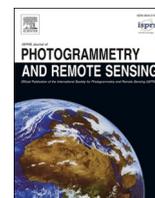


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Downscaling land surface temperature: A framework based on geographically and temporally neural network weighted autoregressive model with spatio-temporal fused scaling factors

Jinhua Wu^a, Linyuan Xia^{a,b,*}, Ting On Chan^{a,b}, Joseph Awange^c, Bo Zhong^d

^a School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China

^b Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou 510275, China

^c School of Earth and Planetary Sciences, Discipline of Spatial Sciences, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

^d State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Keywords:

Land surface temperature (LST)
Downscaling
Spatio-temporal fusion
Neural network
Non-stationarity
Spatial autocorrelation

ABSTRACT

Downscaling land surface temperatures (LST) from satellite imagery is essential for many fine-scale applications. However, the accuracy of the downscaling is often limited by different environmental and geographical conditions. In this work, a novel LST downscaling framework is proposed to improve the accuracy, especially for heterogeneous areas with varying land covers and complex terrains. The framework focuses on downscaling the MODIS LST from 1 km to 100 m, using the proposed geographically and temporally neural network weighted autoregression (GTNNWAR) model with spatio-temporal fused scaling factors derived from Landsat 8 imagery and digital surface models (DSM). To tackle the issues of the non-stationarity of the scaling factors in heterogeneous areas, a region-adaptive parameterization approach is first applied. Then, the GTNNWAR invokes a two-stage deep neural network to estimate the regression coefficients, resulting in the adaption of varying weights for the scaling factors to raise the prediction performance. Moreover, the GTNNWAR is incorporated with a spatial autoregressive model which intakes the neighbor effects so that the overall accuracy can be further improved. Prior to the actual downscaling with the GTNNWAR, a filter-based fusion method is applied to ensure the spatio-temporal consistency of scaling factors is high enough for the neural networks to converge. The results suggest that the proposed framework exhibits high accuracy at the boundaries of different land covers and complex terrains. Compared with several other downscaling algorithms in three case study areas (Beijing and Zhangye in China, Netherlands–Germany in Europe), the proposed framework outperforms with a 28% improved R-squared (R^2) and a root mean square error (RMSE) of 1.02 K. In addition, the downscaled LST has R^2 over 0.63 for the UAV observations (Guangdong). It is concluded that the proposed framework has high reliability and robustness to provide LST datasets with high spatio-temporal resolutions in a wide range of land types.

1. Introduction

Land surface temperature (LST) has become one of the most critical environment covariates for scientific research in recent years spanning over a wide range of fields. LST derived from satellite thermal infrared bands is widely utilized for thermal environmental monitoring (Weng et al., 2014), health risks assessment (Estoque et al., 2020), ecological parameter estimation (Awange et al., 2016), and climate change studies (Yang et al., 2021). For environmental monitoring at local scales, high spatio-temporal resolution of the LST data is always desired. Such high-resolution data cannot directly be measured by the sensors due to the

long satellite revisit cycles or cloud contamination, but it can be retrieved by using effective LST reconstruction techniques (Mo et al., 2021).

In the past two decades, two main streams of techniques were developed to reconstruct continuous high-resolution LST: the spatio-temporal image fusion and the LST downscaling. The spatio-temporal fusion technique uses coarse–fine clear-sky image pairs to reconstruct fine images at prediction dates (Gao et al., 2006; Weng et al., 2014; Wu et al., 2015). As an alternative approach, the LST downscaling technique can integrate multiple sources of extra information to decompose LST, with different downscaling models at the subpixel level, e.g.,

* Corresponding author at: School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China.

E-mail address: xialiny@mail.sysu.edu.cn (L. Xia).

<https://doi.org/10.1016/j.isprsjprs.2022.03.009>

Received 26 October 2021; Received in revised form 4 March 2022; Accepted 11 March 2022

Available online 21 March 2022

0924-2716/© 2022 The Author(s). Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

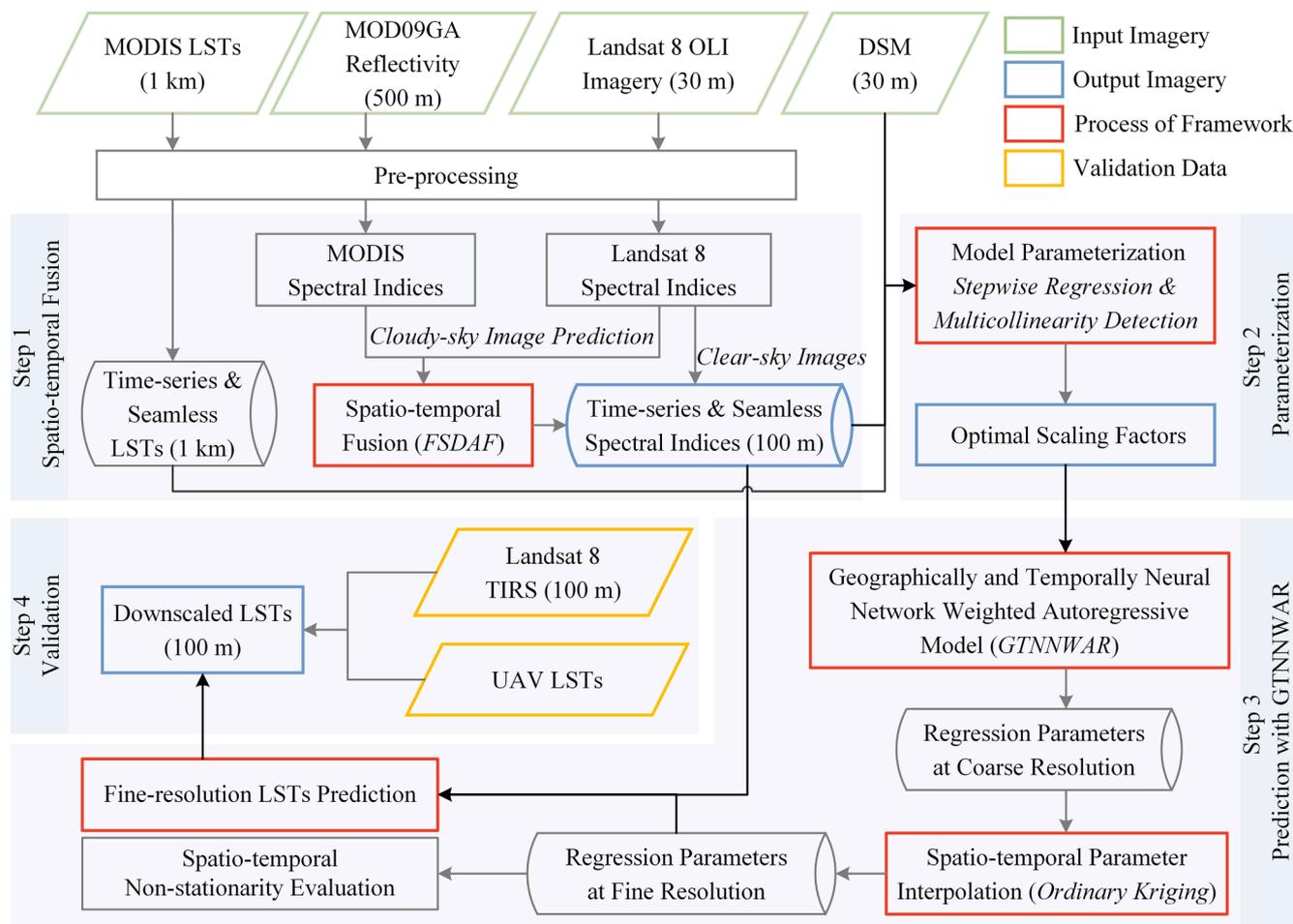


Fig. 1. The proposed GTNNWAR downscaling framework.

modulation-based models (Guo and Moore, 1998), classification-based models (Tonooka, 2005), unmixing-based models (Wang et al., 2020) and statistical-based models (Dong et al., 2020). The statistical-based models are the most popular as they have relatively high simplicity and operability. The models combine the original coarse LST with relevant environmental variables (i.e., scaling factors) to create high-resolution LST datasets (Deilami et al., 2018) based on the assumption of scale invariance. For example, Kustas et al. (2003) proposed the disaggregation procedure for radiometric surface temperature (DisTrad) by establishing a relationship between the radiometric temperature and the normalized difference vegetation index (NDVI). Agam et al. (2007) developed the thermal image sharpening (TsHarp) algorithm by integrating the fractional vegetation cover (FVC) to the regression model instead of NDVI. Bindhu et al. (2013) proposed a non-linear disaggregation method (NL-DisTrad) to downscale MODIS LST for evapotranspiration monitoring based on the relationship between LST and NDVI. These methods performed better in homogeneous areas compared to heterogeneous areas as only a single set of information (e.g., vegetation index (VI)) was considered. For heterogeneous areas, various indices have been proposed and applied in the field (Dominguez et al., 2011), e.g., the normalized difference building index (NDBI), the modified normalized difference water index (MNDWI), the bare soil index (BSI), the normalized multi-band drought index (NMDDI), and other indices associated with digital elevation models (DEM), albedo, land cover, etc. (Tran et al., 2017). As the number of scaling factors increases, how to adapt them to appropriate LST downscaling models becomes predominant and popular in the field.

A variety of methods have been proposed to quantify the two-scale relationship between LST and scaling factors and remedy overfitting

in most cases. The conventional methods include; e.g., linear regression (LR; (Yang et al., 2017)) and polynomial regression (PR; (Dominguez et al., 2011)), while machine learning regression (MLR) methods adapted to the field include; e.g., random forest (RF; (Hutengs and Vohland, 2016)), support vector machine (SVM; (Keramitsoglou et al., 2013)), artificial neural network (ANN; (Li et al., 2019)), and extreme gradient boosting (XGB; (Ghosh and Joshi, 2014)). MLR has commonly been considered to be superior to LR due to its higher applicability in different land covers, especially in the case of the RF (Dong et al., 2020). Most MLR methods can deal with high-dimensional data and effectively reduce the risk of overfitting at global scales (Mo et al., 2021). Furthermore, once the spatial non-stationarity (heterogeneity) and autocorrelation (interdependency) of the observed values are both considered, the sensitivity of MLR models to outliers will be reduced at local scales.

Considering the non-stationary effects of scaling factors on LST, some methods based on geographically weighted regression (GWR) have been developed. Wu et al. (2019) established a spatial non-stationary model between LST and multiple factors for downscaling based on GWR (MFGWR). To characterize non-linear relationships and spatially non-stationary, Xu et al. (2021) developed a multi-factor geographically weighted machine learning (MFGWML) algorithm by fusing XGB, MARS, and the Bayesian ridge regression (BRR) to downscale the Landsat 8 LST to 10 m. These studies considered only the spatial heterogeneity of LST but ignored temporal non-stationarity. It is necessary to consider spatio-temporal non-stationarity to improve the downscaling models. Peng et al. (2019) proposed a geographically and temporally weighted regression (GTWR) model for LST downscaling and achieved promising results in urban regions. However, downscaling of

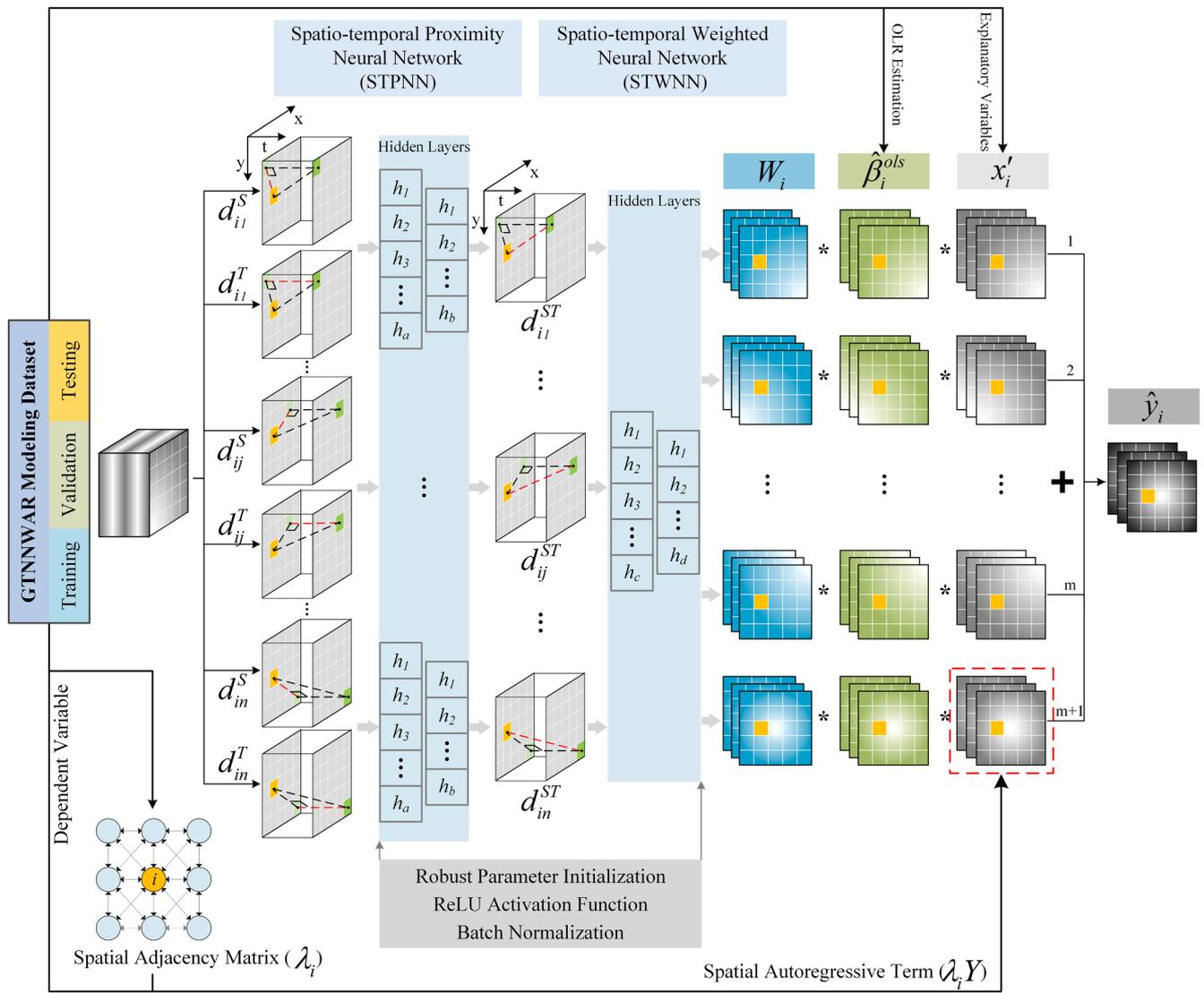


Fig. 2. Architecture of the GTNNWAR model.

some heterogeneous areas with highly varying topographical and environmental conditions remains challenging.

With the recent advance in the development of neural network and spatial–temporal fusion techniques, this paper presents a novel LST downscaling framework based on a new geographically and temporally neural network weighted autoregressive (GTNNWAR) model embedded with spatio-temporal fused scaling factors. The framework blends the LST downscaling and the spatio-temporal image fusion technique, downscaling MODIS LST from 1 km to 100 m, to improve the prediction accuracy in highly heterogeneous areas. There are two major advantages for the proposed framework: (1) a two-stage deep neural network incorporated with an autoregressive model is employed to accurately estimate the non-stationary effects for the scaling factors. (2) a fusion mechanism is adopted to enhance the spatio-temporal consistency of scaling factors to improve the overall accuracy. The results are compared with those of GTNNWR, GTWR, RF, GWR, and TsHarp algorithms in three case study areas (i.e., Beijing, Zhangye, and the Netherlands–Germany) using Landsat 8 LST and UAV data for validation. The paper is organized as follows. Section 2 introduces the four main steps of the framework; Section 3 presents study areas and results; Sections 4 and 5 respectively discuss and conclude the study.

2. GTNNWAR downscaling framework

The proposed GTNNWAR framework (Fig. 1) is implemented to downscale the MODIS LST from 1 km to 100 m with the Landsat 8 imagery and elevation data. The four steps are detailed as follows.

2.1. Step 1: Spatio-temporal fusion of scaling factors

The spatio-temporal fusion is first performed to increase the consistency of the spectral indices (i.e., scaling factors) retrieved from different band reflectance to ensure the high continuity of those spectral indices. Considering the difference between the temporal revisit cycles of the Landsat 8 (16 days) and MODIS (1 day), two methods are merged to generate seamless spectral indices. Firstly, the Savitzky–Golay (SG) filter (Chen et al., 2004) is employed to interpolate the MODIS spectral indices due to its adaptability for static and uniform continuous data. Secondly, the flexible spatio-temporal data fusion (FSDAF) method (Zhu et al., 2016) is utilized to predict the high-resolution indices due to its high robustness in heterogeneous areas, even with large data sparsity. The prediction value $\hat{F}_2(x_{ij}, y_{ij}, b)$ of band b is obtained by weighting the gradients of similar pixels with the space distance (D_k), and then summing up with the observation value $F_1(x_{ij}, y_{ij}, b)$ of the initial date:

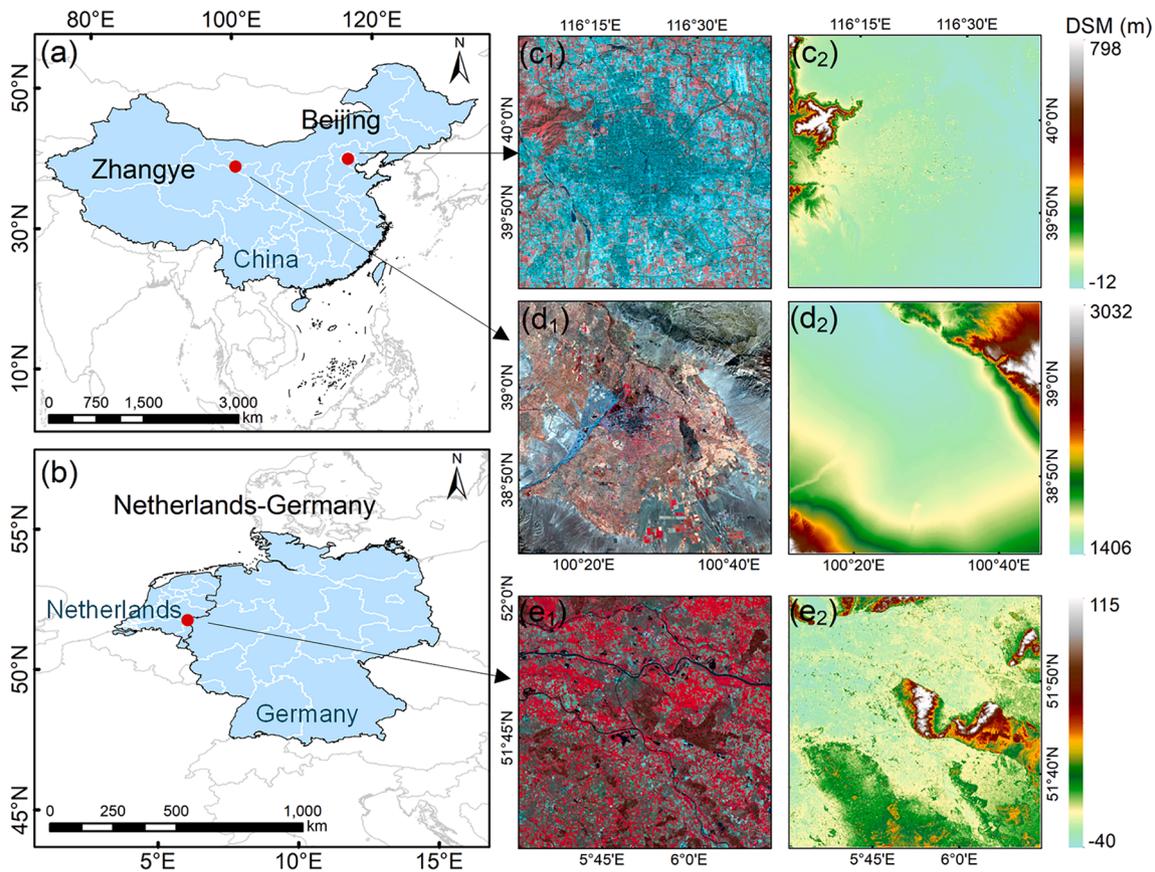


Fig. 3. Case study regions. Subscript 1 represents the false-color image, and Subscript 2 represents DSM.

Table 1
Data employed in this study.

Data Source	Band Name	Resolution (m/day)	Sampling Period
MOD09GA	VNIR (Band 1–4) and SWIR (Band 6 and 7)	500/1	2018–2019
MOD11_L2	LST	1000/1	
Landsat 8	VNIR (Band 2–5) and SWIR (Band 6 and 7) (Band 10)	30/16	
ALOS/AW3D30	DSM	100/16	
UAV	TIR	0.18	01/06/2022

Table 2
The acquisition date and cloud cover of Landsat 8 data.

ID	Acquisition Date/ Cloud Cover (%)		
	Beijing 39° 41' 57"–40° 9' 8" N 116° 5' 8"–116° 40' 21" E	Zhangye 38° 41' 51"–39° 9' 1" N 100° 11' 18"–100° 46' 4" E	Netherlands–Germany 51° 32' 38"–51° 59' 32" N 5° 31' 18"–6° 14' 43" E
T1	08/01/2019	09/21/2018	03/20/2018
T2	08/17/2019	10/07/2018	04/05/2018
T3	09/02/2019*	10/23/2018*	04/21/2018*
T4	09/18/2019	11/08/2018	05/07/2018
T5	10/04/2019	11/24/2018	05/23/2018

* indicates the prediction date.

$$\hat{F}_2(x_{ij}, y_{ij}, b) = F_1(x_{ij}, y_{ij}, b) + \sum_{k=1}^n \left[\frac{1/D_k}{\sum_{k=1}^n (1/D_k)} \right] \cdot \Delta F(x_k, y_k, b) \quad (1)$$

$$D_k = 1 + \sqrt{(x_k - x_{ij})^2 + (y_k - y_{ij})^2} / (w/2) \quad (2)$$

where w is the size of the moving window, $\Delta F(x_k, y_k, b)$ is the total change value of the k th similar pixel.

2.2. Step 2: Model parameterization

Over-parameterization of the regression model for the downscaling will cause an increase in degrees of freedom and multicollinearity. Hence, it is necessary to determine a set of optimal variables for the regression model. In this study, a three-stage procedure is applied to implement model parameterization. For the first step, inspired by (Deilami et al., 2018), we propose using NDVI, NDBI, and MNDWI to enrich the information content for the land covers. The soil moisture is characterized by NMDI, and albedo is estimated by using the narrow-band to broadband (NTB) conversion model (Liang, 2001). In addition, the digital surface model (DSM) is used to characterize the surface elevation of geographical features.

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (3)$$

$$NDBI = \frac{\rho_{SWIR1} - \rho_{NIR}}{\rho_{SWIR1} + \rho_{NIR}} \quad (4)$$

$$MNDWI = \frac{\rho_{GREEN} - \rho_{SWIR1}}{\rho_{GREEN} + \rho_{SWIR1}} \quad (5)$$

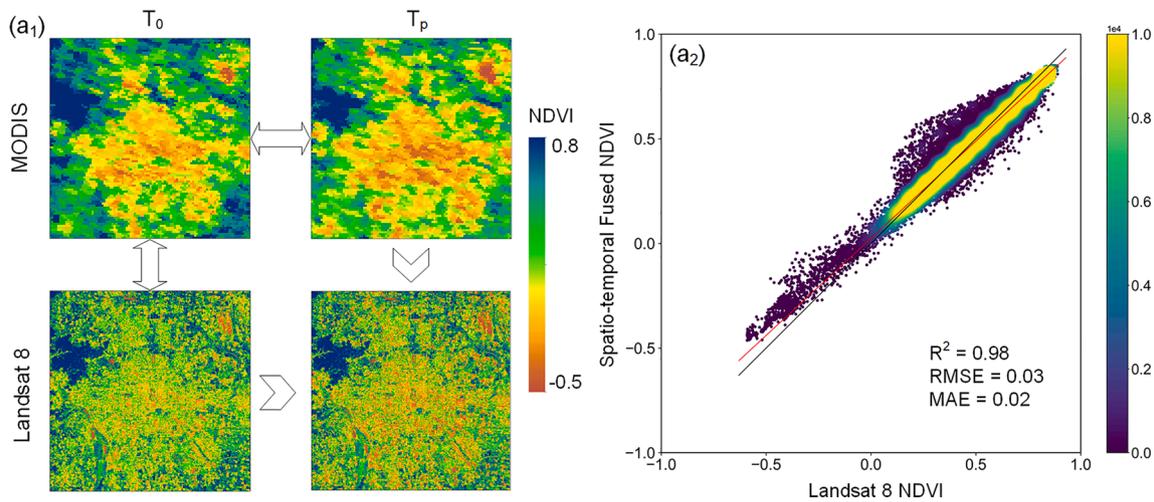


Fig. 4. Validation of spatio-temporal fusion for scaling factors in Beijing. (a₁) Fusion process with FSDAF method. T₀: September 2, 2019, T_p: September 18, 2019. (a₂) The scatter plot of fused NDVI and Landsat 8 NDVI.

$$NMDI = \frac{\rho_{BLUE} - (\rho_{SWIR1} - \rho_{SWIR2})}{\rho_{BLUE} + (\rho_{SWIR1} - \rho_{SWIR2})} \quad (6)$$

$$Albedo = 0.356 \cdot \rho_{BLUE} + 0.130 \cdot \rho_{RED} + 0.373 \cdot \rho_{NIR} + 0.085 \cdot \rho_{SWIR1} + 0.072 \cdot \rho_{SWIR2} - 0.0018 \quad (7)$$

where $\rho_{BLUE}, \rho_{GREEN}, \rho_{RED}, \rho_{NIR}, \rho_{SWIR1}$ and ρ_{SWIR2} are the reflectances of corresponding bands.

For the second step, the multicollinearity and interpretability of the variables are considered to improve the model accuracy. The stepwise regression is first used to select variables and deal with multicollinearity since it can preserve the physical meaning of scaling factors for subsequent analysis (Gevrey et al., 2003). New scaling factors are introduced to the model (i.e., forward selection), and redundant scaling factors are eliminated (i.e., backward elimination) by using statistical tests. (3) For the last step, the variance inflation factor (VIF) (Miles, 2014) and condition index (CI) (Kim, 2019) are used to detect multicollinearity and select the optimal set of variables for the model. When the two indices are less than 10, the combination of independent variables is considered non-collinear.

2.3. Step 3: Model prediction by GTNNWAR

(a) principle of GTNNWAR

Natural eco-environmental processes usually contain significant spatio-temporal non-stationarity, which can be quantified with neural network regressive models due to the high robustness of non-linear fitting (Wu et al., 2021). Furthermore, considering spatial autocorrelation in the model can benefit the prediction of LST due to its strong frequency field. We propose the GTNNWAR model (Fig. 2) to predict finer LST by considering both the spatio-temporal non-stationarity and autocorrelation of LST. the estimated value $i(u_i, v_i, t_i)$ ($i = 1, 2, \dots, n$) is:

$$\begin{aligned} \hat{y}_i &= \sum_{k=0}^m \hat{\beta}_{ik} x_{ik} + \hat{\beta}_{i(m+1)} \lambda_i Y = \sum_{k=0}^{m+1} \hat{\beta}_{ik} x'_{ik} \\ &= (x'_i)^T \hat{\beta}_i = (x'_i)^T W_i \hat{\beta}_i^{ols} \\ &= (x'_i)^T W_i \left[(X^T X)^{-1} X^T Y \right] \end{aligned} \quad (8)$$

where m is the number of independent variables, $\hat{\beta}_i$ is the weighted ordinary least squares (OLS) regression coefficients. λ_i is the spatial adjacency matrix of point i (with the neighbor point $j = 1, 2, \dots, n$):

$$\lambda_i = (\lambda_{i1} \ \lambda_{i2} \ \lambda_{i3} \ \dots \ \lambda_{in}), \lambda_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are adjacent} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

X is the explanatory variable matrix with spatial autoregressive (SAR) term, W_i is a space–time weight matrix:

$$\begin{aligned} X &= \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1m} & \lambda_1 Y \\ 1 & x_{21} & \dots & x_{2m} & \lambda_2 Y \\ 1 & x_{31} & \dots & x_{3m} & \lambda_3 Y \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{nm} & \lambda_n Y \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}, W_i \\ &= \begin{bmatrix} w_{i0} & 0 & 0 & 0 \\ 0 & w_{i1} & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & w_{i(m+1)} \end{bmatrix} \end{aligned} \quad (10)$$

Similar to GTNNWR (Wu et al., 2021), a spatio-temporal proximity neural network (STPNN) is firstly utilized to characterize complex non-linear spatio-temporal interactions:

$$d_{ij}^{ST} = STPNN(d_{ij}^s, d_{ij}^t) \quad (11)$$

Then, a spatio-temporal weighted neural network (STWNN) is adopted to calculate spatio-temporal weights:

$$W_i = STWNN([d_{i1}^{ST}, d_{i2}^{ST}, \dots, d_{in}^{ST}]^T) \quad (12)$$

where STPNN can combine spatial distance (d_{ij}^s) with temporal distance (d_{ij}^t) for any two points. STWNN can quantify the optimal weights of each neighboring point ($1, \dots, n$) to the target point (i) from the spatio-temporal distance (d_{ij}^{ST}). The two neural networks are integrated to construct a deep neural network with multiple hidden layers for model optimization. when they are substituted in (8), the estimated values can be expressed as:

$$\hat{y}_i = (x'_i)^T STWNN([STPNN(d_{i1}^s, d_{i1}^t), \dots, STPNN(d_{in}^s, d_{in}^t)]) (X^T X)^{-1} X^T Y \quad (13)$$

The estimation precision of GTNNWAR is contingent on the performance of a two-stage deep neural network (STPNN and STWNN) with multiple hidden layers denoted by h in Fig. 2. Furthermore, robust parameter initialization, rectified linear unit (ReLU) activation function, and batch normalization techniques (Wu et al., 2021) are used to facilitate the training process. The estimated value, \hat{y}_i , is retrieved by

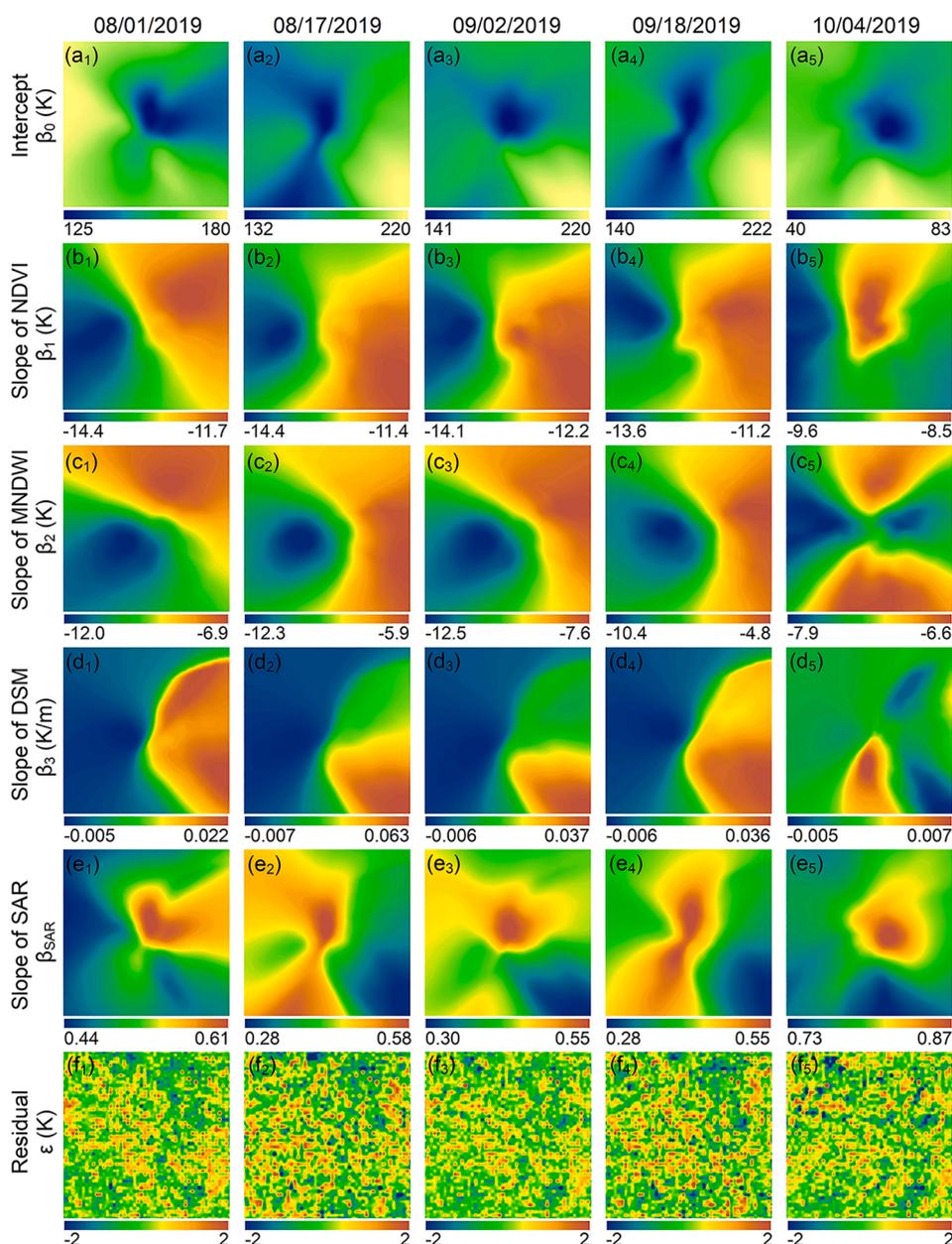


Fig. 5. The spatio-temporal variations of the interpolated GTNNWAR coefficients in Beijing.

the product sums of the weights, coefficients, and the corresponding explanatory variables (as well as the autoregressive term).

(b) LST downscaling with GTNNWAR

The LST downscaling is achieved by first modeling the non-stationarity between coarse LST and scaling factors. The selected scaling factors are introduced into the GTNNWAR to model the spatio-temporal non-stationarity. The regression coefficients are estimated at the coarse resolution with the two-stage neural network; Then, the spatio-temporal parameter interpolation is implemented based on the assumption of scale invariance between temperature and related scaling factors. The regression coefficients of the GTNNWAR model are interpolated into a finer spatial resolution to minimize the prediction bias using the ordinary Kriging interpolation method, which can be used to quantify proximity and geographical connection (Oliver and Webster, 1990); Finally, high-resolution LST is predicted by introducing the interpolated regression coefficients (\tilde{C}) to the GTNNWAR model to

estimate the high-resolution LST. The regression coefficient reflects the non-stationary effects posed by the scale factors (SF) on LST, facilitating the prediction of fine LST which can be expressed as:

$$LST_F = LST_C + f_{\tilde{C}}(SF_F) - f_C(SF_C) \tag{14}$$

where subscripts F and C represent the fine and coarse resolution, respectively.

2.4. Step 4: Validation of the GTNNWAR downscaling framework

Several datasets which include satellite-based and UAV-based thermal observations are used to verify the proposed framework. The GTNNWAR downscaled results from MODIS LSTs are compared with Landsat 8 LSTs due to the acceptable temporal consistency of the two-source data. The calibration method proposed by (Li et al., 2008) is adapted to compensate for the systematic biases for the downscaled LST compared with UAV LST. The verification results are given in Section 3.2.4.

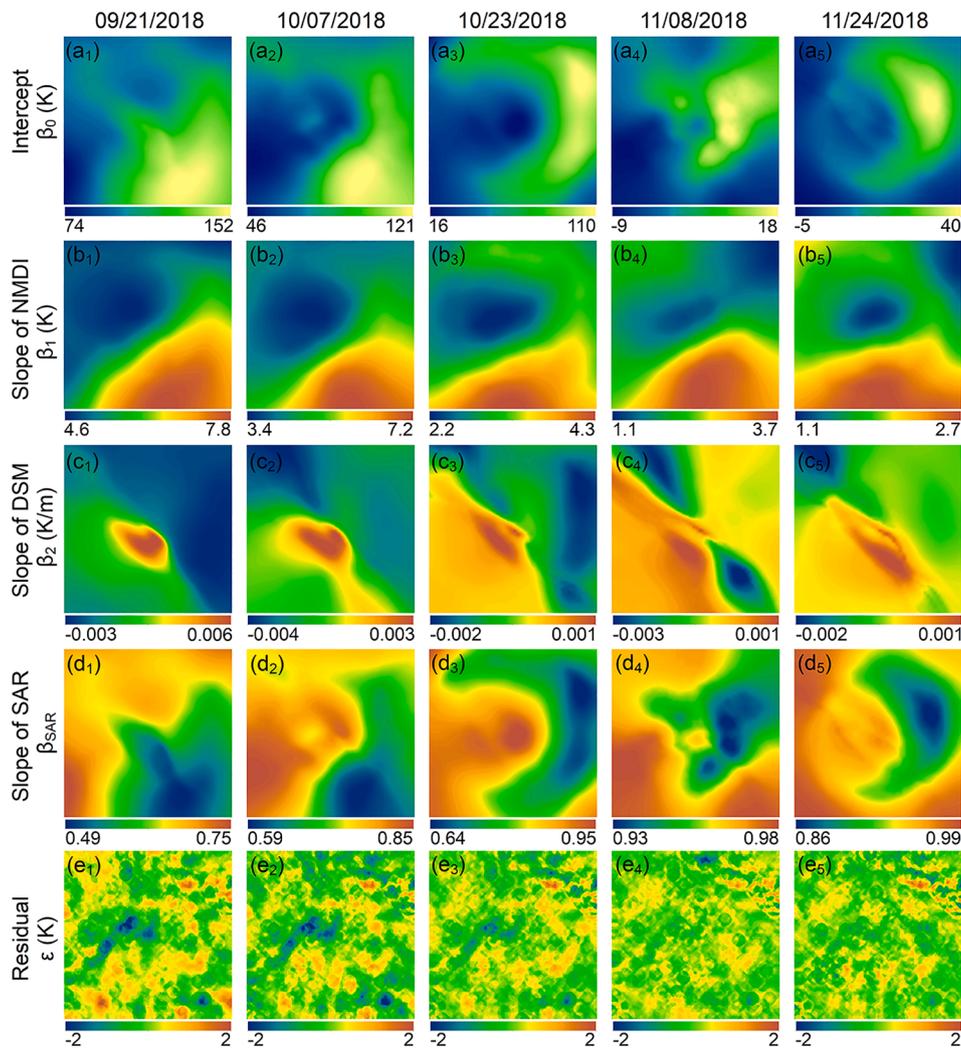


Fig. 6. The spatio-temporal variations of the interpolated GTNNWAR coefficients in Zhangye.

3. Study areas and results

3.1. Case study regions

Three case study areas (Fig. 3) are selected for testing the proposed downscaling framework: (i) the Beijing region consisting of a warm temperate continental monsoon climate, with an annual temperature of 11–13 °C, and annual precipitation of 400–800 mm. There are dense buildings in the central urban areas with mountainous vegetations in the suburbs; (ii) the Zhangye region, which has a cold temperate arid climate, with an annual temperature of 4.1–8.3 °C, and annual precipitation of 112.3–354.0 mm. There are many buildings and farmlands in the central area surrounded by deserts; (iii) the Netherlands–Germany region, which is on the border between the Netherlands and Germany. It has a broad marine temperate leaf forest climate, with an annual temperature of ~10.5 °C and annual precipitation of ~824 mm. There are intricately distributed buildings and farmlands.

Multi-source data are input into the proposed LST downscaling framework (Table 1). The MOD11_L2 swath products are used to provide clear-sky 1 km LSTs for the downscaling model with an accuracy better than 1 K (Wan and Dozier, 1996). The LST products post-processed by (Shiff et al., 2021) with an overall accuracy of 0.93 are used to supplement the missing data of cloudy sky. The Landsat 8 and MODIS images with an overpass time of approximately thirty minutes are employed to implement the fusion. The four base image pairs near the prediction date are integrated to perform the downscaling model

(Table 2). The MODIS visible and near-infrared (VNIR) and shortwave infrared (SWIR) bands of MOD09GA products (500 m) are used to generate cloudy-sky Landsat-like spectral indices by performing the spatio-temporal fusion. The Operational Land Imager (OLI) and Thermal Infrared (TIR) images of Landsat 8 are used for retrievals of the clear-sky spectral indices and validation of the downscaled results, respectively. The OLI imagery is pre-processed to surface reflectance by performing radiometric calibration and atmospheric correction. The TIRS images (100 m) are used to retrieve LST with the mono-window (MW) algorithm (Qin et al., 2001) for verification (accuracy of ~ 1.4 K).

The digital surface model (DSM) was obtained from the Advanced Land Observing Satellite (ALOS) world 3D-30 m (AW3D30) with an accuracy of less than 5 m (Takaku et al., 2014). These images are resampled to 1 km and 100 m for calculating coarse and fine resolution environment variables, respectively. All satellite images are converted to WGS84 UTM projection and georeferenced. The airborne observations are collected using an Unmanned Aerial Vehicle (UAV) (DJI M300) equipped with a thermal camera (Zemuse H20T). The thermal images are calibrated and *ortho*-mosaic matched to an LST map by DJI Thermal Analysis Tool and the Pix4D software. The post-processed UAV LST data was retrieved at an elevation of 200 m within an area of 0.33 km² (at ~10:30, 01/06/2022; in Guangdong, China), to evaluate the down-scaled results.

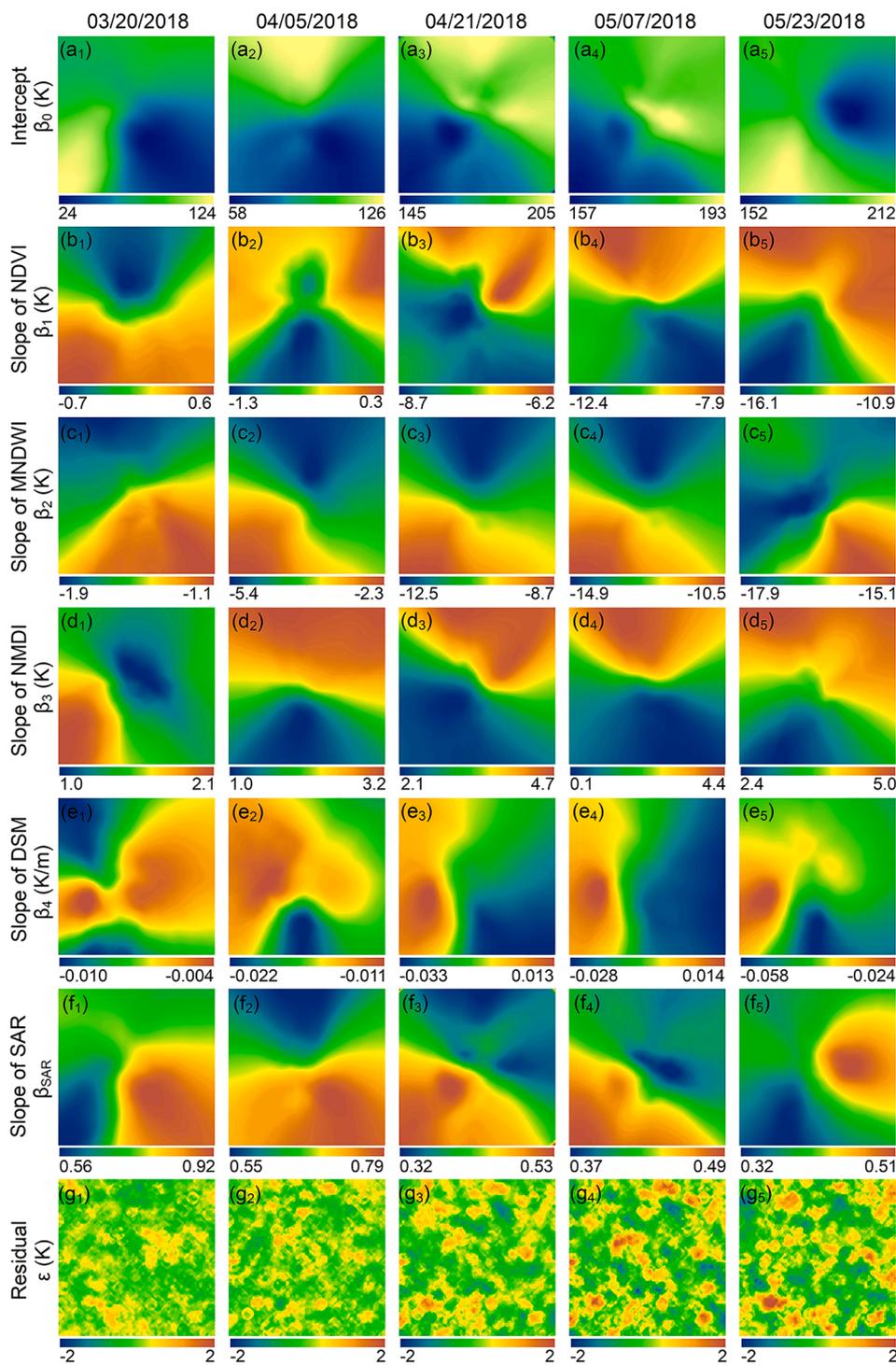


Fig. 7. The spatio-temporal variations of the interpolated GTNNWAR coefficients in the Netherlands–Germany.

3.2. Case study results

3.2.1. Spatio-temporal fusion and model parameterization

Based on the FSDAF fusion, the predicted fine NDVI reveals land covers, showing that the temporal variation is consistent with vegetation phenology (Fig. 4a₁). As shown in Fig. 4a₂, the spatio-temporal fused NDVI is highly correlated with NDVI retrieved from Landsat 8, possessing R² of 0.98, RMSE of 0.03, and MAE of 0.02.

Since the scaling factors are reconstructed by the spatio-temporal fusion, the optimal variables are determined by the results of stepwise regression (adjusted R²) and multicollinearity detection (VIF and CI):

NDVI, MNDWI, and DSM (VIF of 2.725, 2.694, and 1.099; CI of 5.917; adjusted R² of 0.734) for Beijing region, NMDI and DSM (VIF of 1.018, and 1.018; CI of 3.375; adjusted R² of 0.583) for Zhangye region, and NDVI, MNDWI, NMDI and DSM (VIF of 9.004, 1.480, 8.060 and 1.043; CI of 8.124; adjusted R² of 0.677) for the Netherlands–Germany region. The optimal scaling factors of these three study areas are different, indicating that the interpretation ability of different scaling factors to LST is discrepant. Interpolating multiple scaling factors into the model will improve the goodness of fit but increase the risk for multicollinearity to occur.

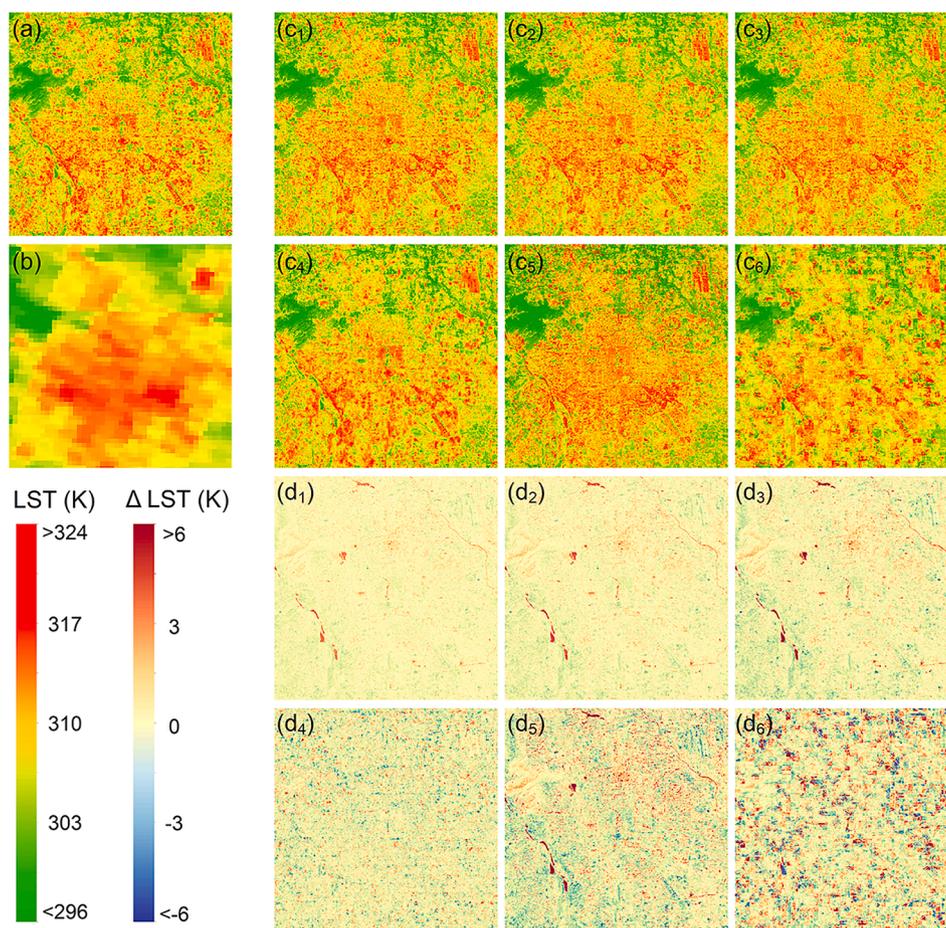


Fig. 8. Comparisons in Beijing. (a) Landsat 8 LST; (b) MODIS LST; (c-d) the downscaled LST and bias. Subscripts 1–6 represent results of GTNNWAR, GTNNWR, GTWR, RF, GWR, and TsHarp, respectively.

3.2.2. Analysis of the GTNNWAR model

Regression parameters of different dates are calculated at local scales to further evaluate the ability to quantify non-stationarity in the GTNNWAR model.

The regression coefficients of Beijing change significantly from the West to the East. NDVI and MNDWI have negative effects on LST on different days, with the slopes varying from -14.4 to -8.5 K and from -12.5 to -4.8 K, respectively. In August, the influence of the West is stronger than that of the East (Fig. 5b₁ and 5c₁). The areas with negative impacts expand from the southwest to the northwest, which is related to phenology changes as the defoliation of forests decreases the interpretability of NDVI exerted to LST. The slopes of DSM show a significant negative impact on LST in the West and a positive impact on LST in the East (Fig. 5d). This is because LST decreases with the elevation on the western mountains, while it increases with building heights in the eastern urban. The high slope of SAR shows more high values of aggregation in the central urban area, indicating that the LST is greatly affected by its autocorrelation.

The difference in the regression coefficients is quite significant between the urban and suburban areas of Zhangye, as shown in Fig. 6. NMDI significantly contributes to LST (from 1.1 K to 7.8 K), with an enormous impact in the southeast desert area and a small impact in the central urban area. From September to November, the impact of NMDI to LST decreased gradually (Fig. 6b), which is related to the change of climatological conditions. The DSM exerts a significant positive effect on LST in the central urban area but a negative effect in the suburbs from September to October. The increasing slope of SAR implies the effect of autocorrelation on LST becomes increasingly significant (from 0.49 to 0.99).

In the Netherlands–Germany region, four covariates are significant and integrated into the model (Fig. 7). The slopes of NDVI, MNDWI, and NMDI fluctuate from the South to the North. The influence of the variables on LST increases with time, which is mainly attributed to crop phenology and hydrological changes. From April to May, the influence of NDVI and NMDI on LST in the South is more significant than that in the North, while MNDWI varies oppositely. There are negative impacts of DSM on LST in other areas due to the large-scale vegetation and farmlands. The slopes of SAR indicate that LST is more affected by its autocorrelation in the southern area than the northern river area. The test results in urban, rural, and arid areas show that the regression coefficients of the GTNNWAR model have significant spatiotemporal non-stationarity, and the randomness of residual distribution indicates the robust fitting performance of the model. Moreover, the regression coefficients show the contribution of dominant factors to LST, which promotes the analysis of the local thermal environment.

3.2.3. Comparisons of different methods using Landsat 8 LST

The downscaling accuracy of GTNNWAR and some classical algorithms (TsHarp, GWR, RF, GTWR, and GTNNWR) are evaluated by cross-validation with LST retrieved from Landsat 8. The TsHarp algorithm in the Beijing region (Fig. 8) performed mediocly with many outliers in the denseness building area because it only considers the vegetation factor. The RF algorithm has a random error distribution with less extreme values than the TsHarp. The GWR algorithm overestimates LST in mountainous areas. This is likely because the GWR model only considers spatial proximity, limiting the accuracy at the boundaries. Compared to the GTWR and GTNNWR methods, the error distribution of the proposed framework appears smoother near rivers (Fig. 8d₁). This is

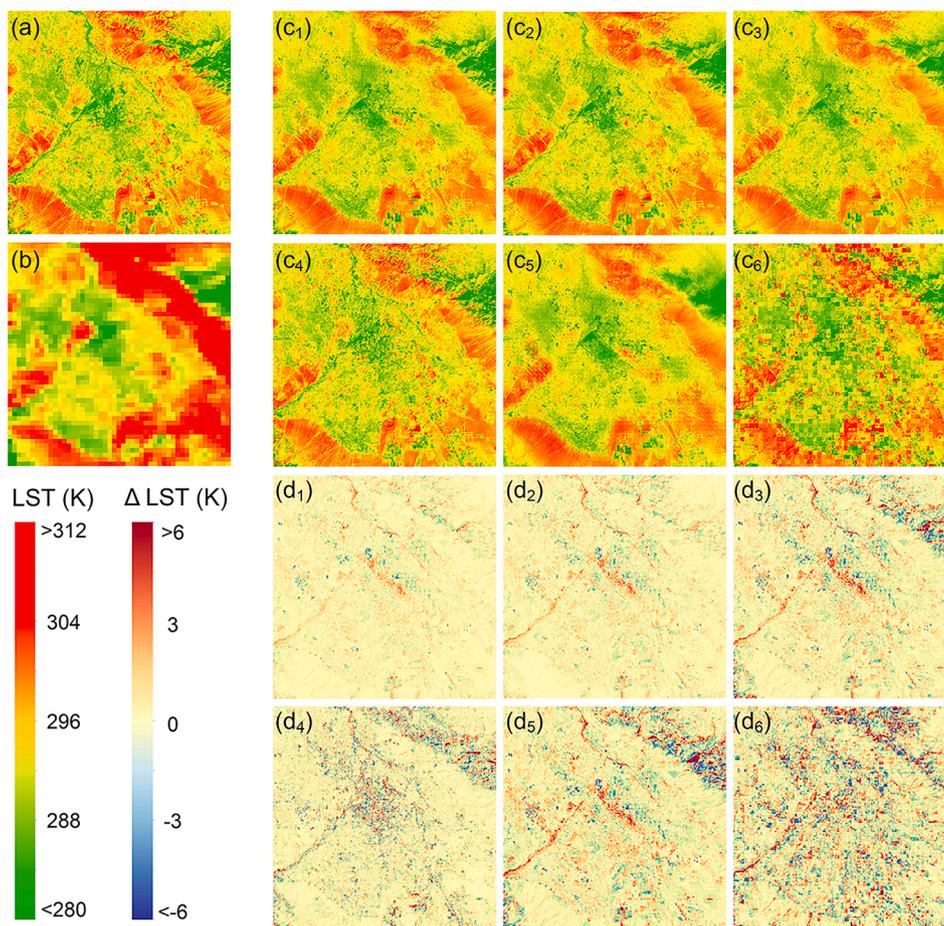


Fig. 9. Comparisons in Zhangye. (a) Landsat 8 LST; (b) MODIS LST; (c–d) the downscaled LST and bias. Subscripts 1–6 represent results of GTNNWAR, GTNNWR, GTWR, RF, GWR, and TsHarp, respectively.

because the GTNNWAR method has considered both spatio-temporal non-stationarity and spatial autocorrelation.

Similarly, the downscaled LSTs obtained from different methods for Zhangye are shown in Fig. 9. The result of TsHarp shows the most extreme outliers, and it exhibits distinct rectangle-shaped artefacts at the boundaries of different land covers. The GWR algorithm underestimates LST on the sunny side of the mountain, while it overestimates LST on the shady side. In addition, the GWR algorithm has many abnormally high values near the rivers (Fig. 9d₅). The accuracy of GTWR is higher than that of GWR as the temporal information of LST in the areas with vegetation and the desert. The global errors of RF are larger than that of GTWR, which indicates that RF has limitations for complex terrain areas (Fig. 9d₄). The GTNNWAR algorithm has the highest accuracy in the mountainous areas because the optimal local weights are used to reduce the errors caused by the arid climate.

For the Netherlands–Germany region (Fig. 10), The result of TsHarp shows blurring boundaries between vegetations and impervious layers. RF performs better on homogeneous surfaces than that on complex surfaces. GWR and GTWR are erroneous at the boundaries of different land covers. GTWR and GTNNWR perform better on the farmlands and rivers (Fig. 10d₂ and Fig. 10d₃) as they effectively capture the temporal variability of LST. To address the phenological changes, the GTNNWAR framework considers the autocorrelation of LST, and the parameterization is optimized with a minimum number of errors at the global scale.

It can be seen from Table 3 that the GTNNWAR framework outperforms its compatriots for different land covers and terrains. In particular, the GTNNWAR framework provides improvements, with R^2 increasing 3–28% and RMSE decreasing 0.12–1.02 K over other methods. As shown in Fig. 11, all the fitting models pass the 0.1%

significance test (p-value less than 0.001), and they possess high linearity. The yellow error ellipses in the scatter plots of GTNNWAR are flatter than that of other methods. This indicates that there is relatively less bias of GTNNWAR. As the box graphs show, the downscaling result of GTNNWAR is closest to that of the Landsat 8 LST compared to those of other methods.

3.2.4. Validation using airborne UAV data

The downscaled LST of the GTNNWAR framework is further verified with the UAV thermal observations for a rural area in Guangdong Province of China. The systematic bias (~3.2 K) between the two datasets is estimated by performing a calibration procedure for comparisons. The results show a fair consistency between the downscaled and UAV LST (Fig. 12) with an acceptable relative error (R^2 over 0.63 and RMSE of 0.27 K).

4. Discussion

4.1. Issues related to model parameterization in the framework

The scaling factors often suffer from missing data caused by cloud contamination, defective sensors, and scanning gaps, which hinders LST downscaling. To overcome these limitations, the fusion mechanism can effectively fill the missing scaling factors primarily caused by cloud pollution. The spatio-temporal fusion is used to obtain scaling factors for the cloudy-sky days; otherwise, the reflectances of Landsat 8 are applied for clear-sky days. The scaling factors of the procedures treating both the clear-sky and cloudy days possess high consistency. Furthermore, the temporal continuity of scaling factors should be emphasized in the

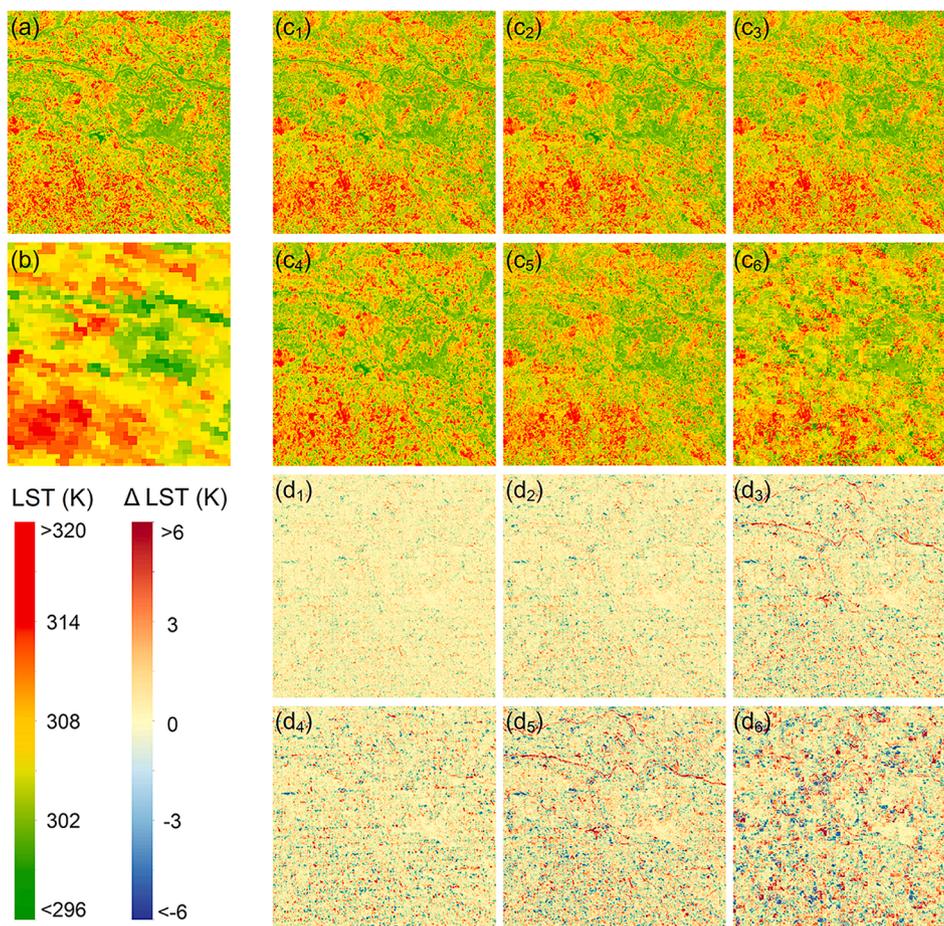


Fig. 10. Comparisons in Netherlands–Germany. (a) Landsat 8 LST; (b) MODIS LST; (c-d) the downscaled LST and bias. Subscripts 1–6 represent results of GTNNWAR, GTNNWR, GTWR, RF, GWR, and TsHarp, respectively.

Table 3
Statistical indicators of different algorithms in the three regions.

Model	Beijing			Zhangye			Netherlands–Germany		
	R ²	RMSE (K)	MAE (K)	R ²	RMSE (K)	MAE (K)	R ²	RMSE (K)	MAE (K)
GTNNWAR	0.88	0.95	0.72	0.85	1.09	0.78	0.86	1.16	0.88
GTNNWR	0.85	1.05	0.80	0.81	1.23	0.89	0.83	1.29	0.98
GTWR	0.79	1.24	0.92	0.75	1.40	0.99	0.74	1.58	1.19
GWR	0.67	1.54	1.11	0.64	1.69	1.19	0.64	1.89	1.42
RF	0.83	1.14	0.88	0.76	1.53	0.97	0.77	1.52	1.14
TsHarp	0.61	1.82	1.31	0.56	2.21	1.56	0.58	2.24	1.66

fusion mechanism for the framework. The temporal interval and frequency of the sampling can be set flexibly. Sampling within a long-time interval will introduce biases to the GTNNWAR model due to the variation of climate and land covers. In contrast, sampling within a short time interval will weaken the ability of the model to capture gradually changing patterns (e.g., vegetation phenology). Sampling with higher frequency is essential to achieve better performance in areas with disturbance events (e.g., cropland harvesting and forest fire), and different temporal intervals and frequencies of sampling can be investigated for parameter optimization in different scenarios.

4.2. Role of the GTNNWAR model in the framework

Compared with spatio-temporal fusion, the downscaling technique can integrate multiple bio-geophysical factors (e.g., land cover and elevation) to decompose LST. The influence of bio-geophysical factors on LST has rarely been analyzed at local scales in previous research. The

GTNNWAR model estimates the local optimal spatio-temporal weights with neural networks and quantifies the spatio-temporal distribution of local regression coefficients, revealing the sensitivity of LST to different scaling factors and its autocorrelation. The regression coefficients vary spatially with the land covers and terrains, and they vary temporally with vegetation phenology and hydrologic conditions. This spatio-temporal variability shows that the spatio-temporal patterns of the local thermal environment can be adequately modeled with the GTNNWAR.

Taking fine satellite LST as the reference, the sensor errors between the downscaled MODIS LST and Landsat 8 LST can be reduced by diurnal temperature correction models in future work. The uncertainty of using airborne observations to validate LSTs downscaled from satellite retrievals can be further assessed by considering the thermal radiation directional effects (Duan et al., 2019). Compared with the spatial downscaling method, the spatio-temporal downscaling methods produce more reliable results in cloudy-sky and heterogeneous areas, partly

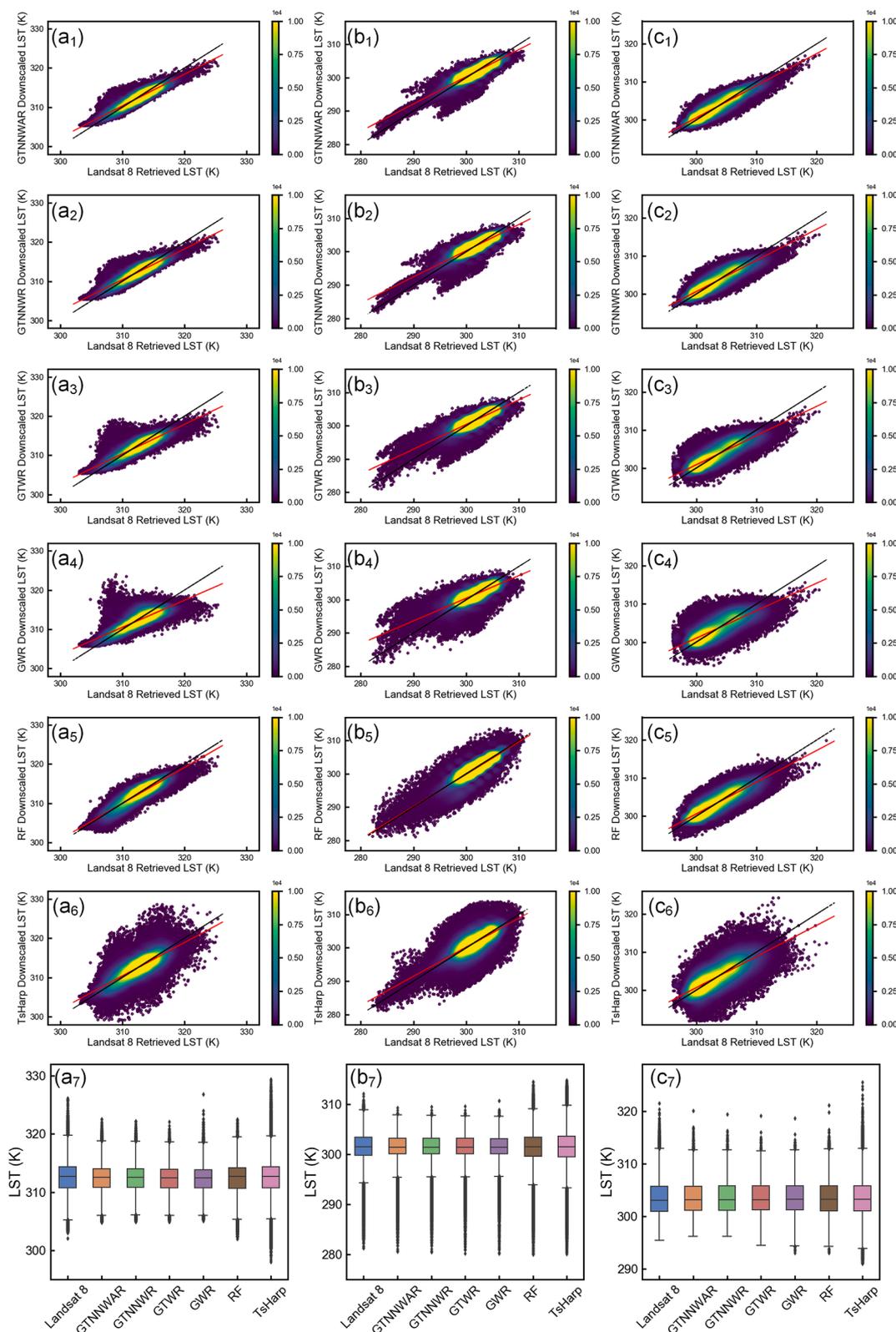


Fig. 11. The accuracy evaluation of six downscaling models in different regions. Rows 1–6 show the scatterplots of the GTNNWAR framework, GTNNWR, GTWR, GWR, RF, and TsHarp, respectively, and Row 7 shows boxplots. Columns 1–3 represent Beijing, Zhangye, and the Netherlands–Germany regions, respectively. The red line is the best-fit line for points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

addressing the limitations of temporal or spatial downscaling methods (Fu and Weng, 2016). The GTNNWAR method can better deal with significant temperature differences in areas with complex land covers and terrains due to the accurate quantification of both spatio-temporal

non-stationarity and spatial autocorrelation of LST. The performances of the GWR and GTWR methods are restricted by the need for manual selection of the appropriate kernel functions for complex geographical properties so that their estimation of the non-stationarity is hindered.

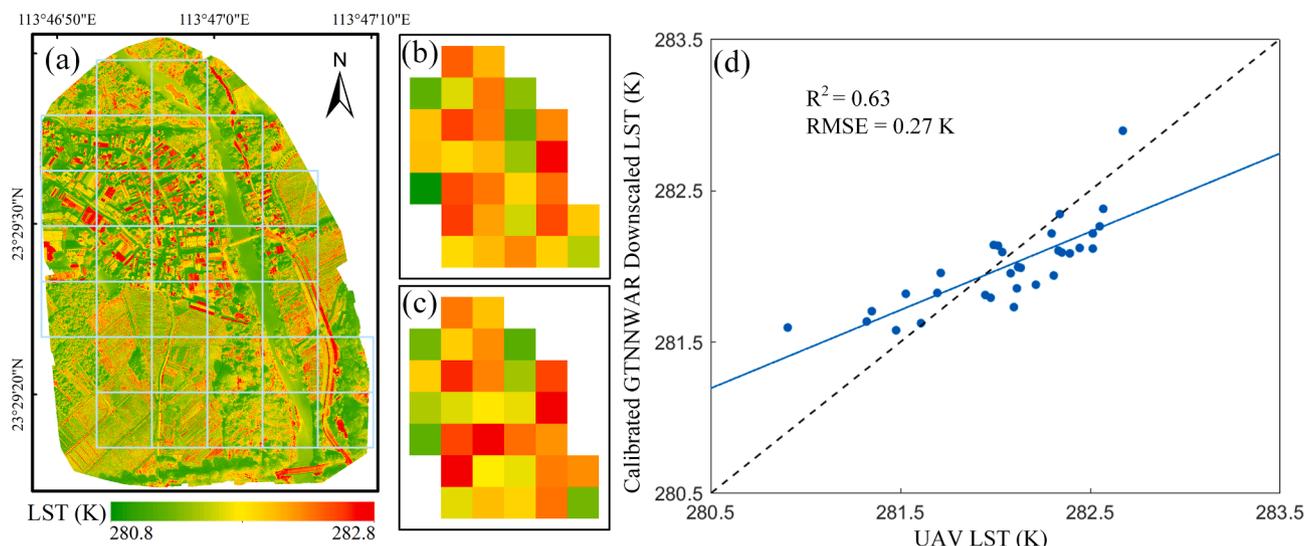


Fig. 12. Comparisons of the calibrated GTNNWAR downscaled LST and UAV LST in Guangdong. (a) the UAV LST; (b-c) the aggregated 100 m UAV LST and downscaled LST; and (d) the goodness of fit.

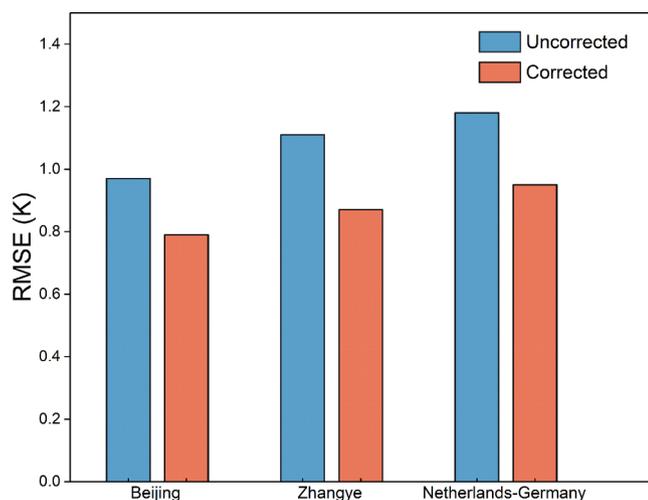


Fig. 13. Comparison of the RMSEs for GTNNWAR framework with uncorrected and corrected scale effect.

GTNNWR can adopt kernels automatically to improve the estimation of spatio-temporal non-stationarity but ignores the spatial autocorrelation of LST. The GTNNWAR can better handle temperature gradients than the RF in areas with complex land covers and terrains; nevertheless, the downscaled LSTs of GTNNWAR are equivalent to those of RF in homogeneous land surfaces.

4.3. Analysis of scale effect in the GTNNWAR model

The proposed GTNNWAR LST downscaling framework is based on the assumption that the relationship between LSTs and scaling factors is invariant at different spatial resolutions. In this study, a comparison strategy is used to analyze the uncertainty caused by the scale effect in the GTNNWAR framework based on the availability of high-resolution LSTs. The coarse images of LST and scaling factors are replaced by the corresponding fine images to train the model and acquire fine parameters on the prediction date. The fine LST can be estimated as:

$$LST_F = LST_C + f_F(SF_F) - f_C(SF_C) \tag{15}$$

The explanation of Eq. (15) is analogous to that of Eq. (14). The scale

effect of the GTNNWAR framework is quantified by comparing the accuracies (Fig. 13) retrieved by Eqs. (14) and (15). The results show that the accuracy of the proposed framework can be further improved (~20%) in different study areas with the scale effect corrected.

4.4. Prospects of the proposed framework

The proposed framework for LST downscaling can be applied to several fields for the following reasons: 1) High compatibility of other datasets: although only the Landsat 8 and MODIS imagery was demonstrated for the framework, other remote sensing products with different spatial scales, texture structures, and noise levels (e.g., Sentinel 2 imagery) can be potentially adapted for the framework; 2) Precise estimation of LST at local scales: the framework allow satellite-observed LST being able to be applied for environmental monitoring at local scales. The framework can be applied to disaggregate LST with higher resolution (e.g., Landsat 8 LST); 3) Availability of extra information: with the GTNNWAR, each downscaled LST unit can be obtained with a set of coefficients related to scaling factors as by-products, so that the downscaled LST is more favorable to environmental analyses. 4) Supporting for wide applications: the downscaled LSTs could provide spatio-temporal continuous and elaborate thermal information by using block processing at large-scale areas. Our framework can effectively provide LST for analyses related to evapotranspiration, heat island effect, drought, forest fire, etc.

5. Conclusion

It has been challenging to obtain high spatial and temporal resolution LST for heterogeneous areas. To solve this problem, a novel downscaling framework based on GTNNWAR with spatio-temporal fused scaling factors is proposed in this paper to downscale MODIS LST from 1 km to 100 m. Moreover, the performance of the downscaling framework was compared with the other five methods in three case study regions and validated by Landsat 8 LST as well as UAV observations. The main contributions of this work are summarized as follows: (1) A new LST downscaling framework is proposed. It consists of a novel neural network model, namely GTNNWAR, which outperformed the other five methods (GTNNWR, GTWR, RF, GWR, and TsHarp) at the boundaries of different land covers and complex terrains. This is due to the optimal estimation of spatio-temporal weights with neural networks and the integration of autocorrelation effect by introducing the

autoregressive model to the regression. Statistical metrics analysis showed that the GTNNWAR framework downscaled results provided improvements with R^2 up to 28% over its compatriots and were consistent with the UAV-observed LST with R^2 over 0.63. (2) The framework first integrates a spatio-temporal fusion mechanism into the GTNNWAR in such a way that the time-series of scaling factors can effectively be reconstructed for the downscaling model to produce high accuracy output. The fusion outputs have high spatio-temporal consistency (R^2 over 0.98) with the referenced image. (3) Within the framework, a region-adaptive model parameterization scheme is proposed to select the optimal sets of scaling factors with strong interpretability of LST. The regression coefficients show significant non-stationarity, which facilitates the analysis of the bio-geophysical mechanism for the thermal environment. Our results demonstrate the innovation for developing a neural network model and the integration of fusion models to solve a practical LST downscaling problem. As more data is retrieved from satellites and ground measurements, image downscaling should be further explored based on the proposed framework in the future.

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 supported by the Key Research and Development Program of Guangdong Province (No. 2020B0101130009), the National Key Research and Development Program of China (No. 2017YFB0504103), the National Key R&D Program of China (No. 2019YFC1510400), the Key Science and Technology Planning Project of Guangdong Province (No. 2015B010104003), and the Key Science and Technology Planning Projects of Guangzhou (No. 201604046007).

References

- Agam, N., Kustas, W.P., Anderson, M.C., Li, F., Neale, C.M.U., 2007. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sens. Environ.* 107 (4), 545–558.
- Awange, J.L., Mpelasoka, F., Goncalves, R.M., 2016. When every drop counts: Analysis of Droughts in Brazil for the 1901–2013 period. *Sci. Total Environ.* 566, 1472–1488.
- Bindhu, V.M., Narasimhan, B., Sudheer, K.P., 2013. Development and verification of a non-linear disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale of Landsat thermal data to estimate evapotranspiration. *Remote Sens. Environ.* 135, 118–129.
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sens. Environ.* 91 (3–4), 332–344.
- Deilami, K., Kamruzzaman, M., Liu, Y., 2018. Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. *Int. J. Appl. Earth Obs. Geoinf.* 67, 30–42.
- Dominguez, A., Kleissl, J., Luvall, J.C., Rickman, D.L., 2011. High-resolution urban thermal sharpener (HUTS). *Remote Sens. Environ.* 115 (7), 1772–1780.
- Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Li, H., Huang, F., Tamang, S.K., Zhao, L., 2020. Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. *ISPRS J. Photogramm. Remote Sens.* 169, 44–56.
- Duan, S.-B., Li, Z.-L., Li, H., Götsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., Coll, C., 2019. Validation of Collection 6 MODIS land surface temperature product using in situ measurements. *Remote Sens. Environ.* 225, 16–29.
- Estoque, R.C., Ooba, M., Seposo, X.T., Togawa, T., Hijioka, Y., Takahashi, K., Nakamura, S., 2020. Heat health risk assessment in Philippine cities using remotely sensed data and social-ecological indicators. *Nat. Commun.* 11, 1–12.
- Fu, P., Weng, Q., 2016. Consistent land surface temperature data generation from irregularly spaced Landsat imagery. *Remote Sens. Environ.* 184, 175–187.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 44, 2207–2218.
- Gevrey, M., Dimopoulos, I., Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol. Modell.* 160 (3), 249–264.
- Ghosh, A., Joshi, P.K., 2014. Hyperspectral imagery for disaggregation of land surface temperature with selected regression algorithms over different land use land cover scenes. *ISPRS J. Photogramm. Remote Sens.* 96, 76–93.
- Guo, L.J., Moore, J.M., 1998. Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery. *Int. J. Remote Sens.* 19 (13), 2477–2491.
- Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sens. Environ.* 178, 127–141.
- Keramitsoglou, I., Kiranoudis, C.T., Weng, Q., 2013. Downscaling geostationary land surface temperature imagery for urban analysis. *IEEE Geosci. Remote Sens. Lett.* 10 (5), 1253–1257.
- Kim, J.H., 2019. Multicollinearity and misleading statistical results. *Korean J. Anesthesiol.* 72 (6), 558–569.
- Kustas, W.P., Norman, J.M., Anderson, M.C., French, A.N., 2003. Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. *Remote Sens. Environ.* 85 (4), 429–440.
- Li, W., Ni, L., Li, Z.-L., Duan, S.-B., Wu, H., 2019. Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12 (7), 2299–2307.
- Li, Z., Gu, X., Dixon, P., He, Y., 2008. Applicability of land surface temperature (LST) estimates from AVHRR satellite image composites. northern Canada.
- Liang, S., 2001. Narrowband to broadband conversions of land surface albedo I: Algorithms. *Remote Sens. Environ.* 76 (2), 213–238.
- Miles, J., 2014. Tolerance and variance inflation factor. *Wiley StatsRef Stat. Ref. Online.*
- Mo, Y., Xu, Y., Chen, H., Zhu, S., 2021. A Review of Reconstructing Remotely Sensed Land Surface Temperature under Cloudy Conditions. *Remote Sens.* 13, 2838.
- Oliver, M.A., Webster, R., 1990. Kriging: a method of interpolation for geographical information systems. *Int. J. Geogr. Inf. Syst.* 4 (3), 313–332.
- Peng, Y., Li, W., Luo, X., Li, H., 2019. A geographically and temporally weighted regression model for spatial downscaling of MODIS land surface temperatures over urban heterogeneous regions. *IEEE Trans. Geosci. Remote Sens.* 57 (7), 5012–5027.
- Qin, Z., Karnieli, A., Berliner, P., 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *Int. J. Remote Sens.* 22 (18), 3719–3746.
- Shiff, S., Helman, D., Lensky, I.M., 2021. Worldwide continuous gap-filled MODIS land surface temperature dataset. *Sci. Data* 8, 1–10.
- Takaku, J., Tadono, T., Tsutsui, K., 2014. GENERATION OF HIGH RESOLUTION GLOBAL DSM FROM 462 ALOS PRISM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 2.
- Tonooka, H., 2005. Resolution enhancement of ASTER shortwave and thermal infrared bands based on spectral similarity. In: *Image Processing and Pattern Recognition in Remote Sensing II. International Society for Optics and Photonics*, pp. 9–19.
- Tran, D.X., Pla, F., Latorre-Carmona, P., Myint, S.W., Caetano, M., Kieu, H.V., 2017. Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS J. Photogramm. Remote Sens.* 124, 119–132.
- Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* 34, 892–905.
- Wang, J., Schmitz, O., Lu, M., Karssenber, D., 2020. Thermal unmixing based downscaling for fine resolution diurnal land surface temperature analysis. *ISPRS J. Photogramm. Remote Sens.* 161, 76–89.
- Weng, Q., Fu, P., Gao, F., 2014. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sens. Environ.* 145, 55–67.
- Wu, J., Zhong, B.o., Tian, S., Yang, A., Wu, J., 2019. Downscaling of urban land surface temperature based on multi-factor geographically weighted regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12 (8), 2897–2911.
- Wu, P., Shen, H., Zhang, L., Götsche, F.-M., 2015. Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature. *Remote Sens. Environ.* 156, 169–181.
- Wu, S., Wang, Z., Du, Z., Huang, B.o., Zhang, F., Liu, R., 2021. Geographically and temporally neural network weighted regression for modeling spatiotemporal non-stationary relationships. *Int. J. Geogr. Inf. Sci.* 35 (3), 582–608.
- Xu, S., Zhao, Q., Yin, K., He, G., Zhang, Z., Wang, G., Wen, M., Zhang, N., 2021. Spatial Downscaling of Land Surface Temperature Based on a Multi-Factor Geographically Weighted Machine Learning Model. *Remote Sens.* 13, 1186.
- Yang, J., Zhou, M., Ren, Z., Li, M., Wang, B., Liu, D.L., Ou, C.-Q., Yin, P., Sun, J., Tong, S., Wang, H., Zhang, C., Wang, J., Guo, Y., Liu, Q., 2021. Projecting heat-related excess mortality under climate change scenarios in China. *Nat. Commun.* 12 (1).
- Yang, Y., Li, X., Pan, X., Zhang, Y., Cao, C., 2017. Downscaling land surface temperature in complex regions by using multiple scale factors with adaptive thresholds. *Sensors* 17, 744.
- Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sens. Environ.* 172, 165–177.