

Review Van Schrier et al: "European daily dataset of soil moisture from in situ meteorological observations"

Major comments

I admire the authors for taking on the challenge to optimize the use of existing observations of meteorological information for estimation of soil moisture for climate research.

However, I was overall somewhat confused as to what the aims of this paper were because it mainly focusses on the development of a new soil moisture dataset, but it also provides an analysis of historical drought indices and a climate assessment. Perhaps the paper would have been better split into two papers, so that more detail can be provided about the models and their data inputs (e.g. vegetation parameters? LAI, root zone depth etc.).

Based on the (i) verification analyses and related performance scores provided, and (ii) the at times tenuous interpolation methods, I am not sure that this dataset provides better soil moisture content (in fact, it could be worse) than what could be obtained from a similar analysis of fully coupled high-resolution weather/climate models, e.g. with ECMWF's IFS that provides products such as ERA-Land. ERA-Land produces sub-daily data of all water balance fluxes based on a mixture of observational (in-situ and remote sensing data) and widely-used and tested equations relating to atmospheric physics and land surface science. In this case all atmospheric variables, and related land surface water- and energy balance state variables and fluxes are coherently linked. See e.g.: <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=download>.

Can you show me somehow that your data do a better job than e.g. ERA-Land soil moisture data?

In my opinion, in this study, both SWAP and SoilClim seem to have been employed in sub-optimal, at times naïve, ways. If SWAP has the option to deal with groundwater then why not invoke this? Also, information regarding the model inputs is lacking, in particular regarding vegetation parameters.

Based on this, and more detailed comments below, unfortunately I do not feel that I this paper (and its dataset) should be accepted in its current form, and I am not sure what could be done to make this dataset more reliable. However, I have decided not to reject it outright, but to give the authors another chance.

Detailed comments

Lines 40-41: These drought indices need to be given in full, and I don't agree that SPEI "makes an attempt to capture the loss of moisture from the soils to atmosphere too". Certainly not an explicit attempt. For starters it uses potential evapotranspiration, not actual ET.

Line 132 Figure 1: You need to explain what the abbreviations rr, tg and qq mean in the legend. In the caption for Figure 2 it should be "meteorological in-situ data"

Line 172: It is good to see that you are not using the PM equation, but the Makkink equation instead. Would it have been interesting to see how different your soil moisture dataset would have looked if you would have used Penman-Monteith instead of Makkink?

Line 196: You say: “where a modification of these equations is used near saturation.” What is the nature of this modification and why is it required?

Line 197: How representative is a bottom boundary condition of free drainage for a large percentage of the lowlands of Europe where shallow groundwater plays a significant role? Did you test that for regions where you know that GW is important such as in large parts of the Netherlands where you have many soil moisture stations?

Line 198: What dataset or approach did you use to obtain “daily input of leaf area index, rooting depth and root density”?

Line 199: You mean ‘interception’ instead of ‘inception’? Give a reference (year) for this VHHB scheme? What were the constants used in this equation for the short grassland?

Table 1: Why do you not mention the source of the vegetation information?

Line 201: do you not want to mention crop water stress and how this is determined (too wet, too dry, too saline)?

Overall the description of the SWAP model has a random feel to it, and lacks proper referencing.

Line 211: For the SoilClim model you mention Allen et al. (1998). Does that mean that in this case you did use the Penman-Monteith equation? Why would you do that after the case you made earlier for Makkink? You need to explain the SoilClim model in more detail, you cannot expect the reader to go through the Hlavinka papers first?

Line 214: What soil information does SoilClim require, how does it use it and how did you obtain this?

Line 238: While your groundwater maps are interesting and while you seem to acknowledge that assuming free drainage is not the best approach for certain areas of Europe, you do not mention how this affects your data analysis approach? Will you omit these areas from the verification exercises?

Line 239-255: Field capacity and Wilting point play quite an important role in these indices, but you have not introduced these in your earlier text. How did you obtain these from the SoilGRIDS dataset and how were these used in the models (e.g. in plant water stress equations) other than in your drought indices.

Line 252-253: I have real trouble understanding this sentence: “In the phase space of soil water content values and the fineness of soil texture, possible values of soil water content are limited by the wilting point and the field capacity”, as well as Fig. 4. What parameter/variable did you use to express ‘the fineness of soil texture’?

Line 273-274: I cannot follow this sentence: “The collection of global radiation records and the transformation of sunshine duration to global radiation records, the former often go further back in time.” There are words missing? I also have trouble following the next sentences in relation to the approach by Suehrcke et al. and what I believe is its explanation in the sentence that follows.

Lines 282-284: I believe you are trying to say that you have used the Van Genuchten-Mualem equations (give proper references) with the parameters obtained via the pedotransfer functions detailed in Wosten et al. (1999). Can you show this equation and its parameters, ideally in proper symbols with units? What is the Tort parameter? The tortuosity? Do you mean parameter λ ?

Line 285: Explain first of all what AWC is and then that it needs correcting for gravel. Why do you need AWC?

Line 287-292: You are partly repeating points you have made earlier. And what is the rationale behind initialising the soil moisture profile in SWAP assuming a hydrostatic equilibrium with an initial groundwater level of -2 m, if you assume throughout that groundwater is deep and therefore free drainage applies?

Line 294: what is the meaning of ? in “‘Root-zone soil moisture’ (?)”?

Line 294-302: I thought that normally the surface layer is 0.10 cm thick and the the root zone soil moisture represents the layer below that? I.e. not including the top layer? You mention 0-40 cm for the root zone soil moisture?

Line 238-239: You say “This suggests that the poor correlation might be attributed to meteorological data that are not sufficiently representative for the FluxNet site”. You could have tested this by using the driving data that come with the FLUXNET sites?

Line 229-230: It is obvious that NL-Hor (Hostermeer, a wetland in a Dutch polder) is severely affected by groundwater (hence the high ET_a values), so why would you use this to verify your model that assumed free drainage and no groundwater effects? That explains the poor correlations for these kinds of sites... Also despite its elevation (1500 m) IT-MB0 is also a wetland so not suitable for model verification of a model assuming free drainage. This also explains the shape of the curves in Fig. 6c and especially d. The measured ET is maintained by groundwater, even during the height of summer. It is worrying that you did not pick up on this, but maybe I missed this.

Line 349-353: Any explanation about ISMN (and FLUXNET for that matter), should be done in the methods section, not in the results section.

Line 387, Fig. 7. These plots do not fill me with confidence that this is a reliable soil moisture product.

Line 393, Fig. 8: Why are you not showing an equivalent figure for the surface soil moisture contents?