



Contents lists available at ScienceDirect

# Remote Sensing Applications: Society and Environment

journal homepage: [www.elsevier.com/locate/rsase](http://www.elsevier.com/locate/rsase)

## Considering geographical spatiotemporal attributes for seamless air temperature data fusion with high accuracy

Tailai Huang<sup>a,b,c</sup>, Xiang Zhang<sup>a,b,c,d,\*</sup>, Berhanu Keno Terfa<sup>e</sup>, Won-Ho Nam<sup>f</sup>, Xihui Gu<sup>g,h,i</sup>, Jiangyuan Zeng<sup>d</sup>, Xu Zhang<sup>a,b,c</sup>, Shuzhe Huang<sup>j</sup>, Wenying Du<sup>a,b</sup>, Chao Wang<sup>j</sup>, Jian Yang<sup>k</sup>, Peng Wang<sup>l</sup>, Wenkui Wu<sup>a</sup>, Nengcheng Chen<sup>a,b</sup>, Dev Niyogi<sup>m</sup>

<sup>a</sup> National Engineering Research Center of Geographic Information System, School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan, 430074, China

<sup>b</sup> Hubei Luojia Laboratory, Wuhan, 430079, China

<sup>c</sup> Songshan Laboratory, Zhengzhou, 450046, China

<sup>d</sup> State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100101, China

<sup>e</sup> Institute of Geophysics, Space Science and Astronomy, Addis Ababa University, Addis Ababa, Ethiopia

<sup>f</sup> School of Social Safety and Systems Engineering, Institute of Agricultural Environmental Science, National Agricultural Water Research Center, Hankyong National University, Anseong, Republic of Korea

<sup>g</sup> Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan, 430074, China

<sup>h</sup> Centre for Severe Weather and Climate and Hydro-geological Hazards, Wuhan, 430074, China

<sup>i</sup> School of Geography and the Environment, University of Oxford, Oxford, OX1 3QY, UK

<sup>j</sup> State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University, Wuhan, 430079, China

<sup>k</sup> School of Geospatial Information, Information Engineering University, Zhengzhou, 450052, China

<sup>l</sup> GAEA Space Time Co., Ltd, Wuhan, 430079, China

<sup>m</sup> Department of Earth&Planetary Sciences, Jackson School of Geosciences, and Maseeh Department of Civil, Architecture, and Environmental Engineering, University of Texas at Austin, Austin, TX, 78712, USA

### ARTICLE INFO

#### Keywords:

Temperature fusion

Deep learning

Spatio-temporal factor

Wuhan metropolitan area

### ABOUT

High-resolution, high spatiotemporal continuity, and high-precision temperature data (3H Ta) are essential for understanding local to global climate change and for studying urban heat conditions. Our prior research demonstrated the effectiveness of combining deep learning and point-to-surface scaling to generate 3H Ta. However, the accuracy of the 3H Ta data fusion methods is influenced by the geospatial attributes of temperature itself. In this study, we investigated the effects of geographical spatio-temporal factors on the three aspects of temperature fusion, i.e., data input, deep learning-based temperature fusion, and fused temperature error calibration. The results showed that temperature data had strong geospatial attributes, exhibiting spatiotemporal autocorrelation at the microscale and different clustering characteristics in different macro spatial regions. After incorporating geospatial autocorrelation factors into the temperature fusion model, the  $R^2$  was 0.995, the RMSE was 0.697 °C, and the MAE was 0.527 °C after 10-fold cross-validation. Compared with the model that does not consider spatio-temporal factors, RMSE and

\* Corresponding author. National Engineering Research Center of Geographic Information System, School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan, 430074, China.

E-mail addresses: [1202121936@cug.edu.cn](mailto:1202121936@cug.edu.cn) (T. Huang), [zhangxiang76@cug.edu.cn](mailto:zhangxiang76@cug.edu.cn) (X. Zhang), [berhanu.keno@aau.edu.et](mailto:berhanu.keno@aau.edu.et) (B.K. Terfa), [wonho.nam@hknu.ac.kr](mailto:wonho.nam@hknu.ac.kr) (W.-H. Nam), [guxihui421@163.com](mailto:guxihui421@163.com) (X. Gu), [zengjy@radi.ac.cn](mailto:zengjy@radi.ac.cn) (J. Zeng), [zhangxufm@126.com](mailto:zhangxufm@126.com) (X. Zhang), [duwenying@cug.edu.cn](mailto:duwenying@cug.edu.cn) (W. Du), [szhuang@whu.edu.cn](mailto:szhuang@whu.edu.cn) (C. Wang), [jian.yang@tum.de](mailto:jian.yang@tum.de) (J. Yang), [i3dgis@qq.com](mailto:i3dgis@qq.com) (P. Wang), [349974434@qq.com](mailto:349974434@qq.com) (W. Wu), [chennengcheng@cug.edu.cn](mailto:chennengcheng@cug.edu.cn) (N. Chen), [happy1@utexas.edu](mailto:happy1@utexas.edu) (D. Niyogi).

<https://doi.org/10.1016/j.rsase.2023.101135>

Received 28 July 2023; Received in revised form 3 December 2023; Accepted 27 December 2023

Available online 30 December 2023

2352-9385/© 2024 Elsevier B.V. All rights reserved.

MAE are reduced by 68% and 73%, respectively. The use of geographical spatio-temporal difference analysis (GSTDA) error correction combining spatio-temporal factors compensated for temperature underestimation or overestimation at specific times or locations. After calibrating for fusion temperatures at the four validation sites, the RMSE and MAE decreased from 0.75 °C to 0.64 °C–0.69 °C and 0.58 °C, with RMAE and MAE decreasing by 9.37% and 10.15%, respectively. Finally, we generated 500 m daily 3H Ta data for Wuhan Metropolitan Area in 2019. The result were expended to Austin, Texas and Los Angeles, in USA. Our research results and comparative analysis confirm the necessity of considering geospatial and temporal factors in temperature fusion models, which helps generate 3H Ta.

## 1. Introduction

Air temperature (Ta) is a crucial climate parameter for the hydrological process, as well as in the urbanization climate research (Chen et al., 2017; Zhang et al., 2021). Changes in near-surface air temperature are closely linked to human activities, climate change, and crop growth(Immerzeel et al., 2014; Tewari et al., 2023). There are two main methods for obtaining Ta data. The first method is through ground monitoring stations, which can provide accurate and continuous air temperature data over time(Xu et al., 2021; Zhang et al., 2018). However, due to the limited number of stations, the spatial distribution is often sparse and uneven, leading to an inability to reflect the complex spatio-temporal patterns of air temperature accurately(Li et al., 2018). The second method involves obtaining large-scale assimilation or reanalysis data products, such as the Global Land Data Assimilation System (GLDAS) and the ERA-5 reanalysis product(Decker et al., 2012; Hersbach et al., 2020; Kobayashi et al., 2015). For example, ERA-5 is a dataset produced by assimilating satellite and In-situ data with spatial resolution of 0.1°, which are critical for reconstructing historical climates and predicting the future(Dutra et al., 2020; Sun et al., 2019). However, the coarse resolution and low accuracy of the reanalysis datasets limit the impact assessment and decision by scale differences (Zhang et al. 2015, 2017). Thus, neither method currently provides temperature products with high resolution, high spatio-temporal continuity (i.e., spatially seamless and temporally continuous), and high accuracy which called 3H Ta(Huang et al., 2022). 3H Ta is helpful for the study of urban heat island effect and evapotranspiration(Zhang et al., 2019; Liu and Niyogi, 2020).

Combining the advantages of ground monitoring sites and assimilated data can be achieved through downscaling and data fusion techniques(Huang et al., 2022). This is currently a common method to generate 3H Ta data. Downscaling provides the small scale, high-resolution data needed to assess the impact of regional climate (Chen, et al., 2017; Xu et al., 2020; Singh et al., 2023). This can be achieved through dynamic or statistical downscaling techniques. Dynamic downscaling involves coupling high-resolution regional climate models with numerical weather models to simulate high-resolution climate variables. For instance, Zhou et al., (2021) conducted dynamic downscaling of the daily mean, maximum and minimum temperatures in the Canadian prairie provinces, and the predicted mean temperatures were significantly enhanced. In contrast, statistical downscaling takes advantage of empirical connections between local and large scales, thereby reducing computational costs. For example, Li et al. (2020) studied the performance of multiple machine learning methods and linear statistical methods on long-term daily temperature downscaling in the development of multi-model integration. The results showed that machine learning and statistical techniques performed well in the multi-model integration of daily temperature downscaling, with similar performance and high accuracy, but had difficulties in predicting temperature extremes below -10 °C and above 20 °C. Xu et al., (2021) compared 10 machine learning algorithms in monthly temperature estimation of Qinghai-Tibet Plateau using 11 variables extracted from MODIS data, DEM data, and topographic index data, demonstrating the great potential of machine learning algorithms in temperature estimation. Furthermore, ground observation data is considered the most accurate data(Shen et al., 2020). our idea is to use site data for deviation calibration and generate 3H Ta data by fusing multi-source point-surface heterogeneous monitoring data through deep learning. The deep learning algorithms utilized in this process has potential to simulate the complex nonlinear relationship between temperature and different predictors. Li et al., (2021) generated 1 km resolution Ta by deep learning downscaling framework. While deep learning algorithms capture nonlinear relationships, they still have some limitations when it comes to generating 3HTa. This is mainly due to the lack of consideration of the spatio-temporal heterogeneity of temperature, and its dependence on local features.

According to the second law of geography, Ta is related to meteorological conditions, terrain, spatial location and time. It has geographical spatiotemporal property and shows spatial local heterogeneity on the macro level and correlation on the micro level (Abowarda et al., 2021). The nearer observations are more informative than further ones(Amiri et al., 2009). Although statistical methods are usually used to describe the numerical relationship, they often ignore the geographical spatiotemporal characteristics of air temperature, resulting in insufficient precision performance in some time or space(Wang et al., 2022; Wang et al., 2022). Despite their strong data mining capabilities, most AI methods used in air temperature fusion ignore its key spatio-temporal characteristics (Reichstein et al., 2019; Wang et al., 2023). Therefore, it is worthwhile to investigate whether integrating spatio-temporal information into the fusion model improve its accuracy, given the geographical and spatio-temporal heterogeneity of temperature. In addition, research on fusing other environmental variables proposes input factors that include spatio-temporal information or integrate such information into the fusion model, ultimately improving data fusion accuracy. For instance, Liu et al., (2017) introduced the spatiotemporal autocorrelation factor into the deep belief network model to describe the spatiotemporal autocorrelation of PM2.5, successfully estimating daily PM2.5 values. Additionally, Wei et al., (2019) proposed a method that simultaneously introduced spatiotemporal and temporal information through the Spatiotemporal Extreme Stochastic Tree (STET) model, which fused site data and satellite remote sensing data to generate PM2.5 data. The root mean square error decreased by 30%. When considering temporal and spatial information in the multi-source fusion of other environmental variables.

Therefore, the study analyzed the effects of geographic spatiotemporal factors in the three phases of temperature fusion (input, fusion and calibration). A temperature fusion model is developed combining point-surface data fusion and geographic spatio-temporal factors to generate temperature products with high resolution (500m), high spatio-temporal continuity (no cloud and daily) and high precision ( $\text{RMSE} < 0.7^\circ\text{C}$ ). The results of this study have value for studying the influence of urbanization on climate and urban heat island effect.

## 2. Study area and data

### 2.1. Study area

There is increasing demand for high-quality temperature data in cities but it is difficult to obtain. We focused our research on the typical urban agglomeration of the Wuhan and its eight adjacent cities (Wuhan Metropolitan Area, WMA) as the study area (Gao et al., 2020). The distribution and topography of the city are shown in Fig. 1. In addition, due to social and economic activities caused by the existence of infrastructure such as buildings and roads and human activities, the temperature in urban areas is obviously more heterogeneous in geographical space and time (Zhao et al., 2014). We chose cities and adjacent areas, as the Ta datasets allow the assessment of the impact of urbanization on urban climate by comparing cities and circumjacent areas (e.g. UHI) (Schwarz et al., 2011; Venter et al., 2020; Huang et al., 2023). WMA has experienced dramatic urban expansion over the past 10 years. Wang et al., (2007) believe that urbanization and developed industrial production activities have caused the emergence of the UHI effect. This makes the air temperature in the region have strong temporal and spatial variability. WMA has a north subtropical monsoon climate with an annual average temperature of  $15\text{--}17^\circ\text{C}$ . The terrain is diverse, including mountains, hills and plains (Amiri, et al., 2009; Kamath et al., 2023). The complex terrain also makes the temperature spatial complexity. Therefore, it is appropriate and meaningful to study the influence of geographical and temporal factors on WMA high-quality temperature data fusion.

### 2.2. Data

We used four kinds of data to generate 3H Ta data, including remote sensing products, reanalysis products, in-situ measurements and other auxiliary data.

Remote sensing products including Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data and Shuttle Radar Topography Mission (SRTM) digital elevation products (Lin et al., 2012; Meyer et al., 2016). Vegetation indicators are used to monitor the state of the global vegetation environment and show land cover change (Marzban et al., 2018). These data can be used as input data for global or regional climate and are essential parameters in the Temperature Vegetation Index (TVX), a common temperature inversion model. The degree of variation of NDVI and EVI over a period of time is smaller than that of land surface temperature (LST) (Huete et al., 2002; Marzban et al., 2018; Meyer et al., 2016). We used NDVI and EVI in the MODIS vegetation index product (MOD-13Q1) as the predictors of the data fusion model with a spatial resolution of 250 m and a temporal resolution of 16 days. Elevation is one of the main factors causing spatial regional heterogeneity of temperature. The DEM comes from SRTM with a spatial resolution of 30m.

Reanalysis product from ERA5-land is a  $0.1^\circ$  spatial resolution data set, in which we selected Ta, albedo, wind speed and soil moisture as the predictors for this study (Liu et al., 2012; Zhang et al., 2012). Previous studies have shown that LST has a strong positive correlation with Ta. However, MODIS LST products do not provide long-term spatial seamless LST data (Sun, et al., 2019). Hence, we adopted a 1 km all-weather thermal and reanalysis integrated Medium Resolution Spatial Seamless (TRIMS) LST product (Zhang et al., 2019, 2021).

In-Situ Measurements include the daily near-surface air temperature data for 2019 from the China Meteorological Data Service Center (CMDC). We selected data from 36 sites in the Wuhan metropolitan area and removed outliers.

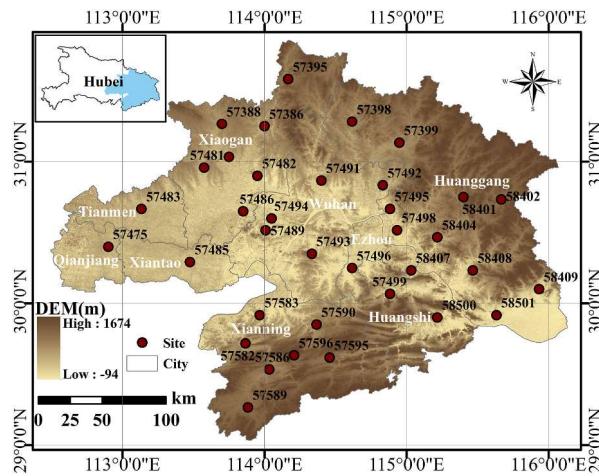


Fig. 1. Geographic location of the 36 air temperature in-situ measurement sites in the WMA.

Other auxiliary data: We chose each day in 2019 as the time data and the spatial coordinates (i.e., latitude and longitude) as the spatial data. Together with in-situ temperature and ERA5 temperature, we calculate geo-spatiotemporal autocorrelation factors.

### 3. Methods

#### 3.1. Methodology

The goal of this paper is to explore the influence of geographical spatio-temporal factors on the fusion of temperature multi-source data, the overall methodology is shown in Fig. 2. We will evaluate the impact of geographic spatiotemporal factors on the accuracy of multi-source data fusion in three stages. Firstly, in the data input stage, we conducted micro-scale spatial autocorrelation analysis and macro-scale spatial heterogeneity analysis of annual mean temperature in the WMA region using Moran's index to confirm the presence of spatial correlation in temperature. After preprocessing the input data through a series of steps such as reprojection, resampling, and extracting station values, we carried out correlation analysis between LST and station temperature with a strong linear relationship at different times or locations. Secondly, we added geospatial autocorrelation factors that consider geographic spatiotemporal factors as inputs to the multi-source temperature fusion model and compared it with the initial model that did not consider geographic spatiotemporal factors. Lastly, we used two methods to correct errors in the fused temperature and compared them. One of the methods considered geographic spatiotemporal factors. We selected the best method to correct the fused results and finally obtained 3H Ta data.

#### 3.2. Using Moran's I to verify the spatial autocorrelation of air temperature

Global Moran's I is a calculation method for verifying spatial autocorrelation, which reflects the degree of dispersion or aggregation of data(Zhang and Lin, 2016). The global index can determine whether clusters or outliers appear in the space, but it does not know where the clusters or outliers occur. The calculation is as follows.

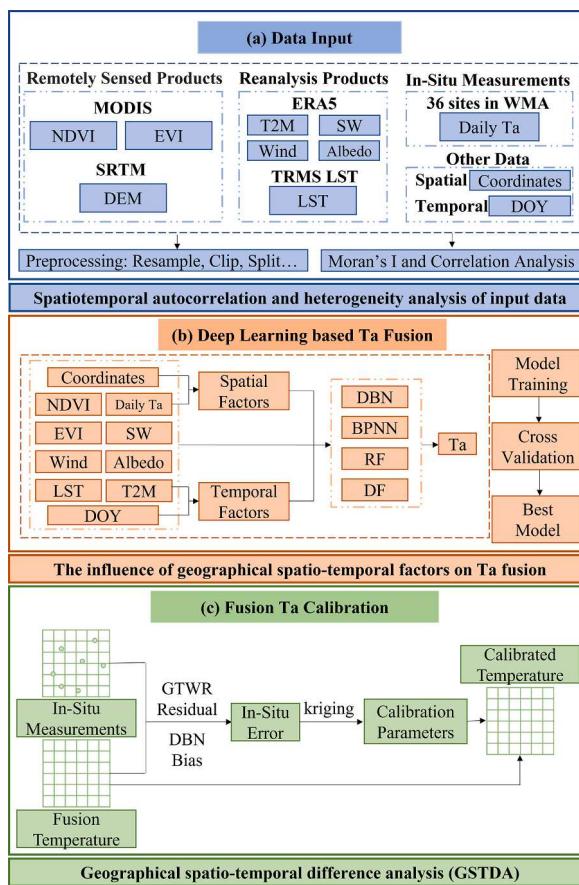


Fig. 2. The methodology of this study. The three steps including (a) data input, (b) deep learning based Ta fusion, and (c) Fusion Ta calibration.

$$I = \frac{N}{W} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (1)$$

Where,  $x$  is the data,  $w$  is the spatial weight, and  $W$  is the sum of all spatial weights.

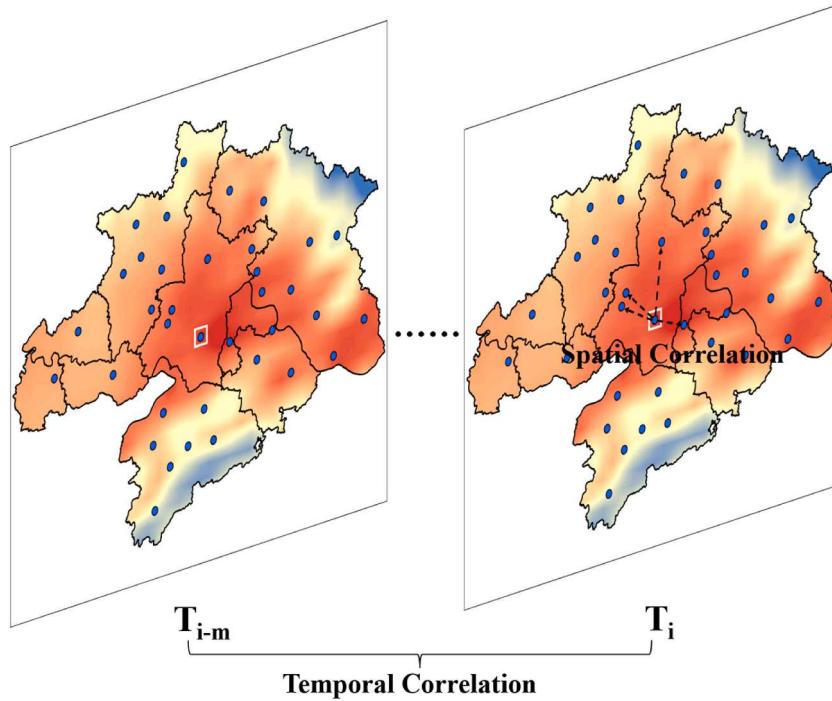
Since the global Moran's I does not reflect where the aggregation occurs, we use the local Moran's I to analyze the heterogeneity of the spatial distribution of air temperature(Chen and Shen, 2020). Local Moran's I assesses where outliers appear or where clusters appear. The local Moran's I range between  $-1.0$  and  $+1.0$ . We will analyze the above two indices in ArcGIS on the ERA5-Land temperature of the Wuhan metropolitan area to evaluate its spatial autocorrelation.

### 3.3. Calculating the spatio-temporal autocorrelation factor

In addition to the input of the basic station air temperature and its related variable data, this study further introduced the spatiotemporal autocorrelation factor based on the idea of geographic spatiotemporal autocorrelation as an additional model input variable(Liu, et al., 2017). In general, there is some temporal and spatial autocorrelation among environmental variables. The expression of the temperature autocorrelation input factor to be constructed here is as follows and is shown in Fig. 3. Among them,  $ds$  and  $dt$  represent the space and time distance,  $m$  and  $n$  represent the number of adjacent days and adjacent grid points, respectively. In this study, both  $m$  and  $n$  are taken as 4.  $S$  and  $T$  represent the spatial and temporal autocorrelation factors of soil moisture, respectively.

$$S - Ta = \frac{\sum_{i=1}^n ws_i V_i}{\sum_{i=1}^n ws_i} \quad ws_i = \frac{1}{ds_i^2} \quad (2)$$

$$T - Ta = \frac{\sum_{j=1}^m wt_i V_i}{\sum_{j=1}^m wt_i} \quad wt_i = \frac{1}{dt_i^2} \quad (3)$$



**Fig. 3.** Schematics for calculating spatio-temporal autocorrelation factors. The blue dot represents the site,  $m$  is the number of days, and the sites in the white box are connected to the four nearest sites by dotted lines.

### 3.4. Deep learning based temperature fusion

Deep learning is already widely used in many fields of research. In this study, we chose the DBN model to fuse multi-source data to obtain 3H Ta(Parajuli et al., 2021; Shen et al., 2020). The DBN is composed of multiple layers of restricted Boltzmann machine (RBM) layers and one backpropagation (BP) layer. The complete DBN model includes input layer, output layer, and hidden layers. The training process of DBN can be seen as an effective unsupervised hierarchical way. This study uses ERA-5 temperature (T2m), soil moisture, wind speed, surface albedo, LST, NDVI, EVI, elevation (Ele), spatial coordinates (i.e., longitude and latitude), spatiotemporal autocorrelation factors, and a temporal parameter (i.e., DOY) as predictors for 3H Ta fusion model. The output variable of the model is the average in-situ temperature. The fusion model can be represented as:

$$3H\ Temperature = f(T2m, SW, Wind, Albedo, LST, NDVI, EVI, Ele, Lat, Lon, S - Ta, T - Ta, DOY) \quad (4)$$

where  $f(\cdot)$  is the 3H Ta fusion model.

The study period covers the whole year of 2019 and there are 36 in-situ stations in WMA. Therefore, all meteorological stations form a sample set of  $36 \times 365 \times 13$ . Then, four stations are randomly selected from the 36 stations for validation, and the data from the remaining stations are divided into training and testing sets as an 8:2 ratio. Finally, the selected best-fitting Ta fusion model was applied to the entire image and fused into 500 m 3H Ta.

### 3.5. Bias calibration for downscaled temperature

Due to the spatio-temporal heterogeneity distribution of the model's predictive factors can affect the fusion accuracy. We used two calibration methods: the geographical difference analysis (GDA) and geographical spatio-temporal difference analysis (GSTDA)(Li, et al., 2021). GDA uses the bias of the temperature fusion model itself for error calibration, and the formula is as follows:

$$P_{corrected}(x) = P_{fusion}(x) \times \sum_{i=1}^n \lambda_i \frac{P_{obs}(x_i)}{P_{fusion}(x_i)} \quad (5)$$

where  $P_{corrected}$ ,  $P_{fusion}$  and  $P_{obs}$  represent the calibrated, original fusion, and observed temperature at location  $x$ , respectively.  $\lambda$  is the weight at location  $x$ . Then, the in-situ bias is interpolated onto the entire plane using Kriging interpolation.

GSTDA is a new error calibration method proposed in this study that combines Geographically and Temporally Weighted Regression (GTWR) and considers geographical and temporal factors. The traditional geographically weighted regression (GWR) only considers the spatial information between samples, which is an extension of the ordinary linear regression model(Li et al., 2020; Tang et al., 2022). For a certain variable  $y$  (such as temperature), there are  $n$  sample points in the research area, then the geographic weighted regression equation is defined as:

$$y_i = \beta_{k0} + \sum_{k=1}^n \beta_{ik}(u_i, v_i) x_{ik} + \epsilon_i \quad i = 1, 2, \dots, n \quad (6)$$

Where  $(u_i, v_i)$  is the latitude and longitude coordinates of the  $i^{th}$  sample point,  $\beta_{ik}(u_i, v_i)$  is the  $k$  regression coefficient on the  $k$  sample point, represents the spatial position function,  $\epsilon_i$  is the residual, and  $x_{ik}$  is the  $k$  independent variable at  $(u_i, v_i)$ .

On this basis, geographically weighted regression has been further developed, and spatiotemporal geographically weighted regression is proposed. Compared with traditional GWR, GTWR also considers the time information between data, and its expression is as follows:

$$y_i = \beta_{j0}(u_i, v_i, t_i) + \sum_{k=1}^n \beta_{ik}(u_i, v_i, t_i) x_{ik} + \epsilon_i \quad i = 1, 2, \dots, n \quad (7)$$

Where  $t$  represents the moment of the sample. By establishing the spatiotemporal geographic weighted regression expression of temperature and its related variables, the relationship between different time and space factors on temperature can be better considered.

Through the above steps, the initial rainfall field based on the statistically constrained point-surface fusion model is obtained. Assuming that the final fused data is  $p_f$ , the initial rainfall field data is  $p_s$ , and the monitoring data at the station is  $p_o$ , the relationship between the three is as follows:

$$p_f = p_s + \epsilon(p_o - p_s) \quad (8)$$

Where  $\epsilon(p_o - p_s)$  refers to the residual between the observed value and the initial field value. First, the accuracy of results obtained by GTWR are compared with the accuracy of the model considering geographical and spatial factors, and the influence of spatial and temporal factors on data fusion is analyzed. Then, by calculating the residual error between the observed value at the station and the initial field, the typical interpolation method is used to extend it to the area domain, and finally correct the error of the initial field of temperature to generate high-quality temperature data.

### 3.6. Model evaluation index

To avoid overfitting, the multiple determination coefficient ( $R^2$ ), mean absolute error (MAE), root mean square error (RMSE) and deviation from the true value were calculated comparisons are made to evaluate model accuracy, as shown in [Table 1](#).

## 4. Results

### 4.1. Geographical spatio-temporal analysis of temperature and LST

We first used the global Moran's I to analyze the spatial autocorrelation of the temperature distribution in ERA5-Land, and the results are shown in the [Fig. 4](#). It can be seen from the result [Fig. 4](#) that the global Moran's I was 0.965, the z-score is 2778.956, and the p-value is 0, indicating that there is a significant spatial autocorrelation in the temperature of the Wuhan metropolitan area, that is, the temperature in this area is clustered as high with high, and low with low. After confirming the existence of spatial autocorrelation in temperature, the geographical heterogeneity of environmental variables such as temperature. Therefore, local spatial autocorrelation analysis is required. We used the local Moran's I to obtain the aggregation of temperature in the Wuhan metropolitan area, and the results are shown in [Fig. 4](#). It can be seen that the temperature shows a high-high aggregation in the middle of the urban area, and the significance test reaches 99%, while there are obvious low-low aggregation and cold spot aggregation in the north and south of the urban area. Combined with the analysis of the topography of the WMA, the high with high agglomerations are mainly distributed in some low areas such as Wuhan, Ezhou, and Xiantao, while the areas with low with low agglomerations are distributed in the higher areas of Huanggang and Xianning. Previous studies have shown that the temperature tends to decrease with the increase of the terrain. Therefore, the temperature is similar to the temperature distribution of the WMA. In addition, the degree of industrialization and anthropogenic activities in areas with high and high agglomeration are higher than those in areas with low and low agglomeration. Urbanization has led to the generation of heat island effect, which increases the temperature. These results show that temperature, has obvious spatial autocorrelation, and is affected by geographical heterogeneity in different regions.

Previous studies have shown that in the classical temperature-vegetation index (TVX) inversion model, LST is the most important input parameter for estimating Ta inversion, and it has a strong linear correlation with temperature([Lin et al., 2010](#)). Although there is a strong positive correlation between the LST of the entire sample and in-situ temperature, this relationship may not be stable under different temporal and spatial conditions. To better assess the spatio-temporal relationship between Ta and LST, we tested their correlation in different spatiotemporal contexts. As shown in [Table 2](#), the LST-Ta relationship was significantly affected by different environments, especially in different months. The correlation between LST and Ta is strongest in October, reaching 0.935. However, in January with the lowest temperature and August with the highest temperature, the correlation between LST and Ta is weakest, especially in the hot August, where the correlation is only 0.455. The difference between the highest and lowest correlations reaches 0.5. Overall, the correlation between LST and Ta is greater in low-altitude areas, high latitudes, and seasons with relatively lower temperatures, which may have a significant impact on the accuracy of temperature fusion.

### 4.2. 3H air temperature data fusion with geographical spatio-temporal autocorrelation factor

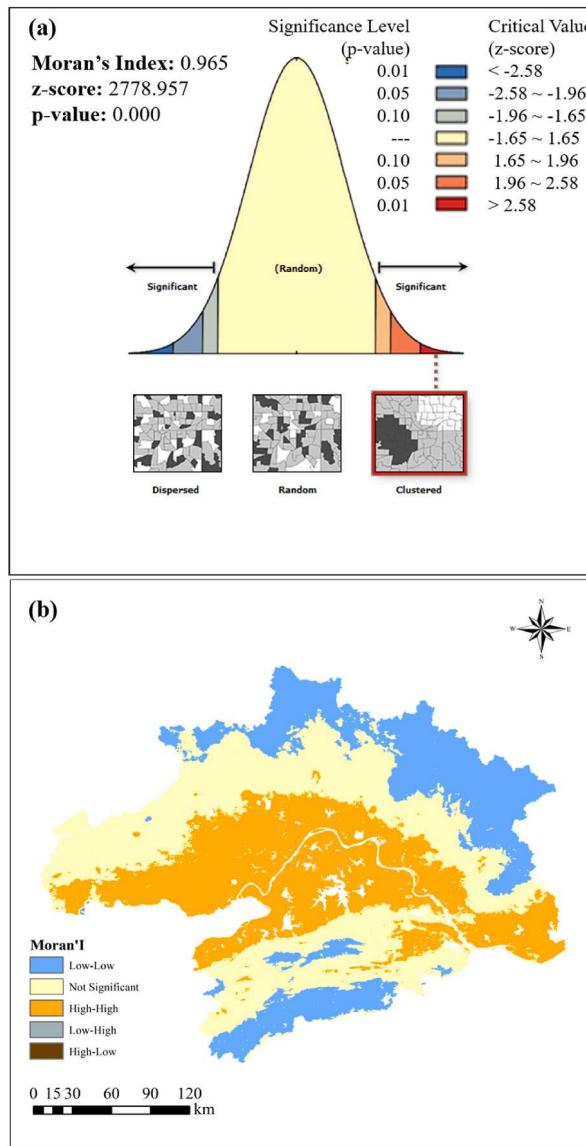
We added spatiotemporal autocorrelation factors to different models to verify the influence of spatiotemporal factors on 3H Ta fusion accuracy. We will use the deep belief network model (DBN), back propagation artificial neural network (BPNN), random forest (RF) compared with four models of Deep Forest (DF). As shown in [Fig. 5](#), the accuracy results of the four models are plotted as scatter density plots. Overall, these four data-driven models have achieved very good results, and all the points are clustered around the diagonal. This shows that the fusion temperature is very close to the true value. The  $R^2$  correlation coefficients of the four models are above 0.992, the RMSE is between 0.597 °C and 0.799 °C, and the MAE is between 0.444 °C and 0.599 °C.

Comparing the DBN and BPNN, the two neural network models; the performance of the DBN model is better than that of the BPNN model, the  $R^2$  correlation coefficient has slightly increased from 0.993 to 0.995, and the RMSE and MAE errors have also decreased from 0.806 °C to 0.608 °C–0.697 °C and 0.527 °C, respectively. First of all, because the structure of BPNN is relatively simple with only three layers, while the structure of DBN is more complex, it initializes itself through unsupervised pre-training. The pre-training

**Table 1**

The equations and detailed information of,  $R^2$ , RMSE, BIAS, MAE and PCC.

Metric	Equation	Unit
$R^2$	$R^2 = 1 - \frac{\sum_{i=1}^n (P_i - M_i)^2}{\sum_{i=1}^n (M_i - \bar{M})^2}$	–
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - M_i)^2}{n}}$	°C
MAE	$MAE = \frac{1}{n} \sum_{i=1}^n  P_i - M_i $	°C
BIAS	$BIAS = \frac{1}{n} \sum_{i=1}^n (P_i - M_i)$	°C
PCC	$PCC = \frac{\sum_{i=1}^n (P_i - \bar{P})(M_i - \bar{M})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2(M_i - \bar{M})^2}}$	–



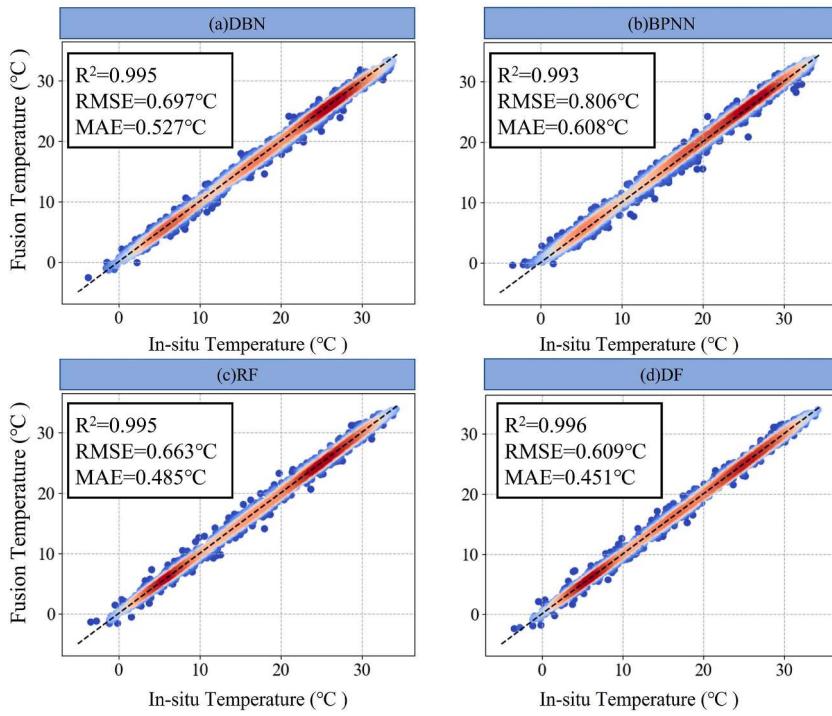
**Fig. 4.** The result of the calculation of the Moran's I. (a) is the result of global Moran's I, (b) is the result of local Moran's I.

**Table 2**

The R between LST and Ta for different months, elevation ranges and latitude ranges.

Month	R	Month	R	Elevation(m)	R	Latitude (°)	R
Jan	0.505	Jul	0.779	<30	0.981	<30	0.934
Feb	0.609	Aug	0.455	30~40	0.969	30~30.5	0.968
Mar	0.874	Sep	0.845	40~90	0.948	30.5~31	0.968
Apr	0.907	Oct	0.935	>90	0.976	>31	0.976
May	0.791	Nov	0.917	/	/	/	/
Jun	0.761	Dec	0.781	/	/	/	/

step of DBN model overcomes the difficult initialization of BPNN model. The shortcomings of falling into local optimum and long training time. The two decision tree models RF and DF have shown excellent performance. Especially for the deep forest model, the  $R^2$  correlation coefficient reached the highest 0.996, and the RMSE and MAE were the lowest 0.609 °C and 0.451 °C. Different from the deep neural network composed of differentiable neurons, the basic component of the deep forest is a non-differentiable decision tree, which makes it have superior performance in modeling tabular data. It has fewer hyperparameters and training efficiency is also better than the neural network model. The results show that the spatio-temporal autocorrelation factor can significantly improve the accuracy of high-quality temperature data fusion, and the deep forest model is suitable for high-quality temperature fusion.

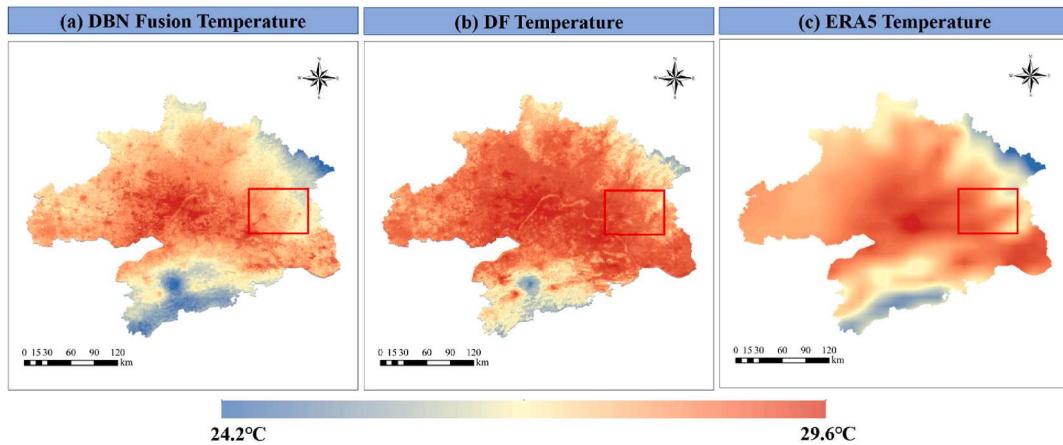


**Fig. 5.** Scatter plot of accuracy results of four fusion models.

Here we compare the fusion temperature obtained by two models: DBN and DF model, with the spatial distribution of T2m. It can be seen from the Fig. 6 that the spatial pattern of fusion temperature data of the two models is similar to that of T2m. The results show that the data-driven fusion model can better capture the spatial distribution of air temperature. In addition, fusion temperatures exhibit more spatial detail than model-based temperatures. For example, in the fusion temperature data, more detailed spatial changes can be observed in the red box area in the Fig. 6 compared to the assimilation temperature product. The fusion temperature obtained by DBN model is closer to the spatial distribution of ERA5 temperature than that obtained by DF model, and the fusion temperature obtained by DF model is higher on the whole. Therefore, we will continue to use DBN model to conduct subsequent experiments. To sum up, a seamless and time-continuous temperature of 500m space is successfully generated after applying the point-surface data fusion method and DBN model and by adding the geographic spatiotemporal autocorrelation factor into it.

#### 4.3. Assessment of the calibrated temperature

Due to imperfect model training, while the previous fusion framework has successfully generated a WMA temperature of 500m, underestimation or overestimation still exists in some time or space. To achieve the goal of 3H, we used two error calibration methods

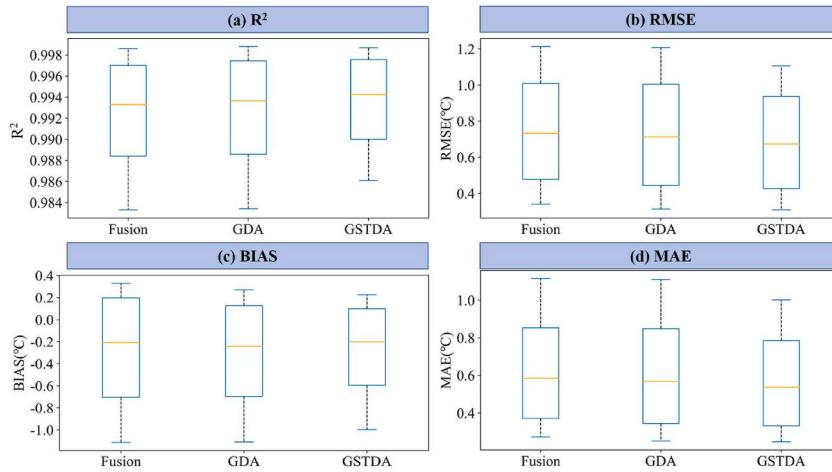


**Fig. 6.** The spatial distribution of DBN fusion temperature, DF fusion temperature and ERA5 temperature for October 1st in WMA. The red box shows the details of the spatial distribution of Ta.

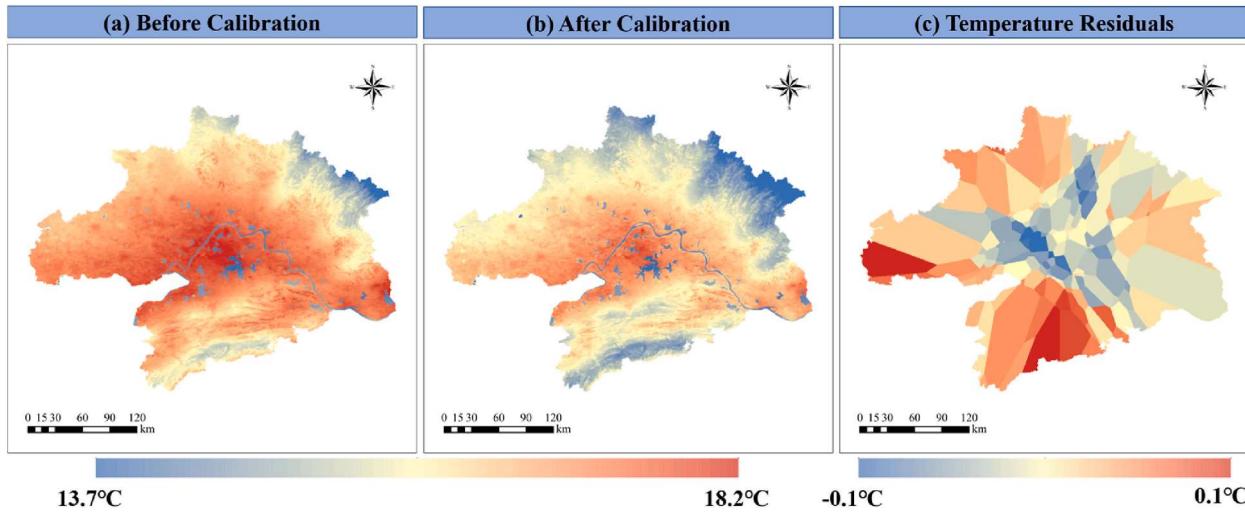
to further improve the accuracy of the fusion temperature: GDA method for DBN model bias and GSTDA method for GTWR model residual. Since temperature can have negative values, we will choose addition instead of comparison in error calibration. The calibration methods for four validation sites were evaluated, the statistical indicators of the four validation sites before and after calibration are shown in Fig. 7. Generally speaking, after bias calibration, the  $R^2$  value increases while RMSE, bias and MAE decrease. Among them, the GSTDA method has a more significant improvement in fusion accuracy than the GDA method. This is because the GSTDA method considers the spatiotemporal changes of temperature, which enables it to better compensate for the underestimation or overestimation of temperature in some time or space. In summary, after GSTDA calibration, the average  $R^2$ , RMSE, and MAE reached 0.993, 0.69 °C, and 0.58 °C respectively, and the errors decreased by 9.37% and 10.15% respectively.

Fig. 8 presents the temperature spatial distribution before and after GSTDA calibration. According to Fig. 8c, the range of residuals between merged temperature and ground observation values is from  $-0.1$  °C to  $0.1$  °C. The results indicate that temperature overestimation is mainly in the areas of Xianning and Qianjiang, while underestimation is mainly in the downtown area of Wuhan.

Based on the above analysis, by combining point-to-area data fusion, deep learning, and bias calibration with consideration of geographic temporal and spatial factors during the process of temperature multi-source data fusion, the best performance was achieved in generating 500 m spatial seamless and time-continuous temperature. After calibration, the average  $R^2$ , RMSE, and MAE reached 0.993, 0.69 °C and 0.58 °C, respectively, ensuring high-precision data and demonstrating the necessity of considering geographic temporal and spatial factors in the fusion process.



**Fig. 7.** The box plots of the statistical metrics for four validation stations.



**Fig. 8.** The spatial distribution of fusion temperature before (a) and after (b) bias calibration. (c) is the distribution of residuals between fusion temperature and in-situ observations.

## 5. Discussions

### 5.1. Evaluation of influence of geographic space-time factors on data fusion accuracy

To evaluate the influence of geographical space-time factors on temperature data fusion, we changed the input variables in the DBN model to conduct comparative experiments (Table 3). We introduce two typical regression models that consider geographical space-time factors, geographical weighted regression model and geographical space-time weighted regression model. To compare with DBN model we have divided into three types, the first is the initial model (DBN-ori), that is, there is no simple spatiotemporal information (longitude, latitude and date) and geographic spatiotemporal autocorrelation factors in the input parameters. The second is only added The simple model (DBN-s) of simple spatio-temporal information. The third is the geographic spatio-temporal model (DBN-st) that includes two kinds of spatio-temporal information in the input. The experimental results are shown in the table below. When the model does not consider geographical space-time factors, it shows the lowest performance, with the lowest  $R^2$  correlation coefficient of 0.983 and the highest RMSE and MAE, which are 1.173 °C and 0.916 °C, respectively. The performance of the GWR model that only considers spatial information is poorer than other models that consider spatiotemporal information, indicating that the temporal heterogeneity of air temperature will affect the fusion accuracy. The DBN-s model, which considers simple spatio-temporal information, performs better than the GTWR model. This is due to the fact that the layer-by-layer pre-training of the deep learning model can better fit the nonlinear relationship and overcome the complexity of temperature prediction. The multicollinearity test in the GTWR model also makes the input selection of the model less, and the prediction ability will also be reduced. From the original model to the geo-intelligent model, the performance has been improved. The DBN-st model with geographic spatiotemporal autocorrelation factors showed the most powerful performance, with the highest  $R^2$  correlation coefficient of 0.995, 0.012 higher than the DBN-ori model, and the lowest RMSE and MAE of 0.697 °C and 0.527 °C, which are 0.476 °C and 0.389 °C lower than the DBN-ori model. It shows that geographical space-time factors impact the generation of high-quality temperature data, and adding geographical space-time autocorrelation factors to the 3H temperature data fusion model based on point-surface heterogeneity in deep learning can greatly improve the accuracy of the model product.

### 5.2. Spatio-temporal evaluation of the fusion accuracy

#### 5.2.1. Temporal evaluation of DBN-st model performance

To investigate the performance of the fusion model after adding spatio-temporal autocorrelation factors in different months, the monthly box plots of the statistical indicators of each verification station, as shown in Fig. 9. Overall, the changes in fusion accuracy vary significantly across different months. Specifically, the temperature fusion error is smaller in summer (June–August) and winter (December–February) than in spring and autumn, with average RMSE values of 0.623 °C and 0.659 °C, respectively. The fusion accuracy is lowest in August, with an average  $R^2$  value of 0.787. Additionally, the uncertainty of temperature fusion results in August is significantly higher than that in other months, with a difference of about 1.157 °C between the maximum and minimum MAE. This finding is consistent with our expectation, as previous studies have shown that the correlation between LST and temperature is lower in hot months than in other months. We also generated a line chart (Fig. 10) depicting the accuracy variation DOY for all stations. From the chart, it is evident that the overall MAE ranges between 0.2 °C and 1.1 °C, with an average value of 0.535 °C. The  $R^2$  values for most DOYs fall within the range of 0.6–1, with only a few days showing lower accuracy, primarily occurring in the autumn and spring periods. In addition, we also generated heatmaps illustrating the temperature deviation between the fused temperature and station temperature for four test stations every five days in 2019. These four test stations are located in the four cardinal directions of the WMA, effectively demonstrating the spatial accuracy of the model. As shown in Fig. 11, the temperature deviations for all four stations remained generally within the range of -1 to 1 throughout the year, confirming the reliability of the fusion model. Station 57595 exhibited larger deviations, attributed to its highest elevation and more complex terrain, consistent with our previous research findings. These findings emphasize the importance of considering temporal factors in the Ta fusion model. To further understand this point, a seasonal temperature fusion model is worth exploring in future work.

To verify the ability to capture the temporal dynamics of fusion temperature, a total of four validation stations were randomly selected. Fig. 12 shows the temporal changes in fusion temperature by the DBN-ori model (blue solid line), on-site ground observations (black dots), and DBN-st model (orange line). Overall, both fusion temperatures are consistent with the on-site measurements. Compared to the on-site observations, the DBN-ori model fusion temperature has more overestimation or underestimation than the DBN-st model fusion temperature, with significant deviations observed at stations 57386 and 57595. This is mainly due to the heterogeneity

**Table 3**

The cross-validation performance of the models.

Model	Input	$R^2$	RMSE (°C)	MAE (°C)
DBN-ori	original	0.983	1.173	0.916
DBN-s	Coordinate	0.988	0.974	0.757
DBN-st	Coordinate DOY	0.995	0.697	0.527
GWR	Geographic spatio-temporal factors	0.985	1.130	0.870
GTWR	Coordinate	0.987	1.067	0.820
	DOY			

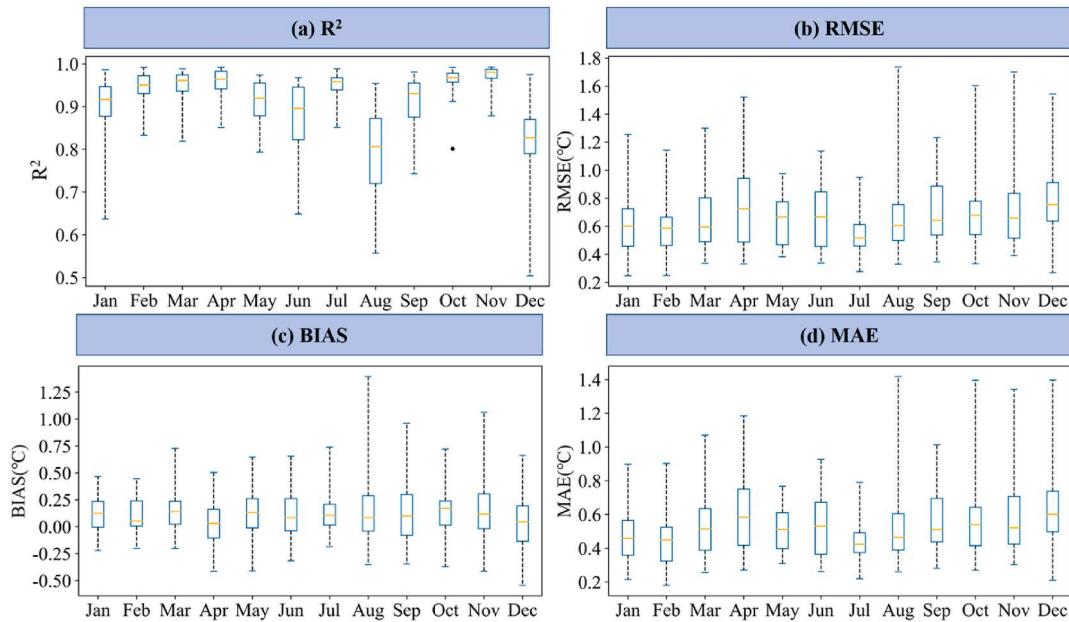


Fig. 9. The temporal variations of  $R^2$ , RMSE, MAR, and bias in each month.

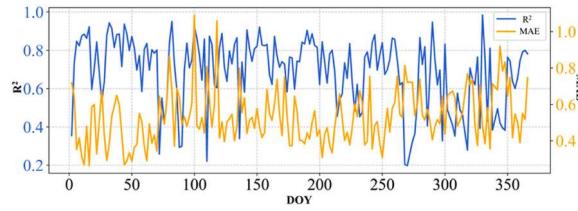


Fig. 10. Line chart illustrating accuracy variation with the day of the year (DOY).

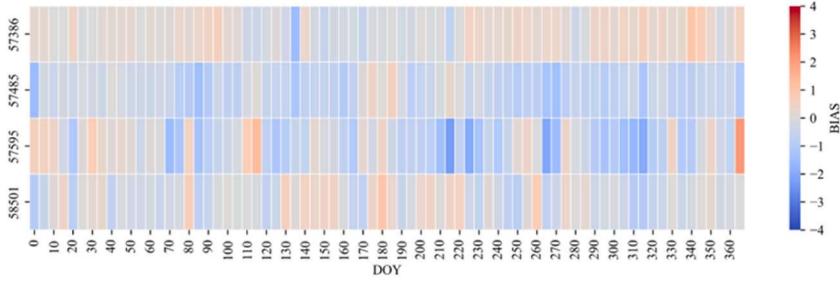
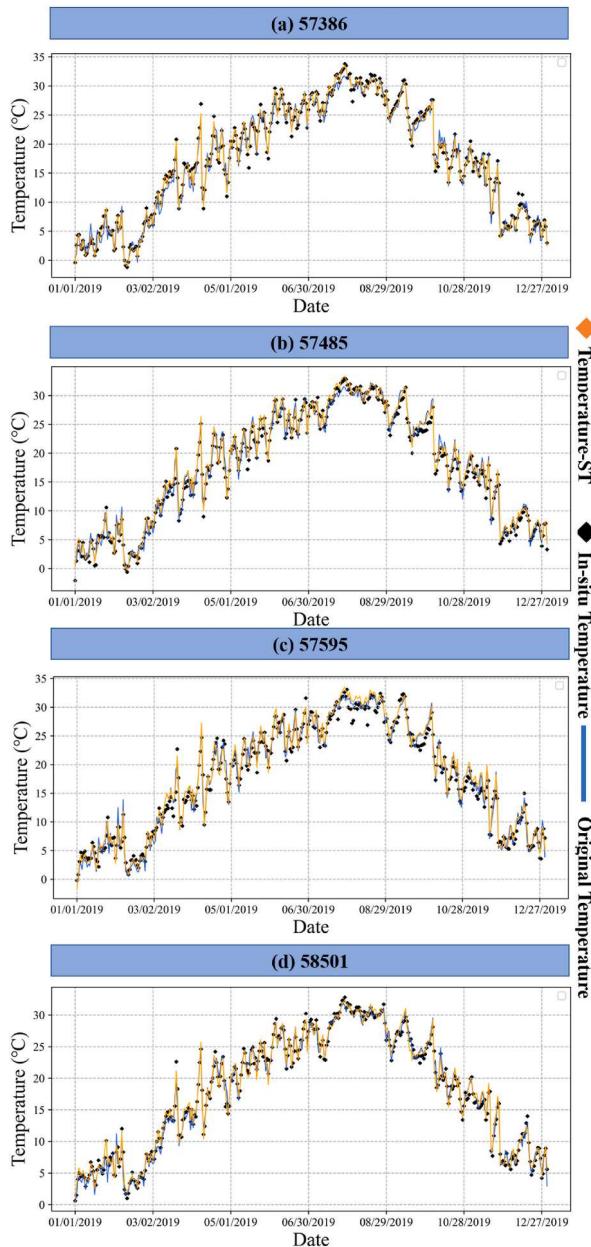


Fig. 11. Heatmap of temperature deviation for four test stations every five days in 2019.

of spatiotemporal scales considered by the DBN-st model. Although there are errors in the fusion of a certain time period with the DBN-st model, it can well capture the overall trend of temperature changes. Based on the results mentioned above, after incorporating the geographical spatial autocorrelation factors into the model, the accuracy of the temperature fusion model has been improved at different times. This further indicates the importance of considering geographical and spatial factors in the multi-source data fusion model for temperature.

#### 5.2.2. Spatial evaluation of DBN-st model performance

To further investigate the spatial patterns and downscaling accuracy of meteorological station statistical indicators, we employed a leave-one-out approach, obtaining four accuracy metrics for the 36 stations. Fig. 13 illustrates the distribution of these indicators, revealing significant spatial variations. Stations in the central and western parts of WMA demonstrate notably higher  $R^2$  values compared to those in the southern region, while RMSE and MAE values are lower in the central and western areas. Station 57389, located in the central part of Xiaogan city, stands out with the highest  $R^2$  value (0.998), along with the lowest RMSE ( $0.363 ^{\circ}\text{C}$ ) and MAE ( $0.291 ^{\circ}\text{C}$ ). This can be attributed to the predominantly flat and open terrain of the Yangtze and Han River alluvial plains in the south-



**Fig. 12.** The temporal variations of the temperature data in 2019 at four randomly selected stations.

ern part of Wuhan. Consequently, stations situated in the low-altitude areas around Wuhan and its surroundings (9–21 m) generally exhibit higher accuracy. On the other hand, Station 57596 in the southern part of Xiantao records the lowest  $R^2$  (0.877) and the highest RMSE ( $3.200\text{ }^\circ\text{C}$ ). This is due to the complex terrain in the southern part of Xiantao, characterized by steep slopes, domes, folds, and faults composed of granite and metamorphic rocks. This complexity contrasts with the flat plains around Wuhan. Previous studies have consistently found lower accuracy in temperature products in areas with significant elevation differences, such as mountains and hills, aligning with our research findings.

Previous studies have shown that land cover type has a significant impact on the relationship between LST and Ta (Wang et al., 2021). The performance of different land surface cover models was compared by calculating accuracy indicators for four validation sites. As shown in Fig. 14a, the DBN-st model performed the best, with lower RMSE and MAE values than DBN-ori for all five different land surface cover types. However, the model errors were different for different land cover types in the temperature fusion models. The difference between the maximum and minimum MAE values for the DBN-st model was about  $0.22\text{ }^\circ\text{C}$ . The MAE for both models was highest on grassland. For sites near cities, construction sites, and forests, the model performance was relatively good. It is worth mentioning that the DBN-st model achieved a significant improvement in fusion temperature accuracy compared to the DBN-ori model at sites close to water sources and grasslands, reducing the MAE from  $0.995\text{ }^\circ\text{C}$  to  $1.030\text{ }^\circ\text{C}$ – $0.390\text{ }^\circ\text{C}$  and  $0.560\text{ }^\circ\text{C}$ , respectively,

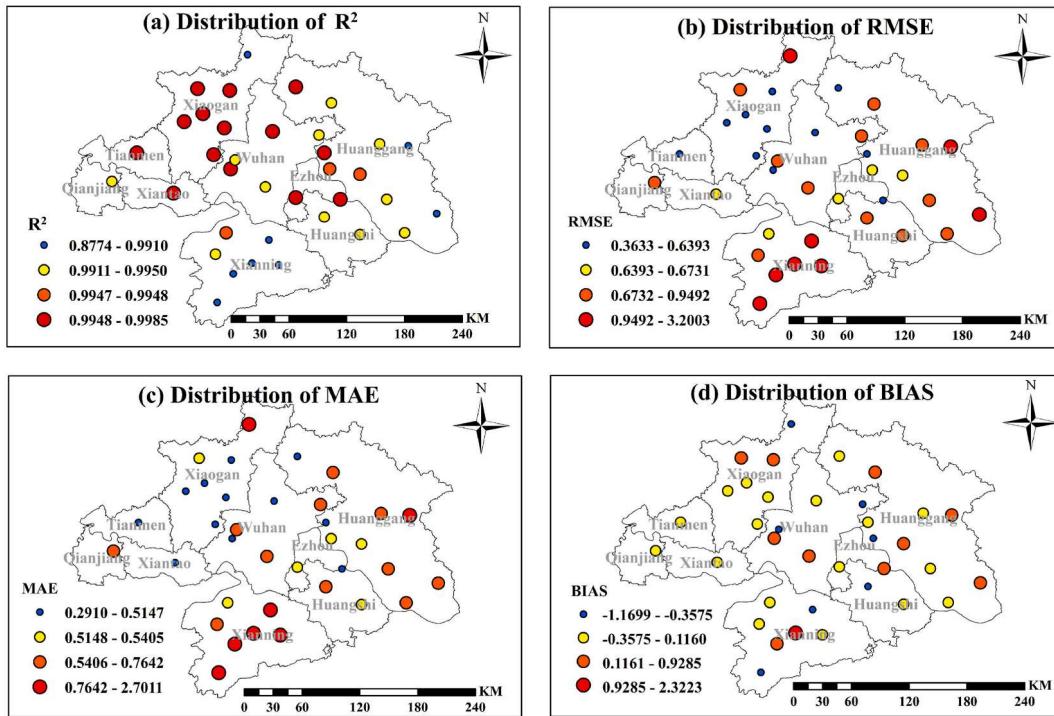


Fig. 13. The spatial distribution of statistical metrics all sites in WMA. (a-d) represent the distribution of  $R^2$ , RMSE, MAE and BIAS.

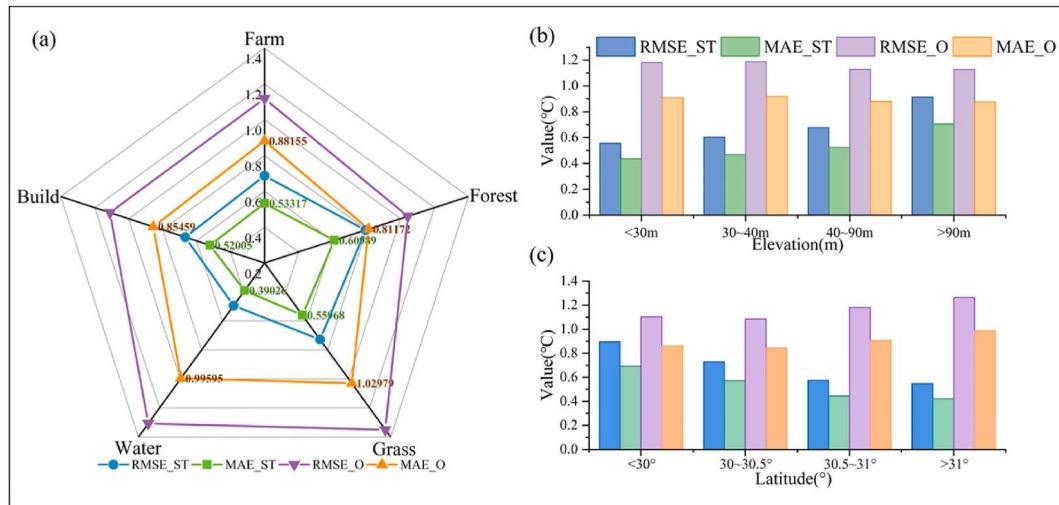


Fig. 14. Model performance for specific (a) land cover types, (b) elevation ranges and (c) latitude ranges.

a reduction of 60% and 47%. In addition to the above land cover types, we also assessed the model's performance in different latitude ranges. As shown in the results in Fig. 14c, the performance of the DBN-st model was better than that of the DBN-ori model across all latitude ranges. The MAE values for the DBN-st model varied between  $0.421\text{ }^{\circ}\text{C}$  and  $0.692\text{ }^{\circ}\text{C}$ . Overall, the higher the latitude, the better the model performance. In areas with latitudes greater than  $31^{\circ}$ , the DBN-st model's accuracy improved the most compared to the DBN-ori model, reducing the MAE from  $0.987\text{ }^{\circ}\text{C}$  to  $0.421\text{ }^{\circ}\text{C}$ , a decrease of 57%.

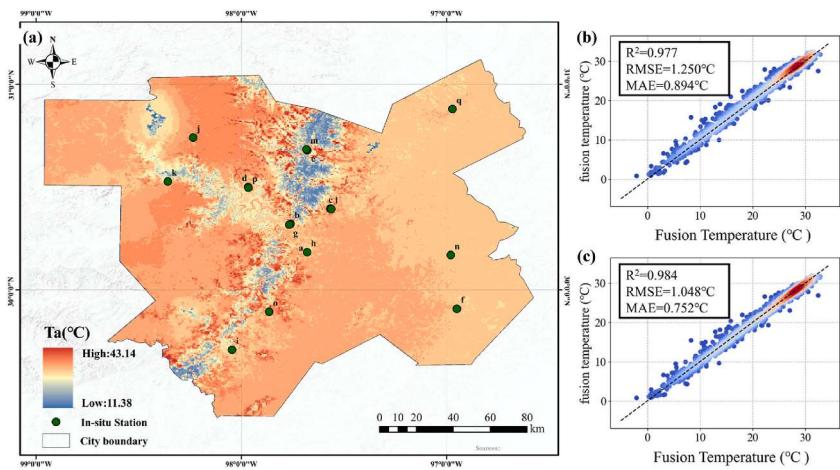
Furthermore, it is generally believed that temperature is highly correlated with altitude. In most cases, temperature decreases as altitude increases. At the same time, the correlation coefficient between LST and Ta varies significantly at different elevation ranges. Therefore, this study also analyzed the performance of the DBN-st and DBN-ori models in different elevation ranges from the perspective of MAE. As shown in Fig. 14b, the model's performance varied at different elevation ranges, with the DBN-st model performing the best, with lower RMSE and MAE values than the DBN-ori model. Generally, the model performance of DBN-st weakened with increasing elevation, with MAE ranging from  $0.436\text{ }^{\circ}\text{C}$  to  $0.706\text{ }^{\circ}\text{C}$ . In the plain area with an elevation below 30m, the DBN-st model

had the most significant improvement compared to the DBN-ori model, reducing the MAE from 0.909 °C to 0.435 °C, a decrease of 52%.

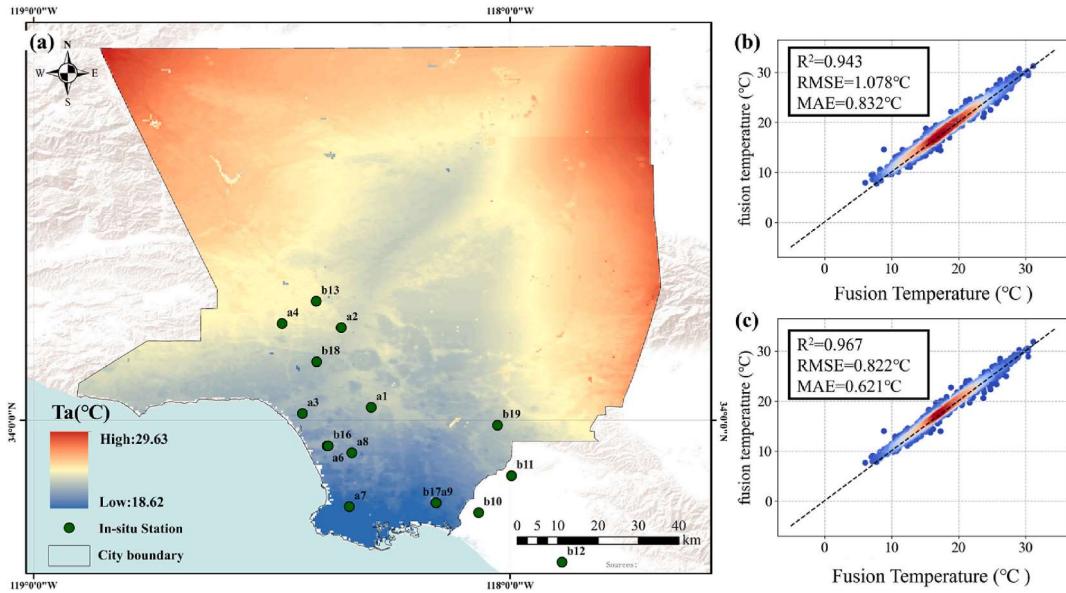
### 5.3. Applicability of DBN-st model in other typical cities

To validate the applicability and spatial transferability of the proposed DBN-st model in this study, supplementary experiments were conducted in two representative cities: the Austin area in Texas, USA, and the Los Angeles area (LA) in California, USA. Austin is situated in the central-southern part of Texas, USA. Positioned in the western hemisphere, it is influenced by the Gulf of Mexico. Austin falls under a subtropical climate with distinct seasons. The temperatures typically range between 25 °C and 30 °C, with summer highs reaching up to 40 °C. LA is situated in the southwestern part of the U.S. in California, along the eastern edge of the Pacific Ocean. It falls under a temperate Mediterranean climate, characterized by mild temperatures throughout the year. The average highest temperature is around 23.3 °C, while the average lowest temperature is approximately 13 °C. These two cities, while situated in the similar latitude zone in the northern hemisphere as WMA, exhibit significant climatic differences. Therefore, selecting the Austin and Los Angeles areas as comparison regions reflects geographical heterogeneity, enabling the examination of the model's applicability and spatial transferability.

We conducted experiments in the respective study areas using both the DBN-ori model and the DBN-st model, and the model performances are illustrated in the figures (Figs. 15c and 16c). The results indicate that the DBN-st model demonstrates favorable perfor-



**Fig. 15.** Experimental results for the Austin area: (a) Distribution of average temperatures in August for Austin, (b) Scatter plot depicting the accuracy performance of the DBN-ori model, and (c) Scatter plot illustrating the accuracy performance of the DBN-st model.



**Fig. 16.** Experimental results for the LA area: (a) Distribution of average temperatures in August for LA, (b) Scatter plot depicting the accuracy performance of the DBN-ori model, and (c) Scatter plot illustrating the accuracy performance of the DBN-st model.

mance in both the Austin and LA areas, with RMSE values of 1.048 °C and 0.822 °C, MAE values of 0.752 °C and 0.621 °C, and R<sup>2</sup> values of 0.984 and 0.967, respectively. Upon comparison with the results of the DBN-ori model ([Figs. 15b and 16b](#)), it is evident that the temperature fusion accuracy of the DBN-st model is superior, with a reduction of 19.2% and 31.1% in RMSE and 18.9% and 34.0% in MAE for the Austin and LA areas, respectively.

Austin experiences long, hot summers and short, cold, and windy winters, leading to diverse climate conditions throughout the year. During the experimental year, temperatures ranged between 6 °C and 36 °C, rarely dropping below –1 °C or exceeding 39 °C. The hottest month in Austin is August, with an average high temperature of 35 °C and an average low of 24 °C. LA, characterized by a Mediterranean climate, has dry summers and humid winters. In August, the average Tmax and Tmin in LA are 29.1 °C and 17.8 °C, making it the warmest month ([Ho et al., 2014; Yoo et al., 2018](#)). Therefore, we generated maps of the average temperatures in August 2019 for Austin and LA after GSTDA correction, as depicted in [Figs. 15a and 16a](#). In the Austin area, the highest temperature reached 43.1 °C in August, with an average temperature of 29.6 °C, consistent with previous studies. In the LA, the highest temperature in August was 29.6 °C, with an average temperature of 24.6 °C, aligning with the local climate patterns. The supplementary research conducted on the two typical cities indicates that the 3H temperature fusion model proposed in this study exhibits favorable fusion accuracy across regions with diverse climates and geographical locations.

#### 5.4. Comparison with related studies

In order to evaluate the effectiveness of our proposed deep learning DBN-st model framework, we conducted a comparison with results from other similar studies, as shown in [Table 4](#). Our framework outperformed linear regression models like the GWR model ([Li, et al., 2018](#)) and machine learning models such as RF and SVM ([Zhang, et al., 2021; Mohsenzadeh Karimi et al., 2018](#)), with significantly higher R<sup>2</sup> values and smaller errors. This suggests that our model is capable of providing more accurate fusion than other commonly used models. Furthermore, in comparison to other studies that also utilized the DBN model ([Huang et al., 2022; Li et al., 2021; Shen et al., 2020](#)), our research took into account geographic spatio-temporal factors that are overlooked in previous studies. This approach allowed us to obtain understanding of how these factors impact model performance. The DBN-st was able to achieve better accuracy by taking into account additional spatial and temporal information, which is critical for Ta fusion accurately. It is worth noting that our research had the highest spatial resolution (500m) and highest time resolution (daily) among similar studies, enabling a more detailed analysis of meteorological patterns at a local level. The high spatial and temporal resolutions of our model allow it to provide a more fine-grained representation of Ta data, capturing the subtle nuances in the data that might not be discernible at lower resolutions. This makes our model more reliable and useful for applications that require spatial disaggregation. Overall, our comparison showed that our DBN-st 3H Ta fusion model framework is superior to previous research results. The high accuracy of the model, combined with its ability to incorporate geographic temporal and spatial factors, makes it an effective tool for fusion Ta.

## 6. Conclusions

In this study, we investigated the impact of geographic and spatiotemporal factors on the accuracy of generating 3H Ta data through deep learning-based fusion of multi-source data. Specifically, we examined the influence of these factors at three stages of temperature fusion: data input, deep learning-based temperature fusion, and temperature fusion error calibration. By integrating remote sensing products, reanalysis products, ground measurement data, and other spatiotemporal data, we obtained daily temperatures at a 500-m resolution for the WMA in 2019. The results were extended for Austin and LA. The results revealed distinct geographic and spatiotemporal characteristics in temperature data, showing micro-level spatiotemporal autocorrelation and different clustering features in macro spatial regions. After introducing geographic and spatiotemporal autocorrelation factors into the temperature fusion model and conducting ten-fold cross-validation, the R<sup>2</sup> correlation coefficient was found to be 0.995, RMSE was 0.697 °C, and MAE was 0.527 °C. Compared with the DBN-ori model that did not consider geographic and spatiotemporal factors, the R<sup>2</sup> increased by 0.012, and RMSE and MAE decreased by 68% and 73%, respectively. The introduction of GSTDA error calibration, incorporating geographic and spatiotemporal factors, enabled compensation for temperature underestimation or overestimation at specific times or locations. The resulting average R<sup>2</sup>, RMSE, and MAE reached 0.993, 0.69 °C, and 0.58 °C, respectively, with a reduction of 9.37% and 10.15% in RMAE and MAE. The supplementary research conducted in other typical city clusters further demonstrates the applicability and portability of the proposed method. Our research findings and comparative analysis affirm the necessity of consider-

**Table 4**  
Comparison between model performances of the DBN-st and other models from previous related studies.

Model or Dataset	Spatial resolution	Temporal resolution	R <sup>2</sup>	RMSE (°C)	MAE (°C)	Literature
ERA5-2m Ta	0.1°	Hourly	\	\	\	\
CLDAS-Ta	0.0625°	Hourly	0.97	0.88	\	\
GWR	1 km	Daily	0.965	2.0	1.4	X. Li et al. (2018)
RF	Site	Monthly	0.968	\	1.456	Mohsenzadeh Karimi et al. (2018)
SVM	Site	Monthly	0.945	\	1.855	Karimi et al. (2018)
DBN	0.1°	Daily	0.972	1.996	1.539	Shen et al. (2020)
GB	1 km	Daily	\	1.7–2.7	\	H. Zhang et al. (2021)
DBN	1 km	Daily	0.966	1.96	1.57	Rui Li et al. (2021)
DBN	500m	Daily	0.986	1.086	0.839	X. Zhang et al., (2022)
DBN-st	500m	Daily	0.993	0.69	0.58	Our study

ing geographic and spatiotemporal factors in temperature fusion models, aiding in the generation of seamless and continuous 3H Ta data across space and time.

### CRediT authorship contribution statement

**Tailai Huang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft. **Xiang Zhang:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft. **Berhanu Keno Terfa:** Writing – review & editing. **Won-Ho Nam:** Validation, Writing – review & editing. **Xihui Gu:** Investigation, Resources, Writing – review & editing. **Jiangyuan Zeng:** Validation, Writing – review & editing. **Xu Zhang:** Writing – review & editing. **Shuzhe Huang:** Methodology, Software, Writing – review & editing. **Wenying Du:** Writing – review & editing. **Chao Wang:** Writing – review & editing. **Jian Yang:** Writing – review & editing. **Peng Wang:** Writing – review & editing. **Wenkui Wu:** Writing – review & editing. **Nengcheng Chen:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. **Dev Niyogi:** Validation, 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.

### Data availability

No data was used for the research described in the article.

### Acknowledgements

This work was supported by grants from the Open Fund of Hubei Luojia Laboratory (220100059), Special Fund of Hubei Luojia Laboratory (220100034), Pre-research Project of SongShan Laboratory (YYYY062022001), Open Fund of State Key Laboratory of Remote Sensing Science (OFLRSS202114), and KartoBit Research Network (KRN2301GK), and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan).

### References

- Abowarda, A.S., Bai, L., Zhang, C., Long, D., Li, X., Huang, Q., Sun, Z., 2021. Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale. *Rem. Sens. Environ.* 255. <https://doi.org/10.1016/j.rse.2021.112301>.
- Amiri, R., Weng, Q.H., Alimohammadi, A., Alavipanah, S.K., 2009. Spatial-temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. *Rem. Sens. Environ.* 113 (12), 2606–2617. <https://doi.org/10.1016/j.rse.2009.07.021>.
- Chen, J., Gao, C., Zeng, X., Xiong, M., Wang, Y., Jing, C., Krysanova, V., Huang, J., Zhao, N., Su, B., 2017. Assessing changes of river discharge under global warming of 1.5 °C and 2 °C in the upper reaches of the Yangtze River Basin: approach by using multiple-GCMs and hydrological models. *Quat. Int.* 453, 63–73. <https://doi.org/10.1016/j.quaint.2017.01.017>.
- Chen, Y.H., Shen, T.J., 2020. Unifying conspecific-encounter index and Moran's I index. *Ecography* 43 (12), 1902–1904. <https://doi.org/10.1111/ecog.05281>.
- Decker, M., Brunke, M.A., Wang, Z., Sakaguchi, K., Zeng, X.B., Bosilovich, M.G., 2012. Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Clim.* 25 (6), 1916–1944. <https://doi.org/10.1175/jcli-d-11-00004.1>.
- Dutra, E., Munoz-Sabater, J., Boussetta, S., Komori, T., Hirahara, S., Balsamo, G., 2020. Environmental lapse rate for high-resolution land surface downscaling: an application to ERA5. *Earth Space Sci.* 7, 5. <https://doi.org/10.1029/2019ea000984>.
- Gao, X., Zhang, A.L., Sun, Z.L., 2020. How regional economic integration influence on urban land use efficiency? A case study of Wuhan metropolitan area, China. *Land Use Pol.* 90. <https://doi.org/10.1016/j.landusepol.2019.104329>.
- Hersbach, H., Bell, B., Berrißford, P., Hirahara, S., Horanyi, A., Munoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellán, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Holm, E., Janiskova, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thepaut, J.N., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146 (730), 1999–2049. <https://doi.org/10.1002/qj.3803>.
- Ho, H.C., Knudby, A., Sirovayk, P., Xu, Y.M., Hodul, M., Henderson, S.B., 2014. Mapping maximum urban air temperature on hot summer days. *Remote Sens. Environ.* 154, 38–45. <https://doi.org/10.1016/j.rse.2014.08.012>.
- Huang, S., Gan, Y., Zhang, X., Chen, N., Wang, C., Gu, X., Ma, J., Niyogi, D., 2023 Feb. Urbanization amplified asymmetrical changes of rainfall and exacerbated drought: Analysis over five urban agglomerations in the Yangtze River Basin, China. *Earth's Future* 11 (2), e2022EF003117.
- Huang, S., Zhang, X., Chen, N., Ma, H., Fu, P., Dong, J., Gu, X., Nam, W.H., Xu, L., Rab, G., Niyogi, D., 2022. A novel fusion method for generating surface soil moisture data with high accuracy, high spatial resolution, and high spatio-temporal continuity. *Water Resour. Res.* <https://doi.org/10.1029/2021wr030827>.
- Huang, S.Z., Zhang, X., Chen, N.C., Ma, H.L., Fu, P., Dong, J.Z., Gu, X.H., Nam, W.H., Xu, L., Rab, G., Niyogi, D., 2022. A novel fusion method for generating surface soil moisture data with high accuracy, high spatial resolution, and high spatio-temporal continuity. *Water Resour. Res.* 58, 5. <https://doi.org/10.1029/2021wr030827>.
- Huang, S.Z., Zhang, X., Chen, N.C., Ma, H.L., Zeng, J.Y., Fu, P., Nam, W.H., Niyogi, D., 2022. Generating high-accuracy and cloud-free surface soil moisture at 1 km resolution by point-surface data fusion over the Southwestern U.S. *Agric. For. Meteorol.* 321. <https://doi.org/10.1016/j.agrformet.2022.108985>.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Rem. Sens. Environ.* 83 (1–2), 195–213. [https://doi.org/10.1016/s0034-4257\(02\)00096-2](https://doi.org/10.1016/s0034-4257(02)00096-2).
- Immerzeel, W.W., Petersen, L., Ragettli, S., Pellicciotti, F., 2014. The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resour. Res.* 50 (3), 2212–2226. <https://doi.org/10.1002/2013wr014506>.
- Kamath, H.G., Martilli, A., Singh, M., Brooks, T., Lanza, K., Bixler, R.P., Couderc, M., Yang, Z.L., Niyogi, D., 2023 Nov 1. Human heat health index (H3I) for holistic assessment of heat hazard and mitigation strategies beyond urban heat islands. *Urban Clim.* 52, 101675.
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K., 2015. The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Japan. Ser. II* 93 (1), 5–48. <https://doi.org/10.2151/jmsj.2015-001>.
- Li, R., Cui, L.L., Fu, H.B., Meng, Y., Li, J.L., Guo, J.P., 2020. Estimating high-resolution PM1 concentration from Himawari-8 combining extreme gradient boosting-geographically and temporally weighted regression (XGBoost-GTWR). *Atmos. Environ.* 229. <https://doi.org/10.1016/j.atmosenv.2020.117434>.
- Li, R., Huang, T., Song, Y., Huang, S., Zhang, X., 2021. Generating 1 km spatially seamless and temporally continuous air temperature based on deep learning over

- Yangtze River basin, China. *Rem. Sens.* 13, 19. <https://doi.org/10.3390/rs13193904>.
- 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. *Rem. Sens. Environ.* 215, 74–84. <https://doi.org/10.1016/j.rse.2018.05.034>.
- Li, X.Y., Li, Z., Huang, W., Zhou, P.X., 2020. Performance of statistical and machine learning ensembles for daily temperature downscaling. *Theor. Appl. Climatol.* 140 (1–2), 571–588. <https://doi.org/10.1007/s00704-020-03098-3>.
- 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>.
- Lin, W.P., Huang, J.F., Hu, X.M., Zhao, M., 2010. Crop yield forecast based on modis temperature-vegetation angel index. *J. Infrared Millim. Waves* 29 (6), 476–480.
- Liu, G., Chen, J.M., Ji, L.R., Sun, S.Q., 2012. Relationship of summer soil moisture with early winter monsoon and air temperature over eastern China. *Int. J. Climatol.* 32 (10), 1513–1519. <https://doi.org/10.1002/joc.2352>.
- Liu, X., Yang, T., Hsu, K., Liu, C., Sorooshian, S., 2017. Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. *Hydrol. Earth Syst. Sci.* 21 (1), 169–181. <https://doi.org/10.5194/hess-21-169-2017>.
- Liu, J., Niyogi, D., 2020 Sep 1. Identification of linkages between urban heat Island magnitude and urban rainfall modification by use of causal discovery algorithms. *Urban Clim.* 33, 100659.
- Marzban, F., Conrad, T., Marzban, P., Sodoudi, S., 2018a. Estimation of the near-surface air temperature during the day and nighttime from MODIS in Berlin, Germany. *Int. J. Adv. Rem. Sens. GIS* 7 (1), 2478–2517. <https://doi.org/10.23953/cloud.ijarsg.337>.
- Marzban, F., Sodoudi, S., Preusker, R., 2018b. The influence of land-cover type on the relationship between NDVI-LST and LST-T-air. *Int. J. Rem. Sens.* 39 (5), 1377–1398. <https://doi.org/10.1080/01431161.2017.1402386>.
- Meyer, H., Katurji, M., Appelhans, T., Muller, M.U., Nauss, T., Roudier, P., Zawar-Reza, P., 2016. Mapping daily air temperature for Antarctica based on MODIS LST. *Rem. Sens.* 8, 9. <https://doi.org/10.3390/rs8090732>.
- Mohsenzadeh Karimi, S., Kisi, O., Porrajabali, M., Rouhani-Nia, F., Shiri, J., 2018. Evaluation of the support vector machine, random forest and geo-statistical methodologies for predicting long-term air temperature. *ISH J. Hydraulic Eng.* 26 (4), 376–386. <https://doi.org/10.1080/09715010.2018.1495583>.
- Parajuli, B., Zhang, X., Deuja, S., Liu, Y.B., 2021. Regional and seasonal precipitation and drought trends in Ganga-Brahmaputra basin. *Water* 13, 16. <https://doi.org/10.3390/w13162218>.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. *Nature* 566 (7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>.
- Schwarz, N., Lautenbacher, S., Seppelt, R., 2011. Exploring indicators for quantifying surface urban heat islands of European cities with MODIS land surface temperatures. *Rem. Sens. Environ.* 115 (12), 3175–3186. <https://doi.org/10.1016/j.rse.2011.07.003>.
- 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. <https://doi.org/10.1016/j.rse.2020.111692>.
- Singh, M., Acharya, N., Jamshidi, S., Jiao, J., Yang, Z.L., Couder, M., Baumer, Z., Niyogi, D., 2023 May 31. DownScaleBench for developing\*\* and applying a deep learning based urban climate downscaling-first results for high-resolution urban precipitation climatology over Austin, Texas. *Comput. Urban Sci.* 3 (1), 22.
- Sun, Z.Y., Zhang, B., Yao, Y.B., 2019. An ERA5-Based model for estimating tropospheric delay and weighted mean temperature over China with improved spatiotemporal resolutions. *Earth Space Sci.* 6 (10), 1926–1941. <https://doi.org/10.1029/2019ea000701>.
- Tang, Y.B., Xie, S.F., Huang, L.K., Liu, L.L., Wei, P.Z., Zhang, Y.B., Meng, C.Y., 2022. Spatial estimation of regional PM2.5 concentrations with GWR models using PCA and RBF interpolation optimization. *Rem. Sens.* 14, 21. <https://doi.org/10.3390/rs14215626>.
- Tewari, K., Tewari, M., Niyogi, D., 2023 Jan 30. Need for considering urban climate change factors on stroke, neurodegenerative diseases, and mood disorders studies. *Comput. Urban Sci.* 3 (1), 4.
- Venter, Z.S., Brousse, O., Esau, I., Meier, F., 2020. Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sens. Environ.* 242. <https://doi.org/10.1016/j.rse.2020.111791>.
- Wang, K., Wang, J., Wang, P., Sparrow, M., Yang, J., Chen, H., 2007. Influences of urbanization on surface characteristics as derived from the Moderate-Resolution Imaging Spectroradiometer: a case study for the Beijing metropolitan area. *J. Geophys. Res.* 112, D22. <https://doi.org/10.1029/2006jd007997>.
- Wang, R., Gao, W., Peng, W., 2021. Spatial downscaling method for air temperature through the correlation between land use/land cover and microclimate: a case study of the Greater Tokyo Area, Japan. *Urban Clim.* 40, 101003. <https://doi.org/10.1016/j.ulclim.2021.101003>.
- Wang, Z., Hu, B., Huang, B., Ma, Z., Biswas, A., Jiang, Y., Shi, Z., 2022a. Predicting annual PM2.5 in mainland China from 2014 to 2020 using multi temporal satellite product: an improved deep learning approach with spatial generalization ability. *ISPRS J. Photogrammetry Remote Sens.* 187, 141–158. <https://doi.org/10.1016/j.isprsjprs.2022.03.002>.
- Wang, Z., Li, R., Chen, Z., Yao, Q., Gao, B., Xu, M., Yang, L., Li, M., Zhou, C., 2022b. The estimation of hourly PM2.5 concentrations across China based on a spatial and temporal weighted continuous deep neural network (STWC-DNN). *ISPRS J. Photogrammetry Remote Sens.* 190, 38–55. <https://doi.org/10.1016/j.isprsjprs.2022.05.011>.
- Wang, S., Zhang, X., Wang, C., Chen, N., 2023 Jan 1. Multivariable integrated risk assessment for cyanobacterial blooms in eutrophic lakes and its spatiotemporal characteristics. *Water Res.* 228, 119367.
- Wei, J., Li, Z., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., Song, Y., 2019. Improved 1-Km-Resolution PM2.5 Estimates across China Using the Space-Time Extremely Randomized Trees. <https://doi.org/10.5194/acp-2019-815>.
- Xu, L., Chen, N., Moradkhani, H., Zhang, X., Hu, C., 2020. Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. *Water Resour. Res.* 56, 3. <https://doi.org/10.1029/2019wr026444>.
- Xu, L., Chen, N., Zhang, X., Moradkhani, H., Zhang, C., Hu, C., 2021. In-situ and triple-collocation based evaluations of eight global root zone soil moisture products. *Rem. Sens. Environ.* 254. <https://doi.org/10.1016/j.rse.2020.112248>.
- 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. Photogrammetry Remote Sens.* 137, 149–162. <https://doi.org/10.1016/j.isprsjprs.2018.01.018>.
- Zhang, F., Zhang, H., Hagen, S.C., Ye, M., Wang, D., Gui, D., Zeng, C., Tian, L., Liu, J., 2015. Snow cover and runoff modelling in a high mountain catchment with scarce data: effects of temperature and precipitation parameters. *Hydrol. Process.* 29 (1), 52–65. <https://doi.org/10.1002/hyp.10125>.
- Zhang, H., Immerzeel, W.W., Zhang, F., de Kok, R.J., Gorrie, S.J., Ye, M., 2021. Creating 1-km long-term (1980–2014) daily average air temperatures over the Tibetan Plateau by integrating eight types of reanalysis and land data assimilation products downscaled with MODIS-estimated temperature lapse rates based on machine learning. *Int. J. Appl. Earth Obs. Geoinf.* 97. <https://doi.org/10.1016/j.jag.2021.102295>.
- Zhang, T.L., Lin, G., 2016. On Moran's I coefficient under heterogeneity. *Comput. Stat. Data Anal.* 95, 83–94. <https://doi.org/10.1016/j.csda.2015.09.010>.
- Zhang, X., Chen, N.C., Chen, Z.Q., Wu, L.X., Li, X., Zhang, L.P., Di, L.P., Gong, J.Y., Li, D.R., 2018. Geospatial sensor web: a cyber-physical infrastructure for geoscience research and application. *Earth Sci. Rev.* 185, 684–703. <https://doi.org/10.1016/j.earscirev.2018.07.006>.
- Zhang, X., Chen, N.C., Li, J.Z., Chen, Z.H., Niyogi, D., 2017. Multi-sensor integrated framework and index for agricultural drought monitoring. *Rem. Sens. Environ.* 188, 141–163. <https://doi.org/10.1016/j.rse.2016.10.045>.
- Zhang, X., Chen, N.C., Sheng, H., Ip, C., Yang, L., Chen, Y., Sang, Z., Tadesse, T., Lim, T.P.Y., Rajabifard, A., Bueti, C., Zeng, L.L., Wardlow, B., Wang, S., Tang, S.Y., Xiong, Z., Li, D., Niyogi, D., 2019. Urban drought challenge to 2030 sustainable development goals. *Sci. Total Environ.* 693. <https://doi.org/10.1016/j.scitotenv.2019.07.342>.
- Zhang, X.D., Zhou, J., Gottsche, F.M., Zhan, W.F., Liu, S.M., Cao, R.Y., 2019. A method based on temporal component decomposition for estimating 1-km all-weather land surface temperature by merging satellite thermal infrared and passive microwave observations. *IEEE Trans. Geosci. Rem. Sens.* 57 (7), 4670–4691. <https://doi.org/10.1109/tgrs.2019.2892417>.
- Zhang, X.D., Zhou, J., Liang, S.L., Wang, D.D., 2021. A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature. *Rem. Sens. Environ.* 260. <https://doi.org/10.1016/j.rse.2021.112437>.
- Zhang, Y.F., Wang, X.P., Pan, Y.X., Hu, R., 2012. Diurnal relationship between the surface albedo and surface temperature in revegetated desert ecosystems, northwestern China. *Arid Land Res. Manag.* 26 (1), 32–43. <https://doi.org/10.1080/15324982.2011.631687>.

- Zhao, L., Lee, X., Smith, R.B., Oleson, K., 2014. Strong contributions of local background climate to urban heat islands. *Nature* 511 (7508), 216–219. <https://doi.org/10.1038/nature13462>.
- Zhou, X., Huang, G.H., Li, Y.P., Lin, Q.G., Yan, D.H., He, X.J., 2021. Dynamical downscaling of temperature variations over the Canadian prairie provinces under climate change. *Rem. Sens.* 13, 21. <https://doi.org/10.3390/rs13214350>.