Reply to reviewers' comments

The GRENE-TEA Model Intercomparison Project (GTMIP) stage 1 forcing dataset

T. Sueyoshi, et al.

We greatly appreciate the anonymous referee #1 for insightful comments and suggestions to the manuscript, thanks to which the authors believe that the manuscript was significantly improved. We have added explanations and modified expressions in the texts, and added two figures to answer and clarify the questions the reviewer raised.

In the following, we have provided our response to the questions and remarks posed by the referee, also showing the changes of the text in case they are needed. Referee's comments are written in blue with italic, in which the part we picked up as questions are underlined, and our answers are in black.

18 November 2015

Tetsuo Sueyoshi

Answer to the Referee Comment #1 (28 Sep 2015)

For Level 0

For level 0 the approach is based on meteorological reanalysis (here ERA-interim) corrected from data based climatology: CRU and GPCP. The method is very similar to what has been done for previous reanalysis-corrected product like for instance WFDEI, WATCH or CRUNCEP. Then my question to the authors is, (1) *why did you produce a new dataset and not directly used one of the existing products? What is the improvement of the data generated compared to these existing datasets? This should be justified.* (2) *More detail about the method should also be provided. In particular the method for temporal interpolation is well defined* (2.1) *but there is no information about the spatial interpolation. Indeed ERA-Interim as a spatial resolution that is lower than CRU or GWSP product for instance. So how did re-projection have been done to have the different products within the same grid? I am also surprised that climatic variables directly taken from ERA-I (e.g. pressure, pressure...) are* (2.2) *just repeated for the different 30' time step and not interpolated.* (3) *This could induce relatively abrupt change on the transition between two 6h periods? LSM for thermic and hydrologic parts are in general very sensitive to such kind of behaviour.* (essdd-8-C329-2015.pdf)

(1) *why did you produce a new dataset and not directly used one of the existing products? What is the improvement of the data generated compared to these existing datasets?*

At the beginning of this project (around 2012), we took considerations which dataset we should use as the starting dataset for the project. There were several requisites coming from the parental GRENE Arctic project (2011-2015), such as the study period to be covered (1979-2013), and locations (the dataset should be ready to cover any GRENE Arctic sites). The most important factor was to have full transparency from a scratch to a product, but the existing dataset is opaque at least in part, in its data creation procedure. The processes of creating site-fitted data constituted an essential part of the collaboration between modelling and field scientists, which was also important for the project. Under those conditions and circumstances, we decided to use reanalysis data (which turned out to be ERA-Interim after pre-examination among other reanalysis data, as described in section 4.1) and make our own dataset. We have added these explanations to the text (mainly in section 1). We believe that the resulting Lv1 data have sufficient quality and provided a rare occasion for close collaboration and mutual understanding between the modelling and field scientists (cf. Miyazaki et al. 2015, GMD).

Revised text in section 1:

> After consideration of the project's requisites for the study period (i.e. data should cover the period from 1979 to 2013) and locations (i.e. the dataset should be ready to cover any GRENE Arctic sites), GTMIP took the approach of preparing its own harmonized dataset from reanalysis data. Another motivation for not using the existing datasets was that the processes of creating site-fitted data constituted an essential part of the collaboration between modelling and field scientists, for both of whom it was important to be part of a transparent data development process. Therefore, we decided to use reanalysis data to create temporally continuous data (basis for level 0 data, see section 3) while utilizing observational data in customizing the level 0 data to fit to each site.

(2) *More detail about the method should also be provided.*

We provided the additional details for the comments (2.1) and (2.2) as the following:

(2.1) *... but there is no information about the spatial interpolation. Indeed ERA-Interim as a spatial resolution that is lower than CRU or GWSP product for instance. So how did re-projection has been done to have the different products within the same grid?*

As for the spatial interpolation, we used the "nearest neighbour" method (p. 710, L. 25 in the discussion paper), assuming that corrections done to make the Lv1 data (i.e., merging of the in-situ observation data) will supersede the spatial interpolations within reanalysis and/or climate datasets. Our preliminary analysis indeed showed that bi-linear interpolation was not always better solution in the arctic or high-latitude sites, due to the lack of reliable observational points in the higher-latitude side of these sites. More specifically for the case of Tiksi, as the site is located in the coastal area, the interpolation results in averaging between land and ocean points, which doesn't make the value more realistic. We have revised the manuscript as follows:

> For the selection of the data grid, the 'nearest neighbour' principle was used without applying any spatial interpolation. From the preliminary analysis, bi-linear interpolation was not always better solution in the arctic or high-latitude sites. Table 4 shows the coordinates of grid centres of the ERA-Interim, CRU, and GPCP datasets for four sites.

(2.2) *... just repeated for the different 30' time step and not interpolated.*

We applied the linear interpolation for the variables with intrinsic diurnal cycle (e.g., temperature and downward solar radiation), but decided not to do so for other variables. Applying temporal interpolation for such variables means to take additional assumptions for the temporal changes, which is not favorable in the light of project's purpose. We have included this explanation in Sect. 4.2.1. as follows:

> For the variables with intrinsic diurnal cycles (e.g. air temperature and downward solar radiation), however, an effort was made to reconstruct the diurnal cycle as much as possible. For other variables, we took the minimalist approach of adding no additional assumptions (i.e. temporal interpolation) before fixing the datasets, unless this caused unexpected disruptions of our examination simulations by LSMs.

(3) *This could induce relatively abrupt change on the transition between two 6h periods?*

We tested those dataset (that is, repetition of the same value for 6 hours) with LSMs before finalizing, which showed no harm results. We have included the following sentence in Sect. 4.2.1.to explain this:

A test run was performed using LSM before finalizing dataset, which showed no harm results.

For Level 1

For level 1, an ad-hoc correction for the different sites is done by correcting the level 0 data. [(4)] *For temperature for instance correction is done only on the diurnal range that is corrected by the monthly diurnal range of CRU data.* [(5)] *Precipitation is corrected from GWSP if sufficient data is available. As there is in situ meteorological data, I don't understand why a systematic regression between Lv0 and in situ data was not applied instead?* [(6)] *This seems to by what have been done for some parameters like pressure, long-wave radiation, but the method used for regression is not sufficiently detailed.* (essdd-8-C329-2015.pdf)

(4) *For temperature for instance correction is done only on the diurnal range that is corrected by the monthly diurnal range of CRU data.*

As for air temperature, the Lv0 daily mean already showed good correspondence to the observation, but the daily range did not, as the reanalysis data is 6-hourly (as described in the first sentence in section 4.3.2). Therefore, only correction to the daily range was applied.

(5) *Precipitation is corrected from GWSP if sufficient data is available. As there is in situ meteorological data, I don't understand why a systematic regression between Lv0 and in situ data was not applied instead?*

As for precipitation data, in-situ observation data are very limited and regression is not applicable. Instead, we used the near-by weather station data for "observation". Although those station data may or may not correlate well with the precipitation at the sites on a short time scales (such as hourly or daily), we assumed that the temporal variations on a longer time scales (monthly or longer) correlate reasonably. Therefore, we have not applied (linear) regressions, but constrained the Lv0 data with monthly and annual amount.

We agree that the explanation in the original manuscript was not enough regarding (4) and (5), so have revised the texts (sections 4.3.2 and 4.3.5, respectively) to include the explanations given above:

p. 714 L. 15-19 (for temperature):

> Comparison with observations shows that the Lv0 temperature data agree well with respect to daily temperature averages, so that Eq. 1 is valid for creation of Lv1 data. The Lv0 data, however, show a smaller daily temperature range than the observed data, mainly because the Lv0 data were produced by interpolation of four points per day. Therefore, we employed two-step evaluations for the following targets (i.e. daily temperature range correction as shown by Eq. 5, and monthly average values correction as shown by Eq. 6, similar to Eq. 1) to produce Lv1 temperature data.

p. 717 L. 10-15 (for precipitation):

> On the other hand, site observation data are also not always reliable, mainly due to instrumental problems. Another issue with site observations is short observation periods. For these reasons, the precipitation data used to constrain the Lv1 data were those collected at nearby weather stations, which we assumed may not necessarily correlate well on hourly or daily scales, but will show similar variations in precipitation amounts on a longer time scale (month or longer).

(6) *This seems to by what have been done for some parameters like pressure, long-wave radiation, but the method used for regression is not sufficiently detailed.*

In order to illustrate how the regression used for longwave radiation and pressure worked, we have added two figures (Figure 6c and 6d, respectively), which we believe is straightforward to explain the detailed nature of the regression. The manuscript was revised as follows:

p. 718 L. 23-p. 719 L.1 (for longwave radiation):

> Comparison between observed values and the ERA-interim output (i.e. Lv0 data, L0) revealed that the difference between the two datasets can largely be accounted for by emission differences of the air layer due to altitude differences between the observation sites and the closest ERA-Interim grid points (see Fig. 6a and 6c as examples for Kevo). The resulting equations for creating Lv1 data, $L1$, are as follows:

p. 719 L. 13-16 (for surface pressure):

> Similar to downward longwave radiation, differences in surface pressure between the observed and reanalysis data (i.e. Lv0 data, $L0$) are largely correlated with altitude differences (see Fig. 6b and 6d as examples for Fairbanks). The resulting correction equations are as follows:

Minor comment

p. 706 L. 28: *missing "d" at Vuichard to reference (Vuichard and Papale 2015)*

Thank you for pointing this out. We have corrected the author name.

Other corrections:

p. 709 L. 14: “Sect. 4.3 and 4.4” should be “Sect. 4.2 and 4.3”

p. 729 Table 4: There were mistakes on the numbers in decimals. They are all corrected.