Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-70-RC1, 2019

Interactive comment on "A Maximum Entropy Production Evaporation – Transpiration Product for Australia" *by* Olanrewaju Abiodun et al.

Anonymous Referee #1

Received and published: 17 September 2019

This paper addresses an important but complex issue, namely to estimate the rates of evaporation from soils and transpiration from vegetation over large areas (here the whole of Australia). However, this contribution suffers from multiple problems. The following comments are intended to highlight those issues and help the Authors to revise their manuscript.

Major comments:

1 Each and every input data set should be systematically characterized in terms of its nature, units of measurements, spatial and temporal resolution, accuracy (error bar), and source. Much of that information is currently lacking for most of the inputs used.

2 Assuming that the heat flux into the ground G is null for vegetated areas is a crude

approximation at best. At the very least, this should be supported by ample empirical evidence.

3 Soil moisture, derived from a mix of space-based instruments, is arguably the most important source of information in this study, and likely the driving factor that guarantees reasonable results. Yet, there is no discussion of the accuracy of that product, nor of the dependency of the output on that accuracy.

4 The Authors correctly point out, in the introduction, the difficulty of mixing and merging (let alone assessing) input data obtained at widely different spatial and temporal scales and resolution. Yet, little or no discussion of this key issue appears in the paper: Satellite-derived soil moisture data comes at a spatial resolution of 25 km. Re-sampling it at 5 km may be convenient programmatically, but that does not mean that this information is suddenly available (and reliable) at this finer resolution. Similarly, saying that field data obtained from flux towers have a footprint "ranging from 100 m^2 up to about 2 km^2 depending on the measuring height of the EC system and vegetation height" (lines 204–205) has to be taken at face value, because no evidence is provided to support such a claim. Hence, comparing those outcomes and claiming that one validates the other is a gross oversimplification of the matter.

5 The paper does not discuss the concept of potential evaporation (PE: the maximum rate of evaporation when water supply is not a limiting factor), nor does it provide a map of the annual mean precipitation over Australia. Yet, both of those variables constitute caps which evapotranspiration (ET) is not supposed to exceed... A map of precipitation for Australia is available from the Bureau of Meteorology at (http://www.bom.gov.au/jsp/ncc/climate_averages/rainfall/index.jsp). Comparing the ET output of MET with that map, it is not obvious that the mean evapotranspiration (first Figure 3 on p. 12) is everywhere lower than the precipitation at the same location...

6 The MEP model mobilizes multiple input sources as well as a large number of equa-

tions or algorithms with (often fixed) empirical parameters. The latter may have been derived for other locations or time periods. It is not clear what is achieved by this complexity, or to what extent each source actually contributes to the final outcome. Could a much simpler model account fo the bulk of the variability? What is gained by the complexity, especially when it involves constant parameters? When developing large models like the MEP, it is essential to document the relative importance of the main inputs and the sensitivity of the outputs to those inputs.

7 Although the study is carried out over a decade (2003–2013), the manuscript does not discuss very much the time evolution of ET during that period. Figure 4 does show some time series over periods ranging from 80 to about 350 days, but it is not clear whether those are for a particular year or averaged over the entire period (and if so, where do these occur in the calendar year?). In any case, why are those time series exhibited over such different periods? For similar reasons, the enigmatic second Figure 3 (on p. 12) barely addresses the issue of confronting the time evolution of ET generated by MEP with actual measurable evidence. Options include comparisons with agricultural output, drought and flood periods, or any other biogeophysical variable that could betray the impact of ET fluctuations during that decade.

8 Lastly the legends of the Tables and Figures are so minimalist as to be largely useless to understand their contents. Please provide essential information to understand and appreciate the nature and contribution of those displays.

Editorial and substantial remarks (in sequential order):

- Lines 56–63: Replace "accepted" by "used": MODIS products are widely used because they are available and accessible, and because standard tools exist to manipulate the large data sets. The word "accepted" implies a vetting of the quality and performance of the product which may or may not apply, as indicated by the Authors themselves (lines 58 to 64).

- Line 68: This model is called 'Maximum Entropy Production', but there is no indication about what is maximized, or what connection may exist with the concept of entropy.

- Line 71: Clarify that MEP requires the specific humidity *of the air*.

- Lines 94 and 102: Indicate explicitly that q_s refers to the air specific humidity: the expression 'surface specific humidity' can be ambiguous.

- Lines 94–95: The text mentions q but the equation uses q_s. Adding to the confusion, the subscript s is used throughout the paper to designate soil variables, while q appears to be an atmospheric variable...

- Lines 95–97: Equations (3) and (4) show that the heat flux into the ground G and the latent heat flux E_s are both proportional to the sensible heat flux from the soil surface to the atmosphere H_s, but there is no indication about how the latter is actually estimated. Is it assumed to be the residual in Equation (1)? If so, what about the sensible heat flux from the plant canopy to the atmosphere H_v? This system of equation does not appear to be complete or energy conserving as stated. What is the accuracy of those equations, and to what extent are they (in particular the inverse Bowen ratio β_s) relevant and applicable to the Australian environment, given the presence of empirical coefficients, which may have been fine-tuned for US conditions? Where are the values of q_s and T_s coming from, what are their accuracy, spatial and temporal resolutions?

- Lines 111–114: Comments on the flowchart in Figure 1:

* This Figure describes how the various input data sets are processed by MEP to generate the desired output, evapotranspiration. Yet, none of those boxes appear to require any of the fluxes H, E or G mentioned earlier: only R_n is discussed or used subsequently. So what is the relation, if any, between the materials and equations described between lines 89 and 108, on one side, and the rest of the paper?

* Figure 1 distinguishes between "soil surface relative humidity" and "soil surface specific humidity", yet the equations or algorithms to derive those variables are nowhere to be seen...

* According to Equation (2), the variable soil sigma (σ_s) is derived from both soil surface temperature T_s and atmospheric specific humidity q_s, yet, the Figure does not show this latter dependency.

- Line 118: What are the source, accuracy, spatial and temporal resolutions of R_n?

- Line 119: FPAR is mentioned here (on page 6), but that product does not appear as an input in Figure 1. Also, it is not clear what its role is at this point (this is only clarified on page 20!–see the comment on Line 292 below).

- Line 121–125: The text appears to use the expressions "vegetation cover" and "vegetation fraction" as synonyms. This is confusing, as "vegetation cover" could be understood as "land cover", which typically refers to the type of ecosystem (forest, savanna, etc.) while "vegetation fraction" hints to the more appropriate concept of "fractional vegetation cover". This latter variable is itself generally poorly defined (and hard to estimate when the Leaf Area Index is less than 3), as is often the case in arid environments like Australia. In any case, if the fractional vegetation cover F_c is derived from FPAR, then Figure 1 should show FPAR in the crossed-out box, and F_c should appear in a separate box.

- Lines 59–61 and 131: There is an internal inconsistency in first decrying the poor quality of the MODIS temperature product and in using it nevertheless. This Reviewer cannot comment on the value of this product, but if the Authors estimate it is incorrect, they should use another one.

- Lines 131–134: The rationale for using the lowest T_s during the month is dubious: what is the accuracy and reliability of that product, generated on an 8-day basis (according to Line 59), if the area of interest happens to be overcast on successive 8-day periods? And even if T_s is always observed at least once a month, that measurement would necessarily occur on a clear (relatively hotter) day, so there is still a bias towards high temperatures during cloudy days. What is the possible impact of that bias?

- Lines 135–149: This whole paragraph appears very confusing, because it amounts to a somewhat disparate assemblage of algorithms and equations found in the literature, using multiple empirical coefficients which may or may not be applicable to Australia. Why mention methods that are not used? And again, those tools depend on additional soil properties (whose accuracy, spatial and temporal resolutions are unknown, by the way) apparently obtained from the Australian Soil Resource Information System (ASRIS), though the latter is not reported as an input in Figure 1. As the Authors rewrite this paragraph, they would be well advised to describe the necessary steps in logical order, to explicitly provide the full equations, to indicate clearly what inputs are needed, where they come from, their spatial and temporal resolutions, accuracy, etc.

- Line 140: The text appears to indicate that the values of the soil water content at wilting point and at field capacity are fixed in space and time to -1.5MPa and -10kPa, respectively. What evidence is there that this is reasonable, given the high diversity of soil types and properties?

- Line 145: What exactly is implied here by "modest data requirement and relative accuracy"? Constant values would be even simpler... This is not an acceptable rationale: the selection of models and parameterizations should be guided by strict requirements in terms of accuracy and performance to achieve a particular objective, not in terms of data volume or approximate value. Besides, how small should a database be to be "moderate", and what is the benchmark to evaluate the "relative accuracy"?

- Line 151: The authors introduce the concept of "distance z above the target surface for which the Monin-Obukhov similarity theory is valid in the formula of the thermal inertia of turbulent air above soil surface", but none of those Equations are provided. A Table of (fixed) values is provided, and a map of z over Australia is produced, but what are the significance and implications of this variable in the MEP model? Since 90% or more of that continent is covered by low vegetation anyway, does it matter to consider this parameter?

- Line 155: What is the scientific basis for assigning those particular values in Table 1? Note also that the text refers to z as a "target distance", while the legend appears to refer to a "target surface".

- Line 158: In Figure 2, why is there a blue block corresponding to z values of 11 m when Table 1 does not show any land cover with that value? And why is the legend to this figure labeling z as the 'target height (z)', which would be normally understood as the height of vegetation, rather than the theoretical concept mentioned above?

- Line 161: The text refers to Eq. 14, but that Equation (on line 197) has nothing to do with "the Hutson and Cass coefficients a and b": this probably refers to an equation in another paper...

- Line 166: Where does the soil density come from? What is its accuracy and spatial resolution?

- Lines 171–181: The ESA CCI Soil Moisture (SM) product is delivered at the spatial resolution of $0.25°$. Resampling it (by duplication or interpolation) on a grid at $0.05°$ may be programmatically convenient or necessary, but that does not change the intrinsic nature or properties of the product! Also, which version of that product is actually used here, since it has already been released 6 times?

- Lines 186–187: Please provide an explicit description of the interpolation procedure as well as the formula to calculate the specific humidity of the atmosphere using the Clausius-Clapeyron equation: any reader should be allowed to duplicate the work without having to guess which tools, techniques or equations should be used.

- Line 200: The text mentions the variables x_n and y_n, but there are no such variables in the equations above, and it is not clear what distinguishes them anyway.

- Line 201: The diacritical marks on top of the Q symbols are barely distinguishable on this line, and almost impossible to differentiate in the Equations above (Lines 193–198). Please use other signs or reformat or enlarge the Equations.

Printer-friendly version

Discussion paper

- Lines 204–205: The wording here may be misleading: The "MEP ET product" may be technically available "at 5 km^2 resolution", but to the extent the space-derived soil moisture data is the dominant input, the actual spatial resolution may be closer to 25 km by 25 km or 625 km^2, as noted in the comment above (Lines 171–181). Similarly, the statement "tower flux data with footprints ranging from 100 m^2 up to about 2 km^2" is purely gratuitous, in the absence of evidence, or algorithms, to support these estimates. Hence comparing these products and claiming that this is a validation is unwarranted, on the basis of the information presented here.

- Lines 229–233: Please note that there are two Figures labeled "Figure 3" on this page.

- Line 231: In the second Figure 3, what is the source of those data? What is the area concerned (the whole of Australia)? Is this the result of an accumulation of data from rain-gauges, or a satellite product, or a reanalysis? What is the accuracy of those precipitation estimates? What is the spatial representativity of the precipitation data? Do they provide a spatial coverage comparable with the gridded data of MEP? And of course, most importantly, what is the error bar associated with the MEP estimates?

- Line 247: The legend of Table 2, or the text (or both), should indicate that the numbers set in bold face in this table point to the best performing method to evaluate the evapotranspiration rate at the various Eddy Covariance (EC) sites, according to the criteria indicated in the table header.

- Line 292: At long last, the text reveals that "The FPAR product [is] used in partitioning net radiation between soil and canopy". However, this is basically incorrect and inappropriate: FPAR is a measure of the effectiveness of the vegetation in absorbing photosynthetically active radiation (PAR), not in any way an indicator of the fractional vegetation cover. A particular value of FPAR derived from satellite observations could be obtained from a wide range of ecosystems with widely varying F_c and Leaf Area Index (LAI), not to mention other possible factors.

- Lines 307 and 335: Table 3 is mentioned in the text before Table 2, and no Table is labeled 3 in this document. However, there is a Table 2 starting on Line 244, and another one starting on Line 335...

- Line 312: The Figure referred to by "as a percentage of rainfall (Fig 2)" exhibits the spatial distribution of z, not precipitations. In fact, there is no precipitation map in the paper, although one would be useful, as noted above.

- Line 346: The statement "The MEP model appears lacking spatial continuity, probably due to the use of pedotransfer functions..." is invalid. Those functions are just mathematical formulae, fitting functions: they cannot by themselves generate spatial discontinuities. If the MEP model outputs appear spatially discontinuous, it must be because the soil moisture input data themselves are discontinuous, or because of model (coding) errors. For this reason too, it would be useful to conduct sensitivity analyses to establish to what extent each input and algorithmic parameter or equation contributes to the outcome. If the discontinuity does arise from the intrinsic variability of the soil moisture data, no amount of tuning of the pedotransfer functions will reduce those discontinuities. A contrario, if the soil moisture data are reasonably homogeneous to start with, then there may indeed be a problem with the way those functions, or other aspects of the model, are implemented. In either case, the conclusion that "Hence, further improvement to the MEP model may be achieved by improving the parameterization of the pedotransfer functions for each soil type." is currently unfounded.

- Line 351: The text states "The low correlation of the MEP model" but does not indicate with what MEP does not correlate well.

- Lines 410–411: The acknowledgment mentions the use of FLUXNET and ERA-Interim reanalysis data: Why are those data sets not mentioned in the text and appear in Figure 1?

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-70, 2019.