

Article

Estimation of Climatologies of Average Monthly Air Temperature over Mongolia Using MODIS Land Surface Temperature (LST) Time Series and Machine Learning Techniques

Munkhdulam Otgonbayar ¹, Clement Atzberger ^{2,*}, Matteo Mattiuzzi ³
and Avirmed Erdenedalai ¹

¹ Institute of Geography and Geoecology, Mongolian Academy of Sciences, Ulaanbaatar 15170, Mongolia; munkhdulamo@mas.ac.mn (M.O.); abirmede@mas.ac.mn (A.E.)

² Institute of Surveying, Remote Sensing and Land Information (IVFL), University of Natural Resources and Life Sciences (BOKU), 1190 Vienna, Austria

³ European Environmental Agency, 1050 Copenhagen, Denmark

* Correspondence: clement.atzberger@boku.ac.at; Tel.: +43-1-47654-85700

Received: 2 September 2019; Accepted: 30 October 2019; Published: 4 November 2019

Abstract: The objective of this research was to develop a robust statistical model to estimate climatologies (2002–2017) of monthly average near-surface air temperature (T_a) over Mongolia using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature (LST) time series products and terrain parameters. Two regression models were analyzed in this study linking automatic weather station data (T_a) with Earth observation (EO) images: partial least squares (PLS) and random forest (RF). Both models were trained to predict T_a climatologies for each of the twelve months, using up to 17 variables as predictors. The models were applied to the entire land surface of Mongolia, the eighteenth largest but most sparsely populated country in the world. Twelve of the predictor variables were derived from the LST time series products of the Terra MODIS satellite. The LST MOD11A2 (collection 6) products provided thermal information at a spatial resolution of 1 km and with 8-day temporal resolution from 2002 to 2017. Three terrain variables, namely, elevation, slope, and aspect, were extracted using a Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM), and two variables describing the geographical location of weather stations were extracted from vector data. For training, a total of 8544 meteorological data points from 63 automatic weather stations were used covering the same period as MODIS LST products. The PLS regression resulted in a coefficient of determination (R^2) between 0.74 and 0.87 and a root-mean-square error (RMSE) from 1.20 °C to 2.19 °C between measured and estimated monthly T_a . The non-linear RF regression yielded even more accurate results with R^2 in the range from 0.82 to 0.95 and RMSE from 0.84 °C to 1.93 °C. Using RF, the two best modeled months were July and August and the two worst months were January and February. The four most predictive variables were day/nighttime LST, elevation, and latitude. Using the developed RF models, spatial maps of the monthly average T_a at a spatial resolution of 1 km were generated for Mongolia ($\sim 1566 \times 10^6$ km²). This spatial dataset might be useful for various environmental applications. The method is transparent and relatively easy to implement.

Keywords: near-surface air temperature; MODIS land surface temperature; terrain parameters; partial least square regression; random forest regression

1. Introduction

Near-surface air temperature (T_a) is a key descriptor of the climate [1]. T_a is a critical variable to the effective understanding of the many physical and biological processes between the atmosphere and land systems [2–4] because it regulates many land surface processes such as photosynthesis, respiration, and evaporation [5]. As air temperature influences nearly all biotic processes [6], climatologies of T_a also permit a good characterization of terrestrial environmental conditions [5,7]. As this variable can change quickly over space, cost-efficient mapping procedures are needed that can depict T_a using high spatial resolution.

Since the early 1980s, various interpolation methods have been used to estimate T_a given adequate sample points [8,9]. The literature shows that the most common interpolation techniques are global interpolators, thin plate smoothing splines and different forms of kriging [10,11], inverse distance weighting [12], and climatologically-aided interpolation [13]. In the comparative study of [14], most interpolation methods gave similar results. However, interpolation errors typically range between 1 and 3 °C [15,16] depending on the spatial and temporal resolution of recorded T_a data and the density of the station network [17].

Direct measurements of T_a at a height of 2 m above ground are only available from a limited number of meteorological stations. In many cases such as Mongolia, the spatial coverage of these measurements is inadequate; in addition, typical T_a time series come with many missing values [6]. On the contrary, satellite-derived land surface temperature (LST) data are continuous in both spatial-temporal coverages and are relatively inexpensive. However, the satellite does not directly measure T_a but only the LST.

Based on the physical linkage between LST and T_a , several authors have offered methods to estimate T_a using remote sensing satellite data [5,18,19]. During past decades, a large body of research has been collected regarding the retrieval of LST from satellite-based thermal infrared (TIR) data [20], in particular that related to the better understanding of emissivity and atmospheric effects [18,21,22]. As a result, LST can be retrieved nowadays relatively accurately from remotely sensed TIR data [23]. Several studies have demonstrated that T_a and LST data are highly correlated [24]. However, as expected, large differences have been noticed [25] which are related, for example, to physical properties and atmospheric conditions [26].

Three major approaches have been used to estimate T_a from LST data [27]:

- (1) energy-balance parameterization based on thermodynamic approaches [15,19,28],
- (2) contextual approaches based on temperature–vegetation index relations (TVX) [5,29,30], and
- (3) statistical approaches using various form of regression techniques [15,16,23,31–34].

Good exemplary studies estimating air temperatures with MODIS LST products using the aforementioned methods can be found for example in Bartkowiak et al. [35], Lu et al. [36], Zhou et al. [37], Janatian et al. [38], Ho et al. [39], Duan et al. [40], and Benali et al. [3].

Within the last two decades, statistical approaches, including simple and advanced regression, (e.g., linear and multiple regression, and machine learning techniques) have been developed to estimate T_a from Moderate Resolution Imaging Spectroradiometer (MODIS) LST products with varying levels of success. More recently, several studies have investigated more complex and advanced approaches to estimate T_a from MODIS LST products, such as geographically weighted regression (GWR) and climate space weighted regression (CSWR) [6], spatiotemporal regression-kriging (STRK) [31], stepwise [38,41], random forest (RF) [39,41–45], generalized boosted model (GBM) [45], cubist [41,45], support vector machine (SVM) [39], ordinary least squares (OLS) [39] and M5 model tree [46]. To ensure high modeling accuracy, several papers have highlighted the usefulness of multivariate and non-parametric algorithms such as RF and STRK. For instance, Kilibarda et al. [31] estimated mean, maximum, and minimum daily T_a with a spatial resolution of 1 km at a global scale using STRK with MODIS 8-day time-series LST products along with elevation, wetness index, and geographical location. The performance of STRK to predict T_a from MODIS LST products was compared with the performance of the linear regression model. The results indicated that the root-mean-square errors (RMSEs) for predicting mean, maximum, and minimum daily T_a

are ± 2 °C for areas with a high density of stations and from ± 2 °C to ± 4 °C for areas with a coarse station density. The lowest accuracy was 6 °C in Antarctica and at locations with high altitudes. Yoo et al. (2018) [43] estimated maximum and minimum daily T_a in two megacities using LST data from MODIS Terra/Aqua and seven auxiliary variables based on the RF machine learning method, resulting in an RMSE of 1.1 °C and 1.2 °C for maximum and minimum T_a , respectively, in Seoul, and an RMSE of 1.7 °C and 1.2 °C for maximum and minimum T_a , respectively, in Los Angeles. Several authors have concluded that machine learning techniques perform better than more conventional methods which provide multi-variables and nonlinear and nonparametric regression and classification [38,39,43,45,47,48]. Machine learning algorithms are particularly useful for cases where no deterministic model is available to solve the problem. Our research objective was to develop a robust empirical model to estimate climatologies of average monthly T_a across Mongolia at 1 km spatial resolution using time-series of MODIS Terra LST products, terrain parameters (elevation, slope, and aspect), and other ancillary information.

2. Study Area

The area studied in this work covers the entirety of Mongolia with a total area of approximately 1566×10^6 km². Mongolia is the eighteenth largest and most sparsely populated country in the world. Mongolia extends between the latitudes 41°35' N–52°09' N and the longitudes 87°44' E–119°56' E, with an average land surface elevation of 1580 m above sea level [49]. The elevation ranges between 524 m to 4320 m above sea level (Figure 1a). Its continentality increases from east to west. The climate conditions are extreme continental with semiarid and arid regions (Figure 1b). The “blue sky” country counts on average 260 sunny days per year and is characterized by a long cold winter and a short dry-hot summer with generally low precipitation [50]. The country’s average annual temperature varies between -8 °C and 6 °C [49] with strong temperature gradients. Annual total precipitation ranges between 50 mm in the desert steppe and desert regions to 500 mm in the high mountain regions. Generally, precipitation gradually increases from south to north.

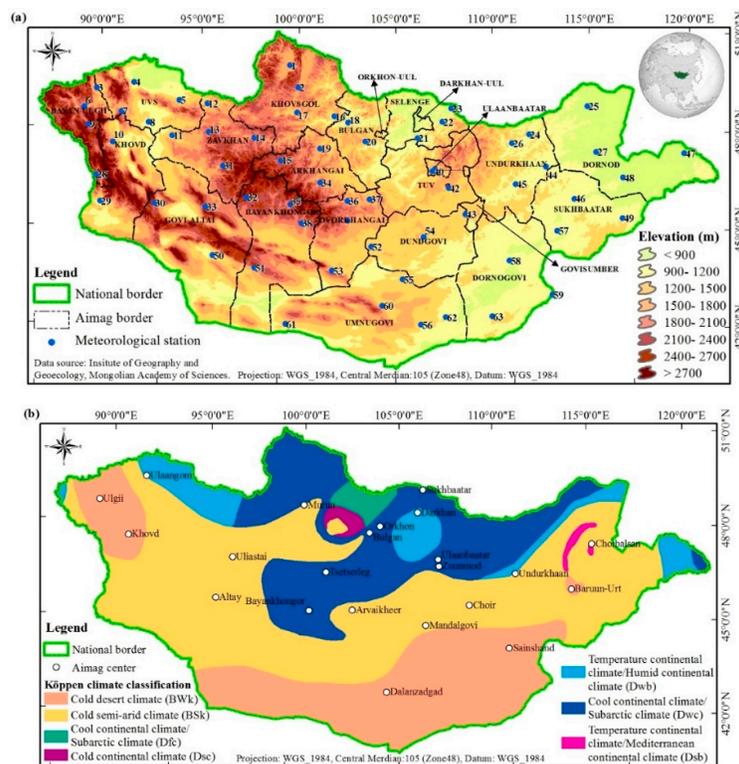


Figure 1. Mongolia: (a) Topography and location of meteorological stations ($n = 63$) for which reference air temperature information from automatic weather stations was available. The digital

elevation model (DEM) was derived from the Shuttle Radar Topography Mission (SRTM-DEM) with a resolution of 90 m. (b) Köppen climate classification of Mongolia based on long-term meteorological conditions (>30 years) [51].

According to the Köppen climate classification, Mongolia can be divided into seven climatic regions (Figure 1b). The most prevalent climates are BWk, BSk, and Dwc. By contrast, Dfc, Dsc, Dwb, and Dsb are less prevalent. Exemplary temperature and precipitation charts for the three most prevalent climates are shown in Figure 2, demonstrating a large variation in precipitation and temperature variation.

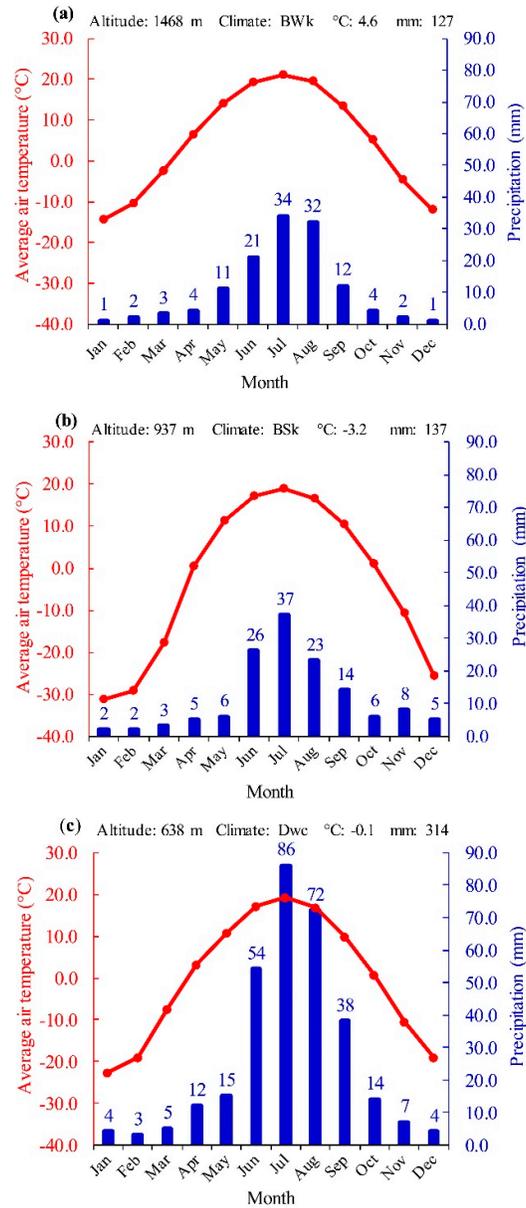


Figure 2. Climate graphs of the three most prevalent climatic zones in Mongolia [52]. (a) Dalanzadgad with cold desert climate (BWk), (b) Ulaangom with cold semi-arid climate (BSk), and (c) Sukhbaatar with cool continental climate/subarctic climate (Dwc). Information about altitude, average annual air temperature, and total precipitation is provided in each sub-plot. The data reflect long-term averages (>30 years).

3. Data and Methods

3.1. Remote Sensing Data

MODIS LST products are distributed by the Land Processes Distributed Active Archive Center (LP DAAC) in a hierarchical data format or HDF file. We used observations from MOD11 from the Terra satellite. MODIS generates two daily observations which involve one for daytime (LSTd) and one for nighttime (LSTn) at approximately 10:30 and 22:30 local time, respectively. The new collection 6 (c006) of MODIS LST products have been used to estimate Ta [53]. This dataset was made available in 2016. It covers the entire period (2002–2017) and data are of higher quality compared to the earlier collection(s), which had been used for previous studies such as [3,17,46,47,54,55]. The LST accuracy of the c006 products is reported as being approximately twice as good as collection 5 (c005) due to the incorporation of the emissivity adjustment model in the MODIS split-window algorithm [53]. For instance, the c006 LST product reduced the RMSE of bare soil sites of the c005 LST product by 1.24 °C during the day and 0.58 °C at night ([53], p. 88).

To cover the land surface of Mongolia, seven tiles of granules with horizontal (h) and vertical (v) title numbers h23v03, h23v04, h24v03, h24v04, h25v03, h25v04, and h26v04 had been used. The MODIS MOD11A2 c006 data were obtained through the online Data Pool at the National Aeronautics and Space Administration (NASA), the LP DAAC, and the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. The retrieved MODIS LST used the generalized split-window algorithm [56] to derive surface temperature from the recorded at-satellite radiances.

For this study, we used the MODIS Terra 8-day LST product (MOD11A2) at a spatial resolution of 1 km, gridded in the Sinusoidal projection intervals, and covering the period 2002–2017. The HDF file for the MOD11A2 product includes 12 different scientific data sets (SDSs), as shown in Table 1. A detailed description of SDSs is given in [57,58]. This MOD11A2 product includes daytime and nighttime LST data (LSTd and LSTn), quality information (QCd and QCn), observation information (DvA, NvA, DvT, and NvT), emissivity data (Em31 and Em32), and clear sky coverage (CsD and CsN). The HDF file for this product also contains associated quality science dataset layers which provide users with information regarding the usability and usefulness of the data products. The MODIS LST quality science dataset layers are binary encoded and bit packed. The quality assurance (QA) layer containing integer values had been converted to a bit binary value for interpretation [59,60] (p. 19). The quality controls (QC) are defined by bit flags such as mandatory quality assessment (QA) flags, data quality flags, emissivity quality flags, and cloud error flags.

To retrieve and pre-process the products, the MODIS R-package (MODIS acquisition and processing package v1.1.4) was used [61]. The package is run in the R software system and environment for statistical computing and graphics [62]. The MODIS R-package allows automatic downloading of data and processing such as changing file format, mosaicking, subsetting, and time-series filtering [63]. Using the package, digital numbers (DN) of MODIS Terra LST products were converted into LST (Table 1). Additionally, three terrain parameters (elevation, slope, and aspect) originating from SRTM-DEM [64] were retrieved. All raster data were re-projected to MODIS sinusoidal projection.

Table 1. Description of Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature (LST) products used in this study (source: Land Processes Distributed Active Archive Center (LP DAAC), 2019).

Variable Type	Acronym	Units	Data Type	Fill Value	Valid Range (VR)	Scale Factor (SF)	Additional Offset (AO)
Daytime LST	LSTd	Kelvin	16 bit	0	7500 to 65,535	0.02	N/A
Nighttime LST	LSTn	Kelvin	16 bit	0	7500 to 65,535	0.02	N/A
Day clear sky coverage	CsD	N/A	16 bit	0	1 to 65,535	0.0005	N/A
Night clear sky coverage	CsN	N/A	16 bit	0	1 to 65,535	0.0005	N/A
View zenith angle of daytime	DvA	Degree	8 bit	255	0 to 130	1.0	−65
View zenith angle of nighttime	NvA	Degree	8 bit	255	0 to 130	1.0	−65

Time of daytime (local solar)	DvT	Hours	8 bit	255	0 to 240	0.1	N/A
Time of nighttime (local solar)	NvT	Hours	8 bit	255	0 to 240	0.1	N/A
Emissivity band 31	Em31	None	8 bit	0	1 to 255	0.002	0.49
Emissivity band 32	Em32	None	8 bit	0	1 to 255	0.002	0.49
Quality controls of day LST	QCd	Bit	8 bit	N/A	0 to 255	N/A	N/A
Quality controls of night LST	QCn	Bit	8 bit	N/A	0 to 255	N/A	N/A

Note: Digital number (DN) = VR × SF – AO.

3.2. In Situ Meteorological Data

Sixty-three synoptic weather stations are present in Mongolia. Their geographical locations are indicated in Figure 1a. The weather stations provide Ta every 3 h, i.e., eight times a day. Air temperature data between 2002 and 2004 were obtained from the Mongolian Information Research Institute of Meteorology, Hydrology, and the Environment (IRIMHE). Data from 2004 to 2017 was downloaded from the “Reliable Prognosis (RP5)” website (<https://rp5.ru>). From the three-hourly meteorological data, the average air temperature was calculated for every 8 days of MODIS LST, taking into account the eight daily observations. This led to a total of 8544 meteorological data-points from 63 automatic weather stations covering the same period as the MODIS LST products, allowing for the development of prediction models between the remotely sensed data and Ta. The frequency distribution of the measured Ta reference data from 63 weather stations for the period 2002–2017 is shown in Figure 3 ($n = 8544$). The value of measured Ta ranged from -36.6 °C to 27.2 °C with a mean value of 0.7 °C and a standard deviation of 14.6 °C.

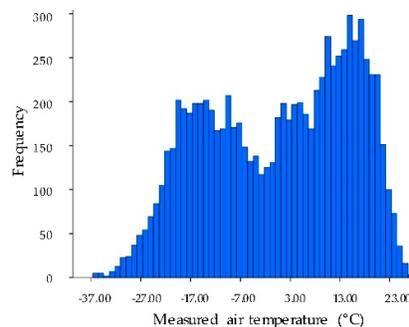


Figure 3. Frequency distribution of measured average 8-day air temperature reference data ($n = 8544$) from the 63 automatic weather stations for the years 2002–2017. Monthly statistics are depicted in Table 2.

3.3. Random Forest and Partial Least Square Regression

RF and PLS models were trained to predict Ta using up to 17 predictor variables. The use of two competing approaches permits the evaluation of the benefits of using non-linear machine learning approaches (e.g., RF) compared to classical linear regression models (e.g., PLS). Twelve of the seventeen variables were derived from LST time-series products of the Terra MODIS for the period 2002–2017 (Table 1). The five remaining variables were elevation, slope, and aspect (extracted from SRTM-DEM), and geographical location (latitude and longitude) of weather stations (extracted from vector data). Summary descriptive statistics of the response and the 17 predictor variables are reported in Table 2. Longitude was included as this indirectly depicts (for Mongolia) distance to sea [65].

Table 2. List of response/predictor variables and corresponding descriptive statistics (period 2002 to 2017). The list includes the measured air temperature (Ta) reference data at the weather station level ($n = 712$ for each of the twelve months) as well as the corresponding seventeen predictor variables extracted from satellite and other geo-data. For the acronyms of the variables, see Table 1.

Variable	No. of Samples (n)	Minimum	Maximum	Mean	Standard Deviation
Ta01	712	-36.60	-6.60	-20.83	5.36
Ta02	712	-35.10	-0.60	-16.60	5.38
Ta03	712	-20.50	4.80	-6.73	4.22
Ta04	712	-7.60	12.40	3.92	3.32
Ta05	712	2.80	19.30	10.75	3.05
Ta06	712	9.20	24.90	16.98	3.06
Ta07	712	11.60	27.20	19.49	3.21
Ta08	712	8.70	25.60	17.16	3.34
Ta09	712	2.60	19.70	10.48	3.07
Ta10	712	-8.20	11.70	1.49	3.08
Ta11	712	-22.70	0.30	-9.62	3.76
Ta12	712	-31.50	-6.10	-17.85	4.40
LSTd	8544	-36.90	48.60	13.43	20.08
LSTn	8544	-42.50	24.40	-5.75	14.47
CsD	8544	0.00	0.13	0.06	0.02
CsN	8544	0.00	0.13	0.07	0.02
DvA	8544	-55.00	62.00	5.04	-52.35
DvT	8544	10.40	12.10	11.82	0.71
Em31	8544	0.96	0.99	0.98	0.50
Em32	8544	0.97	0.99	0.98	0.49
NvA	8544	-65.00	56.00	-0.48	-56.47
NvT	8544	20.80	22.70	21.90	1.29
QCd	8544	2.00	133.00	61.98	16.58
QCn	8544	2.00	145.00	55.14	19.70
Elevation	63	667.00	2255.00	1369.10	411.70
Slope	63	0.08	19.60	N/A	N/A
Aspect	63	6.34	358.10	N/A	N/A
Latitude	63	42.97	51.11	N/A	N/A
Longitude	63	89.93	118.67	N/A	N/A

The correlation matrix (Figure 4) reveals a strong correlation between Ta and daytime/nighttime LST of MODIS, as well as several other correlations and redundancies. Based on these intercorrelations and taking into account that the number of variable sets should be relatively small, the predictor variables were grouped into seven different groupings (Table 3).

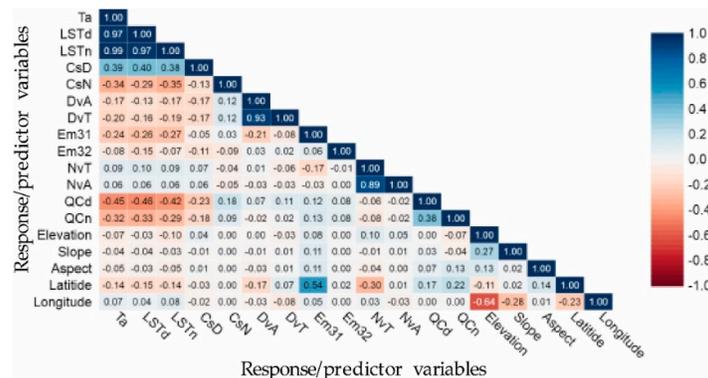


Figure 4. Correlation matrix between response and predictor variables ($n = 8544$). The saturation of the colors indicates the strength of the correlations. Positive correlations are shown in blue and negative correlations in red. In this graph, the air temperature data has been pooled across the twelve months. For the abbreviations, see Table 1.

Table 3. Seven model subsets studied. The seven groups were generated to study the relations between responses and up to 17 predictor variables. N_{var} indicates the number of predictor variables in each group.

	Acronym	Variables	N_{var}
Group 1	G1	LSTd and LSTn	2
Group 2	G2	LSTd, LSTn, and elevation	3
Group 3	G3	Elevation, slope, aspect, latitude, and longitude	5
Group 4	G4	Combined G1 and G3	7
Group 5	G5	CsD, CsN, DvA, DvT, Em31, Em32, NvA, NvT, QCd, and QCn	10
Group 6	G6	Combined G1 and G5	12
Group 7	G7	Combined G1, G3, and G5	17

3.3.1. RF Regression

The well-known random forest regression method [42,44] was chosen as the main approach to model the relation between our response variable (T_a) and the predictor variables listed in Tables 2 and 3 (LST MODIS products plus elevation, slope, aspect, latitude, and longitude). RF is a non-linear statistical ensemble method that leverages uncorrelated decision trees for regression. Developed by Breiman [66], it is capable of modeling discrete and/or continuous data sets [67,68]. RF predictions are obtained by aggregating a large number of individual regression decision trees where each decision tree is built from bootstrapped training samples (as in bagging) and variables are randomly selected at each decision node. The algorithm then randomly selects a subset of the predictors as candidates for splitting [66,69]. To obtain the final regression model, the results of all the individual trees are averaged. Good examples of the benefits and drawbacks of RF are given in [70–72].

The RF algorithm provides out-of-bag error (OOB) estimates and variable importance rankings [73,74], as not all observations are included in the respective bootstraps of the individual trees. In each tree at each split, the enhancement in the split-criterion importance measure is characterized by the splitting variable and aggregates individually all the trees in the forest for each variable [75]. Variable importance is measured by computing the increase in mean square error (MSE) when the OOB data for each variable are again computed but without the left-out variable [66,76]. The variable importance measures can assist in defining which variables are most important in the reduction of prediction error [71]. Two kinds of variable importance measures widely use the “randomForest” package in R [69,77]: (1) percent increase in the mean square error (%IncMSE) and (2) increase in node purity (IncNodePurity). From these, our analysis computed and analysed %IncMSE. However, we also checked the IncNodePurity indicator, but found similar results (not shown). In our research, the basic algorithm shown in Equation (1) was used to build the RF predictor for regression [75]:

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (1)$$

A new bootstrap sample for each decision tree T_b that includes $X = x_1, \dots, x_i$ with responses $Y = y_1, \dots, y_i$; bagging repeatedly (B times) selects a random sample from training data and each unpruned decision tree is increased in the sample. To increase each individual tree T_b , the following steps are repeated at each terminal node of the tree:

- Randomly select m variables from p variables
- Pick the variable that best splits and the corresponding split point
- Split the node into two nodes.

As mentioned above, to implement the RF regression model, two parameters must be set: the number of decision trees (n_{tree}) and the number of variables to select for the best split (m_{try}). For both hyperparameters, standard settings have been chosen. Each decision tree is independently increased to its maximum size, focusing on a new bootstrap sample from the training data (2/3 of samples). The remaining 1/3 of the samples, not used to fit the given decision tree, are referred to as the out-of-bag sample. The OOB sample is used to calculate the OOB error rate and variable importance. For quantifying the OOB error (prediction error) for each RF decision tree, we used Equation (2), i.e.,

$$Error_{OOB} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

where \hat{y}_i is the estimated output of OOB samples, y_i is the actual output, and n is the total number of OOB samples. RF regression is flexible and easy to use in comparison to other machine learning algorithms, even without hyper-parameter tuning.

3.3.2. PLS Regression

For comparison—and to assess the differences between linear and non-linear models—a prominent linear modeling technique was used: partial least square regression. PLS is widely used by the remote sensing community for vegetation analysis [78–80], soil related studies [81–83], and climate and ecological studies [84–87] amongst others.

PLSR is a multivariate linear regression method used to predict a set of dependent variables from a set of independent variables or predictors [88]. PLSR was originally developed for econometrics and chemometrics [89], where commonly a large number of strongly correlated predictor variables exist [90]. PLSR reduces the variables to a smaller set of uncorrelated components and performs least squares regression on these components instead of on the original data. Compared to other techniques, PLSR is more robust and less susceptible to data redundancy and over-fitting [91].

PLSR extracts a set of latent variables that explain the correlation between dependent and independent variables. The optimum number of latent variables for each generated model are implemented using the minimum value of residual mean squared error and the leave-one-out-cross-validation (LOO-CV) methods, e.g., jackknife and bootstrap [92]. To assess which variables are most contributing to the PLS model, we used the variable importance in projection (VIP) method [90], as seen in Equation (3), i.e.,

$$VIP_j = \sqrt{\frac{\sum_{f=1}^F w_{jf}^2 SSY_f J}{SSY_{total} F}} \quad (3)$$

where VIP_j is a measure of the contribution of the j variable in the PLSR model, w_{jf} is the weight value for the j variable and f latent variables (components), SSY_f is the sum of squares of explained variance for the f latent variable and J number of the predictor (independent) variables, SSY_{total} is the total sum of squares explained as the response (dependent) variables, and F is the total number of latent variables. The VIP values determine the contribution of the predictor variables to the PLSR latent variables. A VIP value greater than 0.80 ensures that only relevant variables are considered [90]. In [93] the VIP threshold of predictor variables that were identified as the most relevant variables ranged between 0.83 and 1.21. Predictor variables with ≤ 0.80 VIP values were classified as less important while variables with VIP values ≥ 1.20 were considered the most influential.

3.3.3. Model Evaluation and Statistics

Two widely used statistics were calculated to assess the accuracy of the models [94], including the R^2 and the RMSE. The R^2 describes the percentage of explained variance whereas the RMSE summarizes the deviations of predictions from the one-to-one line.

As both models provide quantitative information about the importance of different variables, we also report these findings. For the RF regression model, we assessed the importance of the individual predictors in Ta estimates focused on the %IncMSE [39]. For the PLS regression model, we used the VIP method [95,90,93].

4. Results

4.1. Comparison of RF and PLS Models: Variable Importance and Prediction Accuracy

The estimated importance of the 17 predictor variables in the RF regression model is shown in Figure 5 for each of the twelve months. Under the top three in each month, LSTn appeared 12 times, LSTd 11 times, elevation 7 times, latitude 5 times and aspect once. The remaining variables were never found under the top three in those rankings. Strong seasonality in the ranking can also be observed. For example, the warm season Ta (April to October) was strongly dependent on elevation. Conversely, the cold season Ta was more heavily affected by latitude. Using all 17 predictor variables led to R^2 values in the range from 0.83 (April) to 0.96 (August), while RMSE were between 0.91 °C (September) to 1.92 °C (February) (see Table 4, column G7).

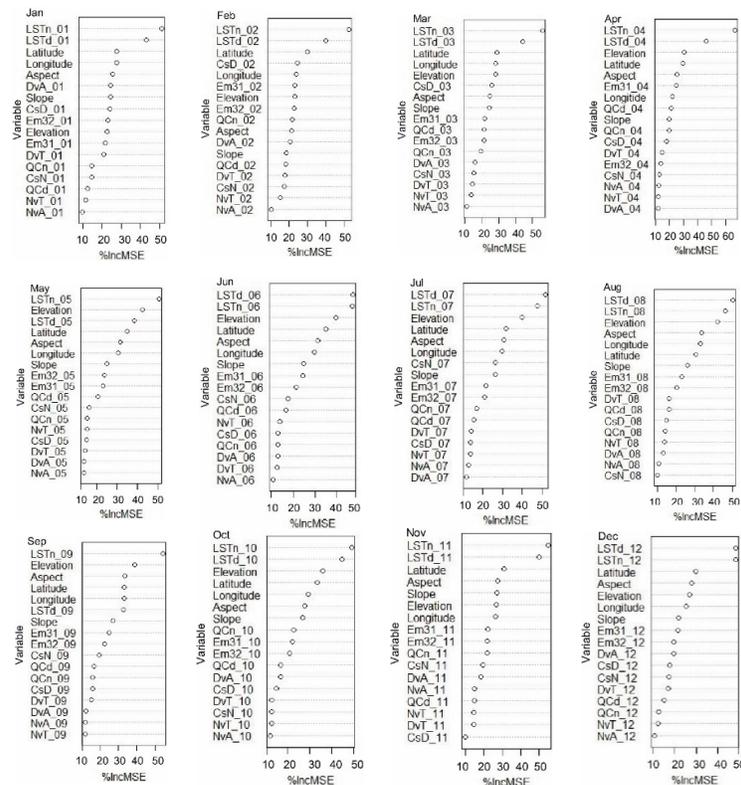


Figure 5. Random forest (RF) variable importance for each month. The importance is here given as the percentage increase in mean square error (%IncMSE).

Similar results were obtained for the PLS models (Appendix A Figure A1). Using PLS regression, the variables most often listed under the top three were LSTn (12 times), LSTd (12 times), elevation (4 times), latitude (4 times), and emissivity (4 times). Again, the ranking was season-dependent. The variables LSTn, LSTd, and latitude were the most important variables for estimating Ta in autumn and winter (September to February). The Ta for spring (March–May) was strongly dependent on LSTn, LSTd, and emissivity. For the summer months (June–August), LSTn, LSTd, and elevation were strongly influenced by the estimation of Ta for summer.

Using the entire set of 17 predictor variables for estimating the monthly average air temperatures, accuracies of PLS models were constantly lower compared to the RF models. The PLS models gave R^2 a measured and estimated monthly Ta between 0.74 and 0.86 and RMSE from 1.20 °C to 2.19 °C (Appendix A Table A2, column G7).

Concerning variable importance, the three variables LSTn, LSTd, and elevation were identical for PLS and RF regression models. This shows that LSTn, LSTd, and elevation play a key role in modeling Ta, with all other variables having a significantly smaller impact.

To further study the impact of the different predictor variables, the seven variable groupings highlighted in Table 3 were analyzed in more detail. Results for each month and the annual average air temperature are shown in Table 4 for the RF models. Compared to the full set of 17 variables

(column G7), the reduced set with only three predictor variables LSTn, LSTd, and elevation (G2) achieved comparable results, again highlighting and confirming the importance of these three predictor variables. None of the other five groupings (G1 and G3 to G5) were able to yield similar model performances. The same findings also hold for the PLS models (Appendix A Table A2 for details) but with constantly lower accuracies compared to the RF models. PLS models and groupings G1 to G7 were, therefore, skipped for the remainder of the study.

Table 4. Modeling results obtained using the RF regression. Reported are the monthly summary statistics (coefficient of determination (R^2) and root-mean-square error (RMSE)) for Ta prediction models for each of the seven groups of variables. For details of groupings G1 to G7, see Table 3.

	n	G1		G2		G3		G4		G5		G6		G7	
		R ²	RMSE												
January	712	0.85	2.07	0.91	1.59	0.52	3.71	0.89	1.77	0.40	4.14	0.88	1.86	0.90	1.71
February	712	0.83	2.24	0.87	1.93	0.49	3.83	0.86	2.05	0.46	3.95	0.85	2.06	0.88	1.92
March	712	0.76	2.09	0.85	1.64	0.55	2.84	0.83	1.77	0.46	3.12	0.81	1.83	0.85	1.67
April	712	0.77	1.59	0.82	1.41	0.47	2.42	0.80	1.47	0.38	2.62	0.79	1.51	0.83	1.40
May	712	0.76	1.48	0.88	1.05	0.75	1.52	0.84	1.07	0.39	2.37	0.81	1.34	0.88	1.08
June	712	0.77	1.45	0.88	1.04	0.76	1.50	0.88	1.07	0.31	2.53	0.81	1.33	0.88	1.07
July	712	0.84	1.29	0.93	0.84	0.81	1.41	0.91	0.97	0.32	2.65	0.87	1.17	0.87	0.93
August	712	0.83	1.38	0.95	0.92	0.83	1.37	0.90	0.97	0.35	2.68	0.86	1.27	0.96	0.92
September	712	0.81	1.32	0.91	0.91	0.84	1.22	0.88	1.05	0.40	2.37	0.85	1.19	0.91	0.91
October	712	0.83	1.28	0.90	0.99	0.70	1.68	0.88	1.08	0.47	2.23	0.85	1.22	0.89	1.03
November	712	0.81	1.65	0.87	1.34	0.54	2.55	0.85	1.47	0.42	2.86	0.84	1.52	0.86	1.39
December	712	0.84	1.76	0.89	1.44	0.60	2.79	0.88	1.55	0.40	3.43	0.86	1.64	0.89	1.49

4.2. Maps of Predicted Air Temperatures Using RF Models with the Reduced Feature Set

Both the results of the variable importance rankings (Figure 5) and the grouping of variables (Table 4) indicate that relatively simple RF prediction models can be built to estimate Ta using only daytime/nighttime LST and elevation information. Scatterplots between measured and estimated monthly average air temperatures using only these three predictor variables are shown in Figure 6. Corresponding maps of modeled air temperatures at 1 km spatial resolution and covering the entire land-mass of Mongolia are shown in Figure 7a (see Figure A2 and Table A3 in the Appendix A for corresponding scatterplots and maps generated using PLS models).

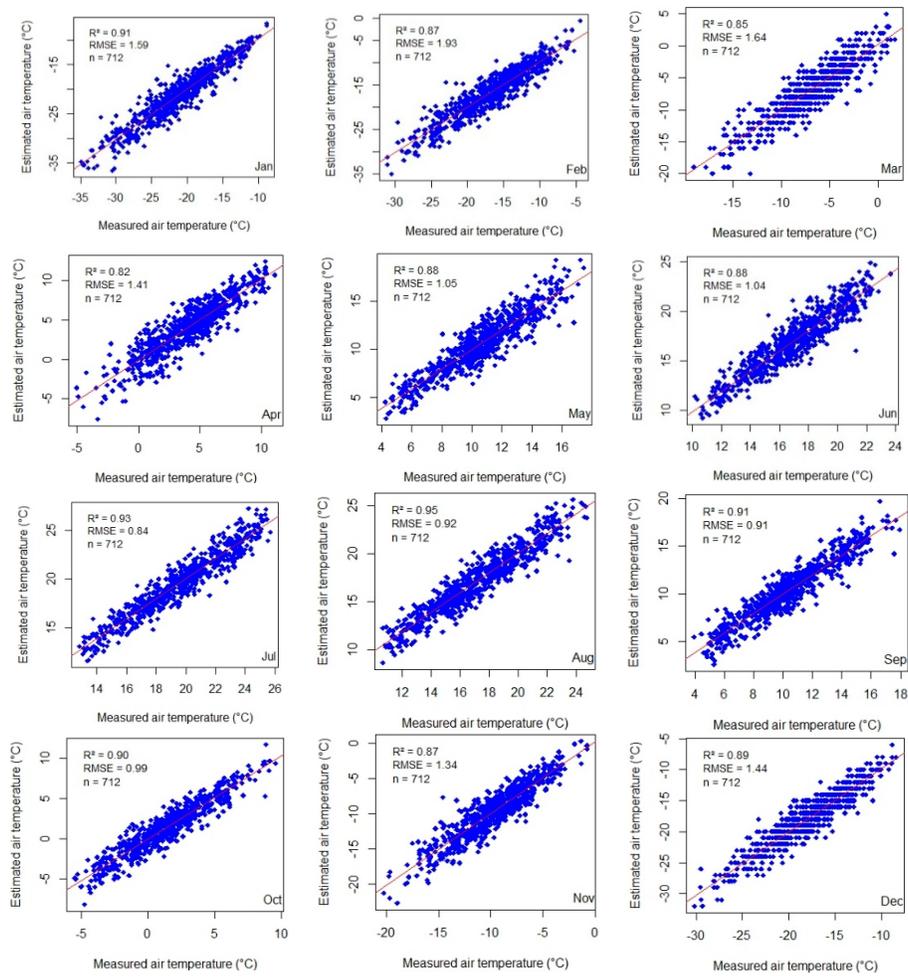


Figure 6. Comparison between measured and estimated monthly average T_a using LSTd, LSTn, and elevation for the RF regression model.

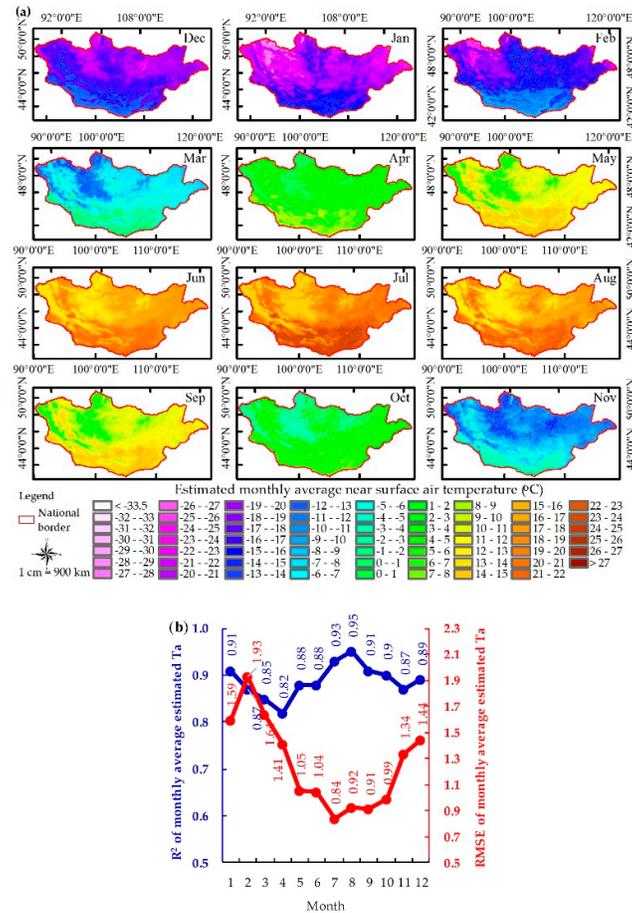


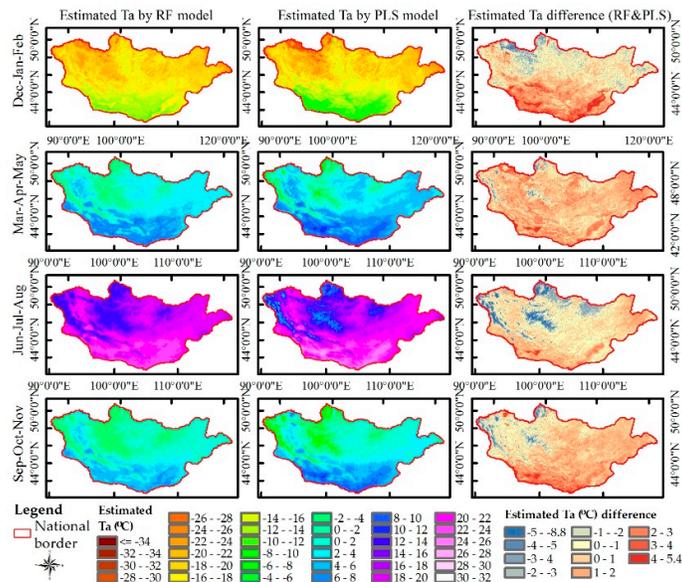
Figure 7. Estimated monthly average Ta based on RF regression model using LSTd, LSTn, and elevation as predictor variables. (a) Spatial maps of estimated monthly average Ta over Mongolia at 1 km spatial resolution. (b) Monthly statistics of R² (blue) and RMSE (red) between observed and predicted air temperature.

The scatterplots in Figure 6 reveal that the RF-predicted Ta is well distributed around the 1-to-1 line, with no apparent systematic deviations. In particular, we do not see any autocorrelation in the errors, nor saturation effects. The errors are generally low and the explained variance (R²) mostly above 0.85. Generally, however, the RMSE increases slightly during the colder months (Figure 7b).

The maps in Figure 7a depict in high spatial detail the model predictions. As expected, the predicted air temperatures decrease with elevation (Figure 1) but reveal additional detail and information. Monthly analyses of the coefficient of determination (R²; in blue) and root mean square error (RMSE; in red) for the period 2002–2017 are shown in Figure 7b. Overall, a good agreement between observed and estimated Ta values was found but reflected again the afore-mentioned seasonal pattern. Large discrepancies were found to occur in transition months, such as the start or end of seasons.

The PLS-generated maps of monthly average air temperatures are generally similar to the maps derived from RF models (see Appendix A Figure A2b). However, a more detailed analysis reveals sometimes larger differences, even if modeled air temperatures are averaged by season. For example, Figure 8 clearly shows that large method-specific differences occur (maps in the third column). The differences show large seasonal fluctuations. The deviations moreover show a clear north–south gradient with generally lower Ta estimated using PLS compared to RF (reddish colors). The deviations are usually strongest during the warmer months. As the RF model outperformed the PLS model when evaluated against the observed air temperatures (Table 4 for RF and Appendix A Table

A2 for PLS), we interpret these findings as mainly being the result of a systematic underestimation of T_a by the PLS model.



to depend on the employed split-window algorithm, cloud cover, and terrain parameters [32]. Nonetheless, estimation of T_a derived from MODIS LST studies using multivariate and non-parametric algorithms is suitable for generating results at high accuracy. For instance, [42] estimated monthly average T_a for the territory of China at a spatial resolution of 1 km using RF regression with MODIS LST, normalized difference vegetation index (NDVI), nighttime light and elevation. Using this dataset, the RMSE of the monthly average T_a ranged between 1.57 °C to 1.99 °C. Our study has showed that monthly average T_a can be accurately estimated using LST_n, LST_d, and elevation with similar RMSE ranging from 0.91 °C to 1.93 °C. The method is relatively easy to implement provided that there is a sufficient amount of training data with corresponding EO time-series observations [96].

6. Conclusions

In this study, PLS and RF regression models were applied to estimate monthly average T_a in Mongolia for the period 2002–2017 using MODIS LST time-series products and terrain parameters. Meteorological data from 63 automatic weather stations were used to calibrate and validate the PLSR and RF models. Both models were trained to predict T_a using up to 17 variables as predictor variables. Twelve variables were derived from LST time-series products of Terra MODIS and three variables were extracted from an SRTM DEM (elevation, slope, and aspect). The geographical location (longitude and latitude) was used as an additional variable. For training, a total of 8544 meteorological data points from 63 automatic weather stations and corresponding MODIS LST were used. Both datasets covered the period 2002–2017. Using only day/nighttime LST and elevation as predictor variables, the correlation between measured and estimated monthly average T_a RMSE ranged from 1.20 °C to 2.19 °C for the PLSR and 0.84 °C to 1.93 °C for the RF. The significantly lower errors of the RF models confirm the benefits of this machine learning approach compared to traditional (linear) modeling techniques (e.g., PLSR). We therefore recommend the use of RF models for similar studies.

Concerning the MODIS land surface temperature data, we found that this information contributed significantly to the modeling of air temperature. For example, it was not possible to obtain similarly low errors in the modeled air temperature using only terrain parameters as predictors. It is recommended that day- and nighttime LST be used simultaneously as both variables scored high in the feature importance metric.

Both machine learning models (RF and PLSR) represented well seasonal and spatial variations in T_a when time-series of LST were included as predictor variables. Using the models, maps of the monthly average T_a of Mongolia were developed at a spatial resolution of 1 km which were representative for the period 2002–2017. Although errors in the predicted T_a were generally low, the residual errors showed a significant seasonality; the warmer months were generally better modeled compared to the extremely cold winter months. Probably, the increased errors during the winter months reflect a lower accuracy in the input (LST) data. Further research is warranted to better understand the seasonality of the model quality.

Despite these trends, we firmly believe that this spatial dataset may be useful for various environmental applications; for instance, it may be useful for better assessing bioclimatic variations within the huge land-mass of Mongolia. The developed methodology is relatively easy and transparent and can be applied in different geographic regions, provided that enough weather stations are available to permit a model calibration. The spatial resolution of the final map product mainly depends on the ground sampling distance of the employed satellite sensors. As sensor technology advances at a rapid pace, the current 1 km spatial resolution can be further improved in the near future.

Author Contributions: The first author M.O. analyzed the data and performed the experiments; she also runs the statistical analysis in R software and drafted the first version of the manuscript. The second author C.A. contributed to data analysis and the revision on an earlier and on the final version of the manuscript. The third author M.M. contributed MODIS data analysis preprocessing in R software. All authors together developed and discussed the manuscript and wrote the paper.

Funding: This article was funded by the Eurasia-Pacific Uninet, Ernst Mach Grant (reference no. ICM-2018-10238).

Acknowledgments: We are grateful to Anja Klisch and Valentin Pesendorfer, who are researchers at the Natural Resources and Life Sciences University (BOKU), for their assistance in writing R code. Munkhduulam Otgonbayar received funding from Eurasia-Pacific Uninet on behalf of the Austrian Federal Ministry of Science, Research and Economy (BMFWF) during the preparation of the manuscript. We thank all colleagues at the Institute of Geography and Geoecology, Mongolian Academy of Sciences who helped in this study. We would like to thank our reviewers for providing valuable feedback and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

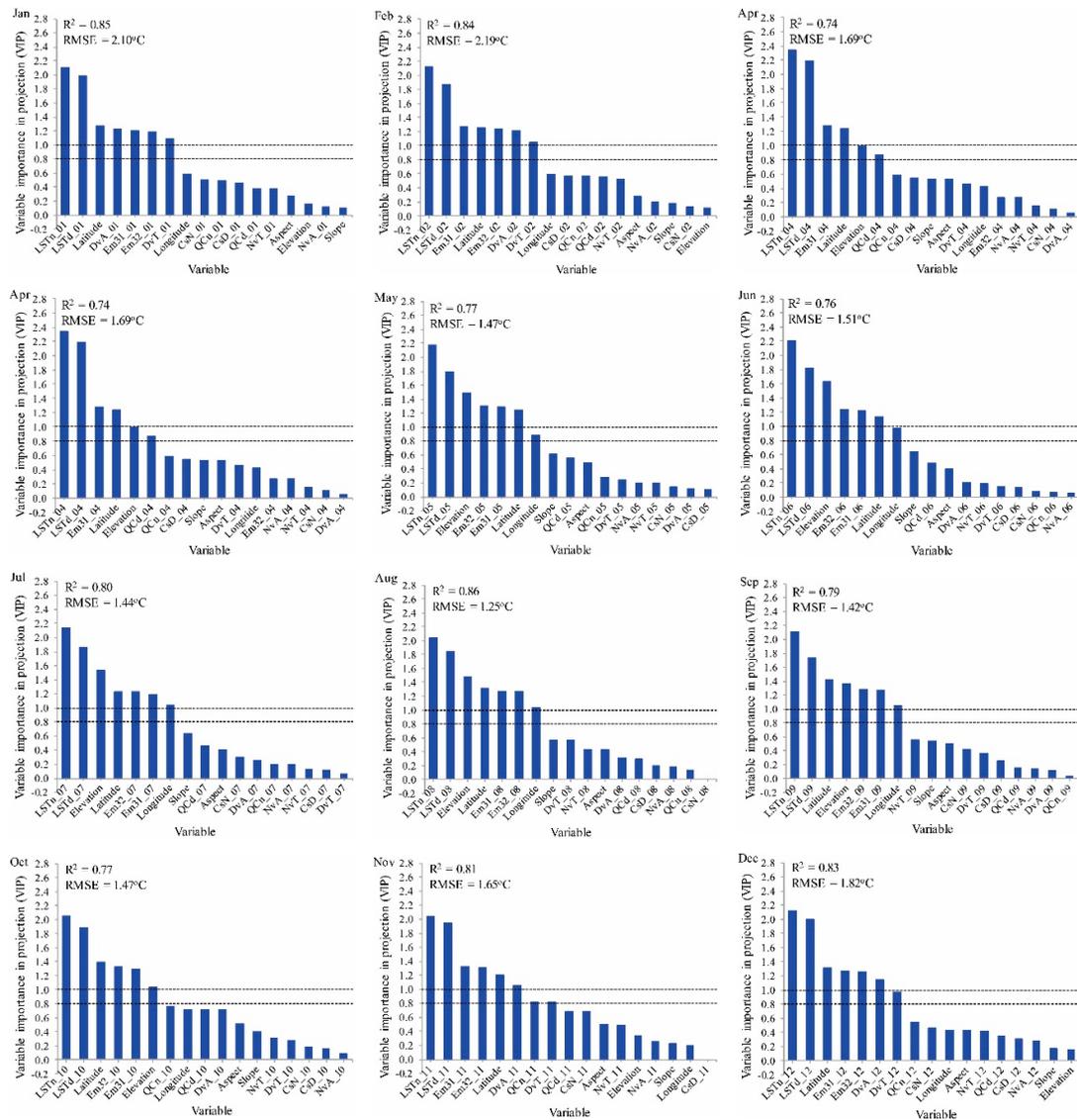


Figure A1. Variable importance in projections (VIPs), R^2 , and RMSE for the 17 predictor variables in the twelve-monthly PLSR models.

Table A1. Model coefficients for the twelve-monthly PLS regression models including all 17 predictor variables. In the last column, the model of the annual average Ta is also included.

Variable	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Intercept	-52.83	-93.62	26.17	55.52	93.68	89.37	97.27	45.11	68.76	62.92	53.54	22.17	82.62
LSTd	0.255	0.219	0.208	0.185	0.159	0.146	0.137	0.178	0.150	0.128	0.197	0.234	0.319
LSTn	0.420	0.457	0.425	0.265	0.223	0.232	0.227	0.284	0.210	0.196	0.344	0.407	0.464
CsD	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
CsN	-0.014	-0.003	0.007	0.000*	0.001	0.001	0.002	0.000*	0.004	0.000*	0.000*	-0.011	0.000*
DvA	0.015	0.015	0.007	-0.002	0.005	0.008	0.011	0.021	0.006	0.022	0.003	0.004	-0.007
DvT	0.094	0.109	-0.015	-0.130	-0.072	-0.041	-0.019	0.003	-0.105	0.061	-0.044	-0.021	0.095
Em31	0.069	0.104	0.021	-0.152	-0.133	-0.126	-0.131	-0.050	-0.127	-0.125	-0.047	0.006	-0.079
Em32	0.101	0.158	0.003	-0.012	-0.208	-0.198	-0.209	-0.096	-0.199	-0.195	-0.087	-0.009	0.007
NvA	-0.019	-0.034	-0.034	-0.012	-0.009	-0.002	-0.007	-0.017	0.007	-0.004	-0.017	-0.028	-0.038
NvT	0.031	0.100	-0.114	0.045	0.053	0.050	0.034	0.029	0.169	0.083	-0.030	-0.038	-0.413
QCd	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
QCn	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
Elevation	0.000*	0.000*	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001	0.000*	-0.001
Slope	-0.021	0.028	-0.009	-0.061	-0.059	-0.062	-0.065	-0.047	-0.051	-0.037	-0.025	-0.031	0.026
Aspect	-0.002	-0.001	-0.001	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002	-0.002	-0.003	-0.003	0.003
Latitude	-0.411	-0.401	-0.248	-0.268	-0.236	-0.214	-0.246	-0.244	-0.260	-0.248	-0.241	-0.393	0.336
Longitude	0.099	0.106	0.039	0.023	0.042	0.046	0.052	0.065	0.048	0.032	0.016	0.062	0.012

* <0.0001.

Table A2. PLS regression results. Summary statistics (R² and RMSE) for the monthly Ta prediction models, including six groups of variables.

	n	G1		G2		G3		G4		G5		G6		G7	
		R ²	RMSE												
January	712	0.87	1.96	0.87	1.95	0.34	4.34	0.82	2.26	0.39	4.17	0.86	2.01	0.85	2.10
February	712	0.83	2.19	0.83	2.19	0.32	4.40	0.79	2.45	0.37	4.26	0.85	2.05	0.84	2.19
March	712	0.80	1.89	0.74	1.94	0.36	3.40	0.77	2.04	0.39	3.32	0.81	1.84	0.81	1.83
April	712	0.79	1.51	0.79	1.53	0.48	2.39	0.77	1.58	0.32	2.73	0.75	1.67	0.74	1.69
May	712	0.76	1.48	0.79	1.41	0.74	1.54	0.80	1.37	0.31	2.52	0.76	1.50	0.77	1.47
June	712	0.78	1.44	0.80	1.38	0.75	1.52	0.79	1.41	0.26	2.63	0.78	1.45	0.76	1.51
July	712	0.83	1.33	0.86	1.20	0.79	1.48	0.84	1.28	0.28	2.71	0.82	1.34	0.80	1.44
August	712	0.84	1.36	0.87	1.23	0.81	1.44	0.85	1.28	0.30	2.80	0.83	1.37	0.86	1.25
September	712	0.81	1.35	0.84	1.24	0.82	1.30	0.83	1.26	0.32	2.53	0.80	1.37	0.79	1.42
October	712	0.83	1.27	0.83	1.26	0.68	1.73	0.82	1.31	0.41	2.36	0.81	1.32	0.77	1.47
November	712	0.83	1.54	0.83	1.57	0.37	2.97	0.79	1.70	0.40	2.92	0.82	1.59	0.81	1.65
December	712	0.86	1.68	0.86	1.67	0.36	3.53	0.82	1.89	0.34	3.58	0.85	1.74	0.83	1.82

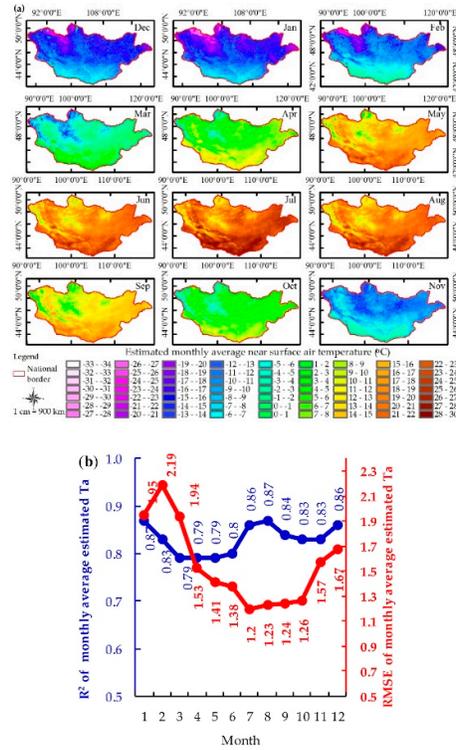


Figure A2. Estimated monthly average Ta based on the PLS regression model and using LSTd, LSTn, and elevation as predictor variables. (a) Spatial maps of estimated monthly average Ta over Mongolia at 1 km spatial resolution. (b) Monthly statistics of R² (blue) and RMSE (red) between observed and predicted air temperature.

Table A3. Model equations obtained from the PLS regression models using three variables: LSTd, LSTn, and elevation. The months are numbered from 01 to 12.

Regression Models	R ²	RMSE
Ta ₀₁ = -2.137 + 0.347 × LSTd + 0.497 × LSTn + 0.00033 × elevation	0.87	1.95
Ta ₀₂ = -3.037 + 0.297 × LSTd + 0.493 × LSTn + 0.00019 × elevation	0.83	2.19
Ta ₀₃ = -1.986 + 0.252 × LSTd + 0.477 × LSTn - 0.001 × elevation	0.74	1.94
Ta ₀₄ = 0.516 + 0.296 × LSTd + 0.424 × LSTn - 0.002 × elevation	0.79	1.53
Ta ₀₅ = 3.863 + 0.272 × LSTd + 0.383 × LSTn - 0.002 × elevation	0.79	1.41
Ta ₀₆ = 7.060 + 0.242 × LSTd + 0.384 × LSTn - 0.002 × elevation	0.80	1.38
Ta ₀₇ = 8.440 + 0.241 × LSTd + 0.398 × LSTn - 0.002 × elevation	0.86	1.20
Ta ₀₈ = 7.644 + 0.253 × LSTd + 0.419 × LSTn - 0.002 × elevation	0.87	1.23
Ta ₀₉ = 5.294 + 0.291 × LSTd + 0.407 × LSTn - 0.002 × elevation	0.84	1.24
Ta ₁₀ = 3.418 + 0.266 × LSTd + 0.406 × LSTn - 0.002 × elevation	0.83	1.26
Ta ₁₁ = -0.912 + 0.271 × LSTd + 0.411 × LSTn - 0.001 × elevation	0.83	1.57
Ta ₁₂ = -2.560 + 0.314 × LSTd + 0.463 × LSTn + 0.00037 × elevation	0.86	1.67

References

- Nieto, H.; Sandholt, I.; Aguado, I.; Chuvieco, E.; Stisen, S. Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula. *Remote Sens. Environ.* **2011**, *115*, 107–116.
- Shamir, E.; Georgakakos, K.P. MODIS Land Surface Temperature as an index of surface air temperature for operational snowpack estimation. *Remote Sens. Environ.* **2014**, *152*, 83–98.
- Benali, A.; Carvalho, A.C.; Nunes, J.P.; Carvalhais, N.; Santos, A. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sens. Environ.* **2012**, *124*, 108–121.

4. Stisen, S.; Sandholt, I.; Nørgaard, A.; Fensholt, R.; Eklundh, L. Estimation of diurnal air temperature using MSG SEVIRI data in West Africa. *Remote Sens. Environ.* **2007**, *110*, 262–274.
5. Prihodko, L.; Goward, S.N. Estimation of air temperature from remotely sensed surface observations. *Remote Sens. Environ.* **1997**, *60*, 335–346.
6. Hooker, J.; Duveiller, G.; Cescatti, A. A global dataset of air temperature derived from satellite remote sensing and weather stations. *Sci. Data* **2018**, *5*, 180246.
7. Peón, J.; Recondo, C.; Calleja, J.F. Improvements in the estimation of daily minimum air temperature in peninsular Spain using MODIS land surface temperature. *Int. J. Remote Sens.* **2014**, *35*, 5148–5166.
8. Lam, N.S.N. Spatial interpolation methods: A review. *Am. Cartogr.* **1983**, *10*, 129–150.
9. Thiébaux, H.J. Statistics and the environment: The analysis of large-scale earth-oriented systems. *Environmetrics* **1991**, *2*, 5–24.
10. Ishida, T.; Kawashima, S. Use of cokriging to estimate surface air temperature from elevation. *Theor. Appl. Clim.* **1993**, *47*, 147–157.
11. Hutchinson, M.F. A New Objective Method for Spatial Interpolation of Meteorological Variables from Irregular Networks Applied to the Estimation of Monthly Mean Solar Radiation, Temperature, Precipitation and Windrun; Need for climatic and hydrologic data in agriculture in South East Asia. In Proceeding UN University workshop, Canberra, 1983; pp. 95–104.
12. Willmott, C.J.; Rowe, C.M.; Philpot, W.D. Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-Point Interpolation and Contouring. *Am. Cartogr.* **1985**, *12*, 5–16.
13. Willmott, C.J.; Robeson, S.M. Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Clim.* **1995**, *15*, 221–229.
14. Burrough, P.A.; McDonnell, R.A. *Principles of Geographical Information Systems*; Oxford University Press: Oxford, UK, 1998.
15. Mostovoy, G.V.; King, R.L.; Reddy, K.R.; Kakani, V.G.; Filippova, M.G. Statistical Estimation of Daily Maximum and Minimum Air Temperatures from MODIS LST Data over the State of Mississippi. *GIScience Remote Sens.* **2006**, *43*, 78–110.
16. Vogt, J.V.; Viau, A.A.; Paquet, F. Mapping regional air temperature fields using satellite-derived surface skin temperatures. *Int. J. Clim.* **1997**, *17*, 1559–1579.
17. Vancutsem, C.; Ceccato, P.; Dinku, T.; Connor, S.J. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sens. Environ.* **2010**, *114*, 449–465.
18. Dash, P.; Göttsche, F.-M.; Olesen, F.-S.; Fischer, H. Land surface temperature and emissivity estimation from passive sensor data: Theory and practice-current trends. *Int. J. Remote Sens.* **2002**, *23*, 2563–2594.
19. Oke, T. The urban energy balance. *Prog. Phys. Geogr. Earth Environ.* **1988**, *12*, 471–508.
20. Chatterjee, R.; Singh, N.; Thapa, S.; Sharma, D.; Kumar, D. Retrieval of land surface temperature (LST) from landsat TM6 and TIRS data by single channel radiative transfer algorithm using satellite and ground-based inputs. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 264–277.
21. Li, Z.-L.; Tang, B.-H.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.F.; Sobrino, J.A. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37.
22. Prata, A.J.; Caselles, V.; Coll, C.; Sobrino, J.A.; Otlé, C. Thermal remote sensing of land surface temperature from satellites: Current status and future prospects. *Remote Sens. Rev.* **1995**, *12*, 175–224.
23. Cresswell, M.P.; Morse, A.P.; Thomson, M.C.; Connor, S.J. Estimating surface air temperatures, from Meteosat land surface temperatures, using an empirical solar zenith angle model. *Int. J. Remote Sens.* **1999**, *20*, 1125–1132.
24. Stoll, M.J.; Brazel, A.J. Surface-air temperature relationships in the urban environment of phoenix, Arizona. *Phys. Geogr.* **1992**, *13*, 160–179.
25. Garand, L.; Buehner, M.; Wagneur, N. Background Error Correlation between Surface Skin and Air Temperatures: Estimation and Impact on the Assimilation of Infrared Window Radiances. *J. Appl. Meteorol.* **2004**, *43*, 1853–1863.
26. Jin, M.; Dickinson, R.E. Land surface skin temperature climatology: Benefitting from the strengths of satellite observations. *Environ. Res. Lett.* **2010**, *5*, 044004.
27. Zakšek, K.; Schroedter-Homscheidt, M. Parameterization of air temperature in high temporal and spatial resolution from a combination of the SEVIRI and MODIS instruments. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 414–421.

28. Sun, Y.J.; Wang, J.F.; Zhang, R.H.; Gillies, R.R.; Xue, Y.Y.C.B.; Bo, Y.C. Air temperature retrieval from remote sensing data based on thermodynamics. *Theor. Appl. Climatol.* **2005**, *80*, 37–48.
29. Zhu, W.; Lú, A.; Jia, S. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sens. Environ.* **2013**, *130*, 62–73.
30. Czajkowski, K.; Goward, S.; Stadler, S.; Walz, A. Thermal Remote Sensing of Near Surface Environmental Variables: Application Over the Oklahoma Mesonet. *Prof. Geogr.* **2000**, *52*, 345–357.
31. Kilibarda, M.; Hengl, T.; Heuvelink, G.B.M.; Gräler, B.; Pebesma, E.; Tadić, M.P.; Bajat, B. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *J. Geophys. Res. Atmos.* **2014**, *119*, 2294–2313.
32. Chen, F.; Liu, Y.; Liu, Q.; Qin, F. A statistical method based on remote sensing for the estimation of air temperature in China. *Int. J. Climatol.* **2015**, *35*, 2131–2143.
33. Xu, Y.; Qin, Z.; Shen, Y. Study on the estimation of near-surface air temperature from MODIS data by statistical methods. *Int. J. Remote Sens.* **2012**, *33*, 7629–7643.
34. Yan, H.; Zhang, J.; Hou, Y.; He, Y. Estimation of air temperature from MODIS data in east China. *Int. J. Remote Sens.* **2009**, *30*, 6261–6275.
35. Bartkowiak, P.; Castelli, M.; Notarnicola, C. Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas. *Remote Sens.* **2019**, *11*, 1319.
36. Lu, L.; Zhang, T.; Wang, T.; Zhou, X. Evaluation of Collection-6 MODIS Land Surface Temperature Product Using Multi-Year Ground Measurements in an Arid Area of Northwest China. *Remote Sens.* **2018**, *10*, 1852.
37. Zhou, W.; Peng, B.; Shi, J.; Wang, T.; Dhital, Y.P.; Yao, R.; Yu, Y.; Lei, Z.; Zhao, R. Estimating High Resolution Daily Air Temperature Based on Remote Sensing Products and Climate Reanalysis Datasets over Glacierized Basins: A Case Study in the Langtang Valley, Nepal. *Remote Sens.* **2017**, *9*, 959.
38. Janatian, N.; Sadeghi, M.; Sanaeinejad, S.H.; Bakhshian, E.; Farid, A.; Hashemini, S.M.; Ghazanfari, S. A statistical framework for estimating air temperature using MODIS land surface temperature data. *Int. J. Climatol.* **2017**, *37*, 1181–1194.
39. Ho, H.C.; Knudby, A.; Sirovyak, P.; Xu, Y.; Hodul, M.; Henderson, S.B. Mapping maximum urban air temperature on hot summer days. *Remote Sens. Environ.* **2014**, *154*, 38–45.
40. Duan, S.-B.; Li, Z.-L.; Tang, B.-H.; Wu, H.; Tang, R.; Bi, Y.; Zhou, G. Estimation of Diurnal Cycle of Land Surface Temperature at High Temporal and Spatial Resolution from Clear-Sky MODIS Data. *Remote Sens.* **2014**, *6*, 3247–3262.
41. Noi, P.T.; Kappas, M.; Degener, J. Estimating Daily Maximum and Minimum Land Air Surface Temperature Using MODIS Land Surface Temperature Data and Ground Truth Data in Northern Vietnam. *Remote Sens.* **2016**, *8*, 1002.
42. Li, L.; Zha, Y. Estimating monthly average temperature by remote sensing in China. *Adv. Space Res.* **2019**, *63*, 2345–2357.
43. Yoo, C.; Im, J.; Park, S.; Quackenbush, L.J. Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *137*, 149–162.
44. Yang, Y.Z.; Cai, W.H.; Yang, J. Evaluation of MODIS Land Surface Temperature Data to Estimate Near-Surface Air Temperature in Northeast China. *Remote Sens.* **2017**, *9*, 410.
45. Meyer, H.; Katurji, M.; Appelhans, T.; Müller, M.U.; Nauss, T.; Roudier, P.; Zawar-Reza, P. Mapping Daily Air Temperature for Antarctica Based on MODIS LST. *Remote Sens.* **2016**, *8*, 732.
46. Emamifar, S.; Rahimikhoob, A.; Noroozi, A.A. Daily mean air temperature estimation from MODIS land surface temperature products based on M5 model tree. *Int. J. Clim.* **2013**, *33*, 3174–3181.
47. Zhang, H.; Zhang, F.; Zhang, G.; He, X.; Tian, L. Evaluation of cloud effects on air temperature estimation using MODIS LST based on ground measurements over the Tibetan Plateau. *Atmos. Chem. Phys. Discuss.* **2016**, *16*, 13681–13696.
48. Keramitsoglou, I.; Kiranoudis, C.T.; Sismanidis, P.; Zakšek, K. An Online System for Nowcasting Satellite Derived Temperatures for Urban Areas. *Remote Sens.* **2016**, *8*, 306.
49. Nanzad, L.; Zhang, J.; Tuvdendorj, B.; Nabil, M.; Zhang, S.; Bai, Y. NDVI anomaly for drought monitoring and its correlation with climate factors over Mongolia from 2000 to 2016. *J. Arid. Environ.* **2019**, *164*, 69–77.
50. Nandintsetseg, B.; Greene, J.S.; Goulden, C.E. Trends in extreme daily precipitation and temperature near lake Hövsgöl, Mongolia. *Int. J. Clim.* **2007**, *27*, 341–347.

51. World Map of the Köppen Climate Classification. Available online: [https://commons.wikimedia.org/wiki/File:World_K%C3%B6ppen_Classification_\(with_authors\).svg](https://commons.wikimedia.org/wiki/File:World_K%C3%B6ppen_Classification_(with_authors).svg) (accessed on 18 October 2019).
52. Climate Data Org Webpage. Available online: <https://en.climate-data.org/asia/mongolia-180/> (accessed on 17 October 2019).
53. Duan, S.-B.; Li, Z.-L.; Wu, H.; Leng, P.; Gao, M.; Wang, C. Radiance-based validation of land surface temperature products derived from Collection 6 MODIS thermal infrared data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *70*, 84–92.
54. Oyler, J.W.; Dobrowski, S.Z.; Holden, Z.A.; Running, S.W. Remotely Sensed Land Skin Temperature as a Spatial Predictor of Air Temperature across the Conterminous United States. *J. Appl. Meteorol. Clim.* **2016**, *55*, 1441–1457.
55. Xu, Y.; Knudby, A.; Ho, H.C. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. *Int. J. Remote Sens.* **2014**, *35*, 8108–8121.
56. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905.
57. Wan, Z. New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. *Remote Sens. Environ.* **2008**, *112*, 59–74.
58. Wan, Z. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens. Environ.* **2014**, *140*, 36–45.
59. Wan, Z.; Zhang, Y.; Zhang, Q.; Li, Z.-L. Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sens. Environ.* **2002**, *83*, 163–180.
60. Wan, Z. *Collection-6 MODIS Land Surface Temperature Products Users' Guide*; ICES: Santa Barbara, CA, USA, 2007.
61. Mattiuzzi, M.; Verbesselt, J.; Klisch, A.; Stevens, F.; Mosher, S.; Evans, B.; Lobo, A.; Detsch, F. MODIS: MODIS Acquisition and Processing Package. R Package Version 1.1.4. Available online: <https://cran.r-project.org/src/contrib/Archive/MODIS/MODIS.pdf> (accessed on 25 October 2018).
62. R Core Team. R: A Language and Environment for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 31 October 2019).
63. Mattiuzzi, M.; Verbesselt, J.; Hengel, T.; Klisch, A.; Stevens, F.; Mosher, S.; Evans, B.; Lobo, A.; Hufkens, K.; Detsch, F. MODIS: MODIS Acquisition and Processing package. R Package Version 1.1.5. Available online: <https://cran.r-project.org/web/packages/MODIS/MODIS.pdf> (accessed on 8 March 2019).
64. Rabus, B.; Eineder, M.; Roth, A.; Bamler, R. The shuttle radar topography mission—A new class of digital elevation models acquired by spaceborne radar. *ISPRS J. Photogramm. Remote Sens.* **2003**, *57*, 241–262.
65. Bulgan, D.; Mandakh, N.; Odbayar, M.; Otgontugs, M.; Tsogtbaatar, J.; Elbegjargal, N.; Enerel, T.; Erdenetuya, M. *Desertification Atlas of Mongolia*; Tsogtbaatar, J., Khudulmur, S., Eds.; Admon Press: Ulaanbaatar, Mongolia, 2013; pp. 17–29, ISBN 978-9973-0-065-2.
66. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
67. Mutanga, O.; Adam, E.; Cho, M.A. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 399–406.
68. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222.
69. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
70. Immitzer, M.; Stepper, C.; Bock, S.; Straub, C.; Atzberger, C. Use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. *For. Ecol. Manag.* **2016**, *359*, 232–246.
71. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31.
72. Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Hall, D.E.; Falkowski, M.J. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sens. Environ.* **2008**, *112*, 2232–2245.
73. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Olmo, M.C.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104.
74. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236.
75. Trevor, H.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: Stanford, CA, USA, 2011.

76. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems* **2006**, *9*, 181–199.
77. Ishwaran, H. Variable importance in binary regression trees and forests. *Electron. J. Stat.* **2007**, *1*, 519–537.
78. Atzberger, C.; Guérif, M.; Baret, F.; Werner, W. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric.* **2010**, *73*, 165–173.
79. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Del Frate, F.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58.
80. Inoue, Y.; Sakaiya, E.; Zhu, Y.; Takahashi, W. Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. *Remote Sens. Environ.* **2012**, *126*, 210–221.
81. Gomez, C.; Rossel, R.A.V.; McBratney, A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma* **2008**, *146*, 403–411.
82. Farifteh, J.; Van Der Meer, F.; Atzberger, C.; Carranza, E.; Carranza, E.J. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sens. Environ.* **2007**, *110*, 59–78.
83. Rossel, R.V.; Walvoort, D.; McBratney, A.; Janik, L.; Skjemstad, J.; Rossel, R.V. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75.
84. Fernandes, L.S.; Fernandes, A.; Ferreira, A.; Cortes, R.; Pacheco, F. A partial least squares—Path modeling analysis for the understanding of biodiversity loss in rural and urban watersheds in Portugal. *Sci. Total Environ.* **2018**, *626*, 1069–1085.
85. Liu, Y.; Peng, J.; Wang, Y. Application of partial least squares regression in detecting the important landscape indicators determining urban land surface temperature variation. *Landsc. Ecol.* **2018**, *33*, 1133–1145.
86. Ceglar, A.; Toreti, A.; Lecerf, R.; Van Der Velde, M.; Dentener, F. Impact of meteorological drivers on regional inter-annual crop yield variability in France. *Agric. For. Meteorol.* **2016**, *216*, 58–67.
87. Carrascal, L.M.; Galván, I.; Gordo, O. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* **2009**, *118*, 681–690.
88. Abdi, H. Partial least square regression (PLS regression). *Encycl. Res. Methods Soc. Sci.* **2003**, *6*, 1–7.
89. Wold, H. Soft modeling: The basic design and some extensions. *Syst. Indirect Obs.* **1982**, *2*, 343.
90. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab Syst.* **2001**, *58*, 109–130.
91. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17.
92. Mevik, B.; Cederkvist, H.R. Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR). *J. Chemom.* **2004**, *18*, 422–429.
93. Mkhabela, M.S.; Bullock, P.R.; Sapirstein, H.D. Characterising the most critical climatic parameters that impact the quality of spring-wheat (*Triticum aestivum* L.) on the Canadian Prairies using partial least squares (PLS) analysis. *J. Cereal Sci.* **2018**, *81*, 44–51.
94. Richter, K.; Atzberger, C.; Hank, T.B.; Mauser, W. Derivation of biophysical variables from Earth observation data: Validation and statistical measures. *J. Appl. Remote Sens.* **2012**, *6*, 63557.
95. Farres, M.; Platikanov, S.; Tsakovski, S.; Tauler, R. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *J. Chemom.* **2015**, *29*, 528–536.
96. Otgonbayar, M.; Atzberger, C.; Chambers, J.; Damdinsuren, A. Mapping pasture biomass in Mongolia using Partial Least Squares, Random Forest regression and Landsat 8 imagery. *Int. J. Remote Sens.* **2019**, *40*, 3204–3226.
97. Sun, J.; Passi, K.; Jain, C.K. Improved microarray data analysis using feature selection methods with machine learning methods. In Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China, 15–18 December 2016; pp. 1527–1534.

