

The paper by Tomas-Burguera et al. develops a process to gap fill, homogenize, and grid historical climate data to be used to calculate a weekly 1.1 km gridded estimate of reference evapotranspiration using the modified Penman-Monteith equation. The data is made available in netCDF format via an online repository and the authors provide an online visualization and extraction tool. An overview of the review is provided here with specific comments embedded in the attached annotated manuscript. I was able to access the data via the link provided in the paper and downloaded the evapotranspiration (ET_o) data. It was relatively easy to access the data using MATLAB netCDF tools and the data that I accessed seemed to be usable. The link to the visualization tool provided also worked quite well for viewing and querying the data although I did not use the online tool for downloading.

Overall, the organization of the paper is acceptable and the data and methodology description were understandable. However, the language and paragraph structure needs a lot of work before this paper can be published. I think several key points are confused by language and sentence structure. There are a few key points that I would like to see addressed:

Authors: We would like to thank the reviewer for his valuable comments and his efforts helping the authors to improve this paper.

The use of language has been carefully revised in order to improve the paper, and we addressed all the suggested points, both the three key points as well as all the points in the supplementary file.

1) Please add a paragraph in the introduction to explain what “reference evapotranspiration” actually is and why it is considered a standardized method (with references). I think this is a key point that justifies your methodology and validates your dataset and deserves more than a brief mention.

Authors: We added the following sentences to introduce the concept of reference evapotranspiration:

‘Reference evapotranspiration (ET_o) is a theoretical variable describing the evapotranspiration that would occur from a well-watered reference surface under specific meteorological conditions (Allen et al. 1998). Because well-watered conditions and a reference crop are assumed, both spatial and temporal ET_o variability depends solely on the variability of the meteorological conditions. Hence, ET_o is an accepted proxy for the atmospheric evaporative demand (AED), which is a key variable for understanding both water and energy terrestrial balances and, therefore, relevant to a variety of disciplines, including climatology, hydrology, and agronomy (Espadafor et al. 2011).’

2) I think that your justification for using FAO-PM for calculating ET_o should be stronger and better organized in the Introduction and following this, should be discussed in greater detail in the Discussions and conclusions section.

Authors: We added a sentence in the Introduction to justify the use of FAO-PM, by adding some references to authors who validated the use of FAO-PM against lysimeters data. We also added a paragraph in the Discussions and conclusions section (see answer to point 3).

‘The main advantage of this method is that it is physically based. It has also been tested against lysimeters data obtaining reliable results (Jensen et al. 2000, Itenfisu et al. 2000, Berengena y Gavilan 2005, Trajkovic, 2007).’

New references:

Berengena, J. and Gavilán, P.: Reference evapotranspiration estimation in a highly advective semiarid environment, *Journal of Irrigation and Drainage Engineering*, 131, 147-163

Itenfisu, D., Elliot, R, Allen, R and Walter, I.: Comparison of reference evapotranspiration calculations across a range of climates, in: *Proceedings of the 4th National Irrigation Symposium*, pp. 216-227, St Joseph. asae edn., 2000

Jensen, M., Burman, R. and Allen, R: Evapotranspiration and irrigation water requirements, in: *ASCE manual No. 70*, p.332, New York, asce edn, 1990

Trajkovic, S.: Hargreaves versus Penman-Monteith, *Journal of Irrigation and Drainage Engineering*, 133, 38-42

3) It is my understanding that even though FAO-PM is the recommended methodology for estimating ETo, it does have some issues and limitations. I would like to see this addressed more in the Discussions and conclusions section, especially in regards to climate conditions in Spain and how those limitations may impact your data.

Authors: A paragraph discussing possible effects of using PM has been added to the discussion section:

‘Calculating ETo using PM assumed a well-watered reference surface, which can differ significantly from the actual conditions present in a semiarid region, as is the case across most of our study area. A scarcity of soil moisture can decrease the air humidity and increase the air temperature compared with well-watered conditions due to the effects of the land-atmosphere continuum. Both changes, which especially affect the aerodynamic component of ETo, may have a noticeable effect on ETo, meaning that an overestimation can occur under semiarid conditions (Bouchet 1961, Allen et al. 1998). Such an overestimation would be higher during the warm season when these conditions prevail. The possible overestimation due to the use of PM in a semiarid environment should be considered by potential users of this database.’

Find here the answer to all the *Specific Comments*

Reviewer (R): Page 1. Line 15.- I believe “Budyko” should be capitalized and I think that it should be “the Budyko curve”

Authors (A): We agree. We changed the use of ‘budyko’ for ‘the Budyko curve’

R: P1L20.- Could you add a paragraph to explain what “reference evaporation” actually is and why it is considered a standard (with references). I think this is a key point for justifying the validity of your data.

A: We added the following two sentences to introduce the concept of reference evapotranspiration:

‘Reference evapotranspiration (ET_o) is a theoretical variable describing the evapotranspiration that would occur from a well-watered reference surface under specific meteorological conditions (Allen et al. 1998). Because well-watered conditions and a reference crop are assumed, both spatial and temporal ET_o variability depends solely on the variability of the meteorological conditions. Hence, ET_o is an accepted proxy for the atmospheric evaporative demand (AED), which is a key variable for understanding both water and energy terrestrial

balances and, therefore, relevant to a variety of disciplines, including climatology, hydrology, and agronomy (Espadafor et al. 2011).'

We also added some references to authors who validated the results of ET_0 against lysimeter data:

'The main advantage of this method is that it is physically based. It has also been tested against lysimeters data obtaining reliable results (Jensen et al. 2000, Itenfisu et al. 2000, Berengena y Gavilan 2005, Trajkovic, 2007).'

New references:

Berengena, J. and Gavilán, P: Reference evapotranspiration estimation in a highly advective semiarid environment, Journal of Irrigation and Drainage Engineering, 131, 147-163

Itenfisu, D., Elliot, R, Allen, R and Walter, I.: Comparison of reference evapotranspiration calculations across a range of climates, in: Proceedings of the 4th National Irrigation Symposium, pp. 216-227, St Joseph. asae edn., 2000

Jensen, M., Burman, R. and Allen, R: Evapotranspiration and irrigation water requirements, in: ASCE manual No. 70, p.332, New York, asce edn, 1990

Trajkovic, S.: Hargreaves versus Penman-Monteith, Journal of Irrigation and Drainage Engineering, 133, 38-42

R: P2L4.- I like “especially” better since it is more formal. Suggest changing throughout.

A: We changed ‘*specially*’ for ‘*especially*’ throughout the text.

R: P2L6.- The transition into this paragraph is awkward and repetitive. Merge this with the previous paragraph.

A: See answer to the next question.

P2L8.- The reviewer doesn’t like the use of “... difficult the generation ...”

A: We merged both paragraphs and we changed the whole redaction:

'Although the maximum and minimum air temperature are commonly collected at weather observatories, observations of the other variables are scarce, especially if long time series are required for climate studies (Vanderlinden et al. 2004, McVicar et al. 2007, Irmak et al. 2012, Vicente-Serrano et al. 2014a) or to generate ET_0 grids. The other significant problem facing the generation of ET_0 climate grids is the changing number of observations, which can introduce non-climatic changes in variance (Beguería et al. 2016).

R: P2L11. Methods of what?

A: We changed the expression ‘*use of less demanding methods*’ for ‘*use of methods for calculating ET_0 requiring fewer climatic variables, commonly known as "less demanding methods"*’

R: P2L12. Again, methods of what?

A: In order to clarify, we changed the sentence.
'The use of methods for calculating ET_0 requiring only temperature data'

R. P2L15. This sentence is a bit confusing. Try rewording it

A: We changed the sentence *‘Nevertheless, this strategy is not recommended as methods not using data for all climatic variables are not able to deal with the variability and/or trends of missing variables and they could lead to erroneous conclusions ()’* for:

‘One of the major drawbacks of these methods is that variability and trends in the estimated ET_o values depends only on temperature, regardless of the importance of the other variables ()’

R. P2L20 I think you need a sentence or two here or above explaining what FAO-56 is, with references. You then, in the next sentence tell the reader that the method should be avoided but the reader may not know what they are avoiding. Clarify this.

A: We added some sentences in the introduction to explain what reference evapotranspiration is and why FAO-56 is a recommended method. In this specific line, we also added the following words: *‘in the FAO-56 document, which is the FAO document describing the guidelines for computing ET_o , ...’*

R. P2L23. Which problems, why are they similar?

A: The sentence reads now:

‘First, they use stationary relationships between variables that were empirically derived, which can be problematic in the context of climate change since these relationships may also change. This is in fact the same problem that affects the less demanding methods, which also rely on empirically derived relationships (Tomas-Burguera et al. 2017).’

R. P2L24. Clarify why this is a problem

A: we added *‘limiting the number of locations from which ET_o can be obtained’*

R. P2L32. These next two paragraphs seem disjointed and make for bad flow. They seem to be attached to the “Estimation of missing data” section but should be separate. Also, please re-read these paragraphs for grammar and English.

A: We added a white line to avoid confusion. We re-wrote and joined the two paragraphs, and now they are:

‘The changing number of observations available over time is another relevant problem affecting the generation of ET_o climate grids. To avoid negative effects, usually only the longest climate time series are used to generate climate grids using geostatistical methods, such as universal kriging (UK). Obviously, this strategy diminishes the number of usable climatic observations’

R. P3L12. IC is defined above but I don’t think PM is defined.

A: We added the definition of PM in page 2, Line 20. *‘... Penman-Monteith (PM) is calculated.’*

R. P3L16. Minor thing but it would sound better if you used a different word here.

A: We changed *‘implemented’* for *‘designed’*.

R. P4L15. A couple more sentences here on how you detected bad values would be appreciated. Range detection? Unrealistic jumps? Nearest neighbour?

A: We added some sentences to better explain the quality control:

'The quality of the data were assessed by implementing an automated daily quality control in R. Daily data were tested against two types of controls:codification errors and out-of-range values. The presence of duplicate data or n consecutive days having the exact same values in different observatories were the two most relevant codification errors detected. Out-of-range values mainly detected out-of-physical-range-values and out-of-climate-range values. More details can be found in (Tomas-Burguera et al. 2016)

R. P4L17. These two sentences are clumsy and should be part of the previous paragraph.

A: We changed the sentences in order to clarify the content:

The temporal aggregation of daily data into weekly data was then executed. For all variables, weekly time series were obtained by calculating the mean value of the daily data. Weeks presenting more than one day without data were considered to have no data. This is an adaptation of the WMO rules for monthly data (WMO, 1989)'

R. P5L3. Were there any gap size limits used in this process?

A: All the weather stations available were used in the gap filling process. The last step of the process is a data selection according to the number of original data. In the original version of the manuscript, the data selection was explained in the following paragraph. We mixed the two paragraphs and reword some sentences in order to clarify the process regarding the gap size limits:

'All weather stations available were used in the gap filling process. The last step of the process involved data selection and depended on the amount of original data available. For temperature, only time series accounting for more than 25 years of the original data were used. For the rest of variables, this period was reduced to 15 years due to the low availability of long records (Figure 3) Up to three gap filling loops were implemented for less frequent variables (sunshine duration, dew point temperature, and wind speed). Various steps in the gap filling procedure took advantage of non-overlapping data. This configuration was used previously to generate other databases over Spain (Gonzalez-Hidalgo et al. 2015).

R. P5L13. It would be good to have a sentence or two about this process so that the reader doesn't have to pull the reference to get the idea.

A: We added two sentences to explain the basis of the method:

'This method used as a basis the comparison of the time series to be homogenized, the candidate series, and a reference time series. Reference time series were obtained using the same process used to obtain the gap filling reference time series.'

R. P5L19. This doesn't need to be its own paragraph.

A: Done.

R. P5L29. Redundant.

A: Deleted.

R. P5L29. If this is done before the interpolation, why do you discuss it afterward? Out of order.

A: Obtaining the semi-variogram is the first step of the interpolation process. In order to clarify, we rewrote the sentence:

'As a first step in the interpolation process, a semi-variogram model was generated. This model was unique for each time step and each climatic variable'

R. P5L32. You changed this to UK above. Be consistent.

A: Done.

R. P6L1. Reword.

A: We deleted the expression 'At this point' and we merged it with the previous sentence.

R. P6L11. Attach to previous paragraph.

A: Done.

R. P6L21. What is the advantage and/or utility for splitting the components?

A: We added an example of one specific situation in which having data of the two components separately can benefit the user:

'A variability and trend analysis could benefit from the availability of the two components. For example, wind stilling and solar brightening have opposite effects in ET_o , but studying the two components separately facilitates the study of the impacts of each one on ET_o .'

R: P6L23. This is already stated before.

A: Deleted.

R: P7L17.

A: We changed 'reaching' for 'yielding'.

R: P8L4. Doesn't need to be a new paragraph. Reword these sentences to remove the redundancy.

A: Done. The new paragraph reads:

During the last part of the period (2010-2014), a high number of AWS were installed. A sharp increase in the available RH and W data was observed during this period, compared with the data available from weather stations used to generate the original database (Table 2). The values of these observations and the values of the climate grids were compared directly to obtain the relative humidity and wind speed over the 2010-2014 period using the new stations as an independent dataset.

R: P8L8. Gap Filling section. To me, the order of the paragraphs in this section is reversed. You should first present the results and then discuss why you see them.

A: We moved the first paragraph to the end of the section.

R: P8L10. Important why?

A: We changed 'important' for 'large'.

R: P8L14. Implications?

A: We changed some sentences in order to clarify this point:

‘The wind speed provided the lowest amount of filled data. It was difficult to obtain highly correlated time series to fill in the gaps, which had two major effects in the process: i) the probability of obtaining a reference time series from the neighbors was decreased; and ii) the reconstruction was poor when the reference time series could be obtained. The low correlation of the wind speed time series was a consequence of i) the high spatial and temporal variability of this variable and ii) the low number of observations available.’

R: P8L16. Reword or just remove ‘Evidently’.

A: Removed.

R: P8L19. ...shows an r^2 for the adjustment of ... Also, you start calling the gap filling exercise here an “adjustment”. It’s a bit of a leap in semantics so if you are going to call it an “adjustment”, you should probably lead into this in the previous sentences.

A: Changed:

‘Which showed an r^2 of only ...’

Also, we changed the word ‘adjustment’ for other expressions in the three cases in which we used this word.

R: P8L19. Have ME and PBIAS been defined anywhere?

A: DONE

‘Mean Error (ME) and Percent Bias (PBIAS)’

R: P8L21. Reverse the number and the acronym for this to make more sense, complete throughout

A: Done.

R: P8L24. Speculation as to why?

We hypothesize that the variance increase in gap-filled wind time series is due to the right-skewed nature of this variable. The gap-filling method that we used undergoes standardization of the data to avoid biases when a reference series is computed from neighbor observatories. This method has been tested and works very well with temperature data and other similar variables such as dew point temperature. However, it is possible that in the case of wind speed, which has a right-skewed distribution, the method does not work as expected and generates a slight expansion of the variance. This is an issue that would need further research, though.

R: P8L27. Reword this entire sentence.

A: Done. The text now reads:

‘For wind speed, the most recent decade showed slightly higher R^2 value than the first decades of the period.’

R: P8L30. Do you mean the quantity of data? You should say that.

A: Done. The text now reads:

'The percentage of data affected by the homogenization process exceeded...'

R: P9L4. This whole paragraph is confusing. Reword and be less vague.

A: Done. The new text reads:

'The temporal evolution of the quantity of data detected as inhomogeneous was analyzed (Figure 4), revealing a temporal trend with maximum values at the start of the study period and minimum values at the end. The most likely explanation for this observation is the use of more recent conditions as the standard conditions.'

R: P9L7. Consideration??

A: We changed the sentence:

'Another effect of this assumption is the...'

R: P9L9. This sentence needs some work.

A: We changed the sentence, which now reads:

'The maximum and minimum temperature, which displayed a positive trend in Spain over the study period (DelRio et al. 2012, Gonzalez-Hidalgo et al. 2016), suggested that higher values occurred in the present than in the past. A positive bias was observed in the homogenized data over the first decades. Unlike the maximum and minimum temperature, the wind speed, which displayed a negative trend (Azorin-Molina et al. 2014), was affected by a negative bias during the first decades of the study period.'

R: P10L4. Reaching the recent decades greater values of R2 ???

A: We changed the sentence:

'A temporal analysis of the R^2 values obtained from the spatial validation of the maximum and minimum temperature (Fig. 8) showed slightly better statistics (i. e., closer to one) in recent decades'

R: P10L22. Discussion and conclusions. In your discussion and conclusions section, you discuss limitations in the data for calculating ETo from the gridded climate data but you didn't really discuss any limitations of using PM ETo in applications. I would like to see a paragraph in this section that contains a discussion of any limitations to the ETo database as derived from PM.

A: A paragraph discussing possible effects of using PM has been added to the discussion section:

'Calculating ETo using PM assumed a well-watered reference surface, which can differ significantly from the actual conditions present in a semiarid region, as is the case across most of our study area. A scarcity of soil moisture can decrease the air humidity and increase the air temperature compared with well-watered conditions due to the effects of the land-atmosphere continuum. Both changes, which especially affect the aerodynamic component of ETo, may have a noticeable effect on ETo, meaning that an overestimation can occur under semiarid conditions (Bouchet 1961, Allen et al. 1998). Such an overestimation would be higher during

the warm season when these conditions prevail. The possible overestimation due to the use of PM in a semiarid environment should be considered by potential users of this database.'

R: P10L26. This would be more appropriate in the Introduction rather than in the Discussion section.

A: It has been moved to the introduction.

R: P11L4. '... based in first interpolate' ??

A: We changed the sentence:

'The PM-IC strategy, which consisted of interpolating climatic variables prior to calculating ETo'

R: P11L8. Misused. Not sure if this is the right word.

A: We changed 'misused' for 'not used'.

R: P11L11. Reword.

A: We changed the sentence:

'As 80% of the ETo variability was related to the variability in temperature and radiation (Mendicino and Senatore, 2013; Samani, 2000), using as many temperature observations as possible was important for ensuring the quality of the obtained results.'

R: P12L5. Reword this second part of this sentence.

A: We changed the sentence:

'This dataset was first developed as an input to generate, in combination with the precipitation data, grids of drought indices over the study area (Vicente-Serrano et al. 2017). Due to the relevance of ETo and the high number of possible uses of these data, the ETo climate grid is now being made available to other research groups.'

R: P12L8. I think this should be joined to the previous paragraph in order to make sense.

A: Done.

R: P12L14. Reword these sentences.

A: We reword the previous sentences. Now:

More accurate models of ETo is also useful for rainfed agriculture. Hence, the whole agricultural sector could benefit from this dataset.

R: P12L17. This sentence is very hard to read.

A: The new sentence reads:

'This database could also be used for regional (or global) climate model assessment in the context of climate change studies.'

R. FIG 5.- The lines would be better with color.

A: We added colors to the figure.

R. FIG 8.- These are a little hard to read. How about adding some colour to the plots?

A: We added colors to the figure.

R. Table 1 and 2. Seems like Table 1 and Table 2 could be combined.

A: We have combined both tables into one.

R. Table 3. It would be useful and common practice to define the column headers in the caption.

A: Done.

R. Table 4. Table number format not consistent with other tables.

A: We have checked and homogeneized the number formats in the tables.

R. Table 5. Same comment as for Table 3.

A: Done.