



RESEARCH ARTICLE

10.1002/2013JD020803

Key Points:

- Global spatio-temporal regression-kriging daily temperature interpolation
- Fitting of global spatio-temporal models for the mean, maximum, and minimum temperatures
- Time series of MODIS 8 day images as explanatory variables in regression part

Correspondence to:

M. Kilibarda,
kili@grf.bg.ac.rs

Citation:

Kilibarda, M., T. Hengl, G. B. M. Heuvelink, B. Gräler, E. Pebesma, M. Perčec Tadić, and B. Bajat (2014), Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution, *J. Geophys. Res. Atmos.*, 119, 2294–2313, doi:10.1002/2013JD020803.

Received 27 AUG 2013

Accepted 28 JAN 2014

Accepted article online 3 FEB 2014

Published online 13 MAR 2014

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution

Milan Kilibarda¹, Tomislav Hengl², Gerard B. M. Heuvelink³, Benedikt Gräler⁴, Edzer Pebesma⁴, Melita Perčec Tadić⁵, and Branislav Bajat¹

¹Department of Geodesy and Geoinformatics, Faculty of Civil Engineering, University of Belgrade, Belgrade, Serbia,

²ISRIC-World Soil Information, Wageningen, Netherlands, ³Soil Geography and Landscape Group, Wageningen University, Wageningen, Netherlands, ⁴Institute for Geoinformatics, University of Münster, Münster, Germany, ⁵Meteorological and Hydrological Service, Zagreb, Croatia

Abstract Combined Global Surface Summary of Day and European Climate Assessment and Dataset daily meteorological data sets (around 9000 stations) were used to build spatio-temporal geostatistical models and predict daily air temperature at ground resolution of 1 km for the global land mass. Predictions in space and time were made for the mean, maximum, and minimum temperatures using spatio-temporal regression-kriging with a time series of Moderate Resolution Imaging Spectroradiometer (MODIS) 8 day images, topographic layers (digital elevation model and topographic wetness index), and a geometric temperature trend as covariates. The accuracy of predicting daily temperatures was assessed using leave-one-out cross validation. To account for geographical point clustering of station data and get a more representative cross-validation accuracy, predicted values were aggregated to blocks of land of size 500 × 500 km. Results show that the average accuracy for predicting mean, maximum, and minimum daily temperatures is root-mean-square error (RMSE) = ±2°C for areas densely covered with stations and between ± 2°C and ± 4°C for areas with lower station density. The lowest prediction accuracy was observed at high altitudes (> 1000 m) and in Antarctica with an RMSE around 6°C. The model and predictions were built for the year 2011 only, but the same methodology could be extended for the whole range of the MODIS land surface temperature images (2001 to today), i.e., to produce global archives of daily temperatures (a next-generation <http://WorldClim.org> repository) and to feed various global environmental models.

1. Introduction

Records from near-surface weather stations are the foundation of climate research [Peterson and Vose, 1997]. They are important not only because of their high reliability and accuracy but also because they are the only available records of spatial and temporal variation of climatic variables before the first satellite based observations became available in the 1960s (M. Kilibarda et al., Publicly available global meteorological data sets: sources, representation, and usability for spatio-temporal analysis, *International Journal of Climatology*, in review, 2014).

Station observations are commonly used to predict climatic variables on raster grids (unvisited locations), where the statistical term “prediction” is used here to refer to “spatial interpolation” or “spatio-temporal interpolation” and should not be confused with “forecasting.” In-depth reviews of interpolation methods used in meteorology and climatology have recently been presented by Price et al. [2000], Jarvis and Stuart [2001], Tveito et al. [2006], and Stahl et al. [2006]. The literature shows that the most common interpolation techniques used in meteorology and climatology are as follows: nearest neighbor methods, splines, regression, and kriging, but also neural networks and machine learning techniques.

One of the first global monthly land surface temperature grids at 0.5° (decimal degrees) resolution was produced by Legates and Willmott [1990]. Leemans and Cramer [1991] further generated grids at the same resolution for mean monthly temperature, precipitation, and cloudiness using a triangulation network followed by smooth surface fitting. New et al. [1999, 2000] and Mitchell and Jones [2005] used thin-plate splines to produce global images covering global land areas excluding Antarctica at 0.5° resolution for nine climatic variables. Hijmans et al. [2005] used a thin-plate smoothing spline on a collection of public meteorological data sets of monthly records to produce global (land mass) climatic images at 1 km resolution for the period from 1960 to 1990. Becker et al. [2012] recently mapped monthly precipitation for the whole world using an

empirical interpolation method based on angular distance weighting at resolutions of 0.25°, 0.5°, 1.0°, and 2.5°, using data from the Global Precipitation Climatology Centre.

The first global terrestrial gridded data set of average daily temperature, daily temperature range, and daily precipitation was developed by *Piper and Stewart* [1996] and is intended for use in terrestrial biospheric modeling. Daily station observations for the year 1987 were interpolated to a 1° × 1° grid (longitude, latitude) using a nearest neighbor interpolation technique. Global daily predictions of meteorological variables were produced by *Kiktev et al.* [2003] and *Alexander et al.* [2006], who used an angular distance weighting technique to interpolate extreme daily precipitation and temperature indices onto a 2.5° latitude by 3.75° longitude grid. *Caesar et al.* [2006] mapped daily maximum and minimum temperature anomalies using the same method and the same output resolutions as *Alexander et al.* [2006]. *Haylock et al.* [2008] produced European coverage maps of daily mean, minimum, and maximum temperatures and precipitation with 0.25° and 0.5° resolution using the European Climate Assessment and Dataset Project (ECA&D). These maps were generated by first estimating monthly averages; daily anomalies from those averages were then interpolated using kriging and added back to the monthly estimates [*Haylock et al.*, 2008]. *Van den Besselaar et al.* [2011] mapped sea level pressure for Europe using the same data source and global kriging. *Di Luzio et al.* [2008] presented a method for mapping daily precipitation and temperature across conterminous USA at 2.5 arc min (around 4 km) for the period 1960–2001. Their method also combines interpolation (inverse distance weighting) of daily anomalies from gridded monthly estimates, here estimated using Parameter Elevation Regressions on Independent Slopes Model.

Hofstra et al. [2008] compared six interpolation methods for prediction of daily precipitation mean, minimum, and maximum temperature, and sea level pressure from station data in Europe in relation to the long-term monthly mean (1960–1990). Global kriging on anomalies (from the long-term monthly mean) showed the best overall performing result. Geostatistical interpolation methods (various types of kriging) are considered to be state-of-the-art statistical approaches to spatial and spatio-temporal analysis [*Cressie and Wikle*, 2011]. A method known as “regression-kriging” (RK) (also known as *kriging with external drift* and/or *universal kriging*) has been widely recognized as a flexible and a well-performing technique for unbiased spatial prediction of meteorological and environmental variables [*Carrera-Hernández and Gaskin*, 2007; *Haylock et al.*, 2008; *Hengl et al.*, 2012].

An extension from spatial to spatio-temporal models is a logical evolution of statistical models for mapping spatially and temporally correlated meteorological data. However, fitting of spatio-temporal RK models and prediction using spatio-temporal covariates is more than just the smoothing of station data. The insights obtained from the modeling process and predictions are much richer and allow one to distinguish sources of variability. The model makes the relationships with covariates explicit and can be used to distinguish between purely temporal, purely spatial and spatio-temporal components of variability. Moreover, it predicts values using spatio-temporal observations in spatial (e.g., as daily map), temporal (e.g., trend line at meteorological stations), and for a time series of spatial grids [*Pebesma*, 2012].

To the best of our knowledge, researchers have not yet attempted to interpolate global daily values of meteorological variables using spatio-temporal regression-kriging with a time series of remote sensing-based covariates and at detailed resolution (1 km). Such an attempt poses many challenges, both from a methodological and technological perspective. For example, the global land mass contains about 149 million pixels at 1 km resolution and predicting daily values for 10 years would result in about 4 TB of data. The fitting of geostatistical models with millions of point pairs and accompanied predictions on such HUGE grid stacks can only be accomplished by code optimization to avoid RAM memory overload and to speed up processing through tiling and parallel processing.

In this article, we present an automated mapping framework for producing predictions of daily mean, minimum, and maximum air temperatures using spatio-temporal regression-kriging. The framework is implemented in the R environment for statistical computing and uses the *gstat* [*Pebesma*, 2004] and *stats* [*R Development Core Team*, 2012] packages. As inputs we use a collection of publicly available daily records from the National Climatic Data Center’s (NCDC’s) Global Surface Summary of Day and the European Climate Assessment & Dataset and a time series of MODIS LST 8 day images and topographic layers as covariates. The research reported here focuses on the year 2011 but without modification the methodology can be extended to the whole period for which MODIS LST images are available (2001 to present). The input data

sets and methods are described in section 2. The results of model fitting, cross validation, and validation are presented in section 3, whereas summary results are given in the discussions section (section 4).

2. Methods and Materials

2.1. Measurements at Ground Stations

2.1.1. NCDC's Global Surface Summary of Day

The Global Surface Summary of Day (GSOD) data set is produced and archived at NOAA's National Climatic Data Center (NCDC) under the code NCDC DSI-9618 (<http://www.ncdc.noaa.gov/>). The input data used in building the GSOD are the Integrated Surface Data (NCDC DSI-3505), which includes global hourly data obtained from the U.S. Federal Climate Complex that contains about 27,000 stations.

GSOD is probably the largest publicly available international meteorological station data set. It contains daily measurements for a list of 11 meteorological parameters of several climatic variables (since 1929): mean, minimum, and maximum temperatures (precision of 0.1°F); mean dew point (0.1°F), mean atmospheric pressure and mean sea level pressure (0.1 mb); mean visibility (0.1 mi); mean wind speed, maximum sustained wind speed, and maximum wind gust (0.1 knots); precipitation amount (0.01 in); snow depth (0.1 in); and an indicator (class) for occurrence of fog, rain, or drizzle, snow or ice pellets, hail, thunder, and tornado/funnel cloud. This data set is continuously being updated so that the latest daily summary data are usually available within 1–2 days after the observations were made.

2.1.2. European Climate Assessment and Dataset

The largest European publicly available meteorological data set is the European Climate Assessment and Dataset (ECA&D) project. The idea of the ECA&D project was to provide a uniform analysis methodology for a daily observational series from 62 countries and 6596 European and Mediterranean meteorological stations [Klein Tank et al., 2002].

The ECA&D data set contains measurements of the following meteorological variables and their parameters: minimum, mean, and maximum temperature (precision of 0.1°C), humidity (1 %), mean sea level pressure (0.1 hPa), mean wind speed (0.1 m/s), wind direction (degrees), maximum wind gust (0.1 m/s), precipitation amount (0.1 mm), snow depth (1 cm), cloud cover (oktas), and sunshine duration (0.1 h).

2.1.3. The Global Historical Climatology Network

The Global Historical Climatology Network-Monthly (GHCN-M) temperature data set was developed in the early 1990s [Vose, 1992]. A second version was released in 1997 following extensive efforts to increase the number of stations and length of the data record [Peterson and Vose, 1997]. GHCN-M version 3 released in 2011 focused on improving four issues [Lawrimore et al., 2011]: (a) consolidating duplicate station records, (b) improving station coverage, especially during the 1990s and 2000s, (c) enhancing quality control, and (d) applying a new bias correction methodology that does not require use of a composite reference series.

The current version 3 contains monthly mean temperature, monthly maximum temperature, and monthly minimum temperature.

2.1.4. Merged and Cleaned Global Station Data Set

GSOD and ECA&D daily meteorological data sets were merged to produce a merged global station data set. In case both sources have the same location the station with most observations was kept; in case of the same number of observations GSOD was taken. The final merged data set contains around 9000 stations from GSOD and around 500 from ECA&D.

A large portion of the station data had missing values, but all stations were used in the interpolation procedure. Even though the meteorological services responsible for collecting the data also perform a basic quality control, a small portion of the stations from the merged data set contained obvious gross errors and needed to be cleaned. The gross errors were detected using the following procedure: An initial spatio-temporal model (described in section 2.3) was first fitted using all data, followed by analysis of cross-validation predictions at the station locations. Large cross-validation root-mean-square errors indicated which stations might have gross data errors. This was usually confirmed with abrupt jumps in observation values in time series plots, thereby showing observations against the cross-validation prediction error.

We decided to remove all stations that had an absolute cross-validation residual larger than $\pm 15^\circ\text{C}$, as these clearly contain errors in the data set. Using the 15°C threshold, we removed around 300 stations for each temperature variable (minimum, mean, and maximum temperature). After all data filtering, the final set contained about 9000 stations from merged GSOD and ECA&D data sets.

The number of stations per region are as follows: North and South America approximately 3100 stations, Europe approximately 2500, Africa approximately 750, Asia approximately 1600, Australia approximately 450, and Antarctica approximately 20.

2.2. Covariates: Remote Sensing Images and DEM Derivatives

2.2.1. National Aeronautics and Space Administration

The Moderate Resolution Imaging Spectroradiometer (MODIS) images of the Terra and Aqua Earth Observing System platforms provide retrieval of atmospheric and oceanographic variables using various techniques. There are many data products from MODIS observations describing land (temperature and land cover), oceans (sea surface temperature and optical thickness), and atmosphere (water vapor, cloud product, and atmospheric profiles) that can be used for studies at different spatial scales, ranging from local to global. Land surface temperature (LST) data and images obtained from MODIS thermal bands that are distributed by the Land Processes Distributed Active Archive Center of the U.S. Geological Survey are very often used in meteorological studies [Coll *et al.*, 2009].

In addition to the MODIS data at level 2 or higher, there are also MODIS level 1 data that are distributed through the Level 1 and Atmosphere Archive and Distribution System (LAADS) portal hosted at the Goddard Space Flight Center. The MODIS LST data are available on a daily basis and have spatial resolutions of 1×1 km [Coll *et al.*, 2009]. The nominal accuracy of the MODIS LST product is reported to be ± 1 K. Some validation studies reported accuracy statistics smaller than 1 K in clear sky conditions within the temperature range of -10°C to 50°C [Yoo *et al.*, 2011]. MODIS LST data and/or images are one of the most often used and best documented publicly available remote sensing products in the world.

In this work, we only use level 3 MODIS LST 8 day composite images (MOD11A2) to improve the spatial prediction of mean, minimum, and maximum daily temperature, despite the fact that daily daytime and nighttime MODIS LST images are also available (MOD11A1). The intention was to find covariates that are exhaustively known over the spatio-temporal domain to obtain complete prediction over landmass. Even using the MODIS 8 day composites did not completely fulfill this requirement, so remaining gaps had to be filled using splines.

The original MODIS images still contained 0–15% of missing pixels due to clouds or other reasons. These missing pixels were replaced using the System for Automated Geoscientific Analyses Geographic Information System (SAGA GIS) function *Close Gaps*. This function uses spline interpolation as a robust method for filling gaps in areas with sparse or irregularly spaced data points [Neteler, 2010]. Furthermore, the 8 day images were also disaggregated in the time dimension through the use of splines for each pixel.

The original 8 day MODIS LST images were also converted from Kelvin to degrees Celsius using the formula indicated in the MODIS user's manual [Wan *et al.*, 2004].

2.2.2. DEM Derivatives

The elevation data used in this study were obtained from the *WorldGrids.org* portal. The data set is derived as a combination of the publicly available Shuttle Radar Topography Mission 30+ and ETOPO data sets and is commonly referred to as DEMSRE. DEMSRE comes with an accompanying processing script providing detailed instructions for layer reproduction. DEMSRE is the main source of many geomorphometric layers on the *WorldGrids.org* portal, such as slope, potential incoming solar radiation, topographic wetness index, etc.

The topographic wetness index combines local upslope contributing area and slope, ($\text{TWI} = \ln(A/\tan(\beta))$), where A (m^2) is the contributing area and $\tan(\beta)$ the slope (β is the angle of the raster cell relative to the horizontal plane). It is commonly used to quantify topographic control on hydrological processes [Sorensen *et al.*, 2006] but can also be used as an indicator of cold air accumulation [Bader and Ruijten, 2008]. Methods of computing TWI differ primarily in the way the upslope contributing area is calculated. The *SAGA Wetness Index* is based on a modified catchment area calculation implemented in SAGA GIS [Böhner *et al.*, 2008].

The process of computing the SAGA Wetness Index for global land areas is very time consuming. Computations require several days to complete even if a strong PC configuration is used. To achieve such processing, it is necessary to tile the DEM at continental level, reproject to equal area projection, compute SAGA Wetness Index for each tile and build a mosaic for the global landmass. The Global SAGA Wetness Index used in this study was produced by the authors. The processing script is available via the *WorldGrids.org* data portal.

2.3. Spatio-Temporal Regression-Kriging

Consider the problem of describing the spatio-temporal process of a continuous variable Z , where Z varies over space and time, e.g., temperature varies in space from one location to another and in time from one point in time to another. The statistical model of such a process is typically composed of the sum of a trend and a stochastic residual [Burrough, 1998; Heuvelink and Griffith, 2010; Hengl et al., 2012]:

$$Z(\mathbf{s}, t) = m(\mathbf{s}, t) + \varepsilon'(\mathbf{s}, t) + \varepsilon''(\mathbf{s}, t) \quad (1)$$

where \mathbf{s} , t are the space-time coordinates, m is the trend, $\varepsilon'(\mathbf{s}, t)$ is the spatio-temporal correlated stochastic component, and $\varepsilon''(\mathbf{s}, t)$ is the uncorrelated noise. Z is observed at a finite set of points in space and time. An interpolation technique is required in order to predict Z at an unobserved location or time. Geostatistical interpolation starts by defining the model that describes the degree of variation of the variable of interest in space and time, followed by characterizing its relationship with explanatory variables. These explanatory variables are often denoted as *covariates*.

The global trend of Z is taken as a function of covariates known over the spatio-temporal domain, e.g., part of the variation of temperature can be explained with static covariates such as latitude, elevation and TWI, and time dependent covariates such as day of year, and with space and time dependent covariates such as MODIS LST. It is convenient to represent the relationship between the dependent variable and the covariates using a linear model. The linear trend model is given by

$$m(\mathbf{s}, t) = \sum_{i=0}^p \beta_i f_i(\mathbf{s}, t), \quad (2)$$

where the β_i are unknown regression coefficients, the f_i covariates that must be exhaustively known over the spatio-temporal domain, and p is the number of covariates. Covariate f_0 is taken as unity, resulting in β_0 representing the model intercept.

The linear model for the mean daily temperature is given as a multiple linear regression on the covariate layers (described in section 2.2). In addition to these covariate layers, we assume that the global daily temperature is a function of geometric position of a particular location on Earth and day of the year. We call this a *geometric temperature trend*. The geometric temperature trend for the mean temperature was modeled as a function of the day of year and latitude (ϕ):

$$t_{\text{geom}} = 30.4 \cos \phi - 15.5(1 - \cos \theta) \sin |\phi|, \quad (3)$$

where θ is derived as

$$\theta = (\text{day} - 18) \frac{2\pi}{365} + 2^{1-\text{sgn}(\phi)} \pi. \quad (4)$$

The number 18 represents the coldest day in the Northern and warmest day in the Southern Hemisphere and was derived empirically by graphical inspection of mean daily temperature plots from stations in the Northern and Southern Hemispheres. The sgn denotes the signum function that extracts the sign of a real number. Parameters 30.4°C and 15.5°C of the geometric temperature trend were calculated by least squares fitting on circa 44 million daily temperature observations from 2000 to 2011. These two numbers are, in fact, similar to the mean annual temperature on the equator and the mean global Earth temperature.

The linear model for the minimum daily temperature uses the same covariates as the linear model for mean daily temperature. The geometric temperature trend for minimum daily temperature was

$$t_{\text{geom}} = 24.2 \cos \phi - 15.7(1 - \cos \theta) \sin |\phi|, \quad (5)$$

The geometric temperature trend for maximum daily temperature was

$$t_{\text{geom}} = 37 \cos \phi - 15.4(1 - \cos \theta) \sin |\phi|, \quad (6)$$

Parameters 24.2°C and 15.7°C in equation (5) and 37°C and 15.4°C in equation (6) were also derived by the method of least squares estimation using 11 years of observations (from 2000 to 2011).

In recent years, many covariate layers with fine resolution have become available and regression models can often explain a significant part of the observed variation. However, in practice, the trend cannot explain all variation even though covariates are spatially, temporally, and spatiotemporally varying. The residuals of the regression model might show spatio-temporal dependencies, which indicates that a spatio-temporal variogram may be estimated from the residuals at observation locations and used to interpolate the residuals using kriging. To make this explicit, we write the variable of interest as a sum of the trend and space-time residual:

$$Z(\mathbf{s}, t) = m(\mathbf{s}, t) + V(\mathbf{s}, t), \quad (7)$$

where V is a zero-mean stochastic residual.

To proceed with estimation of the spatio-temporal covariance structure of V , we assume it stationary and spatially isotropic. In other words, we assume that the variance of V is constant and that the covariance of V at points (\mathbf{s}, t) and $(\mathbf{s}+\mathbf{h}, t+u)$ only depends on the separation distance (h, u) , where h is the Euclidean spatial distance $|\mathbf{h}|$. These assumptions might be difficult to fulfil for the random field Z but are more plausible for the residuals. The spatio-temporal covariances are usually described using a spatio-temporal variogram, which measures the average dissimilarity between data separated in the spatio-temporal domain using the distance vector (h, u) defined as

$$\gamma(h, u) = \frac{1}{2} \mathbb{E}(V(\mathbf{s}, t) - V(\mathbf{s} + \mathbf{h}, t + u))^2 \quad (8)$$

where \mathbb{E} denotes mathematical expectation.

The residual V may be thought of as comprising three components: spatial, temporal, and spatio-temporal interaction [Heuvelink *et al.*, 2012]. The sum-metric variogram structure that considers these three components as mutually independent is defined as [Heuvelink *et al.*, 2012]

$$\gamma(h, u) = \gamma_S(h) + \gamma_T(u) + \gamma_{ST}(\sqrt{h^2 + (\alpha \cdot u)^2}), \quad (9)$$

where $\gamma(h, u)$ denotes the semivariance of V with h units of distance in space and u units of distance in time, γ_S , γ_T are purely spatial and temporal components, and γ_{ST} is the space-time interaction component. The spatio-temporal anisotropy ratio α converts units of temporal separation (u) into spatial distances (h). The spatio-temporal sum-metric variogram model can be seen as a surface with ten parameters; three parameters for each variogram component (sill, nugget, and range) and the spatio-temporal anisotropy parameter α . Semivariances (and covariances) can be calculated for any spatio-temporal separation distance (h, u) once these parameters are estimated from the observed residuals. In turn, these can be used in spatio-temporal kriging to compute the best linear unbiased predictor (i.e., with minimum expected mean-squared error) for any space-time point where V (and Z) was not observed. The formulas of kriging in the spatio-temporal domain do not differ fundamentally in a mathematical or statistical sense from those of spatial kriging [Heuvelink *et al.*, 2012]:

$$\hat{V}(\mathbf{s}_0, t_0) = \mathbf{c}_0^T \mathbf{c}^{-1} \bar{\mathbf{V}}, \quad (10)$$

where \mathbf{c} is the $n \times n$ variance-covariance matrix of the residuals at the n observation space-time points, as derived from the spatio-temporal variogram, \mathbf{c}_0 is a vector of covariances between the residuals at the observation and prediction points, \mathbf{T} denotes matrix transpose, and $\bar{\mathbf{V}}$ is a vector of residuals (see equation (7)) at the n observation points.

The final prediction of variable Z at location (\mathbf{s}_0, t_0) is defined as

$$\hat{Z}(\mathbf{s}_0, t_0) = \hat{m}(\mathbf{s}_0, t_0) + \hat{V}(\mathbf{s}_0, t_0). \quad (11)$$

where $\hat{m}(\mathbf{s}_0, t_0)$ is the estimated multiple linear regression trend. The regression coefficients are estimated in the usual way, using where possible generalized least squares or ordinary least squares [Pinheiro and Bates, 2009].

Note that regression-kriging specifically implies that the regression modeling and residual kriging parts are addressed separately: We first produce predictions for the regression part (see equation (2)), followed by

extracting residuals for all observations and next fit a global sum-metric variogram model. The residuals are then interpolated and added to the predicted trend.

Spatio-temporal regression-kriging has made a breakthrough in the past decade with theoretical concepts and various real-world applications [Gething et al., 2007; Heuvelink and Griffith, 2010; Heuvelink et al., 2012; Gräler et al., 2011; Hengl et al., 2012]. Here, we extend the spatio-temporal regression-kriging framework that combines ground observations with MODIS 8 day images, as was presented in a Croatian case study [Hengl et al., 2012], to a global data set and hyper-resolution data. We implemented the statistical methodology in the R environment for statistical computing [R Development Core Team, 2012] by combining functionality of the gstat package (geostatistical modeling), rgdal and raster packages (raster data loading and analysis), and snowfall package (cluster computing). We used the gstat package [Pebesma, 2004] that can work with spatio-temporal data sets defined in the spacetime package [Pebesma, 2012] for space-time variogram model fitting. The sample variograms were estimated with spatial lags of 50 km and time lags of 1 day. Because this is a global point data set, all distances were calculated as great circle distances in the WGS84 coordinate reference system.

2.4. Accuracy Assessment

Two approaches were applied for assessing the accuracy of the predictions made for the daily temperature of the global land surface as obtained with spatio-temporal regression-kriging. These were (1) cross validation and (2) comparison with GHCN-M monthly data. For validation using GHCN-M data, we predicted values at daily resolution and aggregated these predictions to monthly averages. Stations from the GHCN-M data set that were closer than 50 km to any station used in this study were excluded in order to avoid station overlap and to increase the independence of the validation data to obtain more objective results.

For cross validation, we used leave-one-out cross validation. The method works as follows: The model predicts a complete annual time series of daily temperature at all stations using only observations from neighboring stations (i.e., the 35 nearest observations in space were sampled for the current day and the day before and after the current, resulting in a collection of around 105 observations that do not include any data from the target station itself). The predicted values were compared with the actual observations of the target stations to derive cross-validation statistics. The final accuracy of the model is assessed using the root-mean-square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{m} \cdot \sum_{j=1}^m [\hat{T}(s_j, t_j) - T(s_j, t_j)]^2} \quad (12)$$

where $\hat{T}(s_j, t_j) - T(s_j, t_j)$ is the difference between the cross-validation prediction and the observed temperature at space-time point (s_j, t_j) , and m is the number of observations for the station. The derived RMSE per station were then exported to KML and HTML formats to allow for visual exploration of error statistics in space and time. These visualizations can be accessed via the <http://dailymeteo.org> website.

Because stations are heavily clustered (M. Kilibarda et al., in review, 2013), the global RMSE mostly depends on the accuracy in areas with a high station density. In order to obtain a more objective measure of accuracy that accounts for clustering the block aggregated RMSE for 500×500 km blocks of land prepared in sinusoidal equal area projections was also computed and analyzed. The regression-kriging cross-validation statistics were first calculated in the WGS84 coordinate reference system using geodetic line distances and were then reprojected to sinusoidal projection for the block aggregation.

3. Results

3.1. Mean Daily Temperature Interpolation

3.1.1. Linear Regression for Mean Daily Temperature

Figure 1 shows the geometric trend values against observed temperatures for two example stations. Surprisingly, the geometric temperature trend already explains 75% of the daily temperature variation with a standard error of $\pm 5.7^\circ\text{C}$.

Figure 1 shows the disaggregated MODIS LST 8 day layer (MODIS spline) against observed temperature for two stations. The linear regression model using only MODIS LST spline images explains 80% of the variability in mean daily temperature values for the year 2011. Thus, MODIS LST spline images are significant estimators of the daily temperature with an average precision (i.e., standard error) of $\pm 5.2^\circ\text{C}$. Again, this precision is

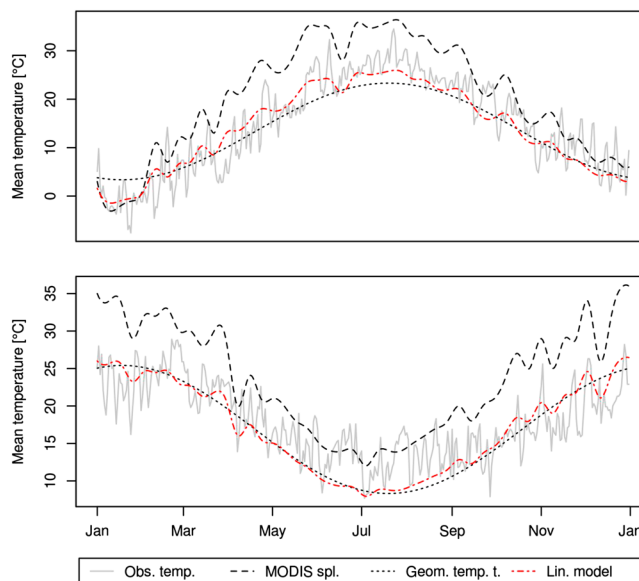


Figure 1. Mean daily temperature observation in 2011 (gray solid line) and multivariate linear model of mean daily temperature (red dashed line) on MODIS LST spline (black dashed line), geometric temperature trend (black dotted line), elevation, and topographic wetness index. (top) Philadelphia, USA ($\lambda = -75, \phi = 39.993$). (bottom) Bunbury, Western Australia ($\lambda = 115.65, \phi = -33.35$).

lower than the one reported by *Wan et al.* [2004] because we use 8 day composites and not daily MODIS LST images in order to reach full land coverage.

The DEM and TWI layers also appeared to be significant covariate layers even though we expected that MODIS LST would account for the variation of temperature with elevation. We suspect that the main reason that part of the elevation dependency was not explained by MODIS LST is due to the fact that this is a cloud-free product: It is likely to underestimate winter temperatures (due to strong radiative cooling in cloud-free situations) and overestimate summer temperatures [*Van De Kerchove et al.*, 2013]. As a consequence, during winter days/nights surface observed temperatures would be higher under clouds due to suppressed radiative cooling, and the MODIS LST gap-filling procedure that we employed would probably underestimate temperature in these areas. During summer, under clouds in mountains, the observed

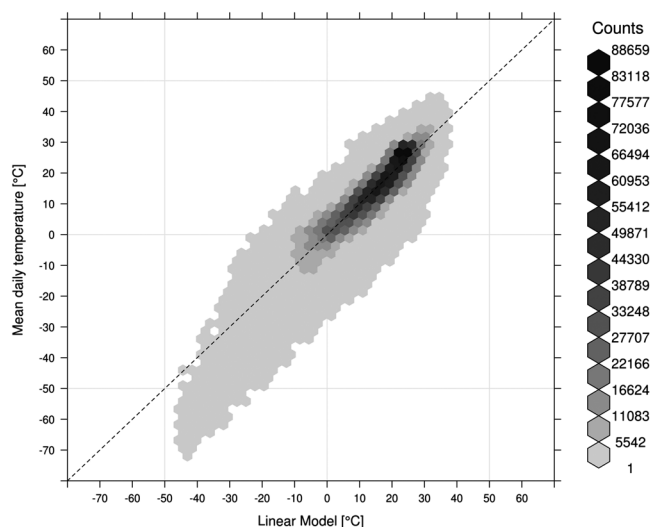


Figure 2. Scatterplot showing the general relationship between mean daily temperature and multivariate linear model prediction of mean daily temperature. The dashed line is the 1:1 relationship.

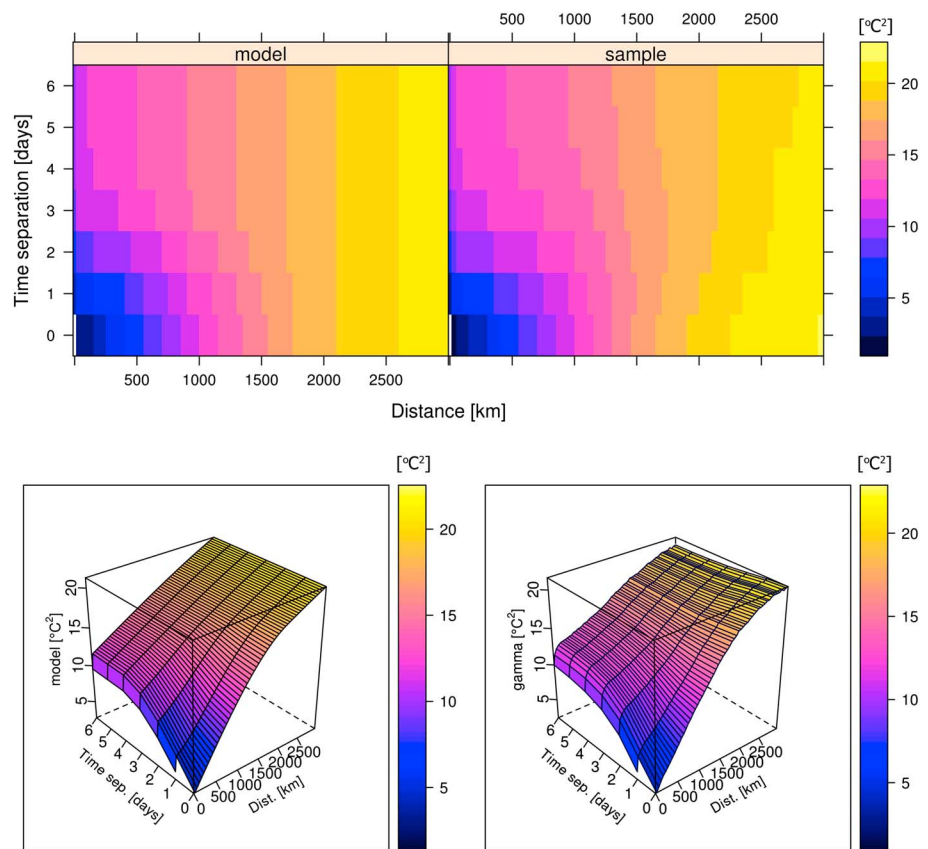


Figure 3. Fitted (left) sum-metric model and (right) sample variogram of residuals from multiple linear regression of mean daily temperature on MODIS, geometric temperature trend, elevation, and topographic wetness index. The variogram surface is presented in (top) 2-D and (bottom) 3-D.

surface temperature would be lower, while the gap-filling procedure would result in higher temperatures. Since these two processes are mainly elevation dependent, this effect can be corrected for with DEM and TWI covariate layers.

The final multiple linear model with four covariates explains 84.2% of the variation and has a residual standard deviation of $\pm 4.6^\circ\text{C}$. Figure 1 shows plots of modeled against observed temperature for the same stations as used in previous figures.

Figure 2 presents the general relationship between the observed temperature and linear model on the full data set used for spatio-temporal modeling. Note that the residuals are in general normally distributed around the regression line and that no heteroscedasticity can be observed.

3.1.2. Spatio-Temporal Variogram Model for Mean Daily Temperature

The right-hand side of Figure 3 shows the 2-D and 3-D sample space-time variogram. The fitted model (10 variogram parameters described in section 2.3 with the *fit.StVariogram* function in *gstat*) is shown in the left-hand side of Figure 3. Table 1 summarizes the parameter estimates of the sum-metric variogram model. Note that all variogram components were modeled as spherical functions.

Table 1. Parameters of the Fitted Sum-Metric Variogram Model for Mean Daily Temperature Regression Residuals, Each Component (See Equation (9)) Was Modeled Using a Spherical Function

	Nugget	Sill	Range Parameter	Anisotropy Ratio
Spatial	1.934	14.13	5903 km	
Temporal	0	0	0 days	
Space-time	0.474	9.065	2054 km	497 km/d

Figure 3 indicates that regression residuals have clear correlations both in space and time, and therefore, spatio-temporal kriging of residuals is certainly applicable. The fitted spatio-temporal variogram parameters of the mean daily temperature

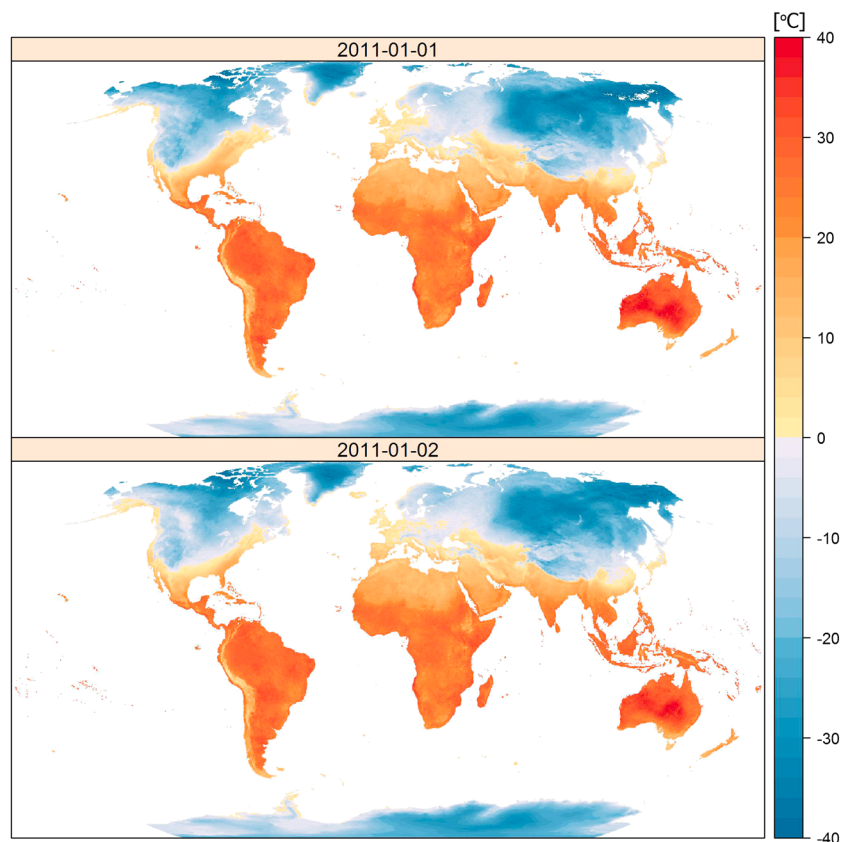


Figure 4. Map of mean daily temperature, interpolated by using spatio-temporal regression-kriging on the GSOD and ECA&D stations observation. The map is Robinson projection.

residuals show a significant purely spatial variogram component, while the purely temporal component is zero and temporal variability is only contained in the space-time interaction component. This suggests that the temporal pattern in mean temperature is probably already sufficiently captured by the regression model. The short-distance variation (nugget effect) in both the purely spatial and spatio-temporal component indicates that the model cannot yield better precision than $\pm 1.5^\circ\text{C}$ globally (for interpolation at daily resolution). The range parameters are very large (especially the pure spatial range) showing that the residuals are correlated along large distances up to 6000 km. In kriging the local neighborhoods must be selected in a way that respects the spatial and temporal variogram ranges. Thus, only a few temporal instances will be selected, while the spatial selection spans several hundreds of kilometers. This is also captured by the spatio-temporal anisotropy, which shows that stations with a temporal lag of 1 day exhibit a similar correlation as stations that are about 500 km apart.

3.1.3. Accuracy Assessment: Mean Daily Temperature

An interpolated map of daily mean temperature for the 1 and 2 January of 2011 is shown in Figure 4. Daily maps for the year 2011 at 1 km spatial resolution in GeoTiff format are available for download via <http://www.dailymeteo.org>. The mean daily temperature map of coterminous USA for the first 4 days in January 2011 is presented in Figure 5.

The cross-validation results on the complete data set yield $\text{RMSE} = 2.4^\circ\text{C}$ for global land areas including Antarctica, with an R-square of 96.6%. The block aggregated RMSE results show that the average accuracy is a bit worse ($\text{RMSE} = 2.8$; see also Figure 6). As mentioned previously, the global block aggregated RMSE gives a more objective global measure of accuracy. Thus, the actual RMSE is half a degree larger than the RMSE calculated as an unweighed mean from all stations.

The monthly RMSE obtained from cross validation of monthly aggregated observations with cross-validation prediction is 1.7°C . This is an important result because it indicates that the model can be used for monthly image production (aggregation of daily gridded data). The yearly RMSE is 1.4°C . The spatial

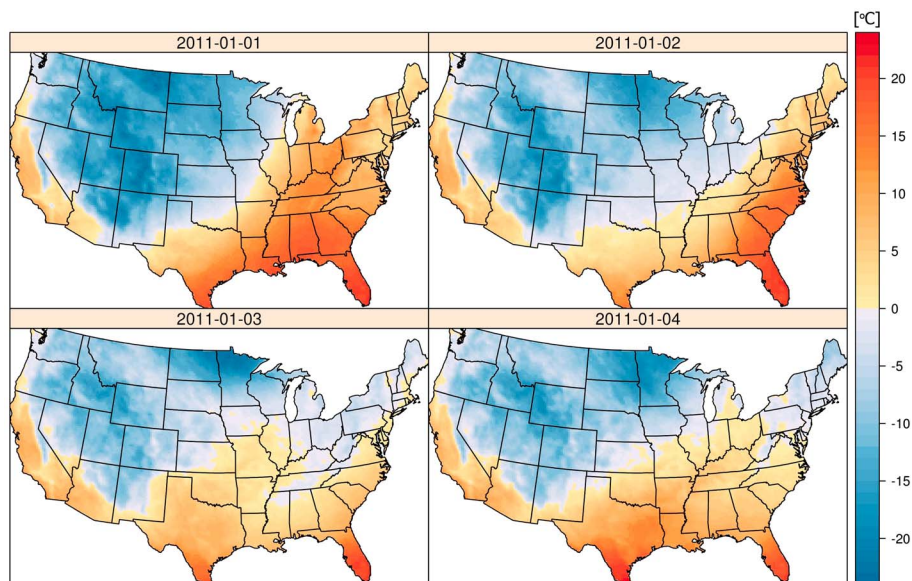


Figure 5. Map of mean daily temperature for the first 4 days in January 2011, covering coterminous USA in Albers equal-area conic projection.

distribution of RMSE calculated per station (yearly average of squared daily cross-validation residuals, which is a daily quality measure) is shown in Figure 7. In this figure, the stations with $RMSE < 2^{\circ}C$ represent 59% of the total number of stations (Figure 7, black dots). Note that 26% of stations have an RMSE between $2^{\circ}C$ and $3^{\circ}C$. Figure 7 is also provided as an interactive map produced with the R package plotGoogleMaps [Kilibarda and Bajat, 2012] and is available via <http://www.dailymeteo.org>.

Observed and cross-validated results for two stations are shown in Figure 8. Considering the fact that cross-validation predictions are made using only 35 neighboring stations in space and three observation days in time) without using any observations from the validation station, it can be concluded that the spatio-temporal regression-kriging model is an accurate tool for filtering missing values in time series of mean daily temperatures.

The spatial distribution of the RMSE can also be aggregated in the spatial domain by region or country. The aggregated results show that the smallest $RMSE = 1.1^{\circ}C$ is achieved in the Netherlands, whereas Europe on average performs with an $RMSE = 1.6^{\circ}C$. Other results for large countries and regions are Russia

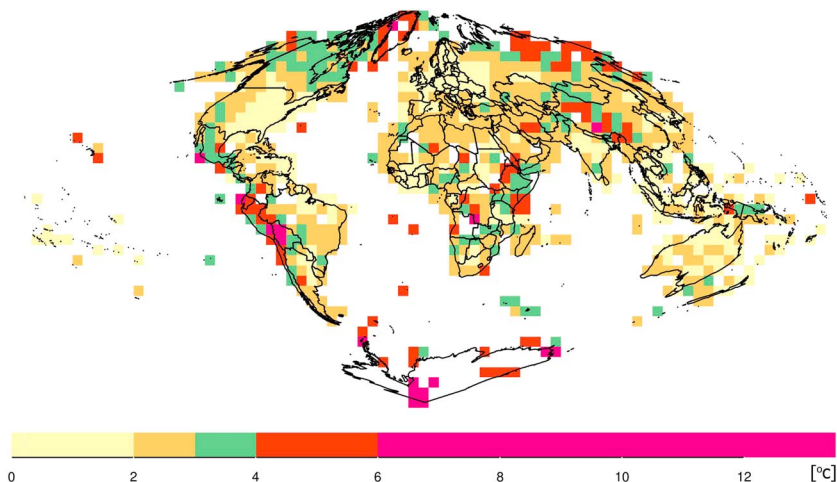


Figure 6. Map of mean daily temperature cross-validation errors (RMSE) aggregated to 500×500 km blocks. Block aggregation is made on equal area sinusoidal projection.

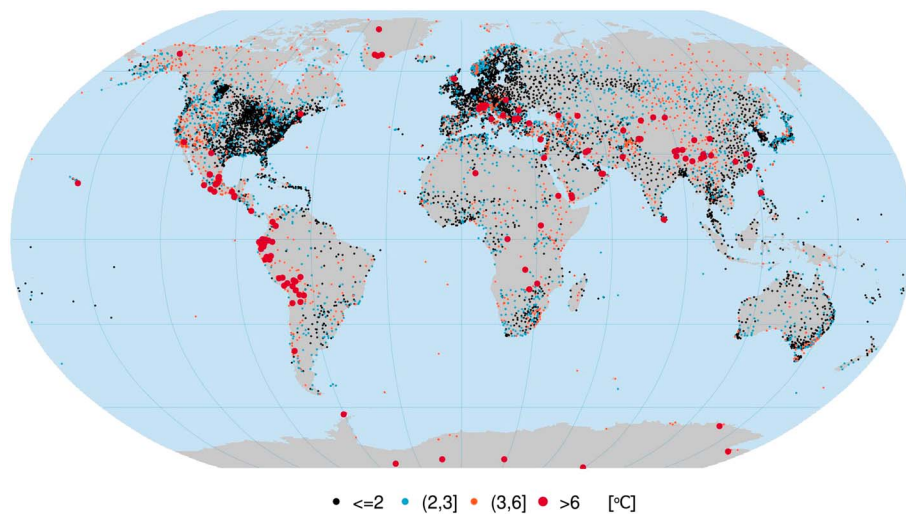


Figure 7. Map of mean daily temperature cross-validation errors (RMSE) averaged per year for each station. Red circles indicate cross-validation outliers with RMSE > 6. Clusters of red circles indicate problematic areas, partly present of gross error in observation time series.

(approximately 2.4°C), USA 1.8°C, South America 3.1°C, while Antarctica has the largest RMSE with 5.9°C. An interactive map of spatially aggregated RMSE at country level is available via <http://www.dailymeteo.org>.

The RMSE on the GHCN-M data is 1.5°C and the spatial distribution of RMSE calculated per year (yearly average of squared monthly validation residuals) for each station is shown in Figure 9 (an interactive map is available at <http://www.dailymeteo.org>). This map shows that 48% of the predicted points have a prediction accuracy smaller than 1°C. GHCN-M stations at a monthly resolution have an accuracy between 1 and 2°C for 40% of the points.

3.2. Minimum Daily Temperature Interpolation

3.2.1. Linear Regression Model for Minimum Daily Temperature

The geometric temperature trend explains about 72% of the minimum daily temperature variations for 2011, with a standard error of ±6°C. Figure 10 shows the geometric trend against observations. The results

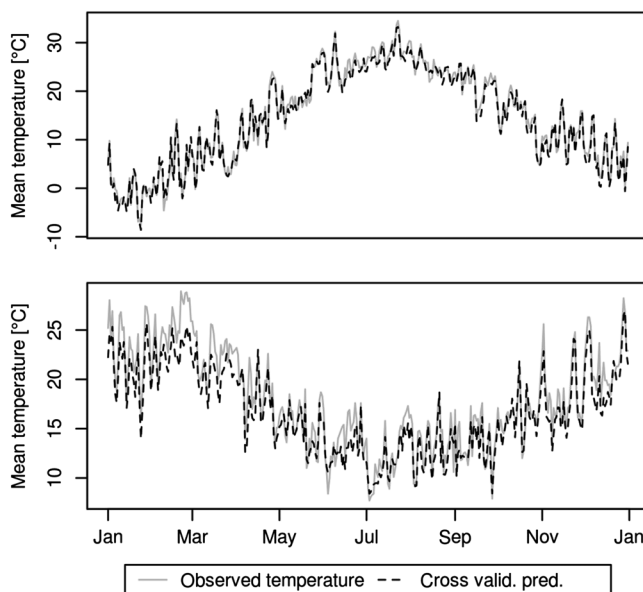


Figure 8. Comparison of mean daily temperature observations in 2011 (gray solid line) and space-time regression-kriging cross-validation prediction of mean daily temperature (black dashed line). (top) Philadelphia, USA ($\lambda = -75, \phi = 39.993$). (bottom) Bunbury, Western Australia ($\lambda = 115.65, \phi = -33.35$).

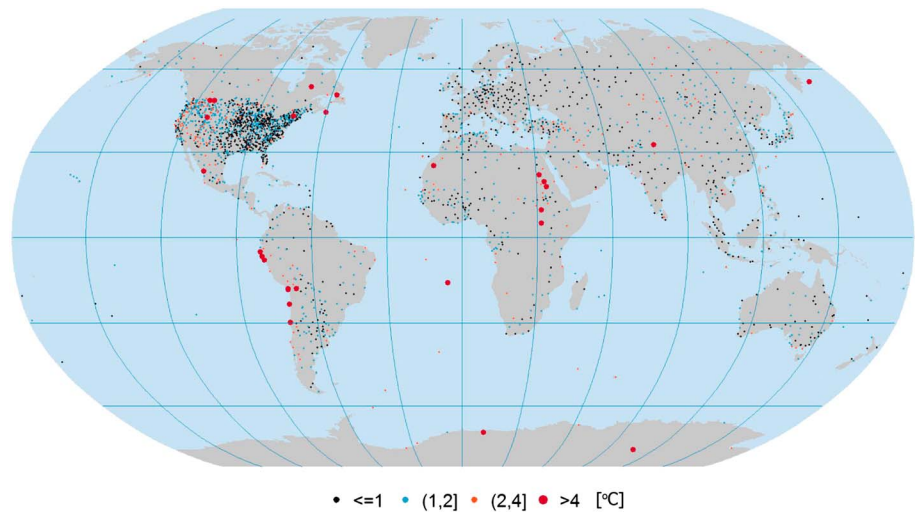


Figure 9. Map of the validation errors (RMSE) averaged per year for each station, calculated by using GHCN-M stations which were not used for model and prediction.

of regression modeling based on MODIS LST spline images explains 70% of the variability in minimum daily temperature values for the year 2011, with an average precision of $\pm 6.3^\circ\text{C}$, thus performing somewhat worse than for the mean temperature.

DEM and TWI layers also showed to be highly significant covariates for minimum daily temperature. The final linear model with four covariates explains 77% of the variation, with a standard error of $\pm 5.5^\circ\text{C}$. Figure 10 shows a plot of the modeled geometric trend for minimum daily temperature and MODIS LST spline values against observed temperature for the same stations.

3.2.2. Spatio-Temporal Variogram Model for Minimum Daily Temperature

The spatio-temporal variogram is modeled in the same way as was described for mean daily temperature. The variogram for minimum daily temperature has similar parameters as those obtained for the mean daily temperature (see Table 2 and Figure 11). Again, there is no purely temporal component. The

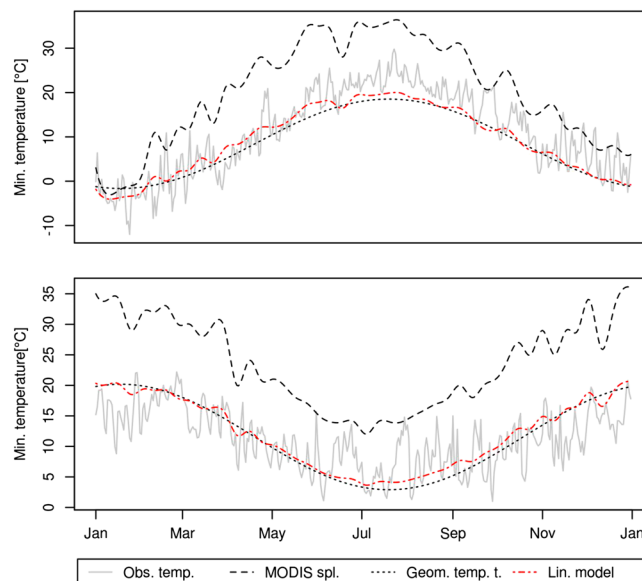


Figure 10. Minimum daily temperature observation in 2011 (gray solid line) and multivariate linear model of minimum daily temperature (red dashed line) on MODIS LST spline (black dashed line), geometric temperature trend (black dot line), elevation, and topographic wetness index. (top) Philadelphia, USA ($\lambda = -75, \phi = 39.993$). (bottom) Bunbury, Western Australia ($\lambda = 115.65, \phi = -33.35$).

Table 2. Parameters of the Fitted Sum-Metric Variogram Model for Minimum Daily Temperature Regression Residuals, Each Component (See Equation (9)) is Modeled Using a Spherical Variogram Function

	Nugget	Sill	Range Parameter	Anisotropy Ratio
Spatial	3.695	22.682	5725 km	
Temporal	0	0	0 days	
Space-time	1.67	9.457	1888 km	485 km/d

purely spatial component has a range of 5725 km at any time separation. Temporal variability of residuals is contained in the spatio-temporal interaction variogram component. The nugget parts of these components sum to around 3.5°C^2 , which is larger than in the mean temperature case and suggests that short-range variability in space and time of minimum temperature regression residuals is significantly larger than for the mean temperature. This suggests that extreme temperatures are being harder to predict.

3.2.3. Results of Accuracy Assessment for Minimum Daily Temperature

The results of cross validation for minimum temperature produced $\text{RMSE} = 2.7^{\circ}\text{C}$ for global land areas including Antarctica, with an R-square of 94.2%. Monthly RMSE obtained from the cross validation of monthly aggregated observations and cross-validation prediction is 2°C , yearly RMSE is 1.7°C . The spatial distribution of RMSE calculated per station (yearly average of squared daily cross-validation residuals, daily quality measure) for each station is shown in Figure 12, where the stations with $\text{RMSE} < 2^{\circ}\text{C}$ represent 40% of the total number of stations (black dots), $2^{\circ}\text{C} < \text{RMSE} < 3^{\circ}\text{C}$ occurs for 35% (blue dots), 23% have $3^{\circ}\text{C} < \text{RMSE} < 6^{\circ}\text{C}$, and 200 stations have an $\text{RMSE} > 6^{\circ}\text{C}$.

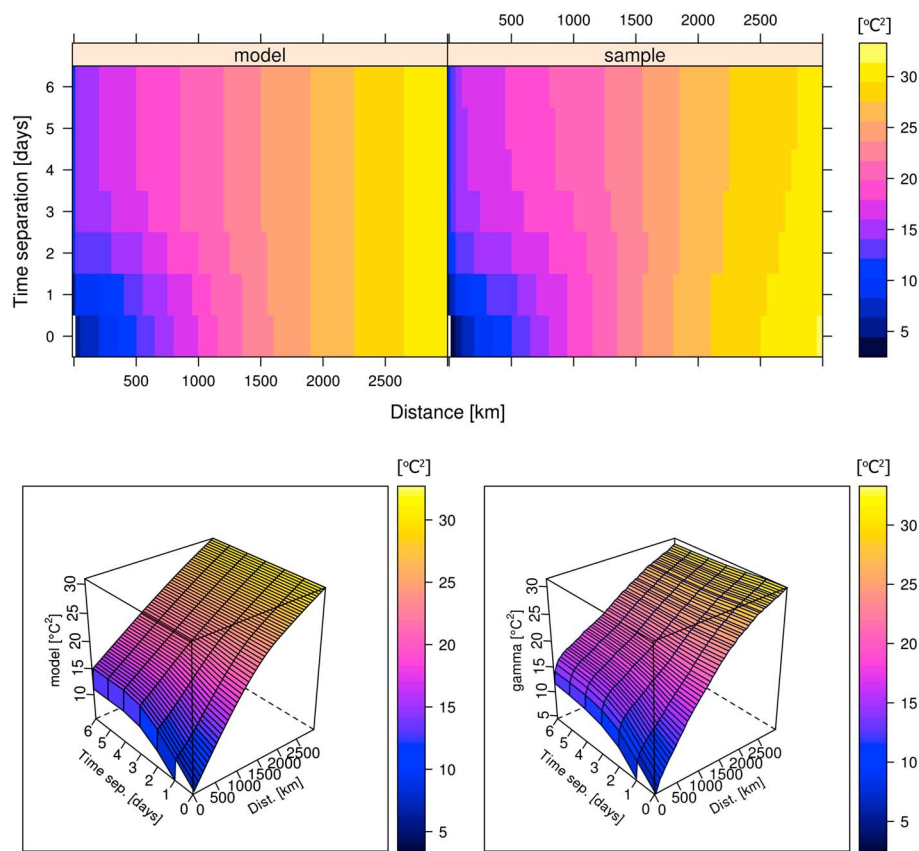


Figure 11. Fitted (left) sum-metric model and (right) sample variogram of residuals from multiple regression of minimum daily temperature observation on MODIS, geometric temperature trend, elevation, and topographic wetness index. The variogram surface is presented in (top) 2-D and (bottom) 3-D manner.

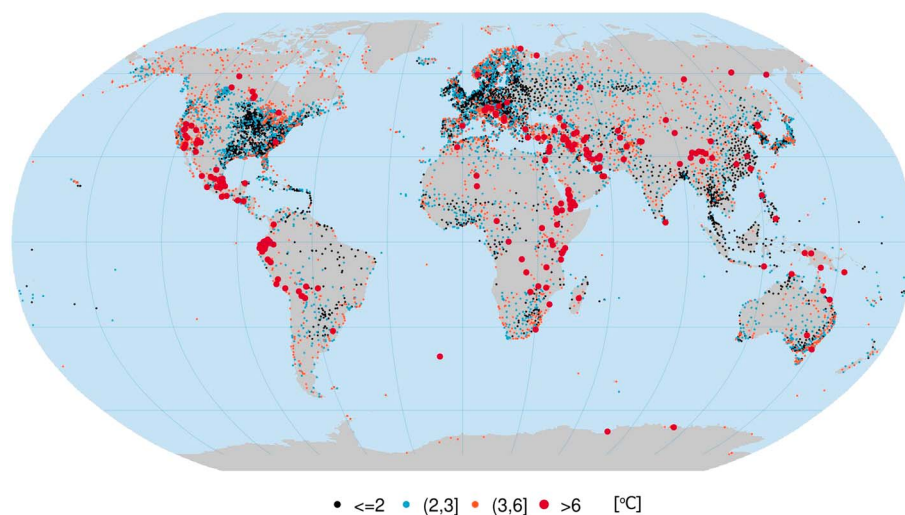


Figure 12. Map of minimum daily temperature cross-validation errors (RMSE) averaged per year for each station. Red circles indicate cross-validation outliers with RMSE > 6. Clusters of red circles indicate problematic areas, partly present of gross error in observation time series.

The spatial distribution of RMSE also shows lower accuracy than predictions of the mean temperature in general. The aggregated results show that the smallest RMSE = 1.4°C is achieved in the Netherlands, Europe without Russia (approximately 2.7°C) has RMSE of around 2.3°C, USA 2.3°C, South America 3.1°C, Antarctica has again the highest RMSE = 4.7°C.

3.3. Maximum Daily Temperature Interpolation

3.3.1. Linear Regression Model for Maximum Daily Temperature

The geometric temperature trend in the linear model explains 75% of maximum daily temperature variation for 2011 with a standard error of ±6.6°C. Figure 13 shows the geometric trend compared against observation. The geometric trend results are comparable to the results of modeling the mean temperature and are hence better than the results for modeling the minimum daily temperature case.

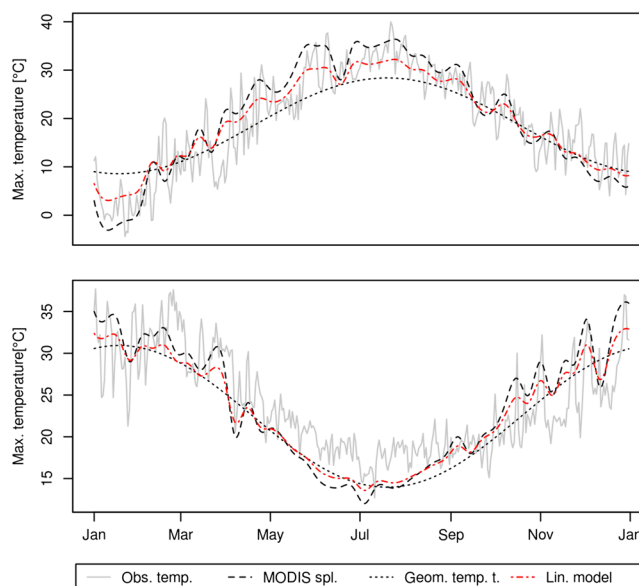


Figure 13. Maximum daily temperature observation in 2011 (gray solid line) and multivariate linear model of maximum daily temperature (red dashed line) on MODIS LST spline (black dashed line), geometric temperature trend (black dotted line), elevation, and topographic wetness index. (top) Philadelphia, USA ($\lambda = -75, \phi = 39.993$). (bottom) Bunbury, Western Australia ($\lambda = 115.65, \phi = -33.35$).

Table 3. Parameters of the Fitted Sum-Metric Variogram Model for Maximum Daily Temperature Regression Residuals, Each Component (See Equation (9)) is Modeled Using a Spherical Variogram Function

	Nugget	Sill	Range Parameter	Anisotropy Ratio
Spatial	2.8722	8.314	4930 km	
Temporal	0	0	0 days	
Space-time	1.750	11.175	2117 km	527 km/d

Regression modeling with only MODIS LST spline images already explains 84.5% of the variability in maximum daily temperature values for the year 2011 with an average precision of $\pm 5.2^\circ\text{C}$. MODIS LST 8 day images are the best predictor for the maximum daily temperature when compared to actual mean and minimum daily temperatures. DEM and TWI layers also were significant covariate layers for the maximum daily temperature. The final linear model with four covariates explains 86.7% of the variation, with a standard deviation of $\pm 4.8^\circ\text{C}$. Figure 13 shows the modeled linear regression line, the geometric trend for the maximum daily temperature, and the MODIS LST spline values against observed temperature for the same stations.

3.3.2. Spatio-Temporal Variogram Model for Maximum Daily Temperature

Table 3 (see also Figure 14) summarizes the parameters of the spatio-temporal variogram model. All components of the variogram are spherical functions. The nugget effect of the purely spatial component is relatively large as was observed before for the minimum temperature, showing that this model cannot achieve a better accuracy than the model for the mean daily temperature.

3.3.3. Results of Accuracy Assessment for Maximum Daily Temperature

Results of cross validation for maximum daily temperature on the complete data set gave an RMSE = 2.6°C for global land areas including Antarctica, with R-square 95.9%. Monthly RMSE obtained from cross

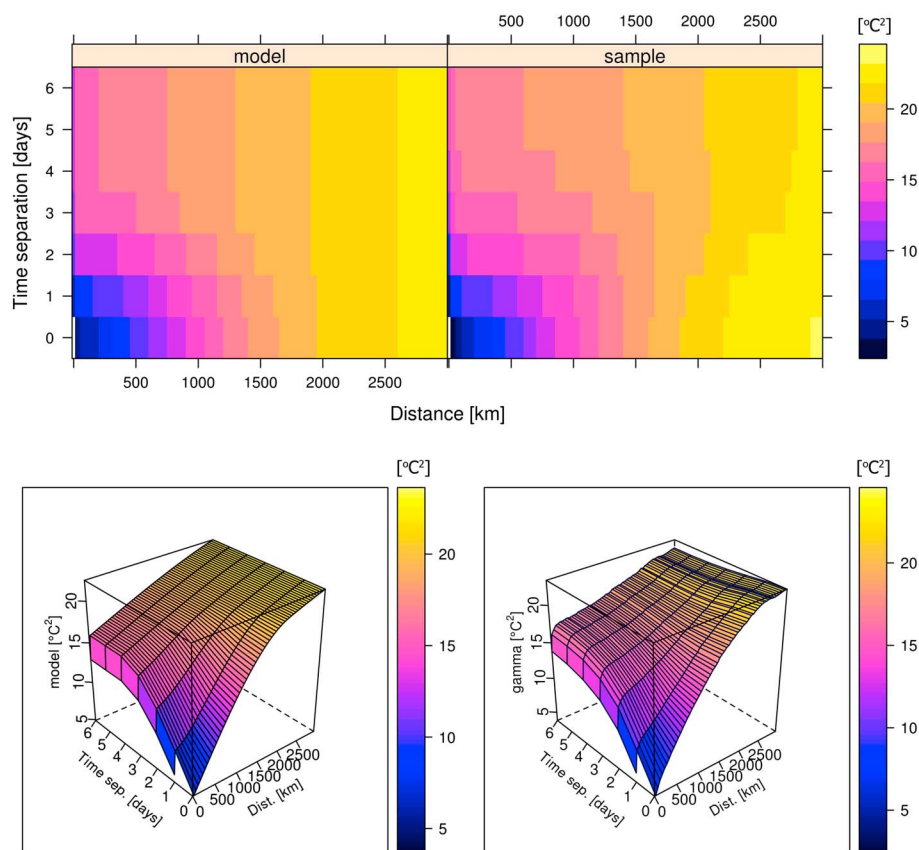


Figure 14. Fitted (left) sum-metric model and (right) sample variogram of residuals from multiple linear regression of maximum daily temperature on MODIS, geometric temperature trend, elevation, and topographic wetness index. The variogram surface is presented in (top) 2-D and (bottom) 3-D.

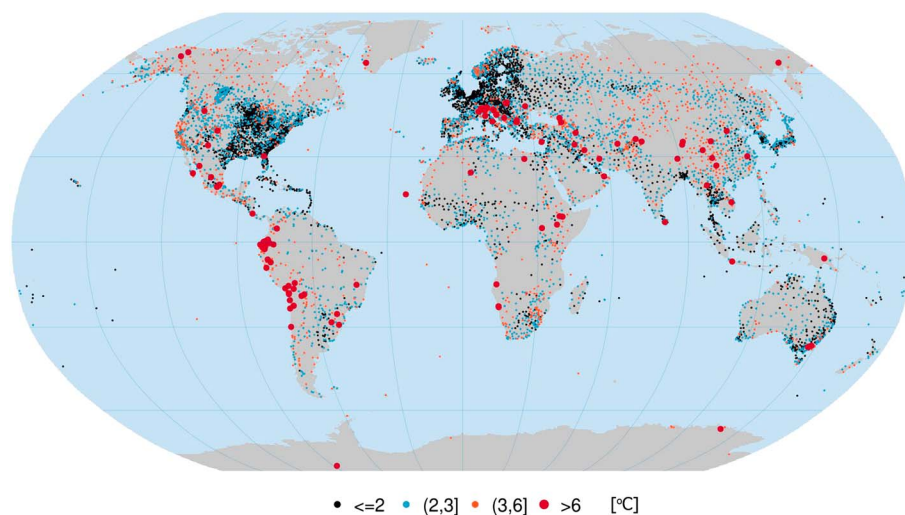


Figure 15. Map of maximum daily temperature cross-validation errors (RMSE) averaged per year for each station. Red circles indicate cross-validation outliers with RMSE > 6. Clusters of red circles indicate problematic areas, partly present of gross error in observation time series.

validation of monthly aggregation of observation and cross-validation prediction is 1.9°C, and yearly RMSE is 1.6°C. The spatial distribution of RMSE calculated per station (yearly average of squared daily cross-validation residuals, daily quality measure) for each station is shown in Figure 15, where the stations with RMSE<2°C are 41% of the total number of stations (black dots), 41% has 2°C< RMSE <3°C (blue dots), 16.6% have 3°C< RMSE <6°C, and 106 stations have RMSE >6°C.

The spatial distribution of RMSE also shows a lower accuracy than for the mean daily temperature. The aggregated results show that the best RMSE = 1.3°C is achieved in the Netherlands. Europe without Russia achieves around 2.1°C (Russia approximately 2.7°C), USA 2.1°C, South America 3.2°C, and Antarctica has the highest RMSE = 5°C.

3.4. Accuracy of the Prediction Models Without Satellite Data

Table 4 shows the results of the accuracy assessment with and without using MODIS LST 8 daily images. Sections 3.1.1, 3.2.1, and 3.3.1 show that MODIS LST is the most important covariate in the linear model, and without using any other covariates it explains the biggest part of the variation. Table 4 also shows that the influence of MODIS LST is significant in the linear model for all dependent variables, despite the fact that the accuracy of the regression-kriging for mean and minimum temperature is almost the same with and without using MODIS. Apparently, the spatio-temporal kriging of the regression residual can compensate for the less accurate linear model. The total accuracy of the spatio-temporal regression-kriging is significantly higher with MODIS for the maximum temperature. The important influence of MODIS in this case is also obvious from visual comparison of Figures 1, 10, and 13, where MODIS values are most strongly correlated with maximum daily temperature. The influence of topography is obvious in the temperature estimates at higher elevations.

Table 4. Accuracy Assessment Parameters for Multiple Linear Regression and Spatio-Temporal Regression-Kriging Model With MODIS LST Images Versus Without MODIS LST Images Used as Covariate

	Regression Part				Spatio-Temporal Regression-Kriging			
	R-Square (%)	Standard Error (°C)	R-Square (Without MODIS) (%)	RMSE (Without MODIS) (°C)	R-Square (%)	Standard Error (°C)	R-Square (Without MODIS) (%)	RMSE (Without MODIS) (°C)
TMEAN	84.2	4.6	75	6.2	96.6	2.39	96.3	2.47
TMIN	77.0	5.5	72.3	6.0	95.9	2.70	94.2	2.77
TMAX	86.7	4.8	75.2	6.3	96.4	2.60	87.5	4.20

4. Discussion and Conclusions

In this paper we have demonstrated how dense publicly available ground station meteorological data together with a time series of remote sensing images and covariates at 1 km resolution can be used to predict mean, minimum, and maximum daily temperature for the global landmass in space and time. The obtained global models for mean, minimum, and maximum temperature were further used to produce gridded images of daily temperatures at high spatial and temporal resolution. We achieved an average global prediction accuracy of about 2–3°C for daily temperature prediction when assessed using cross validation (which confirms the results of some local studies by *Hengl et al.* [2012], *Heuvelink et al.* [2012], and *Neteler* [2010]). This is promising as it indicates that highly accurate maps of daily temperatures can be produced at high spatial resolution using global spatio-temporal models. Figures 7, 12, and 15 also show that the outliers are distinctly grouped in areas that are poorly covered with meteorological stations and in mountainous regions, i.e., areas frequently covered with clouds or snow. This agrees with findings of *Neteler* [2010], who experienced similar difficulties in working with dynamic snow cover on mountain tops.

During the model fitting, we discovered that the GSOD point data sets still contain many artifacts and possible gross errors. We removed a small portion of obvious errors, but it is likely that there are still many gross errors in GSOD. It was beyond the scope of this study to identify and remove all errors. Station data filtering should probably be performed by the organizations that collected the data because they have expert knowledge on the measurements and stations. In that context, a methodology based on cross validation could be used as a tool to detect stations with potential errors. For example, stations with residuals beyond some threshold could be iteratively removed in a series of cross validations until no station with a residual larger than the threshold can be found. Furthermore, by overlaying the point data and WorldGrids.org covariates, we were able to detect stations with inaccurate spatial locations. This is especially important for stations in mountainous regions, which proved to be very important for model building as the error of predicting temperature increases with elevation.

When we tested the models using the GHCN-Daily data (quality controlled data set with 13,395 stations), we obtained similar results for maximum daily temperature with R-square of 95.2% and RMSE = 2.83°C (for details see interactive map at <http://dailymeteo.org>). For minimum daily temperature (13,380 stations), R-square was 94.2% and RMSE = 2.84°C. In this paper, the intention was to analyze three temperature parameters, namely, minimum, maximum, and mean daily temperature. That lead us to choose GSOD over the GHCN-D since GHCN-D still does not have the mean temperature included. But comparing the documentation and quality control procedures for GSOD and GHCN-D, we would probably recommend GHCN-D as the more suited data set for spatio-temporal analysis of the minimum and maximum daily temperature.

Our original idea was to overlay daily air temperature values (minimum, maximum, and mean) with daily LST images (MOD11A1 product) and build calibration models (spatio-temporal regression) for actual temperature above ground. However, the MOD11A1 product contains so many gaps (often >50% of the coverage contains missing pixels) due to cloud cover and atmospheric disturbances, so that we have finally decided to work only with MOD11A2 products. If a remote sensing image contains >50% of missing pixels, then this could lead to geostatistical models which are incomplete and unrepresentative. At this stage we also did not consider using the quality flag (i.e., number of days used to estimate the MOD11A2 LST value) in the model building, so this could be an area where predictions could be further improved. Maybe more elaborate procedure could also be developed to address the problem of missing pixels in the daily LST images due to clouds, but for the moment we do not see much potential in developing global spatio-temporal geostatistical models with so incomplete data.

It is worth noting that the presented global regression-kriging models can also be used to produce maps of associated uncertainty at high spatio-temporal resolution. Moreover, by using a global geostatistical model, unbiased prediction of daily air temperatures for any place on the global land mask can be obtained (at 1 km resolution) and for any day of the year for the period from the beginning of the MODIS mission until today. This is probably impossible using e.g., mechanical interpolators such as splines.

Our results also showed that the geometric temperature trend (equation (3)) is a crucial covariate. Alone, it can explain more than 70% of the temperature variation. This indicates that a similar model without remote sensing images can be made for the daily temperature interpolation for the period for which no MODIS images are available. In that context, the fitted spatio-temporal global models for mean, minimum, and

maximum daily temperature can also be used as a tool for disaggregation of MODIS 8 day images to daily images and for the calibration of land surface temperatures (conversion from land surface to air temperature).

Daily interpolation and spatio-temporal interpolation were performed for several days and produced comparable RMSE. The similarity was confirmed by *Spadavecchia and Williams* [2009], who also did not achieve a significant increase in interpolation performance for spatio-temporal interpolation of temperature compared with mechanical spatial interpolation methods such as splines. However, we believe here that geostatistical techniques are superior to mechanical techniques particularly because they allow for an objective quantification of uncertainty. As *Spadavecchia and Williams* [2009] suggest: “spatio-temporal techniques are useful, particularly in the provision of spatio-temporal uncertainty estimates.”

Another important advantage of the integrated spatio-temporal modeling is that a single regression model and a single semivariogram based on the complete space-time data set are valid for the whole year, while purely spatial models require 365 regression models and 365 semivariograms, using only data records available for a single day. This is much more labor intensive and makes quality checking difficult. Besides the elegance of the spatio-temporal model (estimation of just one model for the whole phenomenon), the spatio-temporal model is also advantageous because temporal trajectories of predictions and simulations look much more realistic (the models described in this paper can be obtained by installing the meteo package for R created and maintained by the authors of this article). If separate spatial interpolation is done for every day, time series at prediction locations may show abrupt jumps that have no physical basis.

The presented computational framework could now be used to produce a global archive of the mean, minimum, and maximum temperature images. The daily maps of temperature could serve as raster files in a similar fashion as climate layers from the WorldClim.org project [*Hijmans et al.*, 2005]. This would require a huge data storage and serving capacities, considering the amount of output pixels (10 years by 365 days by 7–10 meteo variables plus uncertainty maps). Furthermore, globally fitted models of daily temperatures can be used for regional or local studies, e.g., where a limited number of ground observation stations are available so that some reference global model of spatio-temporal variability is required.

Acknowledgments

This study is fully supported by the research projects TR 36035: *Spatial, ecological, energy, and social aspects of settlements' development and climate changes interrelationships*, funded by the Ministry of Education and Science of the Republic of Serbia, WorldGrids.org data portal has been developed as a part of the Global Soil Information Facilities project that provides tools for collating and serving global soil data developed jointly by ISRIC-World Soil Information and collaborators.

References

- Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, *111*, D05109, doi:10.1029/2005JD006290.
- Bader, M. Y., and J. J. Ruijten (2008), A topography-based model of forest cover at the alpine tree line in the tropical Andes, *J. Biogeogr.*, *35*(4), 711–723, doi:10.1111/j.1365-2699.2007.01818.x.
- Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, and M. Ziese (2012), A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, *Earth Syst. Sci. Data Discuss.*, *5*(2), 921–998, doi:10.5194/essdd-5-921-2012.
- Böhner, J., T. Blaschke, and L. Montanarella (Eds.) (2008), *SAGA — Seconds Out*, vol. 19, 113 pp., Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie, Hamb.
- Burrough, P. (Ed.) (1998), *Principles of Geographical Information Systems*, 2nd ed., 220 pp., Oxford Univ. Press, Oxford.
- Caesar, J., L. Alexander, and R. Vose (2006), Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set, *J. Geophys. Res.*, *111*, D05101, doi:10.1029/2005JD006280.
- Carrera-Hernández, J., and S. Gaskin (2007), Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico, *J. Hydrol.*, *336*(3), 231–249.
- Coll, C., Z. Wan, and J. M. Galve (2009), Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, *J. Geophys. Res.*, *114*, D20102, doi:10.1029/2009JD012038.
- Cressie, N., and C. K. Wikle (2011), *Statistics for Spatio-Temporal Data*, Wiley, N. J.
- Di Luzio, M., G. L. Johnson, C. Daly, J. K. Eischeid, and J. G. Arnold (2008), Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous United States, *J. Appl. Meteorol. Climatol.*, *47*(2), 475–497, doi:10.1175/2007JAMC1356.1.
- Gething, P., P. Atkinson, A. Noor, P. Gikandi, S. Hay, and M. Nixon (2007), A local space-time kriging approach applied to a national outpatient malaria data set, *Comput. Geosci.*, *33*(10), 1337–1350, doi:10.1016/j.cageo.2007.05.006.
- Gräler, B., L. Gerharz, and E. Pebesma (2011), Spatio-temporal analysis and interpolation of PM10 measurements in Europe, *ETC/ACM Technical Paper*, 10.
- Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *J. Geophys. Res.*, *113*, D20119, doi:10.1029/2008JD010201.
- Hengl, T., G. B. Heuvelink, M. Perčec Tadić, and E. J. Pebesma (2012), Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images, *Theor. Appl. Climatol.*, *107*, 265–277, doi:10.1007/s00704-011-0464-2.
- Heuvelink, G. B. M., and D. A. Griffith (2010), Space-time geostatistics for geography: A case study of radiation monitoring across parts of Germany, *Geog. Anal.*, *42*(2), 161–179, doi:10.1111/j.1538-4632.2010.00788.x.
- Heuvelink, G. B. M., D. A. Griffith, T. Hengl, and S. J. Melles (2012), *Sampling Design Optimization for Space-Time Kriging*, pp. 207–230, John Wiley, Oxford, doi:10.1002/9781118441862.ch9.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005), Very high resolution interpolated climate surfaces for global land areas, *Int. J. Circuit Theory Appl.*, *25*, 1965–1978, doi:10.1002/joc.1276.

- Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei (2008), Comparison of six methods for the interpolation of daily, European climate data, *J. Geophys. Res.*, *113*, D21110, doi:10.1029/2008JD010100.
- Jarvis, C. H., and N. Stuart (2001), A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method, *J. Appl. Meteorol.*, *40*(6), 1075–1084.
- Kiktev, D., D. M. Sexton, L. Alexander, and C. K. Folland (2003), Comparison of modeled and observed trends in indices of daily climate extremes, *J. Clim.*, *16*(22), 3560–3571.
- Kilibarda, M., and B. Bajat (2012), Plotgooglemaps: The R-based web-mapping tool for thematic spatial data, *Geomatica*, *66*(1), 37–49, doi:10.5623/cig2012-007.
- Klein Tank, A. M. G., et al. (2002), Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol.*, *22*(12), 1441–1453, doi:10.1002/joc.773.
- Lawrimore, J. H., M. J. Menne, B. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, and J. Rennie (2011), An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, *J. Geophys. Res.*, *116*, D19121, doi:10.1029/2011JD016187.
- Leemans, R., and W. P. Cramer (1991), The IIASA database for mean monthly values of temperature, precipitation, and cloudiness on a global terrestrial grid, *International Institute for Applied Systems Analysis*, Laxenburg, Austria.
- Legates, D., and C. Willmott (1990), Mean seasonal and spatial variability in global surface air temperature, *Theor. Appl. Climatol.*, *41*(1–2), 11–21, doi:10.1007/BF00866198.
- Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, *25*(6), 693–712, doi:10.1002/joc.1181.
- Neteler, M. (2010), Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST Data, *Remote Sens.*, *2*(1), 333–351, doi:10.3390/rs1020333.
- New, M., M. Hulme, and P. Jones (1999), Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, *J. Clim.*, *12*, 829–856, doi:10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2.
- New, M., M. Hulme, and P. Jones (2000), Representing twentieth-century space-time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate, *J. Clim.*, *13*, 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.
- Pebesma, E. (2012), Spacetime: Spatio-temporal data in R, *J. Stat. Software*, *51*, 1–30.
- Pebesma, E. J. (2004), Multivariable geostatistics in S: The gstat package, *Comput. Geosci.*, *30*(7), 683–691, doi:10.1016/j.cageo.2004.03.012.
- Peterson, T. C., and R. S. Vose (1997), An overview of the Global Historical Climatology Network temperature database, *Bull. Am. Meteorol. Soc.*, *78*(12), 2837–2849, doi:10.1175/1520-0477(1997)078<2837:AOTGH>2.0.CO;2.
- Pinheiro, J., and D. Bates (2009), *Mixed-Effects Models in S and S-PLUS*, Statistics and Computing, Springer, New York.
- Piper, S. C., and E. F. Stewart (1996), A gridded global data set of daily temperature and precipitation for terrestrial biospheric modeling, *Global Biogeochem. Cycles*, *10*(4), 757–782, doi:10.1029/96GB01894.
- Price, D. T., D. W. McKenney, I. A. Nalder, M. F. Hutchinson, and J. L. Kesteven (2000), A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data, *Agric. For. Meteorol.*, *101*(2), 81–94.
- R Development Core Team (2012), *R: A Language and Environment for Statistical Computing*, 1731 pp., R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Sorensen, R., U. Zinko, and J. Seibert (2006), On the calculation of the topographic wetness index: Evaluation of different methods based on field observations, *Hydrol. Earth Syst. Sci. Discuss.*, *10*, 101–112.
- Spadavecchia, L., and M. Williams (2009), Can spatio-temporal geostatistical methods improve high resolution regionalisation of meteorological variables? *Agric. For. Meteorol.*, *149*(6), 1105–1117.
- Stahl, K., R. Moore, J. Floyer, M. Asplin, and I. McKendry (2006), Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density, *Agric. For. Meteorol.*, *139*(3), 224–236, doi:10.1016/j.agrformet.2006.07.004.
- Tveito, O., M. Wegehenkel, F. van der Wel, and H. Dobesch (2006), Spatialisation of climatological and meteorological information with the support of GIS (Working Group 2), *The Use of Geographic Information Systems in Climatology and Meteorology, Final Report*, pp. 37–172.
- Van De Kerchove, R., S. Lhermitte, S. Veraverbeke, and R. Goossens (2013), Spatio-temporal variability in remotely sensed land surface temperature, and its relationship with physiographic variables in the Russian Altay Mountains, *Int. J. Appl. Earth Obs. Geoinf.*, *20*, 4–19.
- Van den Besselaar, E. J. M., M. R. Haylock, G. van der Schrier, and A. M. G. Klein Tank (2011), A European daily high-resolution observational gridded data set of sea level pressure, *J. Geophys. Res.*, *116*, D11110, doi:10.1029/2010JD015468.
- Vose, R. S., O. R. N. L. E. S. Division., G. C. R. P. (U.S.), U. S. D. of Energy. Office of Health, E. Research., C. D. I. A. C. (U.S.), and I. Martin Marietta Energy Systems (1992), *The Global Historical Climatology Network: Long-Term Monthly Temperature, Precipitation, Sea Level Pressure, and Station Pressure Data*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory; Available to the public from N.T.I.S., Oak Ridge, Tenn.; Springfield, VA.
- Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2004), Quality assessment and validation of the MODIS global land surface temperature, *Int. J. Remote Sens.*, *25*(1), 261–274, doi:10.1080/0143116031000116417.
- Yoo, J.-M., Y.-I. Won, Y.-J. Cho, M.-J. Jeong, D.-B. Shin, S.-J. Lee, Y.-R. Lee, S.-M. Oh, and S.-J. Ban (2011), Temperature trends in the skin/surface, mid-troposphere and low stratosphere near Korea from satellite and ground measurements, *Asia-Pac. J. Atmos. Sci.*, *47*, 439–455, doi:10.1007/s13143-011-0029-4.