

Estimating 1 km gridded daily air temperature using a spatially varying coefficient model with sign preservation

Tao Zhang^a, Yuyu Zhou^{a,*}, Li Wang^b, Kaiguang Zhao^c, Zhengyuan Zhu^b

^a Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA

^b Department of Statistics, Iowa State University, Ames, IA 50011, USA

^c School of Environment and Natural Resources, Ohio Agricultural Research and Development Center, The Ohio State University, Wooster, OH 44691, USA

ARTICLE INFO

Edited by Jing M. Chen

Keywords:

Air temperature
Land surface temperature
Spatially varying coefficient models
Sign preservation
MODIS

ABSTRACT

Near-surface air temperature (Ta) is one of the key variables in a variety of studies such as hydrological modeling, assessment of heat waves, and energy modeling. Among existing methods, statistical algorithms are suitable for integrating auxiliary spatial data with station-based Ta data to produce gridded Ta over large areas. However, existing statistical algorithms (e.g., Geographically Weighted Regression (GWR)) cannot always correctly capture and preserve relationships between Ta and explanatory variables, which may increase uncertainties of relevant applications based on the estimated Ta with abnormal spatial patterns. This issue is mainly caused by the lack of enough observations due to the limited spatial coverage of weather stations, leading to abnormal relationships between Ta and explanatory variables. In order to address this issue, in this study, we introduced a new method named the Spatially Varying Coefficient Models with Sign Preservation (SVCMS-SP) to estimate gridded Ta using gridded land surface temperature (LST) and elevation as explanatory variables with presetting positive and negative signs for coefficients, respectively. Using this method, first, we calculated the preset parameters of the bivariate spline surface. Second, we used the input data at weather stations and constrained least squares regression to obtain the coefficient surface for both the explanatory variables (i.e., elevation and LST) and the intercept. Third, we calculated the gridded Ta using the 1 km gridded LST and elevation data, and the estimated spatially varying coefficient surfaces. We evaluated the model performance for estimating 1 km gridded daily maximum and minimum Ta (i.e., Tmax and Tmin) data in mainland China from 2003 to 2016 using 10-fold cross-validation and compared its performance with the GWR model. The average root mean square error (RMSE) and mean absolute error (MAE) based on the SVCMS-SP are 1.75 °C and 1.22 °C for Tmax, and 1.82 °C and 1.30 °C for Tmin, respectively. The SVCMS-SP method showed better performance than the GWR in terms of accuracy, computing efficiency, and has more interpretable coefficients for explanatory variables to get more realistic spatial pattern of gridded Ta. More important, the sign preservation of the SVCMS-SP method can mitigate the issue of abnormal relationships between Ta and explanatory variables in the traditional methods such as GWR, and therefore will contribute to future studies in developing better gridded air temperature or relevant data products.

1. Introduction

Air temperature (Ta) is an important meteorological parameter for a wide range of applications, such as public health (Lan et al., 2010; Zhang et al., 2019), disease vectors propagating (Lowen et al., 2007; Petrova and Russell, 2018; Wu et al., 2020), epidemic forecasting (Aggarwal et al., 2012; Connor et al., 1998), weather forecasting (Müller et al., 2017; Smith et al., 1988), terrestrial hydrology and phenology (Lin et al., 2012; Ren et al., 2019; Wang et al., 2009), climate and

environment change (Huang et al., 2019; Lamchin et al., 2018; Zhang et al., 2018a). Many methods and techniques exist to measure air temperature across space and over time. Most frequently, Ta is measured at 2 m above the ground at isolated weather stations; despite the high-fidelity accuracies achieved, ground stations are rarely designed to capture detailed spatial variations of Ta over extensive geographic regions. Spatial interpolation, therefore, becomes essential to extend these isolated point measurements to continuous surface estimates.

A wide array of interpolation techniques have been developed for

* Corresponding author.

E-mail address: yuyuzhou@iastate.edu (Y. Zhou).

<https://doi.org/10.1016/j.rse.2022.113072>

Received 15 November 2021; Received in revised form 2 April 2022; Accepted 30 April 2022

Available online 10 May 2022

0034-4257/© 2022 Elsevier Inc. All rights reserved.

generating gridded Ta estimates, but the majority of these traditional spatial interpolation methods are generally suitable for applications with small heterogeneity of Ta and tend to be inaccurate in mountainous areas due to the limited spatial coverage of weather stations and the large heterogeneity of landscape. For example, three most widely used techniques, Inverse Distance Weighting (IDW), Spline, and Kriging interpolation, achieve good performances only when the densities of measurements are high and the heterogeneity of landscape is not large. Even complicated is the spatial heterogeneity of landscape forms and environmental conditions, all of which violate the assumption of many traditional interpolation methods (Shen et al., 2020; Shi et al., 2017). These problems have been widely recognized in previous studies that considered traditional interpolation techniques (Chai et al., 2011; Dodson and Marks, 1997; Li and Heap, 2011; Stahl et al., 2006), arguing for the need to develop new alternative methods. A common remedy is to further incorporate auxiliary spatial datasets to augment point-based station Ta measurements.

Multiple gridded auxiliary datasets have been integrated with weather station data for high-resolution mapping of air temperature. Examples include elevation, latitude, distance from the ocean, distance to water, land surface temperature (LST), and normalized difference vegetation index (NDVI) (Li et al., 2018). These layers can be generally categorized into two types: static geographical datasets vs dynamic temporal layers. Of the static datasets, elevation is the most widely used variable because air temperature decreases with altitude; the physical relationship between air temperature and elevation is strong, especially over local regions, due to the adiabatic lapse processes (Heynen et al., 2016; Zhang et al., 2021a; Zhu et al., 2017). Other static geographical variables, such as latitude and distance from the ocean, are linked to temperature trends over large regions, but they have limited predictability for local variations in Ta. Overall, due to their static nature, geographical variables cannot capture temporal changes of Ta associated with changes in biophysical factors and human activities (e.g., land use change, vegetation growth, and building cooling and heating) (Li et al., 2018). In contrast, temporally-dynamic datasets can overcome the limitations of geographical datasets and have been combined with geographical datasets to estimate Ta (Chen et al., 2015a; Chen et al., 2016; Cristóbal et al., 2008; Kloog et al., 2014; Li et al., 2018; Oyler et al., 2015; Shen et al., 2020). A major source of time-varying observations over space is satellite remote sensing. Among the current satellite observations, LST is the most useful variable because of its strong positive relationship with Ta (Kim et al., 2021; Shen et al., 2020). NDVI has also been used for estimating gridded Ta (Cristóbal et al., 2008; Zhu et al., 2013). However, when both LST and NDVI were used for estimating Ta, the collinearity issue may occur as there are strong negative correlations between LST and NDVI. Despite their usefulness, the relationships between temporal satellite data and Ta typically vary across space and over time—a problem that also needs to be addressed by interpolation algorithms.

Three groups of methods have been used to integrate auxiliary spatial data to grid station-based Ta data, including the Temperature-Vegetation Index (TVX) method, physical models, and statistical models. The TVX method assumes that the temperature of a fully vegetated canopy approximates near-surface Ta within the canopy so that the near-surface Ta can be estimated at the maximum NDVI point through a linear equation between NDVI and LST (Goward and Waring, 1994; Nemani and Running, 1989; Zhu et al., 2013). Physically-based models rely on surface-air interactions, such as those based on energy balance and atmospheric vertical temperature gradient (Bisht and Bras, 2010; Hou et al., 2013; Jock, 2004; Sun et al., 2005; Zhu et al., 2017). The energy balance method was based on the energy balance equation, in which the net radiation is equal to the sum of the surface's sensible, soil, and latent heat fluxes, to estimate Ta based on LST and other corresponding surface environmental parameters (Sun et al., 2005; Zhang et al., 2015). The atmosphere profile-based method was based on the adiabatic lapse rate (ALR) in the troposphere (Bisht and Bras, 2010;

Jock, 2004; Zhu et al., 2017). Statistical methods are generally based on correlative relationships between Ta and other variables, encompassing a host of regression methods. Examples include linear regression, spatial-temporal regression kriging, geographically weighted regression (GWR), Cubist, random forest, artificial neural network, and deep learning (Chen et al., 2015a; Hengl et al., 2012; Hrisiko et al., 2020; Li and Zha, 2019; Li et al., 2018; Rao et al., 2019; Shen et al., 2020; Shi et al., 2017; Yoo et al., 2018).

Of the major groups of methods explored before, statistical methods are most suitable for generating gridded Ta data with a high spatio-temporal resolution for a large spatial extent. Specifically, the TVX method is not suitable for estimating Ta when vegetation cover is low. Physically-based models are too restrictive in terms of model parameter input; in practice, biophysical parameters needed for the physical models are hard to be obtained (Mostovoy et al., 2006; Shen et al., 2020). Atmosphere profile-based methods fail when temperature inversion occurs; another limitation is its relatively coarse spatial resolution (e.g., 5 km) (Zhu et al., 2017). In contrast, statistical methods are less restrictive in data requirements and they can be applied to extensive geographic regions (Noi et al., 2017; Shen et al., 2020; Xu et al., 2018).

Despite the wide availability of statistical algorithms, a key drawback remains that the existing algorithms cannot correctly capture and preserve relationships between Ta and explanatory variables. A manifestation of this problem is that signs of coefficients for explanatory variables are not always stable and sometimes counter-intuitive. For example, Ta is supposed to negatively correlate with elevation. However, the estimated air temperature based on existing statistical models, such as the GWR method, show positive relationships with elevation (Chen et al., 2015a; Li et al., 2018). Similar abnormal relationships between response variable and explanatory variables in regression methods can also be found in other applications such as spatial downscaling of precipitation products (Ma et al., 2017; Xu et al., 2015; Zhang et al., 2018b), and estimating gridded PM_{2.5} (Yang et al., 2020; Zhang and Kondragunta, 2021). Another critical drawback with the existing methods is the computation needed for large-area applications. For example, in the GWR method, calculation of the distance between each grid to all meteorological stations is needed and the process is time-consuming, limiting its usability for large regions and datasets (Chen et al., 2015a; Li et al., 2018).

Here, we reported and tested a class of Spatially Varying Coefficient Models with Sign Preservation (SVCMS-SP) to estimate gridded Ta data. The SVCMS-SP method was recently developed in the statistics community (Kim et al., 2021), but its potential has remained untapped for earth science applications. The SVCMS-SP method was used to address the key limitations listed above in the existing statistical algorithms. The basic hypothesis of the SVCMS-SP method is that the spatial variations of Ta can be expressed as a bivariate spline surface. More importantly, relationships between Ta and an explanatory variable are preset as positive or negative, under the limitations of triangulations. We tested this algorithm to estimate 1 km gridded daily maximum and minimum Ta (i.e., T_{max} and T_{min}) using 10-fold cross-validation in mainland China and further compared its performance with the commonly used GWR method. The remaining of this paper describes in detail the study area and data (Section 2), the methodology of the SVCMS-SP algorithm and accuracy assessment (Section 3), results and discussion (Section 4), and conclusions (Section 5).

2. Study area and data

2.1. Study area

This study focused on mainland China (Fig. 1), because it is a geographically extensive area with complicated terrains and large elevation variations, which lead to strong variations of Ta in space and time (Chen et al., 2015a; Li and Zha, 2018; Yao et al., 2020). Meanwhile, there are hundreds of ground-based weather stations scattered across

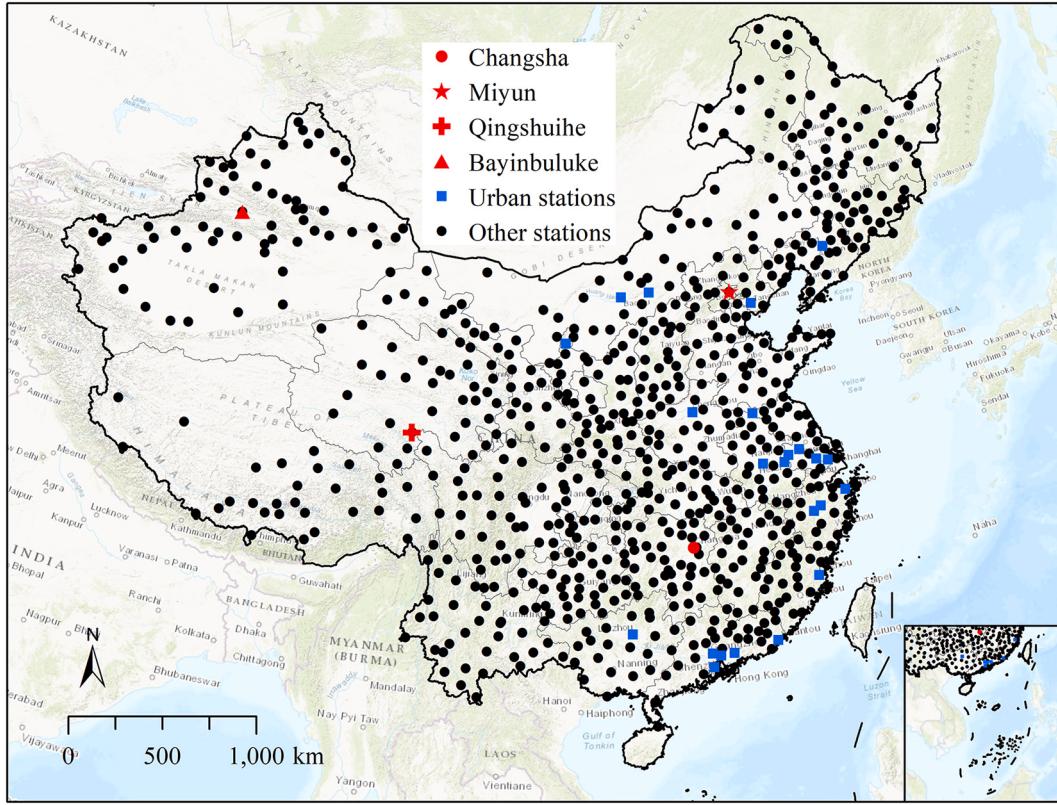


Fig. 1. Study area and locations of weather stations throughout mainland China. Ta from all the weather stations were used for accuracy assessment. Among them, four representative stations (i.e., Changsha, Miyun, Qingshuihe, and Bayinbuluke) were selected for a detailed analysis of the model performance such as daily residuals and coefficients of explanatory variables. Moreover, 23 typical urban stations (i.e., its 3 km surrounding is all urban regions) were selected for evaluating the accuracy of estimated Ta in urban regions.

the country (Fig. 1), providing high-quality Ta measurements to examine the robustness of the algorithm. Therefore, mainland China is an ideal region to test the general applicability of the algorithm for estimating 1 km resolution Ta using ground surface temperature and elevation data.

2.2. Ground-based air temperature measurements

Ground-based Ta measurements for estimating the gridded air temperature and evaluating the performance of our algorithm were obtained from the China Meteorological Data Service Centre (<http://data.cma.cn/en>). These ground-based data have been processed to ensure their high quality before sharing with the public. We chose to use the available daily maximum and minimum air temperature measurements from 2003 and 2016, which contain records of 793 ground weather stations across mainland China (Fig. 1). However, the records we obtained for the locations of these weather stations are not accurate enough due to the omitting of numeric at arc seconds (arcsec) for latitude and longitude. Therefore, we manually calibrated the locations of weather stations in the region with complicated terrains to reduce the uncertainties in the estimated air temperature data in this study (details in Appendix A of the supplement).

2.3. Gap-filled land surface temperature data

A seamless 1 km resolution daily (mid-daytime and mid-nighttime) land surface temperature (LST) dataset from 2003 to 2016 (Zhang et al., 2021b), which is available through Iowa State University's DataShare platform (<https://doi.org/10.25380/iastate.c.5078492>), was used to estimate gridded Ta. This dataset was produced from MODIS daily LST product (MYD11A1/MOD11A1, from Terra and Aqua

satellites) using a spatiotemporal gap-filling framework, with an accuracy of Root Mean Square Error (RMSE) to be 1.88 °C and 1.33 °C, respectively for mid-daytime (1:30 pm) and mid-nighttime (1:30 am) at the global level (Zhang et al., 2022). These mid-daytime (1:30 pm) and mid-nighttime (1:30 am) LST data were used as explanatory variables for estimating gridded Tmax and Tmin, respectively, because LST and Ta of the corresponding time have the smallest time difference with high correlations.

2.4. Elevation data

The digital elevation model (DEM), used for gridded Ta estimation, is a multi-source based 30-arcsec (~1 km) spatial resolution global topography/bathymetry data (i.e., SRTM30_PLUS described by Becker et al. (2009)). The SRTM30_PLUS contains land topography and ocean bathymetry. Land elevation is mainly from the 30-arcsec resolution Shuttle Radar Topography Mission (SRTM30) topography (Hennig et al., 2001; Rosen, 2000) within a latitude of ±55 degrees, GTOPO30 topography (Danielson and Gesch, 2011) in the Arctic, and ICESat derived topography (Dimarzio et al., 2007) in Antarctica.

3. Methodology

3.1. Framework of the SVCMM-SP algorithm

The SVCMM-SP algorithm was proposed by Kim et al. (2021). The value of response variable at a specific location was estimated based on the corresponding explanatory variables at the same location and their coefficients fitted by using the SVCMM-SP algorithm (Eq. (1)).

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{j=1}^m \beta_{ij}(u_i, v_i) \cdot x_{ij}(u_i, v_i) + \varepsilon_i \quad (1)$$

where y_i is the explanatory variable at the location i with the coordinate of (u_i, v_i) , $\beta_{ij}(u_i, v_i)$ is the coefficient for j -th ($j = 1, \dots, m$) explanatory variable $x_{ij}(u_i, v_i)$ at the location i . $\beta_{i0}(u_i, v_i)$ is the corresponding intercept parameter and ε_i indicates the random error at the location i . In the SVCMM-SP algorithm, unknown parameters $\beta_{ij}(u_i, v_i)$ can be constrained to be positive or negative if the explanatory variables are positively or negatively correlated with Ta and no constraints if both positive and negative correlations exist. In other words, it is not necessary to use a positive or negative relationship between Ta and every explanatory variable for implementing the SVCMM-SP method. We can also set no constraints on the coefficient of an explanatory variable using the SVCMM-SP when there is no such positive/negative relationship. Specifically, $\beta_{i0}(u_i, v_i)$ has no constraints as it is an unknown intercept parameter. These unknown parameters were estimated by using a penalized bivariate spline method based on the triangulation technique under constraints (Kim et al., 2021).

We applied the SVCMM-SP algorithm to estimate daily gridded Ta and to build the 1 km spatial resolution Ta dataset in this study. The overall framework is presented in Fig. 2, wherein the input data include ground-based Ta measurements from weather stations, gap-filled 1 km resolution daily MODIS-like LST, and 1 km resolution elevation data. We selected LST and elevation data as explanatory variables since they are consistently positively and negatively correlated with Ta, respectively, and are generally enough to be used for capturing the spatial variations of Ta. We used the data from weather stations to estimate the coefficients of the equation between Ta and explanatory variables (i.e., elevation and LST) and the intercept for each 1 km grid. We then estimated the 1 km gridded Ta values based on the 1 km gridded elevation and LST data and the derived coefficients (Eq. (2)).

$$T_a(i, j) = \beta_0(i, j) + \beta_{elev}(i, j) \cdot Elev(i, j) + \beta_{lst}(i, j) \cdot LST(i, j) \quad (2)$$

where $T_a(i, j)$ is the estimated Ta at the 1 km grid with a central

coordinate of (i, j) , $Elev(i, j)$ and $LST(i, j)$ are the values of elevation and LST at grid (i, j) , and $\beta_0(i, j)$, $\beta_{elev}(i, j)$ and $\beta_{lst}(i, j)$ are the estimated intercept, coefficient values for elevation and LST at the grid (i, j) based on the SVCMM-SP algorithm. Specifically, $\beta_{elev}(i, j)$ and $\beta_{lst}(i, j)$ are constrained to be negative and positive, respectively, by the SVCMM-SP algorithm (details in Appendix B.2 of the supplement).

3.2. Implementing the SVCMM-SP algorithm

There are three steps for implementing the SVCMM-SP algorithm (Fig. 3). First, we constructed a triangulation that covers the whole study area, and used it to compute the basis functions of the bivariate spline surface (Kim et al., 2021; Lai and Wang, 2013; Mu et al., 2018). The basis functions were determined by the triangularization, and the bivariate spline surface can be written as a linear combination of the spline basis functions. Spline basis functions for three sets of points were separately built across the study area using triangulations. The first set of points are the coordinates of centers for grids with a 0.25-degree resolution, which were used for controlling the basic shape of the bivariate spline surface to improve the computing efficiency. The second set of points are the coordinates of weather stations, which were used for fitting the unknown parameters of the bivariate spline surface in the second step. The third set of points are the coordinates of centers of 1 km grids, which are the locations of the output data.

Second, we fitted the bivariate spline surface to obtain 1 km resolution coefficients of explanatory variables. We fitted the coefficient functions of the spline (i.e., spatially varying coefficient functions), which can control the specific shape of the bivariate spline surface, using sign-constrained least square regression based on the spline basis functions and the input data from weather stations. We then calculated the coefficients of explanatory variables (i.e., elevation and LST) at 1 km grids based on the same spatially varying coefficient functions. Third, we estimated 1 km gridded Ta based on 1 km gridded LST and DEM data,

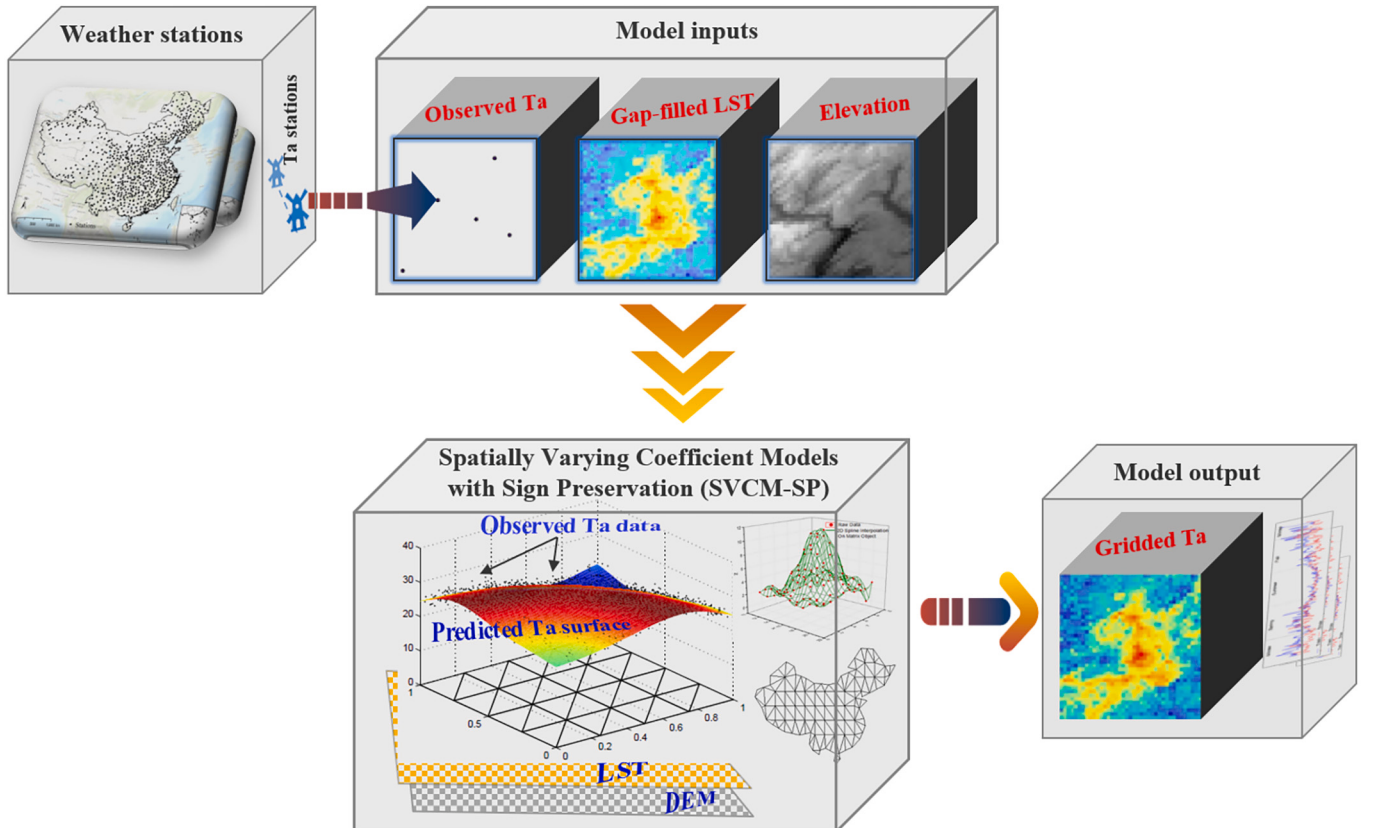


Fig. 2. Framework for estimating the gridded air temperature based on the SVCMM-SP algorithm.

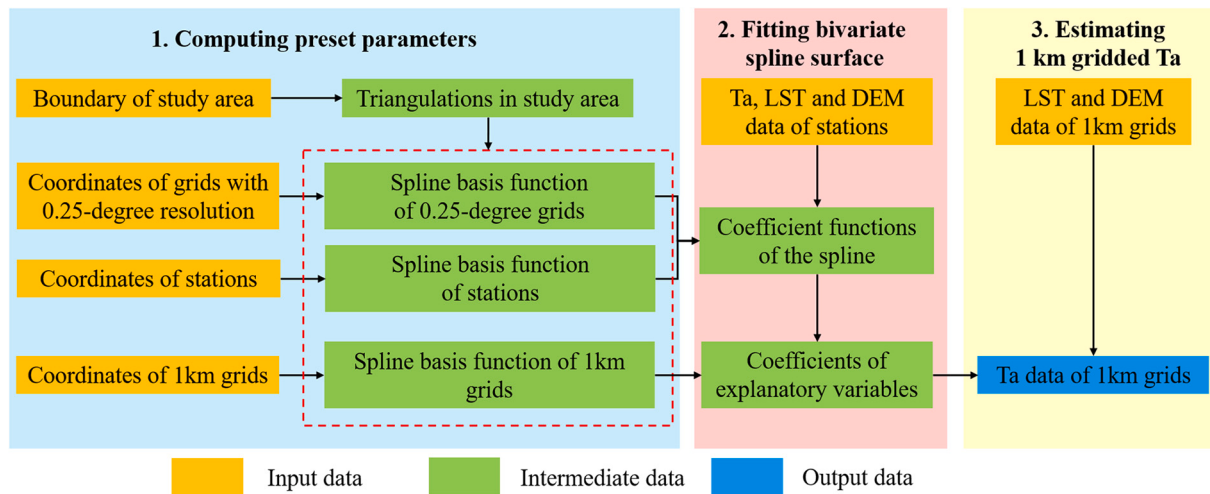


Fig. 3. Flow chart of the SVCN-SP algorithm. The spline basis function represents the preset parameters of bivariate spline surface; coefficient functions of the spline are the coefficients of spline basis functions for the fitted spline surface; explanatory variables include elevation and LST in this study.

and their coefficients.

3.3. Accuracy assessment

We used 10-fold cross-validation to evaluate the accuracy of the SVCN-SP algorithm. That is, for each day, we randomly divided station observations into ten groups, and then used data from 9 groups to estimate Ta for the remaining group. We then compared the estimated Ta with the actual observations of this remaining group to evaluate the accuracy of the estimated Ta. We repeated this evaluation ten times until observations from all groups were used. Finally, the accuracy for each year was calculated based on the mean value of daily accuracies for that year (results are shown in Table 1), and the accuracy for each station was calculated based on observations from all the days of a given year (results of the year 2010 are shown in Fig. 4). In addition, we used the widely used GWR as a benchmark and compared it with the SVCN-SP method using the same accuracy assessment strategy (i.e., 10-fold cross-validation). Moreover, we selected 23 typical urban stations (Fig. 1) to evaluate the model performance in urban regions. A typical urban station was identified if its 3 km surrounding is all urban regions based on the urban extent data (Zhou et al., 2018).

In this study, we used the root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) as the indicators of accuracy.

4. Results and discussion

4.1. Accuracy assessment using cross-validation

The SVCN-SP algorithm performs well and shows higher accuracies than that of the GWR model according to the 10-fold cross-validation of estimated Ta from 2003 to 2016 (Tables 1 and 2). The estimated and observed Ta are strongly correlated with average R^2 values of 0.93 and 0.94 for Tmax and Tmin, respectively in the whole study area (Table 1). The average RMSE and MAE of the SVCN-SP method are 1.75 °C and 1.22 °C for Tmax, and 1.82 °C and 1.30 °C for Tmin, respectively. The accuracies of the GWR model are lower than that of the SVCN-SP method with the average R^2 of 0.90 and RMSE, MAE of 2.03 °C, 1.40 °C for Tmax, and 0.93, 1.99 °C, 1.43 °C for Tmin, respectively. The estimated Ta based on the SVCN-SP in urban areas even performs better with lower RMSE and MAE values than those in the whole study area, possibly because most weather stations are in and around cities.

The accuracy of estimated Ta using the SVCN-SP method varies in space with lower RMSE values for both Tmax and Tmin in Eastern China compared to Western China (i.e., the mountainous areas with sparse weather stations), which is similar to the results from the GWR (Fig. 4). This spatial variation is mainly due to two factors, i.e., topography and weather station density (Li et al., 2018). First, Ta may change dramatically with elevation due to complex topography. Meanwhile, the

Table 1

Ten-fold cross-validation of estimated Ta from 2003 to 2016 in mainland China using the SVCN-SP and GWR.

Year	Tmax						Tmin					
	RMSE (°C)		MAE (°C)		R^2		RMSE (°C)		MAE (°C)		R^2	
	SVCN-SP	GWR	SVCN-SP	GWR	SVCN-SP	GWR	SVCN-SP	GWR	SVCN-SP	GWR	SVCN-SP	GWR
2003	1.74	2.02	1.21	1.4	0.93	0.90	1.79	1.95	1.28	1.41	0.95	0.94
2004	1.73	2.00	1.21	1.38	0.93	0.91	1.83	1.99	1.31	1.44	0.94	0.93
2005	1.78	2.05	1.23	1.41	0.93	0.90	1.83	1.99	1.30	1.43	0.94	0.94
2006	1.75	2.04	1.22	1.42	0.93	0.90	1.85	2.00	1.31	1.43	0.94	0.93
2007	1.71	1.97	1.19	1.37	0.93	0.90	1.80	1.96	1.29	1.41	0.94	0.93
2008	1.71	1.98	1.19	1.36	0.93	0.90	1.80	1.97	1.29	1.42	0.94	0.93
2009	1.76	2.04	1.23	1.42	0.93	0.90	1.83	2.00	1.31	1.45	0.94	0.93
2010	1.80	2.05	1.26	1.43	0.93	0.90	1.85	2.01	1.31	1.44	0.94	0.93
2011	1.78	2.08	1.22	1.41	0.92	0.90	1.82	1.99	1.31	1.44	0.94	0.93
2012	1.76	2.05	1.22	1.41	0.93	0.90	1.79	1.97	1.28	1.42	0.95	0.94
2013	1.76	2.04	1.22	1.41	0.93	0.91	1.83	2.00	1.32	1.45	0.94	0.93
2014	1.74	2.02	1.21	1.39	0.92	0.90	1.81	2.00	1.31	1.44	0.94	0.93
2015	1.76	2.06	1.23	1.42	0.92	0.90	1.81	2.00	1.30	1.45	0.94	0.93
2016	1.76	2.03	1.22	1.41	0.93	0.90	1.81	1.99	1.29	1.44	0.95	0.93
Average	1.75	2.03	1.22	1.40	0.93	0.90	1.82	1.99	1.30	1.43	0.94	0.93

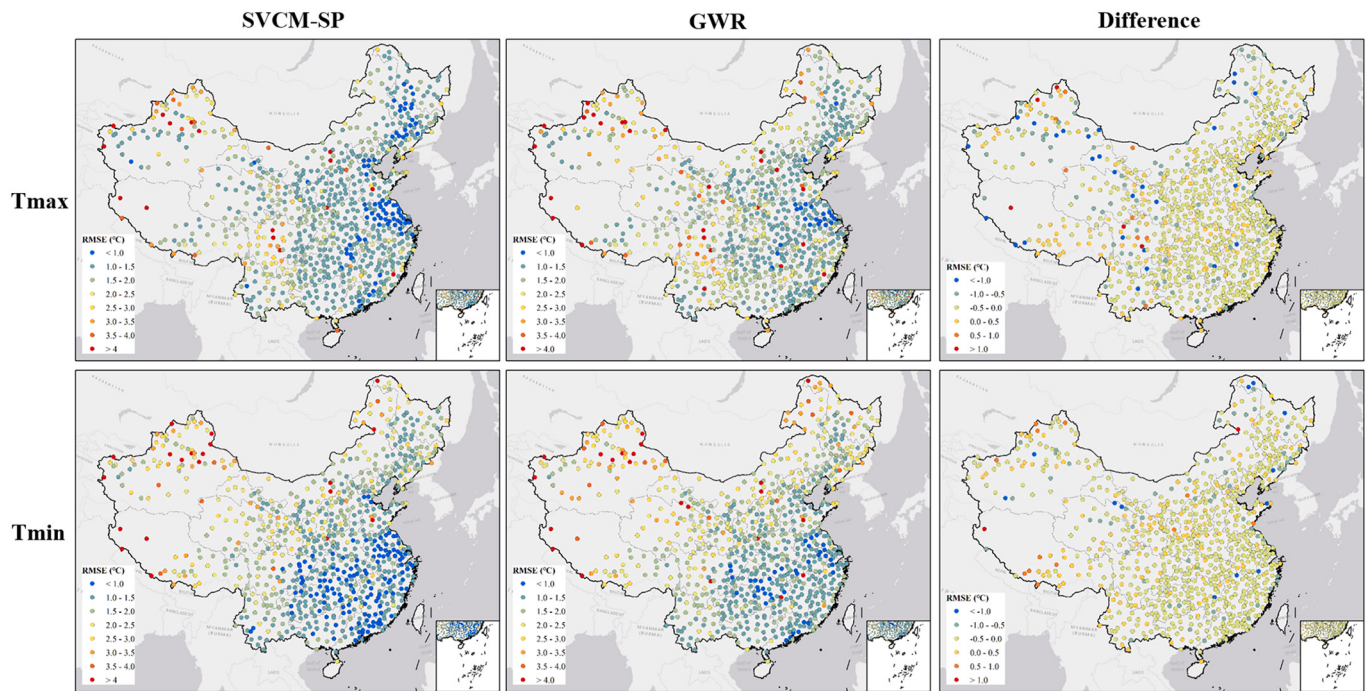


Fig. 4. Spatial patterns of RMSE for Tmax and Tmin at weather stations of mainland China in 2010.

Table 2

Accuracies of estimated Ta in urban stations based on ten-fold cross-validation from 2003 to 2016 in mainland China using the SVCMS-SP and GWR.

Year	Tmax						Tmin					
	RMSE (°C)		MAE (°C)		R ²		RMSE (°C)		MAE (°C)		R ²	
	SVCMS-SP	GWR	SVCMS-SP	GWR	SVCMS-SP	GWR	SVCMS-SP	GWR	SVCMS-SP	GWR	SVCMS-SP	GWR
2003	1.09	1.3	0.83	0.97	0.94	0.92	1.12	1.25	0.86	0.97	0.96	0.95
2004	1.06	1.21	0.81	0.91	0.94	0.91	1.15	1.30	0.88	1.01	0.95	0.94
2005	1.14	1.26	0.87	0.96	0.92	0.90	1.08	1.23	0.83	0.95	0.95	0.94
2006	1.13	1.30	0.86	0.99	0.92	0.90	1.12	1.28	0.86	0.98	0.95	0.94
2007	1.06	1.19	0.81	0.91	0.93	0.91	1.14	1.33	0.88	1.02	0.95	0.94
2008	1.07	1.21	0.82	0.92	0.93	0.91	1.15	1.37	0.88	1.04	0.96	0.94
2009	1.14	1.32	0.87	0.99	0.92	0.89	1.18	1.42	0.90	1.08	0.95	0.94
2010	1.16	1.29	0.88	0.96	0.92	0.90	1.21	1.39	0.91	1.05	0.94	0.93
2011	1.09	1.22	0.83	0.94	0.93	0.90	1.24	1.46	0.94	1.11	0.94	0.93
2012	1.06	1.24	0.81	0.94	0.93	0.90	1.14	1.37	0.88	1.05	0.95	0.94
2013	1.14	1.30	0.87	0.99	0.93	0.91	1.25	1.44	0.96	1.10	0.95	0.94
2014	1.09	1.24	0.85	0.94	0.92	0.90	1.26	1.45	0.97	1.11	0.95	0.94
2015	1.18	1.28	0.89	0.96	0.91	0.89	1.24	1.45	0.94	1.11	0.95	0.94
2016	1.09	1.26	0.84	0.96	0.92	0.90	1.21	1.45	0.93	1.11	0.95	0.94
Average	1.11	1.26	0.85	0.95	0.93	0.90	1.18	1.37	0.90	1.05	0.95	0.94

theoretically decreasing lapse rate of Ta, along with increasing elevation may also vary significantly in space and time (Heynen et al., 2016; Zhang et al., 2021a; Zhu et al., 2017) in the mountains, which leads to higher uncertainties in the estimated Ta in mountainous areas of Western China as compared with plain areas of Eastern China. Second, the density of weather stations in Western China is obviously sparser than that in Eastern China, which may lead to a weaker ability to capture the spatial variations of Ta in both methods. This has been reported in previous studies using the GWR model (Li et al., 2018). As shown in Fig. 4, the accuracy of estimated Ta using the SVCMS-SP method tends to be higher than those of the GWR model since distinctly more stations with lower RMSE for the SVCMS-SP than that of the GWR. The advantage of the SVCMS-SP on accuracy is more obvious in Eastern China as compared with Western China, possibly due to the higher density of weather stations.

The RMSE of the estimated Ta using the SVCMS-SP method shows seasonal variations and is generally lower than that from the GWR

model (Fig. 5). As shown in Fig. 5, the RMSEs of Tmax and Tmin using the SVCMS-SP method show lower values in the Summer and higher values in the Winter, which is similar to those from the GWR model. This seasonal variation may be caused by plant phenology which leads to a closer relationship between Ta and LST in the Summer than that in the Winter (Benali et al., 2012; Cai et al., 2017; Lin et al., 2012). The estimated Tmax shows higher RMSE than that of Tmin in the Summer (Fig. 5), possibly due to the Tmax being less correlated with mid-daytime LST and elevation than the correlations of Tmin with mid-nighttime LST and elevation in the Summer across mainland China (Fig. D1). However, the estimated Tmax shows lower RMSE than that of Tmin in the Winter for both methods (Fig. 5). A possible reason is that the Tmin and mid-night LST are less correlated in local regions than the Tmax and mid-daytime LST due to the longer time lag for the former than that of the latter (Mostovoy et al., 2006; Vancutsem et al., 2010). Meanwhile, RMSE of Tmax and Tmin using the SVCMS-SP method is always lower than those estimated by the GWR model, for most of the days

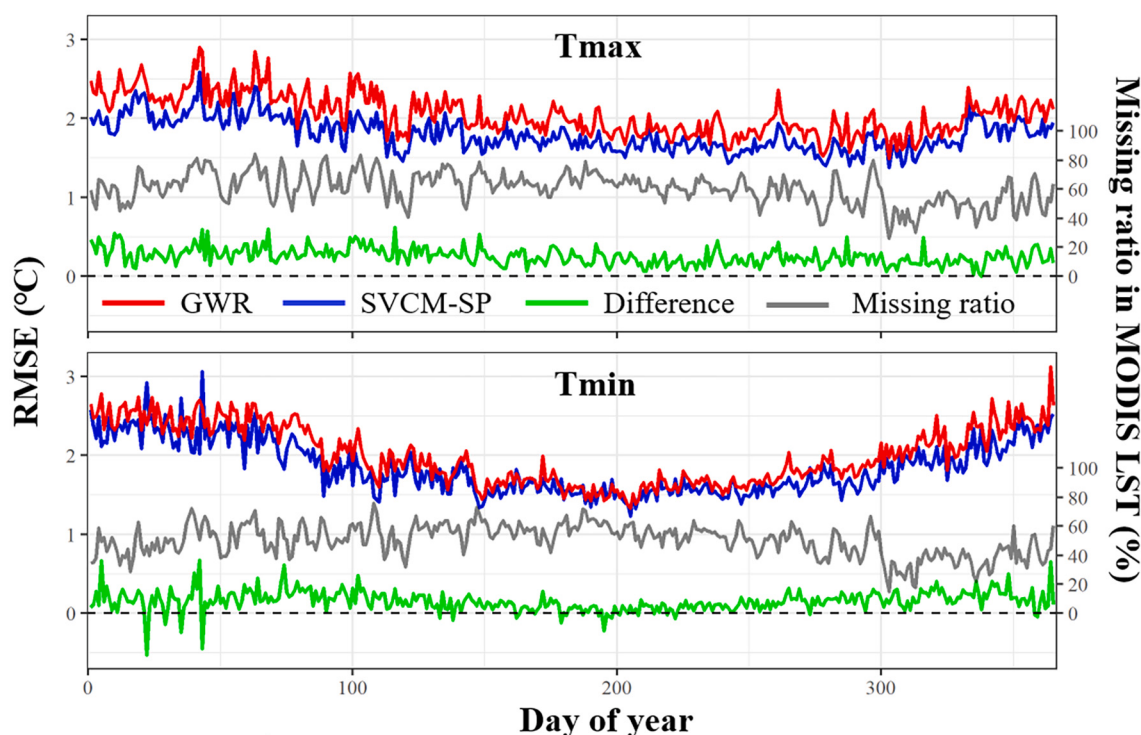


Fig. 5. Daily RMSE of Tmax and Tmin using the SVCMM-SP (blue line) and GWR (red line) methods and their differences (green line), and relevant missing ratios in original MODIS LST, across mainland China in 2010. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in 2010 (Fig. 5).

Accuracies of estimated Ta using the SVCMM-SP can be affected by the quality of LST data to some extent. The RMSE and MAE of estimated Tmax increase with the missing ratio in original MODIS LST data from 1.65 °C to 1.94 °C and 1.11 °C to 1.37 °C, respectively (Fig. 5 and Table F1). Accuracies of estimated Tmin are strongly affected by seasonal factors due to the distinct seasonal trend in RMSE of Tmin but impacts from missing ratios in original MODIS LST data are minor (Fig. 5). A possible reason is that clouds can lead to fluctuations in the relationship between daytime LST and Tmax (i.e., the lower difference between LST and Ta with clouds than that of clear sky condition in the daytime) by reducing solar radiation, while clouds might lead to smaller variations of correlation between nighttime LST and Tmin (i.e., the differences between LST and Ta are small for both cloud and clear sky conditions in nighttime) (Good, 2016). Overall, higher missing ratios in original MODIS LST (mainly caused by clouds) can result in weaker relationships between the gap-filled daytime LST (clear-sky) and Tmax.

The estimated Ta using the SVCMM-SP method illustrates comparable or higher accuracies than those from the GWR model at example weather stations (Fig. 6). Four representative weather stations in different geographical regions (i.e., Changsha, Miyun, Qingshuihe, and Bayinbuluke, in the South, North, Qinghai-Tibetan, and Northwest regions, respectively) spreading across mainland China (Fig. 1) were selected to explore the temporal pattern of residuals, i.e., differences between ground-based measurements and estimated Ta. The results illustrated that residuals of estimated Tmax for the SVCMM-SP and GWR methods show similar temporal patterns, and the residuals of the former show lower absolute values than those of the latter for several periods, especially at Bayinbuluke in the Winter, indicating higher accuracies using the SVCMM-SP method than those derived from the GWR model. Moreover, the residuals show seasonal variations at Bayinbuluke (absolute values are small in the Summer and large in the Winter), which is different from other stations. One possible reason is the lack of enough weather stations, which leads to an overestimation of Ta by

extrapolating the fitted SVCMM-SP (and GWR) model. The Bayinbuluke site is in the intermontane basin of the Tianshan Mountains with different climate characteristics (e.g., larger snow volume and higher snow frequency) and higher elevation compared to neighboring sites (out of the Tianshan Mountains), leading to lower temperature compared to other sites, especially in the Winter. When the SVCMM-SP model was spatially extrapolated to regions where LST and elevation are out of the corresponding numeric ranges in the fitted SVCMM-SP model, it might lead to biases in the estimated Ta since the relationship between Ta and LST/elevation might be different from that was estimated in the SVCMM-SP model. Therefore, most of the residuals are negative (overestimation of Ta based on observed Ta from surrounding sites with higher values) and the accuracy of the estimate Ta at the Bayinbuluke site based on the neighboring sites is lower in the Winter compared to the Summer.

The SVCMM-SP algorithm is robust because the accuracies of estimated Ta using the SVCMM-SP were always higher than those of the GWR with or without additional explanatory variables (i.e., NDVI and/or distance to water body) (Tables B1 and B2). Moreover, the SVCMM-SP model using LST and elevation only (Eq. (2)) and the model with additional explanatory variables (Eqs. (B2)–(B4)) show similar performance (details in Appendix B of the supplement). This is possibly because elevation and LST are always physically negatively and positively correlated with Ta in space and time, which can capture the majority of spatial and temporal variances in Ta. For example, NDVI is generally correlated with LST. The relationships between Ta and other explanatory variables (i.e., NDVI and distance to water body) are more complicated to be properly expressed by using the SVCMM-SP to improve the accuracy of estimated Ta. The SVCMM-SP model using additional NDVI and distance to water body with preserved negative and positive signs of coefficients, respectively, shows better performance than that without sign preservations, but shows similar performance with the SVCMM-SP using only LST and elevation with preserved signs for coefficients (Table B1). Due to the similar performance, we only included

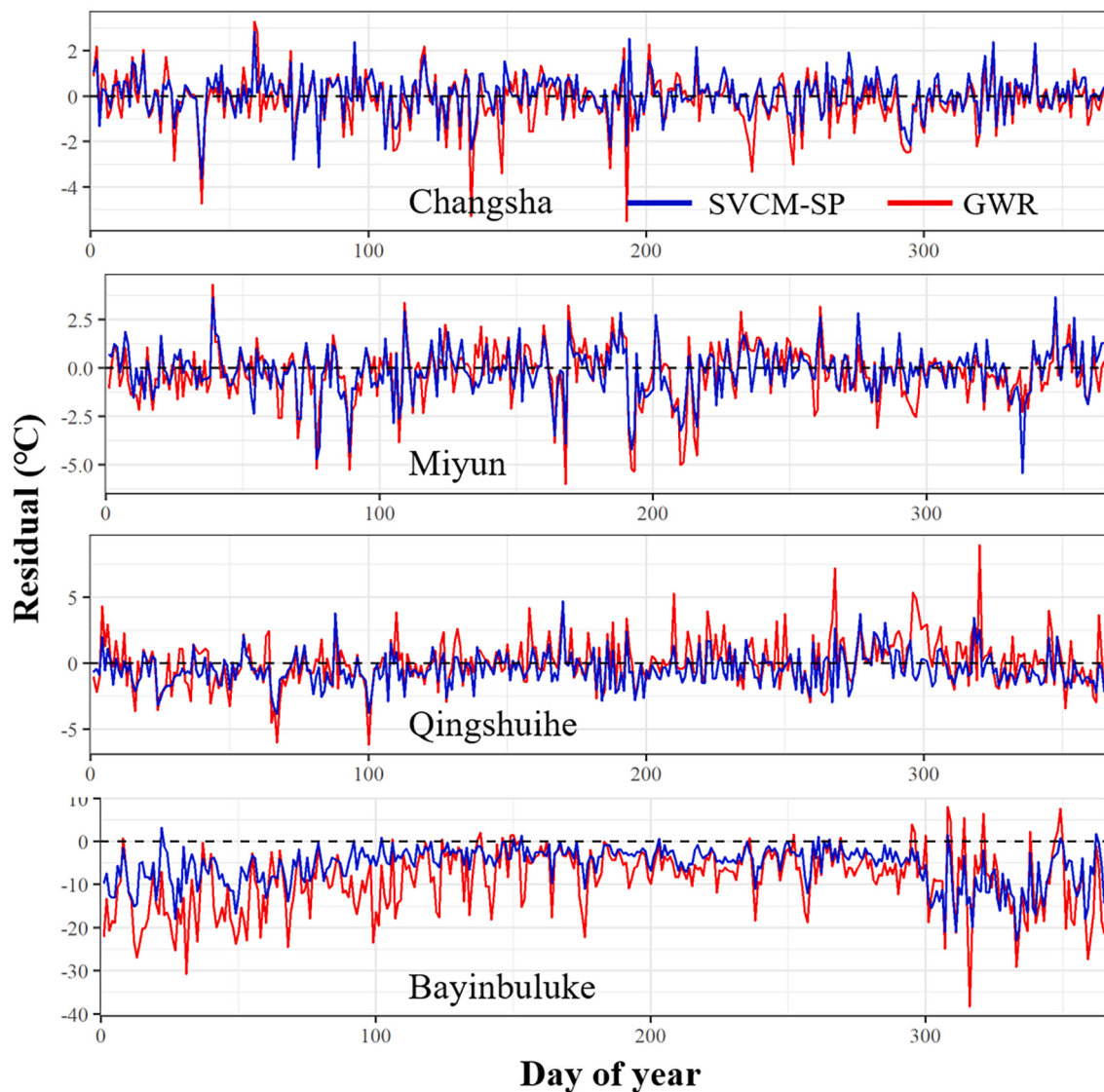


Fig. 6. Daily residuals of Tmax using the SVCMM-SP (blue line) and the GWR (red line) methods at four representative stations in 2010. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

LST and elevation in our SVCMM-SP model in terms of its simplicity and transferability.

Generally, the accuracy of the estimated Ta dataset can meet the needs of most applications (e.g., urban heat island effects and hydrological modeling) over mainland China. In the previous studies in China, Fang et al. (2021) estimated 0.1° daily Ta with the RMSE ranges from 0.78 to 2.09°C based on ground observations from 699 weather stations; Chen et al. (2021) estimated 1 km resolution daily mean Ta over mainland China with the RMSE ranges from 1.615 to 1.957°K based on ground observations from 2384 weather stations. Our study used Ta observations from 793 weather stations and obtained 1 km resolution daily Tmax and Tmin with the RMSE ranges from 1.71 to 1.85°C (Table 1). Considering the higher spatiotemporal resolution and the limited number of weather stations we used in this study, the accuracy is reasonable and acceptable. Specifically, there are some locations (e.g., the Bayinbuluke site) with low accuracies because of the low density of weather stations and the large variations of environmental conditions, which is consistent with the results in the study by Chen et al. (2021). When we estimated the gridded air temperature data, observations from all stations were used and the uncertainties in the estimated air temperature data should be lower compared to the evaluation.

4.2. Coefficients of explanatory variables

The SVCMM-SP fitted model can consistently and correctly preserve the negative relationship between Ta and elevation as indicated by the negative coefficient of elevation across time (Fig. 7). As shown in Fig. 7, the coefficients of elevation at four representative weather stations using the SVCMM-SP method are consistently negative. However, the coefficients of elevation using the GWR model are consistently negative at Qingshuihe, but they are positive for Changsha, Miyun, and Bayinbuluke stations on some days. According to the adiabatic lapse rate, Ta generally decreases along with the increase of elevation, i.e., a negative relationship (Heynen et al., 2016; Zhang et al., 2021a; Zhu et al., 2017). However, the GWR model cannot consistently preserve the negative relationship between Ta and elevation.

The SVCMM-SP fitted model can consistently and correctly preserve the positive relationship between Ta and LST as indicated by the positive coefficient of LST across time (Fig. 8). As shown in Fig. 8, the coefficients of LST at four representative weather stations using the SVCMM-SP method are consistently positive. However, the coefficients of LST using the GWR model are positive or negative for different periods in the four stations. Ta generally increases along with the increase of LST, i.e.,

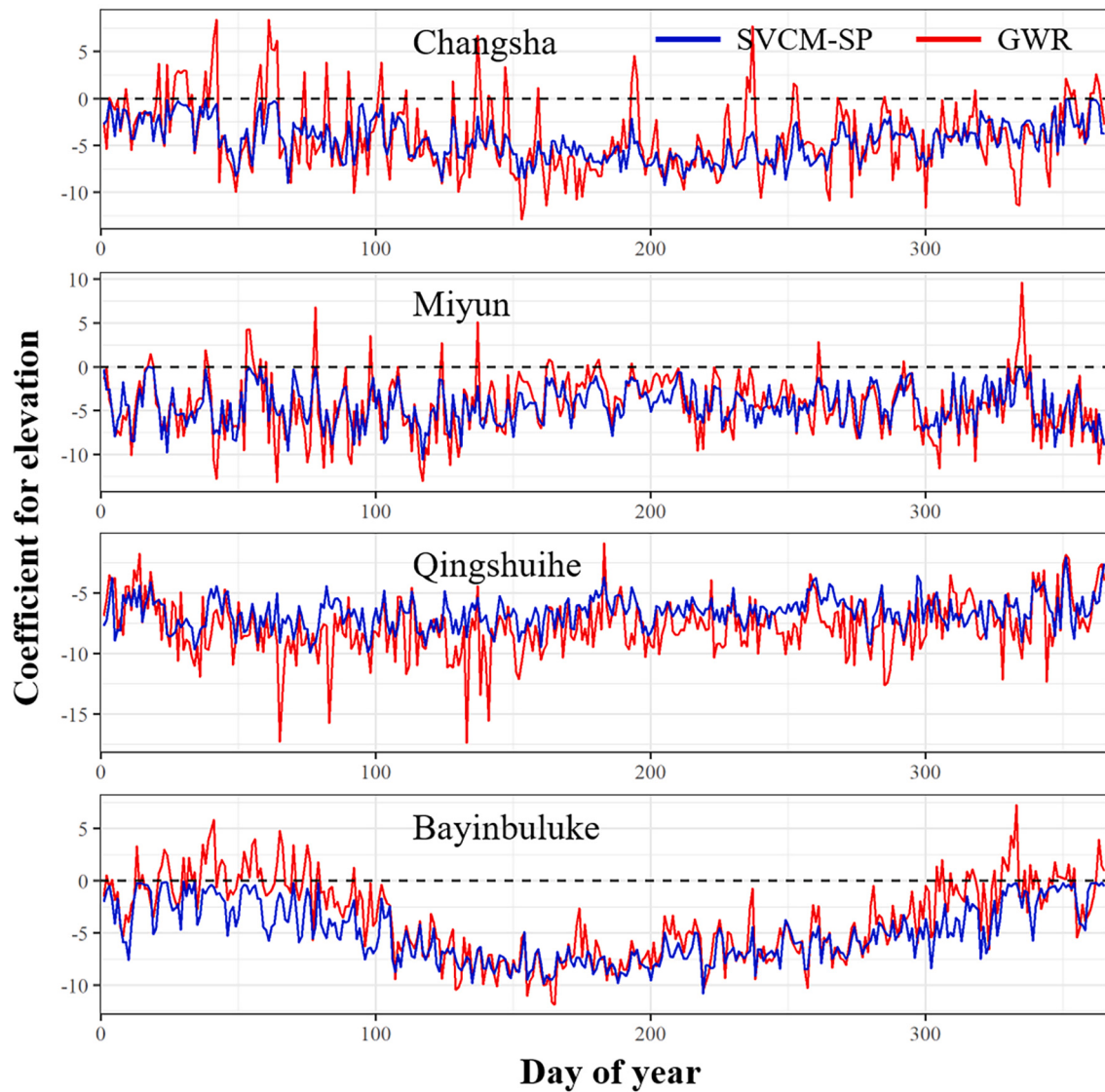


Fig. 7. Coefficients of elevation in estimating T_{max} using the SVCMM-SP (blue line) and the GWR (red line) methods at four representative stations in 2010. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a positive relationship (Kim et al., 2021; Shen et al., 2020). However, the GWR model cannot consistently preserve the positive relationship between T_a and LST. Specifically, coefficients of LST are close to zero on some days, reflecting that the effect of LST on T_a is weak compared with that of elevation, possibly due to the correlations between LST and T_a at surrounding stations are not strong on these days.

The SVCMM-SP model can consistently preserve the negative relationship between T_a and elevation and the positive relationship between T_a and LST across space (Figs. 9 and 10). As shown in the example day, coefficients for elevation are consistently negative in estimating T_{max} and T_{min} (Fig. 9). The largest absolute values for coefficients of elevation, which indicates the strongest impacts of elevation on T_a , mainly occurred in south China, Tibetan Plateau, and Tianshan Mountains, which are mainly covered with mountains. It means that elevation is one of the key factors influencing the estimated T_a using the SVCMM-SP model in mountainous regions. Coefficients for LST are consistently positive in estimating T_{max} and T_{min} (Fig. 9). The largest values for coefficients of LST, which indicate the strongest impacts of LST on T_a , mainly occurred in Eastern China (for T_{max}) and western Gobi deserts (for T_{max} and T_{min}), where the change of elevation is small. It means that the LST is the main factor influencing the estimated T_a using the SVCMM-SP method

for these regions without the large effects of elevation. On the contrary, the GWR model cannot consistently and spatially preserve the correct relationship between T_a and explanatory variables (i.e., elevation and LST), as reported in the literature (Chen et al., 2015a). The consistently and correctly preserved relationships between T_a with LST and elevation using the SVCMM-SP method are shown in detail in representative cities (Fig. 10). However, in the GWR method, the β_{LST} values are negative in Nanjing, Wuhan, and Hefei, contributing to opposite spatial patterns between estimated T_a (cold island) and LST (heat island) in urban areas (Section 4.4).

4.3. Computing efficiency

The computing efficiency of the SVCMM-SP is significantly higher compared to the GWR model (Table 3). We compared the computing time of the SVCMM-SP and GWR methods in a test region (360,000 grids of 1 km resolution) using R codes. We also tested the GWR module in ArcGIS 10.7 software for the computing time of GWR in mainland China (7.65 MODIS tiles, each tile contains 1200×1200 grids of 1 km resolution). The computing time using R codes in the test region was converted for mainland China based on the area proportion between the two

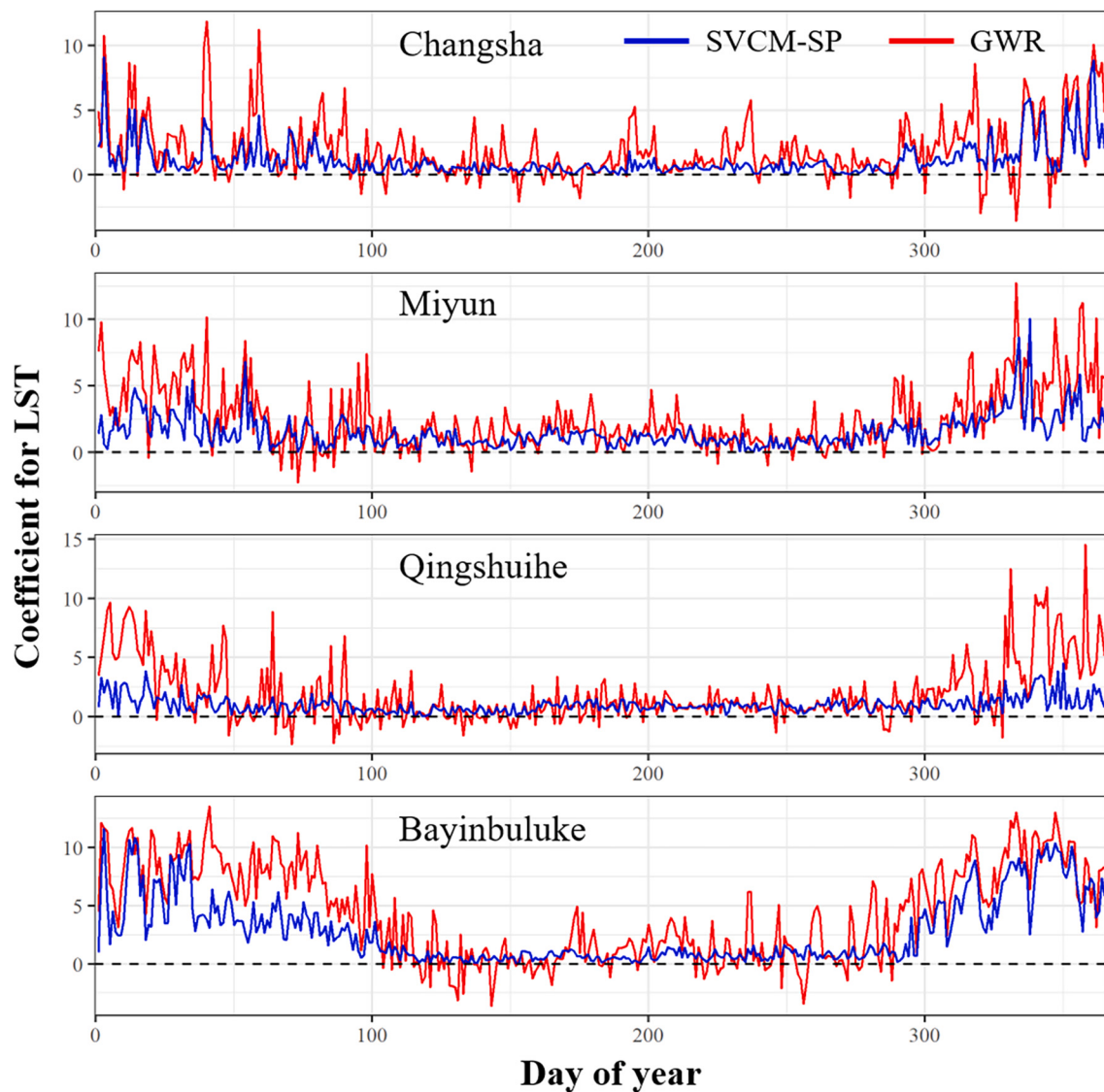


Fig. 8. Coefficients of LST in estimating T_{max} using the SVCM-SP (blue line) and the GWR (red line) methods, for the four stations in 2010. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

regions (i.e., 1:30.6). Results illustrate that the SVCM-SP method needed 5.6 min and 20.4 h for estimating T_a for one day and one year (365 days), respectively, for mainland China, while the GWR (R codes) needed 8.3 h and 126.8 days, about 89 and 149 times slower than that of the SVCM-SP method. The GWR module in ArcGIS 10.7 needed 79.1 min and 20.0 days for estimating T_a for one day and one year, respectively, about 14 and 23.5 times slower than that of the SVCM-SP method. The improvement of the computing efficiency for the SVCM-SP method is larger for the whole year compared to a single day because the computing time for the GWR method linearly increases with the number of days. The main reason for the high computing efficiency of the SVCM-SP method is that its parameters can be pre-calculated only once for a specific region. On the contrary, the parameters of the GWR model need to be calculated for each day (e.g., the best bandwidth) and each grid (e.g., the distances between the target grid and its neighboring stations within a specific best bandwidth), leading to a time-intensive computing process.

4.4. Spatial pattern of air temperature

The estimated air temperature data show significant spatial

variations in mainland China (Fig. 11). Taking estimated T_a in one Summer day as an example, both T_{max} and T_{min} decrease from southeast to northwest region, i.e., the characteristics of air temperature change with latitude and elevation (Fig. 11). T_a in the Tibetan Plateau and the Tianshan Mountains are the lowest because the average elevation of these regions exceeds 4000 m (Chen et al., 2015a; Li and Zha, 2018; Yao et al., 2020). T_a in the northwestern China (e.g., Tarim Basin) shows the highest values since these regions are covered with Gobi deserts (Yao et al., 2020).

The estimated T_a using the SVCM-SP method shows a reasonable spatial pattern in mountainous regions, while there are issues with the GWR model (Fig. 12). Results of an example day in eastern Tibetan (mountainous region) illustrated that the estimated T_a using the SVCM-SP method were higher in the valleys than those on the surrounding hillsides, but an opposite spatial pattern was found in part of the example region in the estimate T_a using the GWR model (Fig. 12). Such abnormal spatial pattern using the GWR model tends to occur in the regions such as mountains with complex topography changes and sparsely distributed ground stations (Chen et al., 2015a), highlighting the difficulty of the GWR in capturing the accurate spatial variations of T_a in such regions also without adequate weather stations. A reliable

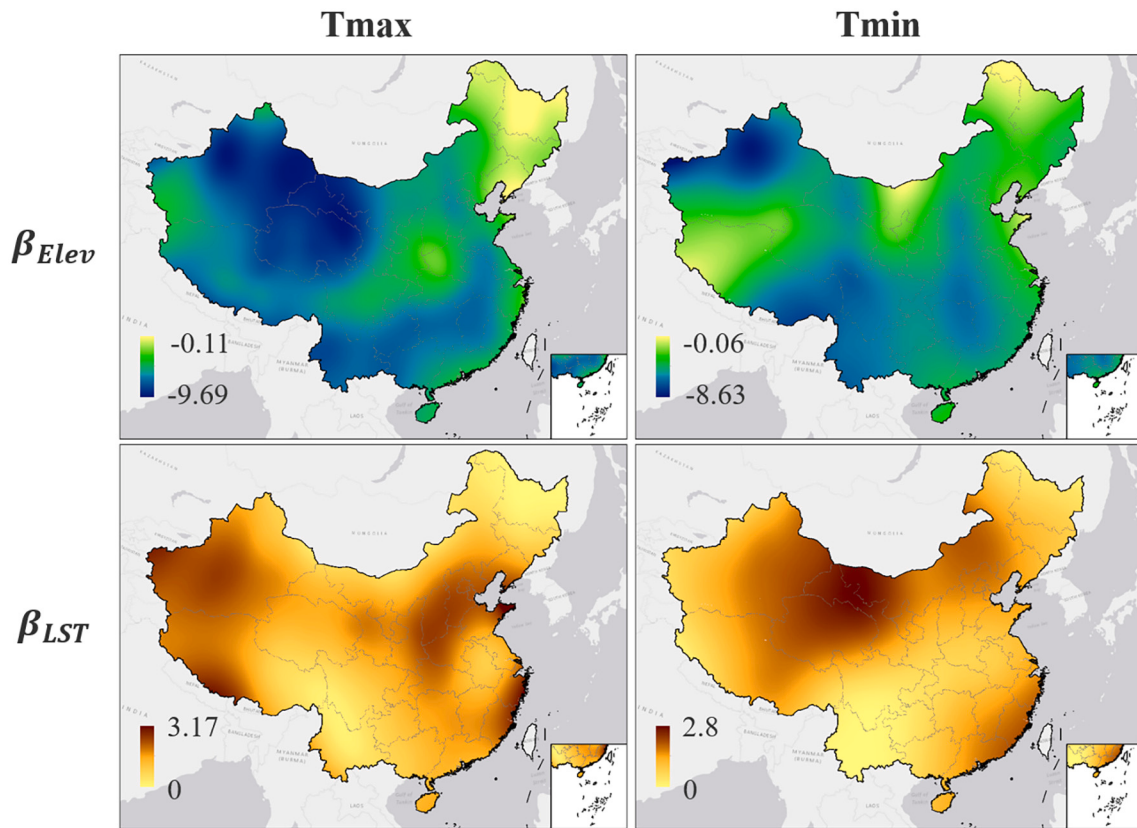


Fig. 9. Spatial patterns for standardized regression coefficients of elevation (β_{elev}) and LST (β_{LST}) in mainland China in an example day of 200 in 2010.

method such as the SVCMM-SP can mitigate such limitations in estimating Ta in mountainous regions.

The estimated Ta using the SVCMM-SP method can well represent urban heat island (UHI) effect (i.e., higher temperature in urban region than that of the surrounding rural area) (Li et al., 2019; Zhou et al., 2019b), while the estimated Ta using the GWR method shows a different pattern in some cases (Fig. 13). The estimated Ta using the GWR method shows UHI effect in Beijing, Shijiazhuang, and Zhengzhou, but it is lower in urban areas than those of the surrounding rural areas in Nanjing, Wuhan, and Hefei in the example day, although they are located at plains with small elevation variations. These abnormal spatial patterns of estimated Ta using the GWR method might lead to uncertainties for quantifying canopy UHI intensity, even though they may only occur on specific days in a few of the cities. Overall, the estimated Ta using the SVCMM-SP method can provide more reasonable Ta data for urban thermal environment studies compared to the GWR model due to the reason discussed in Section 4.2, i.e., the former has the advantage of consistently and correctly preserved relationships between Ta with LST and elevation (expressed by signs of coefficients for LST and elevation in Eq. (2)) compare to the latter (Fig. 10). This is consistent with the quantitative results in Table 2, i.e., higher accuracies in urban regions using the SVCMM-SP than those of the GWR model.

The above-mentioned issues in the estimated Ta using the GWR model are possibly caused by the limited spatial coverage of weather stations (i.e., mainly in plains or river valleys, and sparse in mountainous regions), leading to abnormal relationships between Ta and explanatory variables. Although Ta is positively correlated with LST (Kim et al., 2021; Shen et al., 2020) and negatively correlated with elevation (Heynen et al., 2016; Zhang et al., 2021a; Zhu et al., 2017), these relationships can only preserve if other factors such as latitudes do not have large impacts on Ta in neighboring weather stations (e.g., winds, clouds, snow, and land cover types) (Cai et al., 2017; Good, 2016). In Fig. 12, Ta from weather stations at plains or valleys show an

increasing trend with elevation and this trend was applied to sites without weather stations using the GWR model. Similarly, in Fig. 13, the abnormal spatial patterns of Ta in Nanjing, Wuhan, and Hefei are mainly caused by the negative relationship between Ta and LST estimated using the GWR model. However, the SVCMM-SP algorithm does not have this issue by borrowing features from other sites, i.e., implementing the penalized least squares method with smoothness and inequality constraints based on the sign preservation (Kim et al., 2021).

The spatial pattern of Ta using the SVCMM-SP and GWR models might be different at the local scales in some limited days (Figs. 12 and 13) but did not lead to large difference in accuracy (Table 1), for the following possible two reasons. First, the accuracy in Table 1 is an annual average from daily accuracies for the entire region (mainland China), while these abnormal spatial patterns of Ta due to abnormal signs of coefficients for explanatory variables using the GWR only account for a small ratio (Table E1). Second, the weather stations we used cannot capture all the spatial variations of estimated Ta at the local scale because they are mainly located in flat regions such as plains and river basins, and a limited number of weather stations were used for accuracy assessment. For example, the results in Fig. 12 generated from the GWR model show that “Ta at higher altitudes are higher than those at lower altitudes” on mountain slopes and cannot be captured by weather stations (Fig. 12). Therefore, the spatial variations of Ta were corrected but the reported accuracies of estimated Ta were not largely improved using the SVCMM-SP compared to the GWR model (Fig. 4). That is also the reason to show the advantages of the SVCMM-SP by illustrating the spatial differences of estimated Ta using the two models in Figs. 12 and 13.

Our results using the GWR model are consistent with Chen et al. (2015a), as indicated in their Fig. 7, showing the coefficients of elevation “Coefficients b(u)” for the GWR are positive mainly in the northwest of mainland China (including the Qilian mountains and Tianshan Mountains), indicating a positive relationship between Ta and elevation. Unfortunately, Chen et al. (2015a) did not show the estimated Ta of

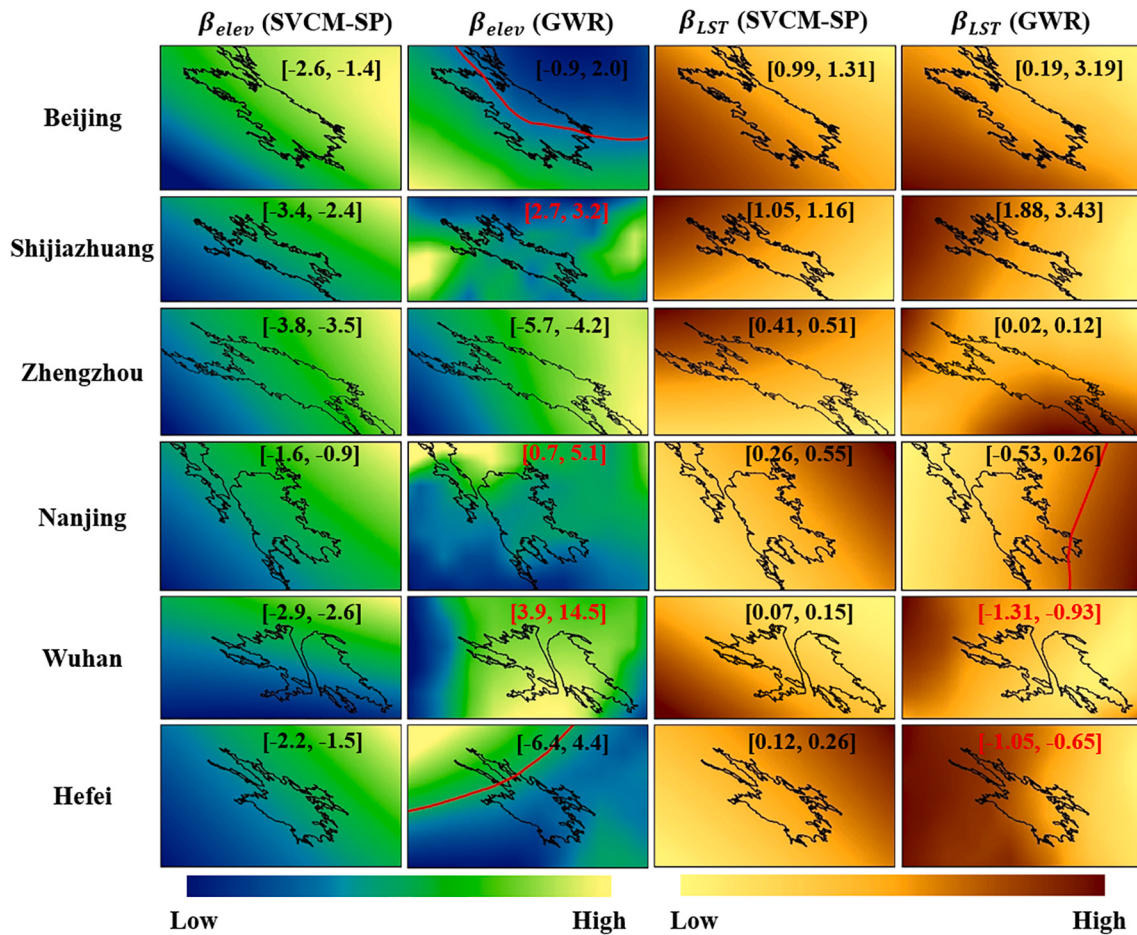


Fig. 10. Spatial pattern for standardized regression coefficients of elevation (β_{elev}) and LST (β_{LST}) for estimating T_{max} using the SVCMM-SP and GWR methods in representative cities on the day of 194 in 2010. Red lines are boundaries between regions with positive and negative values. Black lines are boundaries of urban regions extracted from Li et al. (2020). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Estimated computing time for estimating T_a for the entire area of mainland China (7.65 MODIS tiles) using the SVCMM-SP and GWR methods.

	SVCMM-SP (R code)	GWR (R code)	GWR (ArcGIS 10.7)
One day	336 s (5.6 min)	29,990 s (8.3 h)	4744 s (79.1 min)
One year	73,657 s (20.4 h)	10,946,514 s (126.8 days)	1,731,560 s (20.0 days)

Note: Tested on Intel® Core™ i7-9850H CPU @ 2.60GHz 2.59 GHz, 64bit computer with RAM of 32.0 GB.

specific days at the local scale, and the monthly T_a across mainland China in Fig. 2 of their paper did not show the issue of the GWR that was found in this study.

4.5. Potential applications of the SVCMM-SP algorithm

The SVCMM-SP method offers great potentials in relevant applications of existing statistical methods. Taking the GWR method as an example, it is also widely used to estimate gridded variables such as PM10 (You et al., 2015), PM2.5 (Hu et al., 2013; Ma et al., 2014; Yang et al., 2020; Zhang and Kondragunta, 2021), and population density (Wang et al., 2018), to spatially downscale remote sensing products such as precipitation (Chen et al., 2015b; Xu et al., 2015), soil moisture (Wen et al., 2020), and LST (Luo et al., 2021), and to explore relationships between influential factors of items such as land cover changes (Pineda Jaimes

et al., 2010; Ren et al., 2020), PM2.5 (Liu et al., 2020; Wang and Wang, 2020; Zhou et al., 2019a), and COVID-19 incidence (Han et al., 2021; Mollalo et al., 2020). The physical unrealistic relationships between response variables and their explanatory variables using the GWR model have also been found in these applications such as the spatial downscaling of precipitation (Xu et al., 2015; Zhang et al., 2018b) and PM2.5 estimation (Yang et al., 2020; Zhang and Kondragunta, 2021). Taking spatial downscaling of precipitation as an example, precipitation at the hill slope increases with elevation until reaching a maximum precipitation altitude (Barry, 2008; Zhang et al., 2018b); however, the commonly used statistical models such as the GWR and the Cubist methods cannot well capture this relationship below the maximum precipitation altitude (Ma et al., 2017; Xu et al., 2015; Zhang et al., 2018b). The SVCMM-SP method can possibly be used for spatial downscaling of precipitation to obtain precipitation data with a high spatial resolution (e.g., 1 km) below the maximum precipitation altitude. Similarly, the SVCMM-SP method can also be used to obtain better estimations of variables such as PM2.5 and more accurate relationships for analyzing driving factors of variables such as land cover changes.

4.6. Uncertainties and future work

Despite the improved accuracies in estimating T_a using the SVCMM-SP method, there are still uncertainties in the estimated T_a , depending on the availability of station observations, the accuracy of the gap-filled LST, and the relationship between T_a with elevation and LST. First, the estimated T_a in the Western China, especially in the mountainous

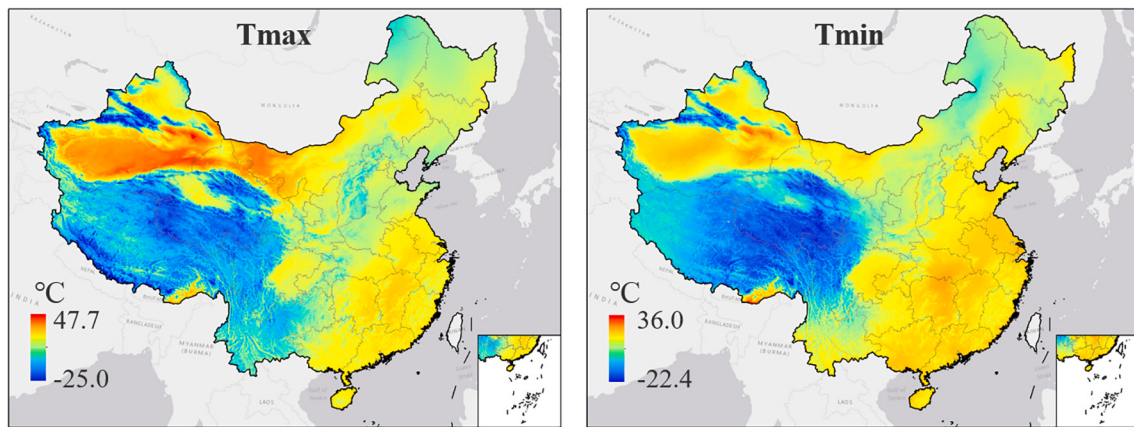


Fig. 11. Spatial pattern of estimated Ta in mainland China in an example day of 200 in 2010.

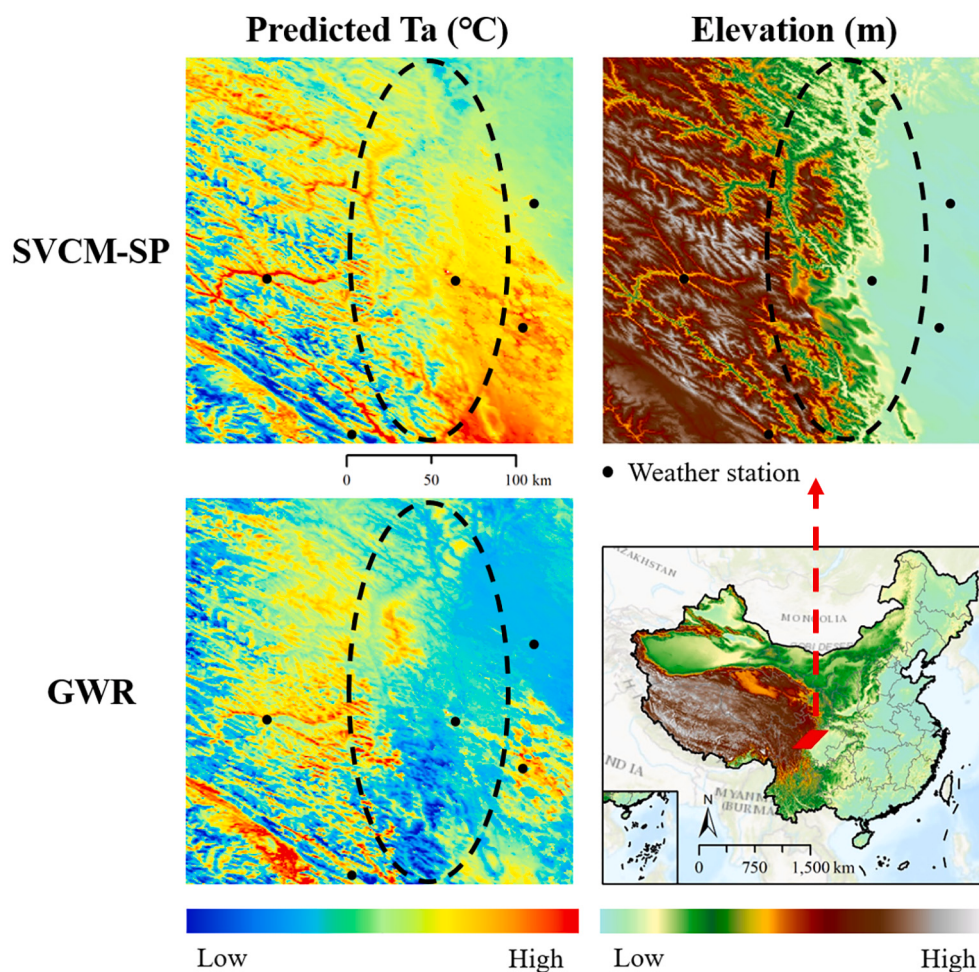


Fig. 12. Spatial pattern of maximum Ta in mountainous areas on the day of 004 in 2010 using the SVCN-SP and GWR method. The projections of elevation and Ta data are same as the MODIS projection. The projection of mainland China (right bottom figure) is the China Lambert Conformal Conic projection.

areas, shows lower accuracy because of the complex topography and sparsely distributed weather stations. The fitted Ta surface might not capture adequately the complex change of Ta in mountainous regions of the Western China due to the limitation of the sparse weather stations. Using other gridded explanatory variables with proper consideration of their relationships with Ta might further improve the performance of the SVCN-SP method (possibly be further improved) in these areas. Second, the uncertainties in the gap-filled LST values, especially in the largely

missing areas of the original MODIS LST data, could be propagated to the estimated Ta, especially for Tmax. Although the estimated Ta using the SVCN-SP method has already achieved reasonably good accuracy in this study, further improvements in gap-filled LST data may also improve the accuracy of the Ta data. Third, the estimated Ta might depend heavily on the LST itself, leading to highly similar spatial patterns as LST. However, other factors such as clouds, winds, snow, and land cover types could also affect Ta (Cai et al., 2017; Good, 2016). The

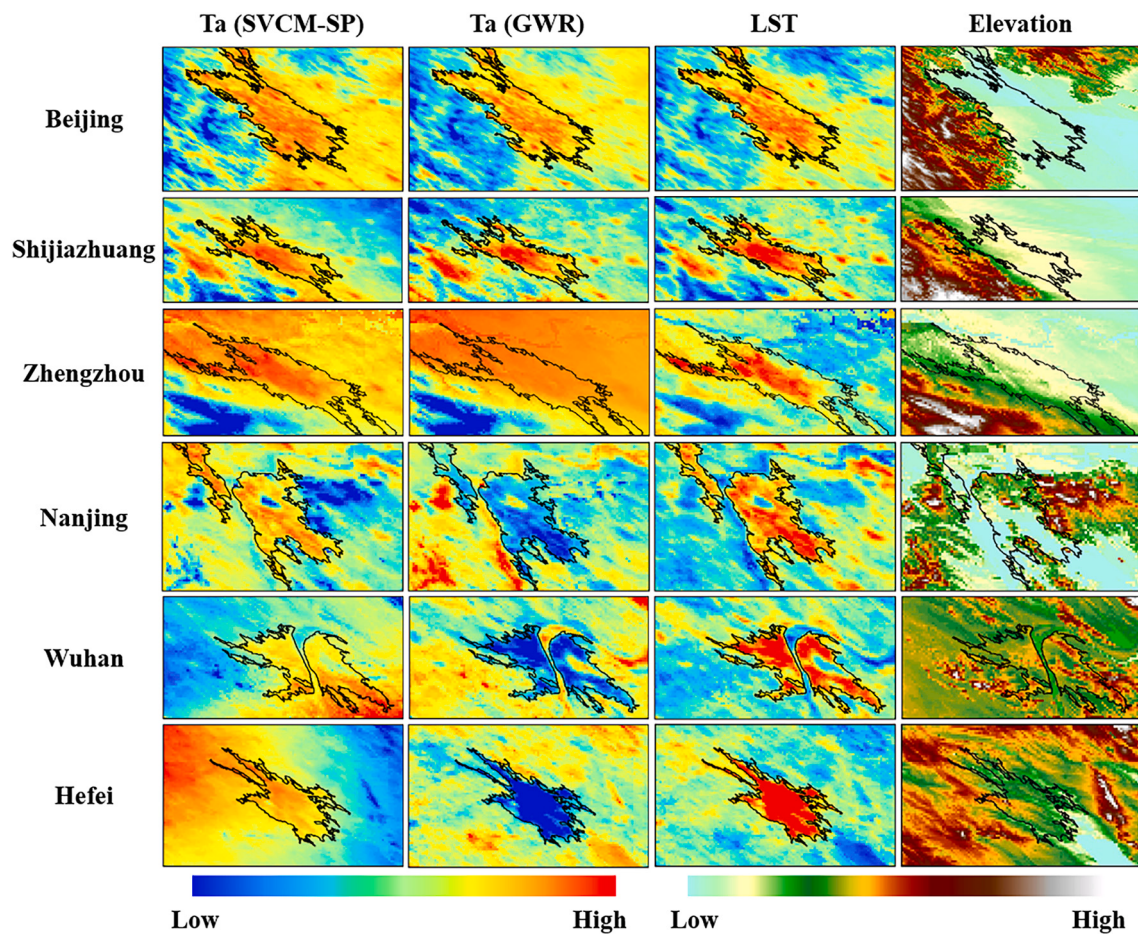


Fig. 13. Spatial pattern of estimated maximum Ta using the SVCMM-SP and GWR methods, and corresponding LST and elevation data in representative cities on the day of 194 in 2010. Black lines are the boundaries of urban regions extracted from Li et al. (2020).

limited consideration of these factors might lead to some biases in the spatial pattern of estimated Ta compared to the actual Ta.

This study opens several promising research avenues for the future. For example, based on the good performance and computing efficiency of the SVCMM-SP method in mainland China, a global 1 km gridded Ta dataset is of great use for broader applications. Another promising area of research is the application of the SVCMM-SP method for estimating and producing other gridded geophysical variable datasets. Moreover, using more gridded explanatory variables with proper consideration of their relationships with Ta may further improve the performance of the SVCMM-SP method.

5. Conclusions

A key limitation of existing statistical algorithms (e.g., Geographically Weighted Regression (GWR)) for estimating air temperature (Ta) is that they cannot always correctly capture and preserve relationships between Ta and explanatory variables due to limited spatial coverages of observations, possibly leading to abnormal relationships between Ta and explanatory variables, especially in largely heterogeneous areas. Accordingly, the estimated Ta with abnormal spatial patterns may provide misleading information in relevant studies such as quantifying urban thermal environment and hydrological modeling in mountains at specific times. In this study, we introduced the Spatially Varying Coefficient Models with Sign Preservation (SVCMM-SP) method to assess relationships between the air temperature and explanatory variables (i.e., negative for elevation and positive for LST), for estimating Ta and developing 1 km gridded Ta product in mainland China. The choice of

the study area was based on the availability of a comprehensive network of weather stations, and diverse geographical landscape features (i.e., desert and mountains) for evaluating the accuracy and robustness of the SVCMM-SP method. Using this method, we estimated the 1 km gridded daily maximum and minimum Ta in mainland China from 2003 to 2016. The estimated Ta can preserve correctly and consistently the negative relationship between Ta and elevation, and the positive relationship between Ta and LST, which are represented with negative coefficients for elevation and positive coefficients for LST, respectively, across space and time. The SVCMM-SP method showed better performance than the GWR method based on several evaluation criteria including accuracy, signs of coefficients for explanatory variables, and computing efficiency. The estimated Ta using the SVCMM-SP can serve as a key data source for quantifying urban thermal environment and it offers potential for improving studies such as hydrological modeling in mountainous regions. The future work will focus on creating a 1 km global gridded Ta dataset using the SVCMM-SP method for broader applications. Future studies on the improvement of the GWR model by preserving signs of coefficients can be also useful for relevant applications since GWR is a simple and widely used model.

CRediT authorship contribution statement

Tao Zhang: Methodology, Software, Validation, Investigation, Writing – original draft. **Yuyu Zhou:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Li Wang:** Writing – review & editing. **Kaiguang Zhao:** Writing – review & editing. **Zhengyuan Zhu:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Science Foundation, United States (CBET- 2041859, DMS-1916204, and DMS-2135493) and the College of Liberal Arts and Science's (LAS) Dean's Emerging Faculty Leaders award at the Iowa State University, United States. We would also like to thank the three anonymous reviewers for their valuable comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2022.113072>.

References

- Aggarwal, A., Mehta, S., Gupta, D., Sheikh, S., Pallagatti, S., Singh, R., Singla, I., 2012. Clinical & immunological erythematosus patients characteristics in systemic lupus Maryam. *J. Dent. Educ.* 76, 1532–1539. <https://doi.org/10.4103/jimr.IJMR>.
- Barry, R.G., 2008. Mountain Weather and Climate, Third edition, pp. 1–506. <https://doi.org/10.1017/CBO9780511754753>. Mt. Weather Clim. Third Ed. 9780521862.
- Becker, J.J., Sandwell, D.T., Smith, W.H.F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., von Rosenberg, J., Wallace, G., Weatherall, P., 2009. Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30 PLUS. *Mar. Geod.* 32, 355–371. <https://doi.org/10.1080/01490410903297766>.
- Benali, A., Carvalho, A.C., Nunes, J.P., Carvalhais, N., Santos, A., 2012. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sens. Environ.* 124, 108–121. <https://doi.org/10.1016/j.rse.2012.04.024>.
- Bisht, G., Bras, R.L., 2010. Estimation of net radiation from the MODIS data under all sky conditions: southern Great Plains case study. *Remote Sens. Environ.* 114, 1522–1534. <https://doi.org/10.1016/j.rse.2010.02.007>.
- Cai, Y., Chen, G., Wang, Y., Yang, L., 2017. Impacts of land cover and seasonal variation on maximum air temperature estimation using MODIS imagery. *Remote Sens.* 9. <https://doi.org/10.3390/rs9030233>.
- Chai, H., Cheng, W., Zhou, C., Chen, X., Ma, X., Zhao, S., 2011. Analysis and comparison of spatial interpolation methods for temperature data in Xinjiang Uygur Autonomous Region, China. *Nat. Sci.* 03, 999–1010. <https://doi.org/10.4236/ns.2011.312125>.
- Chen, F., Liu, Y., Liu, Q., Qin, F., 2015a. A statistical method based on remote sensing for the estimation of air temperature in China. *Int. J. Climatol.* 35, 2131–2143. <https://doi.org/10.1002/joc.4113>.
- Chen, C., Zhao, S., Duan, Z., Qin, Z., 2015b. An improved spatial downscaling procedure for TRMM 3B43 precipitation product using geographically weighted regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 4592–4604. <https://doi.org/10.1109/JSTARS.2015.2441734>.
- Chen, Y., Quan, J., Zhan, W., Guo, Z., 2016. Enhanced statistical estimation of air temperature incorporating nighttime light data. *Remote Sens.* 8, 1–23. <https://doi.org/10.3390/rs8080656>.
- Chen, Y., Liang, S., Ma, H., Li, B., He, T., Wang, Q., 2021. An all-sky 1km daily land surface air temperature product over mainland China for 2003–2019 from MODIS and ancillary data. *Earth Syst. Sci. Data* 13, 4241–4261. <https://doi.org/10.5194/essd-13-4241-2021>.
- Connor, S.J., Flasse, S.P., Thomson, M.C., Perryman, A.H., 1998. Environmental information systems in malaria risk mapping and epidemic forecasting. *Disasters* 22, 39–56. <https://doi.org/10.1111/1467-7717.00074>.
- Cristóbal, J., Ninyerola, M., Pons, X., 2008. Modeling air temperature through a combination of remote sensing and GIS data. *J. Geophys. Res. Atmos.* 113, 1–13. <https://doi.org/10.1029/2007JD009318>.
- Danielson, J.J., Gesch, D.B., 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). Open-File Rep. <https://doi.org/10.3133/ofr20111073>.
- Dimarzio, J., Brenner, A., Fricker, H., Schutz, R., Shuman, C., Zwally, H., 2007. GLAS/ICESat 500m Laser Altimetry Digital Elevation Model of Antarctica, Version 1 [WWW Document]. URL. <https://nsidc.org/data/NSIDC-0304/versions/1>.
- Dodson, R., Marks, D., 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Clim. Res.* 8, 1–20. <https://doi.org/10.3354/cr008001>.
- Fang, S., Mao, K., Xia, X., Wang, P., Shi, J., Batani, S., Xu, T., Cao, M., Heggy, E., 2021. A dataset of daily near-surface air temperature in China from 1979 to 2018. *Earth Syst. Sci. Data Discuss.* 1–37. <https://doi.org/10.5194/essd-2021-309>.
- Good, E.J., 2016. An in situ-based analysis of the relationship between land surface “skin” and screen-level air temperatures. *J. Geophys. Res. Atmos.* 121, 8801–8819. <https://doi.org/10.1002/2016JD025318>.
- Goward, S.N., Waring, R.H., 1994. In: Goward, Samuel N., Waring, Richard H., Dye, Dennis G., Yang, Jingli (Eds.), *Ecological Remote Sensing at OTTER: Satellite Macroscale Observations*, 4. Published by: Ecological Society of America Stable, pp. 322–343. URL: <http://www.jstor.org/stable/1941937>. EC. Ecol. Appl.
- Han, Y., Yang, L., Jia, K., Li, J., Feng, S., Chen, W., Zhao, W., Pereira, P., 2021. Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. *Sci. Total Environ.* 761, 144257. <https://doi.org/10.1016/j.scitotenv.2020.144257>.
- Hengl, T., Heuvelink, G.B.M., Tadić, M.P., Pebesma, E.J., 2012. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theor. Appl. Climatol.* 107, 265–277. <https://doi.org/10.1007/s00704-011-0464-2>.
- Hennig, T.A., Kretsch, J.L., Pessagno, C.J., Salamonowicz, P.H., Stein, W.L., 2001. The shuttle radar topography mission. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 2181, 65–77. https://doi.org/10.1007/3-540-44818-7_11.
- Heynen, M., Miles, E., Ragettli, S., Buri, P., Immerzeel, W.W., Pellicciotti, F., 2016. Air temperature variability in a high-elevation Himalayan catchment. *Ann. Glaciol.* 57, 212–222. <https://doi.org/10.3189/2016AoG71A076>.
- Hou, P., Chen, Y., Qiao, W., Cao, G., Jiang, W., Li, J., 2013. Near-surface air temperature retrieval from satellite images and influence by wetlands in urban region. *Theor. Appl. Climatol.* 111, 109–118. <https://doi.org/10.1007/s00704-012-0629-7>.
- Hrisko, J., Ramamurthy, P., Yu, Y., Yu, P., Melecio-Vázquez, D., 2020. Urban air temperature model using GOES-16 LST and a diurnal regressive neural network algorithm. *Remote Sens. Environ.* 237, 111495. <https://doi.org/10.1016/j.rse.2019.111495>.
- Hu, X., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M., Quattrochi, D.A., Sarnat, J.A., Liu, Y., 2013. Estimating ground-level PM2.5 concentrations in the southeastern U.S. using geographically weighted regression. *Environ. Res.* 121, 1–10. <https://doi.org/10.1016/j.envres.2012.11.003>.
- Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T.F., Peng, S., Berry, J.A., Wang, K., Mao, J., Alkama, R., Cescatti, A., Cuntz, M., De Deurwaerder, H., Gao, M., He, Y., Liu, Y., Luo, Y., Myneni, R.B., Niu, S., Shi, X., Yuan, W., Verbeeck, H., Wang, T., Wu, J., Janssens, I.A., 2019. Air temperature optima of vegetation productivity across global biomes. *Nat. Ecol. Evol.* 3, 772–779. <https://doi.org/10.1038/s41559-019-0838-x>.
- Jocik, A.A.M., 2004. *Estimate Ambient Air Temperature at Regional Level Using Remote Sensing Techniques*. International Institute for Geo-Information Sci- ence and Earth Observation (ITC), The Netherlands.
- Kim, M., Wang, L., Zhou, Y., 2021. Spatially varying coefficient models with sign preservation of the coefficient functions. *J. Agric. Biol. Environ. Stat.* 26, 367–386. <https://doi.org/10.1007/s13253-021-00443-5>.
- Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2014. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the northeastern USA. *Remote Sens. Environ.* 150, 132–139. <https://doi.org/10.1016/j.rse.2014.04.024>.
- Lai, M.J., Wang, L., 2013. Bivariate penalized splines for regression. *Stat. Sin.* 23, 1399–1417. <https://doi.org/10.5705/ss.2010.278>.
- Lamchin, M., Lee, W.K., Jeon, S.W., Wang, S.W., Lim, C.H., Song, C., Sung, M., 2018. Long-term trend and correlation between vegetation greenness and climate variables in Asia based on satellite data. *Sci. Total Environ.* 618, 1089–1095. <https://doi.org/10.1016/j.scitotenv.2017.09.145>.
- Lan, L., Lian, Z., Pan, L., 2010. The effects of air temperature on office workers' well-being, workload and productivity-evaluated with subjective ratings. *Appl. Ergon.* 42, 29–36. <https://doi.org/10.1016/j.apergo.2010.04.003>.
- Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. *Ecol. Inform.* 6, 228–241. <https://doi.org/10.1016/j.ecoinf.2010.12.003>.
- Li, L., Zha, Y., 2018. Mapping relative humidity, average and extreme temperature in hot summer over China. *Sci. Total Environ.* 615, 875–881. <https://doi.org/10.1016/j.scitotenv.2017.10.022>.
- Li, L., Zha, Y., 2019. Estimating monthly average temperature by remote sensing in China. *Adv. Sp. Res.* 63, 2345–2357. <https://doi.org/10.1016/j.asr.2018.12.039>.
- Li, X., Zhou, Y., Asrar, G.R., Zhu, Z., 2018. Developing a 1 km resolution daily air temperature dataset for urban and surrounding areas in the conterminous United States. *Remote Sens. Environ.* 215, 74–84. <https://doi.org/10.1016/j.rse.2018.05.034>.
- Li, X., Zhou, Y., Yu, S., Jia, G., Li, H., Li, W., 2019. Urban heat island impacts on building energy consumption: a review of approaches and findings. *Energy* 174, 407–419. <https://doi.org/10.1016/j.energy.2019.02.183>.
- Li, X., Gong, P., Zhou, Y., Wang, J., Bai, Y., Chen, B., Hu, T., Xiao, Y., Xu, B., Yang, J., Liu, X., Cai, W., Huang, H., Wu, T., Wang, X., Lin, P., Li, Xun, Chen, J., He, C., Li, Xia, Yu, L., Clinton, N., Zhu, Z., 2020. Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environ. Res. Lett.* 15, 094044. <https://doi.org/10.1088/1748-9326/ab9be3>.
- Lin, S., Moore, N.J., Messina, J.P., DeVisser, M.H., Wu, J., 2012. Evaluation of estimating daily maximum and minimum air temperature with MODIS data in East Africa. *Int. J. Appl. Earth Obs. Geoinf.* 18, 128–140. <https://doi.org/10.1016/j.jag.2012.01.004>.
- Liu, Q., Wu, R., Zhang, W., Li, W., Wang, S., 2020. The varying driving forces of PM2.5 concentrations in Chinese cities: insights from a geographically and temporally weighted regression model. *Environ. Int.* 145, 106168. <https://doi.org/10.1016/j.envint.2020.106168>.
- Lowen, A.C., Mubareka, S., Steel, J., Palese, P., 2007. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* 3, 1470–1476. <https://doi.org/10.1371/journal.ppat.0030151>.

- Luo, X., Chen, Y., Wang, Z., Li, H., Peng, Y., 2021. Spatial downscaling of MODIS land surface temperature based on a geographically and temporally weighted autoregressive model. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 7637–7653. <https://doi.org/10.1109/JSTARS.2021.3094184>.
- Ma, Z., Hu, X., Huang, L., Bi, J., Liu, Y., 2014. Estimating ground-level PM_{2.5} in China using satellite remote sensing. *Environ. Sci. Technol.* 48, 7436–7444. <https://doi.org/10.1021/es5009399>.
- Ma, Z., Shi, Z., Zhou, Y., Xu, J., Yu, W., Yang, Y., 2017. A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed. *Remote Sens. Environ.* 200, 378–395. <https://doi.org/10.1016/j.rse.2017.08.023>.
- Mollalo, A., Vahedi, B., Rivera, K.M., 2020. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci. Total Environ.* 728, 138884. <https://doi.org/10.1016/j.scitotenv.2020.138884>.
- Mostovoy, G.V., King, R.L., Reddy, K.R., Kakani, V.G., Filippova, M.G., 2006. Statistical estimation of daily maximum and minimum air temperatures from MODIS LST data over the State of Mississippi. *GISci. Remote Sens.* 43, 78–110. <https://doi.org/10.2747/1548-1603.43.1.78>.
- Mu, J., Wang, G., Wang, L., 2018. Estimation and inference in spatially varying coefficient models. *Environmetrics* 29, 1–19. <https://doi.org/10.1002/env.2485>.
- Müller, M., Batrak, Y., Kristiansen, J., Koltzow, M.A.O., Noer, G., Korosov, A., 2017. Characteristics of a convective-scale weather forecasting system for the European Arctic. *Mon. Weather Rev.* 145, 4771–4787. <https://doi.org/10.1175/MWR-D-17-0194.1>.
- Nemani, R.R., Running, S.W., 1989. Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *J. Appl. Meteorol. Climatol.* 28, 276–284. [https://doi.org/10.1175/1520-0450\(1989\)028%3C0276:EOBSRT%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1989)028%3C0276:EOBSRT%3E2.0.CO;2).
- Noi, P.T., Degener, J., Kappas, M., 2017. Comparison of multiple linear regression, cubist regression, and random forest algorithms to estimate daily air surface temperature from dynamic combinations of MODIS LST data. *Remote Sens.* 9. <https://doi.org/10.3390/rs9050398>.
- Oyler, J.W., Ballantyne, A., Jencso, K., Sweet, M., Running, S.W., 2015. Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. *Int. J. Climatol.* 35, 2258–2279. <https://doi.org/10.1002/joc.4127>.
- Petrova, V.N., Russell, C.A., 2018. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* 16, 47–60. <https://doi.org/10.1038/nrmicro.2017.118>.
- Pineda Jaimes, N.B., Bosque Sendra, J., Gómez Delgado, M., Franco Plata, R., 2010. Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. *Appl. Geogr.* 30, 576–591. <https://doi.org/10.1016/j.apgeog.2010.05.004>.
- Rao, Y., Liang, S., Wang, D., Yu, Y., Song, Z., Zhou, Y., Shen, M., Xu, B., 2019. Estimating daily average surface air temperature using satellite land surface temperature and top-of-atmosphere radiation products over the Tibetan Plateau. *Remote Sens. Environ.* 234, 111462. <https://doi.org/10.1016/j.rse.2019.111462>.
- Ren, S., Qin, Q., Ren, H., 2019. Contrasting wheat phenological responses to climate change in global scale. *Sci. Total Environ.* 665, 620–631. <https://doi.org/10.1016/j.scitotenv.2019.01.394>.
- Ren, Y., Lü, Y., Fu, B., Comber, A., Li, T., Hu, J., 2020. Driving factors of land change in china's loess plateau: quantification using geographically weighted regression and management implications. *Remote Sens.* 12. <https://doi.org/10.3390/rs12030453>.
- Rosen, P.A., 2000. Synthetic aperture radar interferometry. *Proc. IEEE* 88, 333–380. <https://doi.org/10.1109/5.838084>.
- Shen, H., Jiang, Y., Li, T., Cheng, Q., Zeng, C., Zhang, L., 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. *Remote Sens. Environ.* 240, 111692. <https://doi.org/10.1016/j.rse.2020.111692>.
- Shi, Y., Jiang, Z., Dong, L., Shen, S., 2017. Statistical estimation of high-resolution surface air temperature from MODIS over the Yangtze River Delta, China. *J. Meteorol. Res.* 31, 448–454. <https://doi.org/10.1007/s13351-017-6073-y>.
- Smith, W.L., Leslie, L.M., Diak, G.R., Goodman, B.M., Velden, C.S., Callan, G.M., Raymond, W., Wade, G.S., 1988. The integration of meteorological satellite imagery and numerical dynamical forecast models. *Philos. Trans. - R. Soc. Lond. A* 324, 317–323. <https://doi.org/10.1098/rsta.1988.0022>.
- Stahl, K., Moore, R.D., Floyer, J.A., Asplin, M.G., McKendry, I.G., 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, 224–236. <https://doi.org/10.1016/j.agrformet.2006.07.004>.
- Sun, Y.J., Wang, J.F., Zhang, R.H., Gillies, R.R., Xue, Y., Bo, Y.C., 2005. Air temperature retrieval from remote sensing data based on thermodynamics. *Theor. Appl. Climatol.* 80, 37–48. <https://doi.org/10.1007/s00704-004-0079-y>.
- Vancutsem, C., Ceccato, P., Dinku, T., Connor, S.J., 2010. Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sens. Environ.* 114, 449–465. <https://doi.org/10.1016/j.rse.2009.10.002>.
- Wang, M., Wang, H., 2020. Spatial distribution patterns and influencing factors of PM_{2.5} pollution in the Yangtze River Delta: empirical analysis based on a GWR model. *Asia-Pacific J. Atmos. Sci.* <https://doi.org/10.1007/s13143-019-00153-6>.
- Wang, L., Koike, T., Yang, K., Yeh, P.J.F., 2009. Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin. *J. Hydrol.* 377, 21–34. <https://doi.org/10.1016/j.jhydrol.2009.08.005>.
- Wang, L., Wang, S., Zhou, Y., Liu, W., Hou, Y., Zhu, J., Wang, F., 2018. Mapping population density in China between 1990 and 2010 using remote sensing. *Remote Sens. Environ.* 210, 269–281. <https://doi.org/10.1016/j.rse.2018.03.007>.
- Wen, F., Zhao, W., Wang, Q., Sanchez, N., 2020. A value-consistent method for downscaling SMAP passive soil moisture with MODIS products using self-adaptive window. *IEEE Trans. Geosci. Remote Sens.* 58, 913–924. <https://doi.org/10.1109/TGRS.2019.2941696>.
- Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., Du, M., Liu, M., 2020. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci. Total Environ.* 729, 1–7. <https://doi.org/10.1016/j.scitotenv.2020.139051>.
- Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., Niu, Z., 2015. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* 162, 119–140. <https://doi.org/10.1016/j.rse.2015.02.024>.
- Xu, Y., Knudby, A., Shen, Y., Liu, Y., 2018. Mapping monthly air temperature in the Tibetan plateau from MODIS data based on machine learning methods. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 11, 345–354. <https://doi.org/10.1109/JSTARS.2017.2787191>.
- Yang, Q., Yuan, Q., Yue, L., Li, T., 2020. Investigation of the spatially varying relationships of PM_{2.5} with meteorology, topography, and emissions over China in 2015 by using modified geographically weighted regression. *Environ. Pollut.* 262, 114257. <https://doi.org/10.1016/j.envpol.2020.114257>.
- Yao, R., Wang, L., Huang, X., Li, L., Sun, J., Wu, X., Jiang, W., 2020. Developing a temporally accurate air temperature dataset for Mainland China. *Sci. Total Environ.* 706, 136037. <https://doi.org/10.1016/j.scitotenv.2019.136037>.
- Yoo, C., Im, J., Park, S., Quackenbush, L.J., 2018. Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data. *ISPRS J. Photogramm. Remote Sens.* 137, 149–162. <https://doi.org/10.1016/j.isprsjprs.2018.01.018>.
- You, W., Zang, Z., Zhang, L., Li, Z., Chen, D., Zhang, G., 2015. Estimating ground-level PM₁₀ concentration in northwestern China using geographically weighted regression based on satellite AOD combined with CALIPSO and MODIS fire count. *Remote Sens. Environ.* 168, 276–285. <https://doi.org/10.1016/j.rse.2015.07.020>.
- Zhang, H., Kondragunta, S., 2021. Daily and hourly surface PM_{2.5} estimation from satellite AOD. *Earth Sp. Sci.* 8. <https://doi.org/10.1029/2020EA001599>.
- Zhang, R., Rong, Y., Tian, J., Su, H., Li, Z.L., Liu, S., 2015. A remote sensing method for estimating surface air temperature and surface vapor pressure on a regional scale. *Remote Sens.* 7, 6005–6025. <https://doi.org/10.3390/rs70506005>.
- Zhang, Z., Chang, J., Xu, C.Y., Zhou, Y., Wu, Y., Chen, X., Jiang, S., Duan, Z., 2018a. The response of lake area and vegetation cover variations to climate change over the Qinghai-Tibetan Plateau during the past 30 years. *Sci. Total Environ.* 635, 443–451. <https://doi.org/10.1016/j.scitotenv.2018.04.113>.
- Zhang, T., Li, B., Yuan, Y., Gao, X., Sun, Q., Xu, L., Jiang, Y., 2018b. Spatial downscaling of TRMM precipitation data considering the impacts of macro-geographical factors and local elevation in the Three-River Headwaters Region. *Remote Sens. Environ.* 215, 109–127. <https://doi.org/10.1016/j.rse.2018.06.004>.
- Zhang, F., de Dear, R., Hancock, P., 2019. Effects of moderate thermal environments on cognitive performance: a multidisciplinary review. *Appl. Energy* 236, 760–777. <https://doi.org/10.1016/j.apenergy.2018.12.005>.
- Zhang, W., Zhang, B., Zhu, W., Tang, X., Li, F., Liu, X., Yu, Q., 2021a. Comprehensive assessment of MODIS-derived near-surface air temperature using wide elevation-spanned measurements in China. *Sci. Total Environ.* 800, 149535. <https://doi.org/10.1016/j.scitotenv.2021.149535>.
- Zhang, T., Zhou, Y., Zhu, Z., Li, X., Asrar, G.R., 2021b. A global seamless 1 km resolution daily land surface temperature dataset (2003–2020). *Iowa State Univ.* <https://doi.org/10.25380/iastate.c.5078492>.
- Zhang, T., Zhou, Y., Zhu, Z., Li, X., Asrar, G.R., 2022. A global seamless 1 km resolution daily land surface temperature dataset (2003–2020). *Earth Syst. Sci. Data* 14, 651–664. <https://doi.org/10.5194/essd-14-651-2022>.
- Zhou, Y., Li, X., Asrar, G.R., Smith, S.J., Imhoff, M., 2018. A global record of annual urban dynamics (1992–2013) from nighttime lights. *Remote Sens. Environ.* 219, 206–220. <https://doi.org/10.1016/j.rse.2018.10.015>.
- Zhou, Q., Wang, C., Fang, S., 2019a. Application of geographically weighted regression (GWR) in the analysis of the cause of haze pollution in China. *Atmos. Pollut. Res.* 10, 835–846. <https://doi.org/10.1016/j.apr.2018.12.012>.
- Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Froliking, S., Yao, R., Qiao, Z., Sobrino, J.A., 2019b. Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sens.* 11, 1–36. <https://doi.org/10.3390/rs11010048>.
- Zhu, W., Lu, A., Jia, S., 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sens. Environ.* 130, 62–73. <https://doi.org/10.1016/j.rse.2012.10.034>.
- Zhu, W., Lü, A., Jia, S., Yan, J., Mahmood, R., 2017. Retrievals of all-weather daytime air temperature from MODIS products. *Remote Sens. Environ.* 189, 152–163. <https://doi.org/10.1016/j.rse.2016.11.011>.