

This discussion paper is/has been under review for the journal Earth System Science Data (ESSD). Please refer to the corresponding final paper in ESSD if available.

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

T. Sueyoshi^{1,2}, K. Saito², S. Miyazaki^{1,2,a}, J. Mori^{1,2}, T. Ise³, H. Arakida⁴, R. Suzuki², A. Sato⁵, Y. Iijima², H. Yabuki^{1,2}, H. Ikawa⁶, T. Ohta⁷, A. Kotani⁷, T. Hajima², H. Sato², T. Yamazaki⁸, and A. Sugimoto⁹

¹National Institute of Polar Research, Tachikawa, Japan

²Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

³Kyoto University, Field Science Education and Research Center, Kyoto, Japan

⁴RIKEN Advanced Institute for Computational Science, Kobe, Japan

⁵National Research Institute for Earth Science and Disaster Prevention, Snow and Ice Research Center, Nagaoka, Japan

⁶National Institute for Agro-Environmental Sciences, Tsukuba, Japan

⁷Nagoya University, Graduate School of Bioagricultural Sciences, Nagoya, Japan

⁸Tohoku University, Graduate School of Science, Sendai, Japan

⁹Hokkaido University, Faculty of Environmental Earth Science, Sapporo, Japan

^anow at: Sonic Corporation, Tachikawa, Japan

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Received: 14 May 2015 – Accepted: 29 June 2015 – Published: 24 August 2015

Correspondence to: T. Sueyoshi (sueyoshi.tetsuo@nipr.ac.jp)

Published by Copernicus Publications.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

ESSDD

8, 703–736, 2015

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



scribes the method to prepare the forcing datasets used in GTMIP stage 1, and their comparison with observational data.

To evaluate the validity of and variations in performance of a model, and to conduct comparisons between the model's outputs and observational data at specific sites, it is essential to drive the model with such data that reflect realistic conditions at those sites. It is very difficult, however, to create a complete dataset for driving a model solely from observations, which are prone to data gaps due to instrument problems, human mistakes, or other reasons (Nakai et al., 2013; Kodama et al., 2007; Watanabe et al., 2000; Iwahana et al., 2014; Ohta et al., 2001, 2008, 2014; Kotani et al., 2013; Lopez et al., 2007; Nakai et al., 2008; Sato et al., 2001). Despite efforts to maintain and improve our observational network, data gaps are unavoidable, as the availability of instrumentation is often limited by logistical or funding issues. Many of the Arctic sites are located in remote areas with limited access and power supply (see Fig. 1 for GRENE-TEA observation sites; cf. site distributions of other national/international Arctic research projects, such as PAGE21 (<http://portal.inter-map.com/#mapID=49&groupID=308&z=1.0&up=-310027.6&left=2001105.4>; accessed on 20 April 2015) or map of stations used for the Climate Research Unit (CRU) data set (<http://climate.unur.com/cru-station-records/cru-stations-map.html>; accessed on 20 April 2015).

On the other hand, reanalysis data are often used to drive LSMs, as they are always continuous without missing values and physically consistent with regard to all necessary elements. Although there are several good reanalysis datasets available now (e.g. NCEP, ERA, JRA, MERRA), these products, which are essentially model outputs at large horizontal scales (typically on the order of 100 km), are no match for a direct comparison with observational data. Previous efforts of the modelling and empirical research community have produced several datasets available as inputs into LSMs such as the Global Meteorological Forcing Dataset for land surface modelling (Sheffield et al., 2006) and WATCH (Weedon et al., 2010), or FLUXNET observational data (Vuichar and Papale, 2015) for debiasing of reanalysis data (ERA-Interim). After consideration of the general problem, our particular circumstances, and methodologies

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



used in creating the previous products, GTMIP took the approach of preparing a harmonized dataset, which relies on reanalysis data to create temporally continuous data (basis for level 0 data, see Sect. 3) while utilizing observational data in customizing the level 0 data to fit to each site. We produced a dataset with high consistency relative to the observational data and without gaps (level 1 data).

In this paper, after summarizing information about the study sites and observational period in Sect. 2, we describe the specification of the data in Sect. 3, and the details of the procedure used to produce the two datasets in Sect. 4. The procedure is explained for each variable in the dataset, and differences between the sites, when present, are described. Further implications of this work are described in Sect. 5, and access to the data is documented in Sect. 6.

2 Study sites and observational periods

Stage 1 of GTMIP produced a harmonized forcing dataset for six sites in Arctic and sub-Arctic regions, which have been, at least partially, studied as part of the GRENE-TEA project (Fig. 1): Fairbanks in Alaska (Nakai et al., 2013), Tiksi (Kodama et al., 2007; Watanabe et al., 2000) and Chokurdakh (Iwahana et al., 2014) in northeast Siberia, Yakutsk in central East Siberia (Ohta et al., 2001, 2008, 2014; Kotani et al., 2013; Lopez et al., 2007), Tura in central Siberia (Nakai et al., 2008), and Kevo in northern Europe (Sato et al., 2001). The longitudes and latitudes of the sites are summarized in Table 1, as well as the observational periods for data collection.

The sites lie in different geographical zones with different ecoclimatic characteristics. Fairbanks is located in the discontinuous permafrost zone with open black spruce forest, Tiksi and Chokurdakh are in the continuous permafrost zone, the former with tundra vegetation and the latter in the transition between tundra and taiga, Yakutsk and Tura are in the continuous permafrost zone with larch forest, and Kevo is in the non-permafrost (seasonal frozen ground) zone with mixed forest.

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



These sites were selected from “super sites” among the GRENE-TEA observation sites, where automatic weather stations were installed that measure a number of meteorological variables, and enough site information about soil and vegetation conditions were available. Geographical coverage was also considered, to ensure that the sites were distributed over the entire Arctic, and that they covered a combination of tundra and forest.

Although we are planning to expand this network of “harmonized dataset sites”, we believe that these six initial sites, already covering typical geographical conditions in the Arctic, are a good starting point for a model intercomparison aiming to improve our understanding of physical and ecosystem land surface processes in the Arctic and sub-Arctic regions.

3 Dataset specifications

The data created in this study cover a 35 year period, from 1 September 1979 to 31 December 2013, local time, for all sites. Temporal resolution was set to 30 min, considering the requirements of a majority of the physical land surface models. The start date was set to 1 September, when there should be no snow cover, in order to minimize any potential bias for snow depth at the initial condition.

The GTMIP dataset consists of three sets of data, each of which has eight meteorological elements so that a contemporary land surface model can be driven with the dataset. The first set is called level 0 (Lv0), consisting of reanalysis data with minimum correction, corresponding to the existing observation-based global dataset. The second set is called level 1 (Lv1), with additional corrections applied to Lv0 data, based on the GRENE-TEA site observations. The third set is a detrended dataset (DT), which essentially consists of Lv1 data with long-term temporal trends removed. Both Lv0 and Lv1 are forcing datasets. The primary dataset for the GTMIP activity is Lv1, while Lv0 was prepared to enable an evaluation of the effect of the observation-based corrections. The DT is intended for use with ecosystem models with long spin-up times. We

also provided satellite-based photosynthetically active radiation (fPAR) and leaf area index (LAI) data for models that need these data as prescribed conditions. A schematic diagram for the creation of the Lv0 and Lv1 data is shown in Fig. 2.

The included eight variables and their units are as follows: surface air temperature [K] and specific humidity [kg kg^{-1}] at a reference height (set to 2 m in most models), precipitation [$\text{kg m}^{-2} \text{s}^{-1}$], wind speed [m s^{-1}] and wind direction (at 10 m) [clockwise in degrees from north], pressure at the surface [hPa], and downward shortwave and longwave radiation [W m^{-2}].

4 Methods of forcing data production

This section provides documentation for (1) selection of reanalysis data, (2) construction procedure for Lv0 data, (3) construction procedure for Lv1 data, and (4) construction procedure for additional data (detrended data, and LAI data). The outline of the procedures for Lv0 and Lv1 is given in Table 2. Procedures are explained for each variable in Sect. 4.3 and 4.4. When data were handled differently for different sites, it was also noted.

4.1 Selection of reanalysis data

As the first step of the procedure, we selected the reanalysis dataset used as the base data. A detailed performance comparison was carried out for the following four datasets: Japanese Reanalysis JRA-55, The European centre for medium-range weather forecasts ReAnalysis ERA-Interim (Dee et al., 2011), National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR); NCEP/NCAR, NCEP-Department of Energy (DOE). Climate Research Unit (CRU, East Anglia University) data for air temperature (Harris et al., 2014) and the Global Precipitation Climatology Project (GPCP) dataset for precipitation (Adler et al., 2003), based on weather station observations and/or satellite data, were used as a reference. Spe-

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



cial emphasis was placed on the reproducibility of air temperature and precipitation, as these two variables are known to have a large impact on the performance of LSMs. As a result of the comparison of monthly values, ERA-Interim was chosen as the base data set for GTMIP, reproducing these two variables reasonably with minimum root mean square errors (RMSEs) with respect to the CRU (air temperature) and GPCP (precipitation) datasets (Table 3).

Among the ERA-Interim variables, the following parameters were used in constructing the GTMIP forcing data: 10 m zonal (U) wind component, 10 m meridional (V) wind component, 2 m dew point temperature, 2 m temperature (t_{2m}), surface pressure, surface solar radiation downward, surface thermal radiation downward, and total precipitation. The temporal resolution of the used variables is 6 h, and stamped at 00:00 Z, 00:06 Z, 00:12 Z, and 00:18 Z in coordinated universal time (UTC).

4.2 Construction of level 0 data

4.2.1 General

Construction of Lv0 data (Saito et al., 2014a) from reanalysis data contains the following steps: (1) choose grid and time period (in this case September 1979 to December 2013) from the reanalysis data, (2) apply a common correction for all sites for temperature and precipitation, and (3) determine 30 min values using the 6 hourly data. Providing a template dataset for modellers to prepare their own read-in modules in advance is the primary purpose of this dataset, but as mentioned before, it also serves to evaluate the effects of the site-fitting correction by comparing the performance of LSMs using these two datasets. Minimum corrections of temperature and precipitation using CRU and GPCP are applied so that problems with using unrealistic values can be avoided. Differences between sites were not considered for this correction.

For the selection of the data grid, the “nearest neighbour” principle was used. Table 4 shows the coordinates of grid centres of the ERA-Interim, CRU, and GPCP datasets for four sites.

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Temporal interpolation is a requisite as the temporal resolution of the target dataset is 30 min, while that of the ERA-Interim data is 6 h. A single strategy (e.g. linear interpolation or equal distribution) does not always work, depending on the characteristics of variables and sites. The process of temporal adjustment is described for each variable in the next sections.

4.2.2 2 m air temperature

For Lv0 air temperature (T_{Lv0}), monthly mean values derived from 6 hourly ERA-Interim air temperature [K] was adjusted using the CRU TS3.22 dataset of the month, so that the monthly mean of Lv0 data coincides with the CRU temperature data, T_{CRU}^{month} [$^{\circ}$ C];

$$T_{Lv0}^{6h} = T_{ERA}^{6h} + \left(T_{CRU}^{month} + 273.15 \right) - \overline{\left(T_{ERA}^{6h} \right)}_{month}. \quad (1)$$

Each value was set to the corresponding UTC time, and then linearly interpolated with the next 6 hourly data to produce data with a 30 min resolution (T_{Lv0}^{30min}).

4.2.3 Specific humidity

Specific humidity for the Lv0 data (q_{Lv0}) was determined by a two-step procedure. First, relative humidity was calculated using the 6 hourly ERA-Interim dew points ($DewT_{ERA}^{6h}$) and air temperatures. Second, 30 min specific humidity was calculated from the previously obtained 6 hourly relative humidity data and the 2 m air temperature calculated in Sect. 4.2.2.

$$RH_{ERA}^{6h} = q \left(DewT_{ERA}^{6h} \right) / q_s \left(T_{ERA}^{6h} \right), \quad (2a)$$

$$q_{Lv0}^{30min} = q_s \left(T_{Lv0}^{30min} \right) \times RH_{ERA}^{6h}, \quad (2b)$$

where $q_s(T)$ denotes saturation specific humidity at temperature T .

4.2.4 10 m wind speed and direction

Six-hourly wind speed and direction, calculated from zonal and meridional wind of ERA-Interim, were set as the Lv0 data at the corresponding UTC time, and repeated 11 times to create 30 min interval data in the 6 h range.

4.2.5 Total precipitation

Six-hourly precipitation amounts (P_{Lv0}) were calculated from 6 hourly ERA-Interim precipitation, and a correction was applied in such a way that the monthly amount was proportional to the GPCP precipitation data:

$$P_{Lv0}^{6h} = P_{ERA}^{6h} \times \left(P_{GPCP}^{month} \right) / \sum_{month} \left(P_{ERA}^{6h} \right). \quad (3)$$

To determine how to distribute the 6 h precipitation data among the 30 min time slots, we conducted a preliminary analysis using the following three strategies and evaluated their influence on the simulated energy and water budget keeping other factors and variables fixed: (a) assign the entire amount to the first time slot (i.e. the entire 6 h precipitation amount falls in the first 30 min of the 6 h range), (b) assign the precipitation amounts in ratios of eight, four, two, and one to the first four time slots (i.e. the first two hours), and (c) assign equal amounts to the 12 time slots in the range. The results showed a distorted energy and water cycle for (c), while results were reasonable for the first two cases (not shown). Based on this outcome, we used strategy (a) and assigned the entire precipitation amount to the first time slot of every 6 h range.

4.2.6 Downward shortwave radiation

Shortwave radiation (SW), calculated as theoretical values for latitude, longitude, day, and time for a particular site and day at 30 min intervals, was corrected using daily

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



average cloudiness (Cl_{day}) as defined in the following equation:

$$SW_{\text{Lv0}}^{30\text{ min}} = SW_{\text{theory}}(\text{Lat}, \text{Lon}, \text{Day}, \text{Time}) \times Cl_{\text{day}} \quad (4)$$
$$Cl_{\text{day}} = \frac{\sum_{\text{day}} (SW_{\text{ERA}}^{6\text{ h}})}{\sum_{\text{day}} (SW_{\text{theory}}(\text{Lat}, \text{Lon}, \text{Day}, \text{Time}))}$$

Comparison with observed values shows that thus calculated SW coincides well with observations in terms of daily averages, implying that the total incident solar energy is well represented. Figure 3 shows an example of an annual time series for daily maximum, average, and minimum of 30 min SW for Tiksi.

4.2.7 Downward longwave radiation

ERA-Interim 6 hourly longwave radiation data were set as the Lv0 data at the corresponding UTC time, and repeated 11 times to create 30 min interval data for each 6 h range.

4.2.8 Surface pressure

ERA-Interim 6 hourly pressure data were set as the Lv0 data at the corresponding UTC time, and repeated 11 times to create 30 min interval data.

4.3 Construction of level 1 data

4.3.1 General

For the Lv1 dataset (Saito et al., 2014b), we applied a site-specific correction using observational data. The principle of Lv1 data construction is to derive rules to modify base data (ERA-Interim) to make them as similar to the observed values as possible. The formulations for statistical correction were established through comparison between observations and climate data (that is, ERA-Interim, CRU, and GPCP) for the

respective observational period at each of the four sites (Table 1, Fig. 2). Observational periods are summarized in Table 1, and visualized by grey bars in the middle panel, below “GRENE-TEA site observation”, of Fig. 2. The Lv0 data were readily available with no or minor modifications for some variables, whereas site-specific treatment and adjustment were needed for other variables. Comparisons between observed and produced data, and the difference between Lv0 and Lv1 data are shown for selected sites and variables in Figs. 4–6. The overall quality of fit for all variables is shown in Fig. 7, which shows gradient a of the fitted line of daily mean observations y against Lv0, x ($y = ax + b$), and Pearson’s correlation coefficient r for x and y . The more a and r approach unity, the better the fit. Note that the order of independent (abscissa, x) and dependent (ordinate, y) variables of the fitting line (exemplified in Figs. 3b and 4c–f) is unconventional because a choice was made to show the necessary modification or corrections made to Lv0 to construct site-corrected Lv1.

4.3.2 2 m air temperature

Comparison with observations shows that the Lv0 data agree well with daily averages, but have a smaller daily temperature range than the observed data, mainly because Lv0 data were produced by interpolation of four points per day. Therefore, we had to employ two-step evaluations for the following targets (i.e. daily temperature range correction and monthly average values correction) to produce Lv1 temperature data. The climate data (CD: CRU and/or ERA-Interim) that were used for the diurnal correction differed from site to site, and were chosen based on the reproducibility of the observed data (not shown). The height of near-surface air temperature measurements at each site is summarized in Table 1. We chose observations at the height closest to 2 m, but did not convert the observation values for height because the reference height for air temperature is different among models.

First, the daily temperature range (dtr) was corrected for the ERA-Interim 6 hourly data, employing the CRU diurnal monthly temperature range for Fairbanks ($dtr_{CD} = dtr_{CRU}^{month}$) but daily ERA-Interim data (namely, 2 m temperature maximum and minimum;

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



$dtr_{CD} = (T_{\max}^{\text{day}} - T_{\min}^{\text{day}})_{\text{ERA}}$ for the other three sites (cf. Eq. 5 of Sheffield et al., 2006; Eq. 6 of Weedon et al., 2010):

$$T_{\text{dtr_corrected}}^{6\text{h}} = \left(T_{\text{ERA}}^{6\text{h}} \right) \Big|_{\text{daily}} + \left(T_{\text{ERA}}^{6\text{h}} - \left(T_{\text{ERA}}^{6\text{h}} \right) \Big|_{\text{daily}} \right) \times \left(dtr_{CD} / dtr_{\text{ERA}}^{6\text{h}} \right). \quad (5)$$

Second, monthly mean values were adjusted using $T_{\text{CRU}}^{\text{month}}$ [°C], as shown for the creation of Lv0 data in Sect. 4.2.2, for Fairbanks, Kevo, and Yakutsk:

$$T_{\text{Lv1}}^{6\text{h}} = T_{\text{dtr_corrected}}^{6\text{h}} + \left(T_{\text{CRU}}^{\text{month}} + 273.15 \right) - \left(T_{\text{ERA}}^{6\text{h}} \right) \Big|_{\text{month}}. \quad (6)$$

We did not adjust Tiksi air temperature with CRU data, because the CRU values showed poorer reproducibility of the Tiksi site observations than the ERA-Interim monthly mean values (see Table 3), mostly due to the sparse network of stations in the region. However, the ERA-Interim values at the closest grid point to Tiksi show modest continentality (warmer summers and cooler winters) than the observed values at the coastal site. Therefore, a seasonal correction for Lv1 data ($L1 = T_{\text{Lv1}}^{6\text{h}}$) was applied to $L0'$ ($= T_{\text{dtr_corrected}}^{6\text{h}}$) as follows (see also Fig. 4a):

$$L1 = 0.90 \times L0' + 0.1, \quad (\text{November to March}) \quad (7a)$$

$$L1 = L0'. \quad (\text{April to October}) \quad (7b)$$

Finally, these 6 hourly values ($L1$) were linearly interpolated with the next 6 hourly corrected data to produce data at 30 min resolution for all sites. Improvement from fitting the produced values to the observational data (Fig. 7a–c) is most apparent for the seasonal correction at Tiksi. For other sites, Lv0 data were already close to the observational data.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4.3.3 Specific humidity

Values of Lv1 specific humidity were determined using relative humidity values derived from ERA-Interim (Eq. 2a) and surface air temperature determined in Sect. 4.3.2. Figure 4b shows an example for differences in the relationship between Lv0 and Lv1 with observational data.

$$q_{Lv1}^{30\text{ min}} = q_s(T_{Lv1}^{30\text{ min}}) \times RH_{ERA}^{6h} \quad (8)$$

Improvement at Tiksi is a reflection of improvement in air temperatures as illustrated in Fig. 7d, which shows the daily mean for specific humidity.

4.3.4 10 m wind speed and direction

We determined 6 hourly wind speed and direction data from U wind and V wind of the ERA-Interim dataset and converted the data to the corresponding UTC time. We defined these data as $L0$. Then, we applied a statistical correction based on the relationship between the observed values and $L0$ at Kevo and Tiksi (see Fig. 4c). The observational periods used for the statistical correction were 1999 to 2001 and 1997 to 2012 for Kevo and Tiksi, respectively. The equation for correction of the Lv1 data ($L1$) at each site can be written as follows:

$$L1 = 0.6 \times L0, \quad (\text{Kevo}) \quad (9a)$$

$$L1 = 1.4 \times L0, \quad (\text{Tiksi}). \quad (9b)$$

No correction of Lv0 wind speed was applied to Fairbanks and Yakutsk for the following reasons (Fig. 7k–m): Yakutsk already showed a good match between observed and $L0$ data when outliers and data collected during equipment malfunctions were removed, while Fairbanks showed such a poor match between the two datasets that we decided that no statistic correction would have been appropriate. No correction of Lv0 was applied to the wind direction data at all four sites. The heights of the wind

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



measurements at each site are summarized in Table 1. We chose the observations at the height closest to 10 m above ground (Kevo, Tiksi) or canopy (Fairbanks, Yakutsk), but did not convert the observation values for height because the reference height is different among different models.

4.3.5 Total precipitation

Precipitation is the most challenging part of dataset construction because it can have the largest impact on the simulated results. Reanalysis data (in the case of this study, ERA-Interim), which are virtually a model output, generally contain large uncertainties in precipitation and produce highly averaged and smoothed values because of a six-hour output interval. On the other hand, site observation data are also not always reliable, mainly due to instrumental problems. Analysed precipitation products (in the case of this study, GPCP) have difficulties in high-latitude regions due to a scarcity of available data and poor satellite coverage. With these shortcomings of the data sources in mind, we created site-fitted precipitation data by evaluating the reproducibility of precipitation amounts and variations on a monthly and annual basis.

For Fairbanks, we decided to use the Lv0 data as Lv1 data ($L1$) rather than attempting an unreliable correction, mostly because the two source datasets did not show any systematic relationship or good correlation with monthly observations, which were made at the international airport, located about 50 km from the site. For Kevo, comparison between the annual total of observed precipitation and that of the ERA-Interim dataset showed a good inter-annual correlation, albeit with a quasi-constant underestimation. Since ERA-Interim data also showed a good correlation with intra-annual (monthly) variability, we used the ERA-Interim as the basic dataset $L0'$, that is, without any monthly correction by GPCP. The resulting correction equation is shown below.

$$L1 = 0.76 \times L0' \quad (10)$$

For Tiksi, GPCP provided better monthly variations while ERA-Interim data showed better annual total and inter-annual variations. Therefore, we applied a correction to Lv0

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

T. Sueyoshi et al.

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

using the ratio of annual precipitation of the ERA-Interim data to the annual precipitation of GPCP (Fig. 5a). In contrast, as the ERA-Interim 6 hourly total precipitation amounts agreed well with the observational data both on a monthly and annual scale for Yakutsk, they were directly used as the 6 hourly values without GPCP correction (Fig. 5b).

Finally, the 6 hourly precipitation amounts were assigned to the first time slot of every six-hour range to produce a dataset with 30 min intervals. Note that the correlation of daily values is not particularly high (Fig. 7f) relative to that of other variables shown in Fig. 7. This is a natural result since fitting is done only on a monthly and annual basis. The overall improvement is, however, already visible as Lv1 locates closer to the 1 : 1 gradient line (i.e. $a = 1$) than Lv0.

4.3.6 Downward shortwave radiation

As stated in Sect. 4.2.6, downward shortwave radiation is computed from the theoretical instantaneous radiation value decreased by daily average cloudiness, which corresponds well to the daily mean incident energy at the Earth's surface. Some discrepancies were found between Lv0 and the observational data with regard to daily maximum (Fig. 7g) and/or minimum (Fig. 7i). Possible causes are short-time cloud cover or absence of clouds, shading by forested areas (as is the case for the minimum SW in Kevo), or equipment malfunction. As it is difficult to generalize these situations in a statistical formula, no correction of Lv0 was applied for downward shortwave radiation data for all four sites. Figure 7i does not show data for Fairbanks and Yakutsk since the two stations lie south of the Arctic Circle, and hence the daily minimum is always zero.

4.3.7 Downward longwave radiation

Comparison between observed values and the ERA-interim output (i.e. Lv0 data, L_0) 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 as an example for Kevo). The

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



resulting equations for creating Lv1 data, L_1 , are as follows:

$$L_1 = 1.01 \times L_0 + 15, \quad (\text{Fairbanks}) \quad (11a)$$

$$L_1 = 1.00 \times L_0 + 11, \quad (\text{Kevo}) \quad (11b)$$

$$L_1 = 0.85 \times L_0 + 50, \quad (\text{Tiksi}) \quad (11c)$$

$$5 \quad L_1 = 0.98 \times L_0 + 17. \quad (\text{Yakutsk}) \quad (11d)$$

Figure 7j shows a marked improvement for Tiksi. This improvement is mostly manifested by the location's gradient (0.85), which reflects a seasonal difference due to the continental character of the reanalysis value at the grid point closest to Tiksi against more coastal characteristics of the Tiksi observational site. Since diurnal cycles in downward longwave and surface pressure are not as prominent as in downward short-wave radiation or air temperature, only daily means are shown in Fig. 7.

4.3.8 Surface pressure

Similar to downward longwave radiation, differences in surface pressure between the observed and reanalysis data (i.e. Lv0 data, L_0) are largely correlated with altitude differences (see Fig. 6b as an example for Fairbanks). The resulting correction equations are as follows:

$$L_1 = L_0 + 24, \quad (\text{Fairbanks}) \quad (12a)$$

$$L_1 = L_0 + 21, \quad (\text{Kevo}) \quad (12b)$$

$$L_1 = L_0 + 19, \quad (\text{Tiksi}) \quad (12c)$$

$$20 \quad L_1 = L_0 - 5. \quad (\text{Yakutsk}) \quad (12d)$$

Since changes in the intercept of the fitting lines in Eq. (11) do not change the gradient and correlation coefficient, Fig. 7e does not show any difference between Lv0 and Lv1.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



4.4 Additional data

In addition to the two model driving datasets described in the previous section, we have prepared the following two datasets for standardization and convenience for model integration and spin-up: photosynthetically active radiation (fPAR) and leaf area index (LAI) datasets (Saito et al., 2014c), and a 20 year detrended meteorological driving dataset (Saito et al., 2015). The former dataset was derived from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data (MOD15A2, MYD15A2) according to algorithms developed and validated by Sasai et al. (2011). Monthly averaged time series were processed for each of the six GRENE-TEA sites from the closest grid point values.

The latter detrended driving dataset, produced by removing monthly long-term trends from the Lv1 data for the period of 1980–1999, was provided for spin-up to include initial soil carbon conditions comparable to other biogeochemical models without being affected by warming trends (Saito et al., 2015).

5 Discussion and outlook

Concepts and principles of a correction method for creating a forcing dataset from climate product data (e.g. ERA-Interim, CRU, and GPCP) and site observations were established through a comparison of the two data sources over the observational period as described in this article. The set of observation-based, site-specific data thus produced can be expected to produce a more realistic time series than the climate product data themselves, and will allow a more meaningful comparison to observed target variables at a given site (e.g. evapotranspiration, snow pack, ground temperature and soil moisture, and gross primary products).

Once correction rules (formulae) are defined, a time series can be extrapolated to a period of no observation. Not only can these datasets (Lv0 and Lv1, in this case) be used as common driving data for model intercomparison projects to examine inter-site

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and inter-model variations with regard to performance and uncertainty, they are also useful for investigating long-term ecoclimatic variations by analysing the time series per se that incorporates the knowledge of on-going observations, and by examining and testing the capability of reproduction and hindcast through model simulations driven by the time series. Naturally, the available time period is still limited by the time covered by the product-based data, and the efficacy of the rules used to create the product-based data should be thoroughly tested when these datasets are applied.

Another target of this MIP activity was to facilitate collaboration between modelling and field scientists. Such collaboration can be difficult and is often insufficient due to discrepancies between science targets and approaches, terminology, and different prioritization during data acquisition among the two groups (for example, field scientists tend to seek for unusual data in their area of investigation, while modellers often need typical data for an area; cf. Saito et al., 2014d).

Considering the aforementioned discrepancies, the GTMIP study was conducted by a team of modelling and field scientists, which was especially important when the quality of observed data used for Lv1 creation was questioned (e.g. outliers, missing values, representativeness of the site and nearby station data). Similarly, when conducting model performance checks using observational data, which is an important and unavoidable process for model development, field scientists joined together to discuss issues with regard to modelled and observed data, capability and limits of models depending on the extent and complexity of the implemented processes, often concluding that a precise observation does not necessary coincide with a “correct” model output.

The detailed protocol of GTMIP stage 1 (site simulations) is described in our companion paper (Miyazaki et al., 2015).

6 Data access

All data presented in this paper, as well as an overall outline of the GRENE-TEA Model Intercomparison Project (GTMIP), are available from the Arctic Data archive System

(ADS) website (<https://ads.nipr.ac.jp/>). Specifically, the webpages for the Lv0, Lv1, DT, and fPAR/LAI datasets provide the metadata including readme files and a concise explanation of the data creation procedure.

GTMIP overall: <https://ads.nipr.ac.jp/gtmip/gtmip.html>

Lv0 data: <https://ads.nipr.ac.jp/dataset/A20141009-005>

Lv1 data: <https://ads.nipr.ac.jp/dataset/A20141009-006>

DT data: <https://ads.nipr.ac.jp/dataset/A20150205-001>

fPAR/LAI: <https://ads.nipr.ac.jp/dataset/A20141009-007>

Acknowledgements. This study was supported by the GRENE Arctic Climate Change Research Project, Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, 4, 1147–1167, 2003.

Dee, D. P., Uppalaa, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., H'olm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset, *Int. J. Climatol.*, 34, 623–642, 2014.

Iwahana, G., Takano, S., Petrov, R. E., Tei, S., Shingubara, R., Maximov, T. C., Fedrov, A. N., Desyatkin, A. R., Nikolaev, A. N., Desyatkin, R. V., and Sugimoto, A.: Geocryological characteristics of the upper permafrost in a tundraforest transition of the Indigirka River Valley, Russia, *Polar Science*, 8, 96–113, doi:10.1016/j.polar.2014.01.005, 2014.

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Kodama, Y., Sato, N., Yabuki, H., Ishii, Y., Nomura, M., and Ohata, T.: Wind direction dependency of water and energy fluxes and synoptic conditions over a tundra near Tiksi, Siberia, *Hydrol. Process.*, 21, 2028–2037, 2007.
- Kotani, A., Kononov, A. V., Ohta, T., and Maximov, T. C.: Temporal variations in the linkage between the net ecosystem exchange of water vapour and CO₂ over boreal forests in eastern Siberia, *Ecohydrology*, 7, 209–225, doi:10.1002/eco.1449, 2013.
- Lopez, M. L., Saito, H., Kobayashi, K., Shiota, T., Iwahana, G., Maximov, T. C., and Fukuda, M.: Interannual environmental-soil thawing rate variation and its control on transpiration from *Larix cajanderi*, Central Yakutia, Eastern Siberia, *J. Hydrol.*, 338, 251–260, doi:10.1016/j.jhydrol.2007.02.039, 2007.
- Miyazaki, S., Saito, K., Mori, J., Yamazaki, T., Ise, T., Arakida, H., Hajima, T., Iijima, Y., Machiya, H., Sueyoshi, T., Yabuki, H., Burke, E. J., Hosaka, M., Ichii, K., Ikawa, H., Ito, A., Kotani, A., Matsuura, Y., Niwano, M., Nitta, T., O'ishi, R., Ohta, T., Park, H., Sasai, T., Sato, A., Sato, H., Sugimoto, A., Suzuki, R., Tanaka, K., Yamaguchi, S., and Yoshimura, K.: The GRENE-TEA Model Intercomparison Project (GTMIIP): overview and experiment protocol for Stage 1, *Geosci. Model Dev. Discuss.*, 8, 3443–3479, doi:10.5194/gmdd-8-3443-2015, 2015.
- Nakai, T., Kim, Y., Busey, R. C., Suzuki, R., Nagai, S., Kobayashi, H., Park, H., Sugiura, K., and Ito, A.: Characteristics of evapotranspiration from a permafrost black spruce forest in interior Alaska, *Polar Science*, 7, 136–148, 2013.
- Nakai, Y., Matsuura, Y., Kajimoto, T., Abaimov, A. P., Yamamoto, S., and Zyryanova, O. A.: Eddy covariance CO₂ flux above a Gmelin larch forest in continuous permafrost of central Siberia during a growing season, *Theor. Appl. Climatol.*, 93, 133–147, doi:10.1007/s00704-007-0337-x, 2008.
- Ohta, T., Hiyama, T., Tanaka, H., Kuwada, T., Maximov, T. C., Ohata, T., and Fukushima, Y.: Seasonal variation in the energy and water exchanges above and below a larch forest in Eastern Siberia, *Hydrol. Process.*, 15, 1459–1476, 2001.
- Ohta, T., Maximov, T. C., Dolman, A. J., Nakai, T., van der Molen, M. K., Kononov, A. V., Maximov, A. P., Hiyama, T., Iijima, Y., Moors, E. J., Tanaka, H., Toba, T., and Yabuki, H.: Interannual variation of water balance and summer evapotranspiration in an Eastern Siberian larch forest over a 7 year period (1998–2006), *Agr. Forest Meteorol.*, 148, 1941–1953, 2008.

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Ohta, T., Kotani, A., Iijima, Y., Maximov, T. C., Ito, S., Hanamura, M., Kononov, A. V., and Maximov, A. P.: Effects of waterlogging on water and carbon dioxide fluxes and environmental variables in a Siberian larch forest, 1998–2011, *Agr. Forest Meteorol.*, 188, 64–75, 2014.
- Saito, K., Miyazaki, S., Mori, J., Ise, T., Arakida, H., Sueyoshi, T., Hajima, T., Iijima, Y., Yamazaki, T., and Sugimoto, A.: GTMIP meteorological driving dataset for the GRENE-TEA observation sites (level 0.2), 0.20, Arctic Data archive System (ADS), Japan, available at: <https://ads.nipr.ac.jp/dataset/A20141009-005> (last access: 28 April 2015), 2014a.
- Saito, K., Miyazaki, S., Mori, J., Ise, T., Arakida, H., Suzuki, R., Sato, A., Iijima, Y., Yabuki, H., Iijima, Y., Sueyoshi, T., Hajima, T., Sato, H., Yamazaki, T., and Sugimoto, A.: GTMIP meteorological driving dataset for the GRENE-TEA observation sites (level 1.0), 1.00, Arctic Data archive System (ADS), Japan, available at: <https://ads.nipr.ac.jp/dataset/A20141009-006> (last access: 28 April 2015), 2014b.
- Saito, K., Sasai, T., Miyazaki, S., Mori, J., Ise, T., Arakida, H., Sueyoshi, T., Hajima, T., Iijima, Y., Yamazaki, T., and Sugimoto, A.: GTMIP fraction of Photosynthetically Active Radiation (fPAR) and Leaf Area Index (LAI) for the GRENE-TEA observation sites (level 1.0), 1.00, Arctic Data archive System (ADS), Japan, available at: <https://ads.nipr.ac.jp/dataset/A20141009-007> (last access: 28 April 2015), 2014c.
- Saito, K., Miyazaki, S., Hajima, T., and Sueyoshi, T.: Climate and environmental variations in the Northern terrestrial cryosphere, *Kisho–Kenkyu Note (Meteorological Research Note)*, 230, 196–211, 2014d (in Japanese).
- Saito, K., Miyazaki, S., Mori, J., Ise, T., Arakida, H., Suzuki, R., Sato, A., Iijima, Y., Yabuki, H., Iijima, Y., Sueyoshi, T., Hajima, T., Sato, H., Yamazaki, T., and Sugimoto, A.: GTMIP meteorological driving dataset for the GRENE-TEA observation sites (20 year detrended), 1.00, Arctic Data archive System (ADS), Japan, available at: <https://ads.nipr.ac.jp/dataset/A20150205-001> (last access: 28 April 2015), 2015.
- Sasai, T., Saigusa, N., Nasahara, K. N., Ito, A., Hashimoto, H., Nemani, R. R., Hirata, R., Ichii, K., Takagi, K., Saitoh, T. M., Ohta, T., Murakami, K., Yamaguchi, Y., and Oikawa, T.: Satellite-driven estimation of terrestrial carbon flux over Far East Asia with 1 km grid resolution, *Remote Sens. Environ.*, 115, 1758–1771, doi:10.1016/j.rse.2011.03.007, 2011.
- Sato, A., Kubota, H., Matsuda, M., and Sugiura, K.: Seasonal variation of heat exchange in the boreal forest of Finnish Lapland, in: *Second Wadati Conference on Global Change and the Polar Climate*, extended abstracts, 228–230, 2001.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50 yr high-resolution global dataset of meteorological forcings for land surface modeling, *J. Climate*, 19, 3088–3111, 2006.

Vuichard, N. and Papale, D.: Filling the gaps in meteorological continuous data measured at FLUXNET sites with ERA-Interim reanalysis, *Earth Syst. Sci. Data*, 7, 157–171, doi:10.5194/essd-7-157-2015, 2015.

Watanabe, K., Mizoguchi, M., Kiyosawa, H., and Kodama, Y.: Properties and horizons of active layer soils in tundra at Tiksi, Siberia, *Journal of Japan Society of Hydrology and Water Resources*, 13, 9–16, 2000 (in Japanese with English abstract).

Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: The WATCH forcing data 1958–2001: a meteorological forcing dataset for land surface and hydrological models, *WATCH Tech. Rep. 22*, 41 pp., available at: <http://www.eu-watch.org/publications/technical-reports> (last access: 20 April 2015), 2010.

ESSDD

8, 703–736, 2015

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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 1. Names and coordinates of the observation sites.

Site name	Location	Observation Period	Observation height <i>T</i> ; <i>q</i> ; wind
Chokurdakh	148.264° E,	NA ^a	NA ^a
Fairbanks	70.563° N 147.488° W 65.123° N	1 Jan 2011–31 Dec 2012	1.5 m; 1.5 m; 16 m ^b
Kevo	27.010° E 69.757° N	1 Jan 1995–31 Dec 2001, 1 Jan 1995–31 Dec 2012 (only for precipitation)	2.5 m; 2.5 m; 11.5 m
Tiksi	128.774° E 71.589° N	1 Jan 1997–31 Dec 2012	2 m; 2 m; 10 m
Tura	100.464° E,	NA ^a	NA ^a
Yakutsk	64.209° N 129.618° E 62.255° N	1 Jan 2005–31 Dec 2011	1.8 m; 1.8 m; 32 m ^c

^a Observations at Chokurdakh and Tura were not used since Level 1 data were not created for these sites.

^b Average canopy height is 2.91 m (Nakai et al., 2013).

^c Average canopy height is 18 m (Ohta et al., 2001).

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

T. Sueyoshi et al.

Table 2. Outline of the procedure followed to produce forcing variables for (a) Lv0 and (b) Lv1

(a) Site	Air pressure	Downward longwave radiation	Air temperature (T)	Surface specific humidity (q)	Downward shortwave radiation (SW)	Total precipitation (P)	Wind speed
Any site	ERA Output (constant for 6h)	ERA Output (constant for 6h)	ERA-T + (CRU monthly T - ERA monthly T) linearly interpolated	ERA-relative humidity (= ratio of saturation q at ERA-dew point T to that at ERA-T) multiplied by saturation q at L0 T	Theoretical SW (for lat., lon., day, and time) multiplied by Cloudiness (= ERA daily total SW/Theoretical daily total SW)	ERA 6 hourly P assigned at 1st time slot in every 6h (corrected GPCP monthly P/ERA monthly P)	Calculated from u and v of ERA wind speed (constant for 6h)
(b) Site (Altitude of the site, in ERA) (m)	Air pressure	Downward longwave radiation	Air temperature (T)	Specific humidity (q)	Downward shortwave radiation (SW)	Total precipitation (P)	Wind speed
Fairbanks (210, 390)	L0 data corrected for the altitude diff. L1 = L0 + 24 (2011–2012)	L0 data corrected for the emission diff. due to altitude diff. L1 = 1.01 · L0 + 15 (2011–2012)	L0 data corrected for max. and min. based on CRU dtr (daily temperature range)	ERA-relative humidity multiplied by saturation q at L1 T	L0 data	L0 data	L0 data
Kevo (100, 279)	Same as above L1 = L0 + 21 (1995–2001)	Same as above L1 = 1.00 · L0 + 11 (1995–2001)	L0 data corrected for daily max. and min. based on ERA Max. and Min. T	Same as above	L0 data	Statistical correction on L0' (L0 data w o ⁻¹ GPCP correction) L1 = 0.76 · L0 (1995–2012)	Statistical correction L1 = 0.6 · L0 (for > 1 ms ⁻¹) (1995–2001)
Tiksi (40, 184)	Same as above L1 = L0 + 19 (1997–2012)	Same as above L1 = 0.85 · L0 + 50 (1997–2012)	Statistical correction for season on L0' (ERA T corrected for daily max. and min T based on ERA Max. and Min T), L1 = 0.90 · L0' + 0.1, Nov to Mar L1 = L0', May to Oct	Same as above	L0 data	L0 data corrected for annual total P (ERA annual P/GPCP annual P)	Statistical correction L1 = 1.4 · L0 (for > 1 ms ⁻¹) (1997–2012)
Yakutsk (220, 176)	Same as above L1 = L0 - 5 (2005–2011)	Same as above L1 = 0.98 · L0 + 17 (2005–2011)	L0 data corrected for daily max. and min. based on ERA Max. and Min. T	Same as above	L0 data	L0 data w o ⁻¹ GPCP correction	L0 data

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

T. Sueyoshi et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 3. Root mean square errors for the correlation between monthly reanalysis data and reference data (CRU for air temperature and GPCP for precipitation).

(a) Air temperature at 2 m (°C)				
Site	ERA-Interim	JRA55	NCEP/NCAR	NCEP-DOE
Fairbanks	1.59	3.08	3.44	3.39
Kevo	2.15	2.47	2.86	3.19
Tiksi	6.51	7.87	7.75	8.56
Yakutsk	0.80	2.44	3.20	3.81
(b) Precipitation (mm month⁻¹)				
Site	ERA-Interim	JRA55	NCEP/NCAR	NCEP-DOE
Fairbanks	21.14	28.63	71.33	57.14
Kevo	15.60	17.15	25.50	24.15
Tiksi	18.11	17.92	29.43	24.47
Yakutsk	15.11	13.45	20.42	22.89

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

T. Sueyoshi et al.

Table 4. Names and coordinates of the grid centres of ERA-Interim, CRU, and GPCP data for the corresponding sites shown in Table 1.

Site name	ERA-Interim grid centre	CRU Grid centre	GPCP grid centre
Fairbanks	147.488° W 65.123° N	147.750° W 65.250° N	146.250° W 65.250° N
Kevo	27.010° E 69.757° N	27.000° E 69.750° N	26.250° E 68.750° N
Tiksi	128.774° E 71.589° N	129.000° E 71.250° N	128.750° E 71.250° N
Yakutsk	129.618° E 62.255° N	129.750° E 62.250° N	128.750° E 61.250° N

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



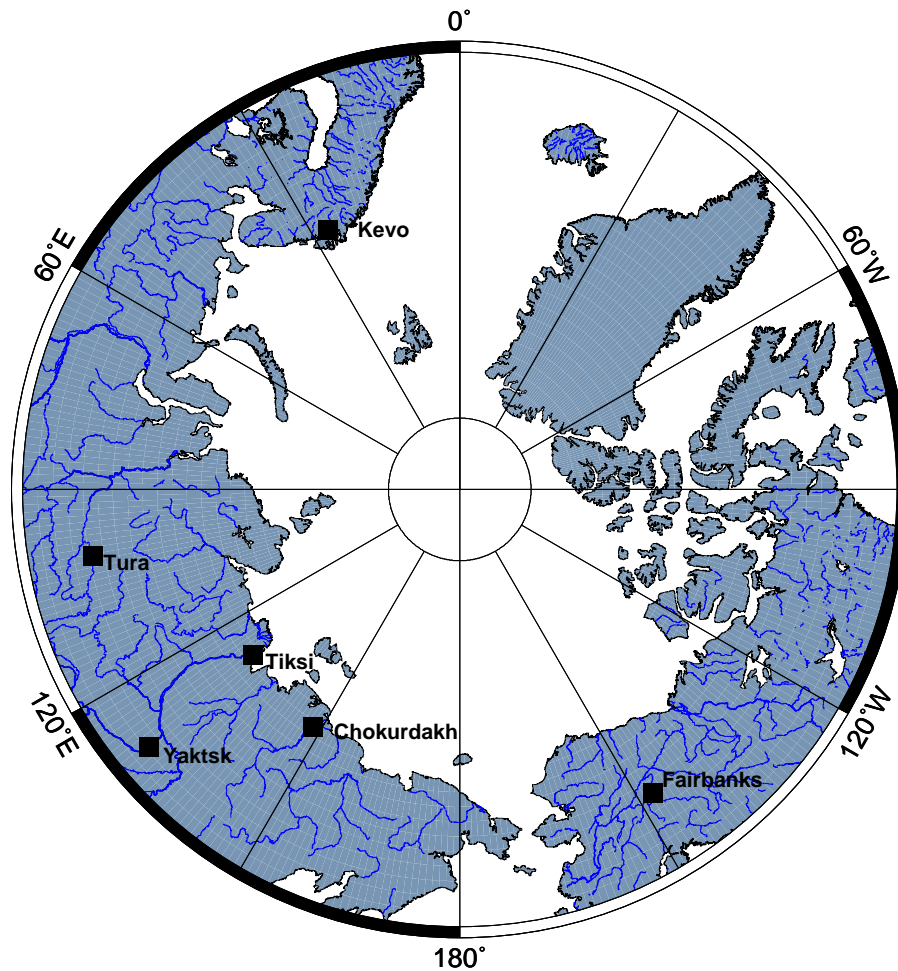


Figure 1. Location of the GTMIP observation sites (after Miyazaki et al., 2015).

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

T. Sueyoshi et al.

[Title Page](#)

[Abstract](#)

[Instruments](#)

[Data Provenance & Structure](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



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

T. Sueyoshi et al.

Schematic of site specific forcing dataset creation for GTMIP

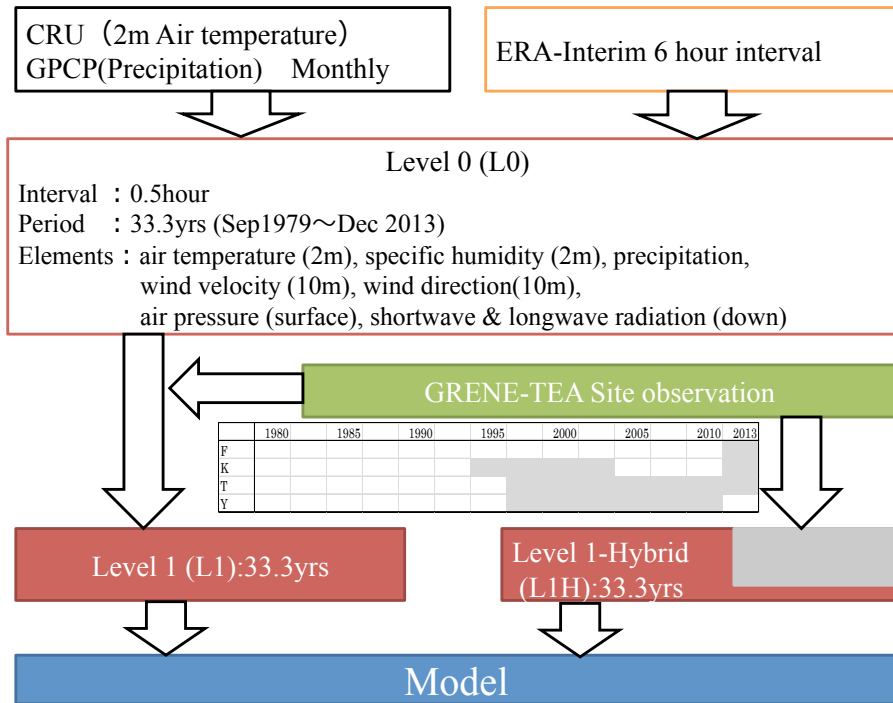


Figure 2. Schematic diagram for data creation during GTMIP stage 1.

[Title Page](#)

[Abstract](#) [Instruments](#)

[Data Provenance & Structure](#)

[Tables](#) [Figures](#)

[◀](#) [▶](#)

[◀](#) [▶](#)

[Back](#) [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



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

T. Sueyoshi et al.

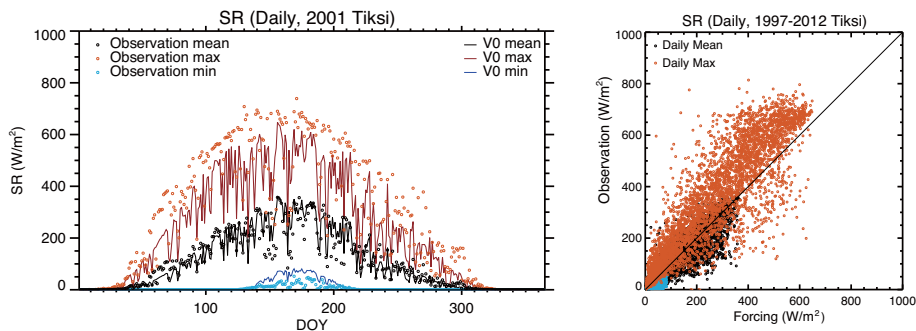


Figure 3. Comparison between Level 0 shortwave radiation and observed values. Example shows one-year time series of daily mean, minima and maxima during 2001 (left), and scatter plot for all daily data (right) at Tiksi.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

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

T. Sueyoshi et al.

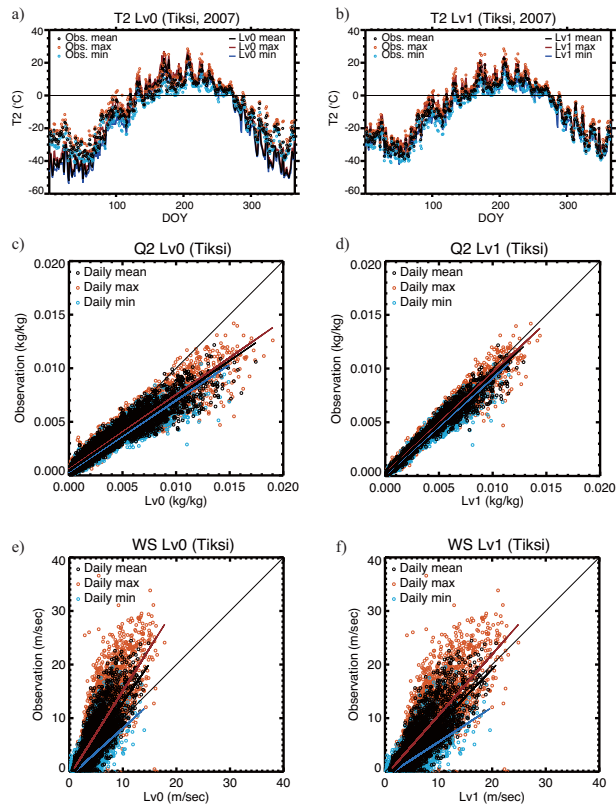


Figure 4. Comparison of GTMIP data with observed values for air temperature (time series in 2007), specific humidity (shown as a scatter plot for all daily data), and wind speed (shown as a scatter plot for all daily data). Examples show the results for Level 0 (top) and Level 1 data (bottom) at Tiksi.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

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

T. Sueyoshi et al.

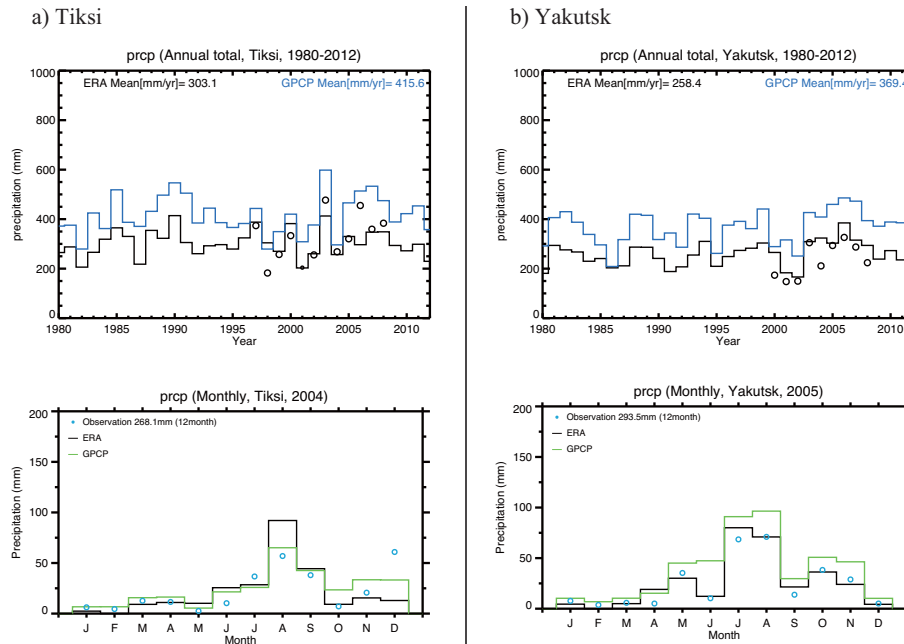


Figure 5. Comparison of annual (upper) and monthly (lower) precipitation amounts between ERA-Interim, GPCP, and observations. Examples show results for Tiksi (left) and Yakutsk (right). Observations are shown in circle. Annual value in 2001 for Tiksi is shown in different symbol (small circle), as there were exceptionally frequent missing data (151 days) in this year.

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

T. Sueyoshi et al.

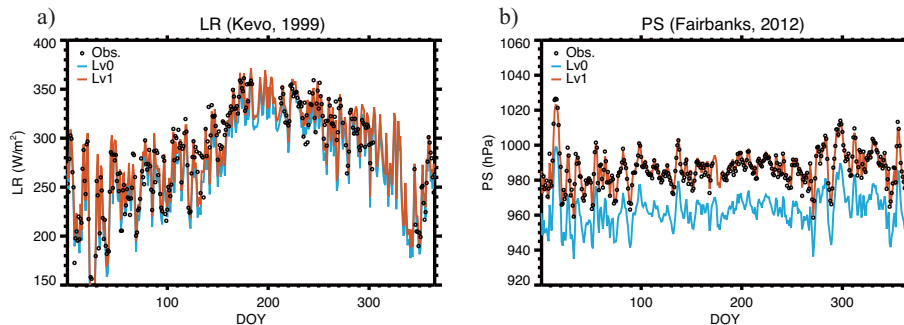


Figure 6. Comparison of GTMIP data and observed values for **(a)** longwave radiation (LR) and **(b)** surface pressure (PS) on a daily basis. Level 0 data are shown in blue and Level 1 data in red. Examples are shown for Kevo (KV) in 1999 (left) and Fairbanks (FB) in 2012 (right).

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

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

T. Sueyoshi et al.

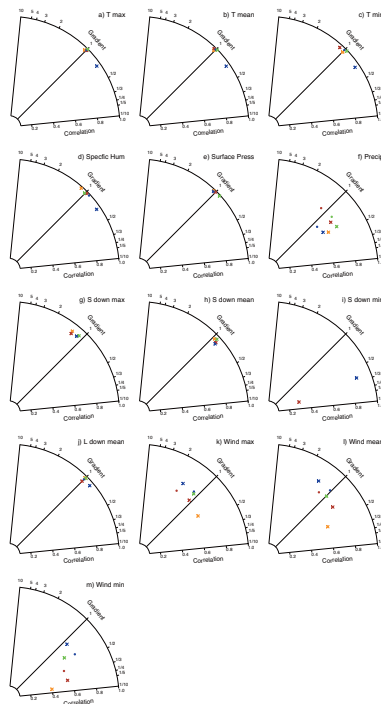


Figure 7. Correlation diagrams for GTMIP data and daily site observation data showing Pearson's correlation coefficient and gradient of the fitted line for **(a)** maximum air temperature, **(b)** mean air temperature, **(c)** minimum air temperature, **(d)** specific humidity, **(e)** surface pressure, **(g)** maximum downward shortwave radiation, **(h)** mean downward shortwave radiation, **(i)** minimum downward shortwave radiation, **(j)** downward longwave radiation, **(k)** maximum wind speed, **(l)**, mean wind speed, and **(m)** minimum wind speed. Level 0 data are shown as crosses and Lv1 data as dots. Sites are shown by colours, i.e. Fairbanks (orange), Kevo (red), Tiksi (blue), and Yakutsk (green).

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)