Earth System
Science
Data

Open Access

Discussions

Interactive comment on "High-resolution daily gridded datasets of air temperature and wind speed for Europe" *by* S. Brinckmann et al.

Anonymous Referee #2

Received and published: 29 October 2015

The authors present a potentially useful gridded dataset for daily surface air temperature and daily mean wind speed on a high spatial resolution. But before the paper/dataset can be considered for publication, major revisions in terms of a rewrite of several parts and additional comparisons are needed, as well as corrections to the netcdf files.

General comments:

1: The time in the netcdf files are not correct. When plotting the one for Tmin for Jan 2001, is seems that the days are 2 January 2001 - 1 February 2001 instead of 1-31 January 2001. The probable reason is that the time value should start with 0 instead of 1 (days since 1 January 2001).

2: The coordinates in the netcdf file are wrong. They are given as degrees N and degrees E, but they are not the regular lat/lon coordinates, but the ones from the rotated grid. No projection information is given in the netcdf files to correct for this.

3: Why are only the IQR values made available and not the total uncertainty values?

4: The results section regarding variance and cross validation should be reduced.

5: What is missing in the manuscript are results regarding the final gridded datasets, such as the effect of the gridding method on extremes and comparison with other existing datasets (e.g. E-OBS, CRU, or even reanalysis in case of wind speed).

6: What is the difference with the current interpolation method compared to other available gridded datasets (except spatial resolution)?

7: Tmin and Tmax from synop are 12h values and not 24h values. In certain conditions, this can result in having not the 'real' daily (=24h) Tmin and Tmax. This should be taken into account when combining with ECA&D data for which Tmin and Tmax are (most of the time) determined over 24h periods. The fact that the Tmin and Tmax grid reflect only 12h extremes should be made more clear (for example in the abstract and/or summary and/or conclusions).

8: Synop values are usually of lower quality than climatological time series such those from e.g. ECA&D, since no quality control is normally applied to synop. What is the reason for not using climatological time series (since these are easily available for 2001-2010 for several parts of Europe) and adding these with synop instead of the other way around (which is used in the manuscript)?

9: In the temperature regression, elevation fields of SRTM are used. But these are not available above 60N (according to the given website). No mention of this is given in the manuscript and Fig 5 does show elevation values above 60N so there is an inconsistenty here.

Specific comments:

1: Make clear in the abstract (and summary/results) that the grids are available on a rotated grid and not on a regular lat/lon grid.

2: p 652, line 15-16: remove "As contribution...period 1961-2010". Also given 2 paragraphs later.

3: p 653, line 18: put the abbreviations behind the parameters.

4: p 654, sect 2.1: It is mentioned that the basic uncertainty for wind speed is relatively high and therefore the potential discrepancies are acceptable. But for temperature, the standard deviation is relatively low (0.21 degreesC), but still twice the basic uncertainty given in line 3, but this is not mentioned.

5: Fig 1a and b: why not use the same month for comparison?

6: Sect 2.3 (qc and assurance) is not very clear, some rewrite needed: What is meant with "assurance" in the section title? How are the outliers determined for Tmin and Tmax? How are the thresholds of 7degreesC (line 25) and 0.65 degreesC/m (line 3 p 657) determined? In Sect. 3.1 smaller values than 0.65 degreesC/m are mentioned. Where do the 19 regions come from? How far away are neighbouring stations (p 657 line 29)?

7: Eq (1) (and others): what is 'x'? Explain all variables in all equations.

8: Actually 4 steps are done in the interpolation instead of 3: Monthly regression, daily regression, monthly kriging, daily kriging.

9: Sections 3+4: not always clear if stations or gridded fields are used.

10: p662, line 5: Why is height correction needed? Observations are already at station height.

11: Fig 7: is this just a schematic are results for a region/month/etc.?

12: p 663, line 12-14: How can you decide this for the final setup (ie full 2001-2010

period) when the 2 models are only performed for 4 test months?

13: Sect 4.1: coastal distance: what is done with coastal distances >100 km? Is ln(coa+1) used in eq(3) or just coa? Why is 5x5km used and not 10x10km as in Walter et al for the exposure?

14: p 667, line 19: In the regression section, no normalization is mentioned!

15: How are IQR determined per day and grid box as there is only 1 uncertainty value per day per grid box.

16: p 668, line 25: It is mentioned that IQR of 1.0 degreesC is high. But almost all areas have values that are even higher. What does that mean?

17: p 669, line 3: why not use 31 July 2010 for wind speed as that day is also used for temperature?

18: p 669, line 10 (and other places): the discussion on nugget in Sect 4.2 is only 1 sentence...

19: p 669, line 13: Fig 11 (instead of 10): are the green grid boxes next to red grid boxes real? These are very large differences over a very small area.

20: p 670, line 20-25: are other winter and summer months checked? Difficult to say something using only 1 (or 2) months...

21: p 671, line 5-10: not sure if the high mean wind speed can be a reason. The results shown are differences between cross-validation and original fields, so the higher mean wind speed is already taken into account by taking this type of anomaly.

22: Fig 16: there is a very sharp drop around March/April for Tmin and Tmax, which is not mentioned at all.

Technical corrections:

1: Be consistent in terminology. E.g. altitude and elevation. Subregion and region.

When 'statistic' is mentioned, different things are meant. Confusing to use 'explained' as well 'unexplained' variance. 'data' is used when 'stations', 'values' or even 'days' are meant.

2: Several typos exist throughout the manuscript.

3: Check the order of the Tables and Figures. Table 3 is mentioned earlier than table 2. Same for figures 7 and 6.

4: From p 669 line 13 onwards, the figure numbering is wrong. They should all be 1 higher than the one given.