. We highlight that there is a high correlation (>0.8 in all our model/years) between topographic patterns and the ESA-CCI soil moisture product, which provides evidence that topography has information that could be used for downscaling soil moisture. This argument could be revised to explain how our machine learning approach can identify non-linearities between satellite-derived soil moisture and topographic variables.

We highlight that our main purpose was to generate a soil moisture downscaled product that is independent of climate or vegetation variables and that provides cross validated hypothesis of a more local (nearly 50% higher) spatial resolution compared with the original satellite soil moisture signal. Finally, we clarify that the use of a soil moisture product where vegetation is not used as predictor (i.e., vegetation variables are independent) could be relevant for decreasing spurious correlations in the further use of soil moisture data (in Earth models and other ecological or geo-scientific analysis).

REVIEWER COMMENT: The presented paper itself also does not analyze the predictive power of the used topographic indices for soil moisture (e.g., the godness-of-fit for the obtained regression, variable importance, etc.). Hence, there is no evidence supporting the reliability of topographic indices as predictor for soil moisture, especially on a global scale.

AUTHORS RESPONSE: We highlight that we have provided information about the predictive power of our models. In Table 1 we report indicators of accuracy and predictive power including the correlation between observed and predicted, the root mean squared error, the number of pixels with soil moisture data available for each year and the best parameters (kernel and k) for each kknn model/year. In our reanalysis the predictive power of our modeling approach increased thanks to the use of a more extensive set of prediction factors. We have also bootstrapped our statistical models and now they provide uncertainty estimates associated with the number of available data for modeling soil moisture patterns.

In a revised version we can include a variable importance analysis (by permutation) to support/generate new hypothesis about the spatial variability of soil moisture, in relation to land surface characteristics represented by multiple sources of environmental information. We can also provide a bootstrapping approach to estimate uncertainty associated with the number of available data for modeling soil moisture patterns.

REVIEWER COMMENT: Even more doubtable is the assumption that the developed regression function can be used to extrapolate soil moisture to regions not covered by the ESA CCI SM, which are mainly the arctic ice sheet and tropical forests. Tropical rainforests, for example, have a quite unique moisture regime that is expected to be largely rainfall dependent. It is very questionable to use a soil moisture - topography relation that is trained over non-tropical regions to predict soil moisture there. Moreover, no in situ measurements are available in these regions to verify the validity of these predictions.

AUTHORS RESPONSE: We agree that soil moisture estimates should be removed from the arctic ice sheet in a revised version.

However, we believe that our approach could be used for predicting soil moisture across tropical forests. We highlight that there are pixels with satellite soil moisture values across the tropical rain forests of the world, including the Amazon and Congo regions that could be used for prediction within our framework (Fig 1).

In a revised version we will update the soil moisture covariate space with information on ecological domains (FAO, 2010) and bioclimatic features (Fick and Hijmans, 2017), soil type variability (Wieder et al., 2014) and geo-spatial information (Møller, et al., 2019) of available data/pixels in order to improve our soil moisture prediction capacity. We believe that comparing and testing multiple modeling approaches across these areas is needed for improving prediction capabilities for downscaling soil moisture. We also believe that working towards data sharing and open source practices will increase the users of soil moisture information, increasing the possibilities to solve uncertainties and explain discrepancies between multiple soil moisture products.

REVIEWER COMMENT: Also, the presented validation does not support the conclusions. First, the statement (L234) "In all cases, the evaluation statistics are equal or better for the downscaled soil moisture predictions based on digital terrain analysis (Table 3) than the original ESA-CCI soil moisture product (Table 2)" is wrong. In fact, results are quite balanced, sometimes the downscaled product is "better", sometimes the original is "better", but most likely results are not actually distinguishable within reasonable confidence limits (which should be estimated). The authors do indeed acknowledge (L252): "The downscaled predictions based on digital terrain analysis are not significantly different compared with the ESA-CCI soil moisture product [...]", but the subsequent conclusions are not supported. Specifically, "[...] but they provide (1) gap free soil moisture-related information ": while they are provided, there is no evidence that they are of any reasonable accuracy (for the earlier discussed regions), and "(2) higher resolution (from 27 to 15 km grids)": This is a mix-up of resolution and sampling.

AUTHORS RESPONSE: We highlight that this is an empirical model (i.e., a statistical learning process driven by a machine learning algorithm) and we used a cross validation to show the accuracy in the change of resolution between the original satellite estimate and the prediction; thus it is not only a spatial re-sampling exercise.

We agree with the reviewer in that the results of the validation against filed information are balanced. There are no significant differences in the relationship between the downscaled estimate and the original product, which demonstrates that the downscaled product is preserving the reliability of the ESA-CCI soil moisture product. This is expected as the information that was used for building predictions of soil moisture were the pixel/values of the ESA-CCI soil moisture products. In the revised version of our data paper we will improve the narrative and the explanation on this comparison to avoid misunderstandings.

REVIEWER COMMENT: Improved resolution would imply that there is different / more information in the downscaled product, but the indistinguishability of performance metrics (see above) suggests that this is not the case.

AUTHORS RESPONSE: We argue that the 'indistinguishability of performance metrics' is a proof of the reliability of our prediction framework. Please note that we are using a data driven model were the selection of best parameters (e.g., distances, kernels, neighbors, predictors) is made by cross validation. Cross validation allows to generate unbiased residuals and identify the linear relationship between observed and predicted data. Therefore, it is expected that our predictions to maintain the general pattern (the similar mean and standard deviation) to the general pattern of the ESA-CCI product, but revealing a physiography modulated (not random) soil moisture spatial pattern across finer grids. The soil moisture variability estimated within each 15km pixel is meant to maintain the numerical integrity of the ESA-CCI product, as we also recognize that 15km grids is still too coarse to identify local controls of soil moisture.

In the revised version of our paper, we propose to increase the reproducibility of the soil moisture prediction framework across multiple scales (country-to-global) and highlight the gain of information between the original satellite and the downscaled predictions. We believe that the value of our downscaled product will be better recognized when the reader see the detail gained within a smaller region of the world (i.e., a country).

REVIEWER COMMENT: Also the comparison in Figure 6 (b and c) shows that the original and the downscaled products exhibit the exact same behaviour with a slightly lower overall variability in the original product. It is, however, not clear whether this different overall magnitude reflects an actual improvement, because soil moisture varaibility and trends are actually supposed to be different at a point scale and at a satellite scale (see e.g. Famiglietti et al. 2008).

AUTHORS RESPONSE: We clarify that we indeed found significant differences between the point scale (1:1), and the ESA-CCI soil moisture (∼27km) trends (comparing

Printer-friendly version

Discussion paper

only the pixels with field stations) (see Figure 6 of submitted manuscript version). We also found that our downscaled predictions (15km) has a better model-data agreement. We clarify that we only predicted at the places with available data from the ISMN. At these places, we found statistically different trends between the ESA-CCI soil moisture product, the downscaled predictions and the ISMN annual averages.

Although these trends are different, they are all negative and significant (as the confidence intervals are not overlapping zero values) in the three datasets. We found that the ISMN is showing the strongest negative trend, followed by the trend of our downscaled predictions and then the ESA-CCI soil moisture product. Thus, the trend reported by our predictions is closer to the trend of field stations compared with the trend from the ESA-CCI.

In a revised version we can include a comparison of the satellite soil moisture mean using both products and different geographical extents in order to enrich the discussion about the scale dependent variance of soil moisture.

REVIEWER COMMENT: Also, the soil moisture mean is supposed to be different at different scales, hence the negative bias between point and satellite measurements cannot be reliably interpreted as error, and a reduction of this bias may as well be a going in the wrong direction with respect to the true areal-average mean. In other words, even though the generated product is sampled on a higher-resolution grid, it can not be concluded that this product contains higher-resolution information. Given the low amount of evidence that topography (alone) is a go od predictor for soil moisture, observed differences may well be a result of the smoothing-nature of the KKNN approach, and any spatial-window resampling approach may lead to a seemingly "higher-resolution" (which is truly only a higher-sampling) product with the same (or even better) performance, but this is not tested.

AUTHORS RESPONSE: This is an important comment and we argue that is a science frontier that still needs active research. In a revised version we could elaborate about

Printer-friendly version

Discussion paper

how different publications address the scale-variance of soil moisture.

We have done a reanalysis of our approach using the recently released ESA-CCI soil moisture version 4.5 and we can confirm a large correlation between the ESA-CCI and our soil moisture predictions (>0.92) at the global scale. Continental to global scales may be consistent in the overall range of values and spatial patterns, however smaller regions may highlight higher differences (Fig 2).

REVIEWER COMMENT:Therefore, I recommend to reject this publication. However, I do believe that topography may well be an important complementary predictor for soil moisture at higher-resolution when combined with other dominant factors. I therefore encourage the authors to pursue this approach addressing the concerns outlined above.

AUTHORS RESPONSE: We are confident that we can address all concerns outlined by the reviewer in a revised version of this manuscript. We believe most of the concerns could be addressed by editing the text to improve clarification and performing new analyses about variable performance, using the recently released version 4.5 of the ESA-CCI, and demonstrate the applicability of our methods at higher temporal resolutions (months) and across smaller areas and spatial extents. Please note that our previous work has demonstrated the effectiveness of our approach at the continental scale of CONUS using 1km grids (Guevara and Vargas, 2019).

References:

Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land area, International Journal of Climatology, 37(12), 4302–4315, doi:10.1002/joc.5086, 2017.

FAO. Global Ecological Zones for FAO Forest Reporting: 2010 Update; FAO: Rome, Italy, 2012.

Guevara, M. and Vargas, R.: Downscaling satellite soil moisture using geomorphometry and machine learning, edited by B. Poulter, PLOS ONE, 14(9), e0219639, doi:10.1371/journal.pone.0219639, 2019.

Mascarro, G., Ko, A. and Vivoni, E. R.: Closing the Loop of Satellite Soil Moisture Estimation via Scale Invariance of Hydrologic Simulations, Scientific Reports, 9(1), doi:10.1038/s41598-019-52650-3, 2019.

Møller, A. B., Beucher, A. M., Pouladi, N. and Greve, M. H.: Oblique geographic coordinates as covariates for digital soil mapping, , doi:10.5194/soil-2019-83, 2019.

Wieder, W.R., J. Boehnert, G.B. Bonan, and M. Langseth. 2014. Regridded Harmonized World Soil Database v1.2. Data set. Available on-line [http://daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. http://dx.doi.org/10.3334/ORNLDAAC/1247.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-191, 2019.

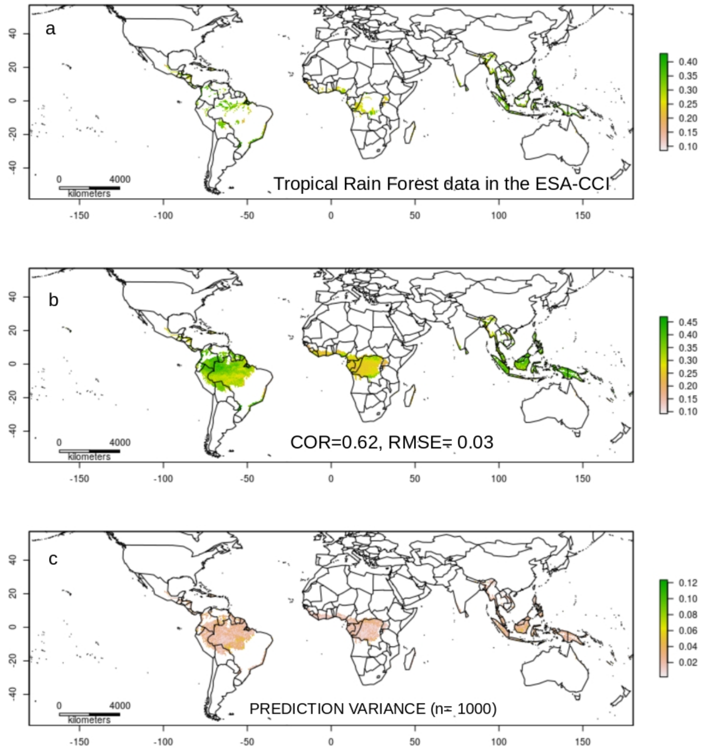

Fig. 1. Soil moisture available data in Tropical Rain Forests (ESA-CCI 4.5, 2018) (a). Soil moisture prediction (b) and soil moisture prediction variance (c).

Printer-friendly version

Discussion paper

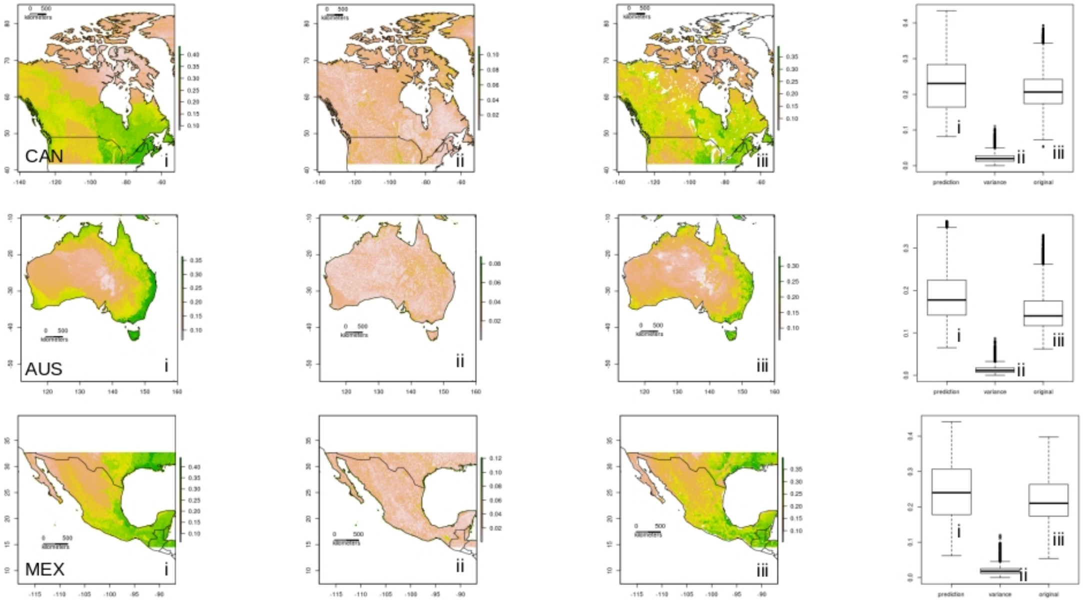

Fig. 2. Example of soil moisture predictions for 2018 (i), variances for 2018 (ii), the training data image for 2018 (iii) and their statistical distribution ('boxplots').

Printer-friendly version

Discussion paper