

JGR Atmospheres

RESEARCH ARTICLE

10.1029/2021JD035542

Key Points:

- A statistical downscaling scheme was designed for weather research and forecasting (WRF) simulations
- To downscale the hourly air temperature from 1-km spatial resolution to 30-m, up to 24 hr in advance in a mountainous area
- Compared to WRF simulation results, root-mean-square error (RMSE) and mean absolute error (MAE) of WRF-downscaled results decreased by 0.87°C and 0.71°C, respectively, at the automatic weather station (AWS) level

Correspondence to:

S. Zhu,
zsyzgx@163.com

Citation:

Zhang, G., Zhu, S., Zhang, N., Zhang, G., & Xu, Y. (2022). Downscaling hourly air temperature of WRF simulations over complex topography: A case study of Chongli District in Hebei Province, China. *Journal of Geophysical Research: Atmospheres*, 127, e2021JD035542.
<https://doi.org/10.1029/2021JD035542>

Received 8 JUL 2021

Accepted 20 JAN 2022

Author Contributions:

Data curation: Guangxing Zhang, Guixin Zhang

Methodology: Guangxing Zhang, Shanyou Zhu, Yongming Xu

Project Administration: Shanyou Zhu

Resources: Nan Zhang

Supervision: Shanyou Zhu

Visualization: Guixin Zhang

Writing – original draft: Guangxing Zhang

Downscaling Hourly Air Temperature of WRF Simulations Over Complex Topography: A Case Study of Chongli District in Hebei Province, China

Guangxing Zhang¹ , Shanyou Zhu¹ , Nan Zhang², Guixin Zhang³, and Yongming Xu¹ 

¹School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science & Technology, Nanjing, China, ²Hebei Meteorological Observatory, Shijiazhuang, China, ³School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing, China

Abstract Accurate and high-resolution air temperature prediction is important in many different applications. Hourly air temperature forecasting in mountainous areas is necessary and important because mountainous areas are becoming increasingly important areas of human activities. At present, scientists successfully employ numerical weather prediction (NWP) models, such as the Weather Research and Forecasting (WRF) model, to achieve reliable forecasts. However, air temperature forecasting and modeling over complex geographical zones are still difficult tasks. The WRF model is a mesoscale model and does not adequately account for the influence of terrain on the air temperature. It is important to downscale larger-scale models to a much finer scale. In this paper, a statistical temperature downscaling method based on geographically weighted regression (GWR) and diurnal temperature cycle (DTC) models is proposed. A statistical downscaling scheme of WRF simulation data is designed to forecast the hourly air temperature from 1-km spatial resolution to 30 m, up to 24 hr in advance. The combined downscaling model's root-mean-square error (RMSE) decreased by 0.87°C at the automatic weather station (AWS) level and 0.62°C over the domain when compared to WRF simulations, and the mean absolute error (MAE) decreased by 0.71°C and 0.51°C, respectively, at these two levels. The results reveal that the combined downscaling model performs very well in correcting and downscaling the air temperature in WRF simulations in the study areas.

1. Introduction

The air temperature is one of the most frequently measured meteorological parameters. Air temperature forecasting is a crucial climatic factor-related task required in many different applications in areas such as agriculture, industry, energy, the environment, and tourism (Abdel-Aal, 2004; Cifuentes et al., 2020). Certain applications include short-term load forecasting for power utilities (Li et al., 2016), protection against freezing injury of various fruits (Chung et al., 2006), adaptive temperature control in greenhouses (Dombayci & Gölcü, 2009), prediction of cooling and energy consumption in residential buildings (Ben-Nakhi & Mahmoud, 2004), and establishment of a planning horizon for infrastructure upgrades, insurance, energy policy, and business development purposes (Smith et al., 2007). Therefore, accurate forecasting of the air temperature with high spatiotemporal resolution has always been an important goal for meteorologists and weather forecasters.

Mountainous areas also represent important areas of human activities, such as fruit cultivation, tourism, and skiing in the winter season. Therefore, hourly air temperature forecasting and high-resolution prediction in mountainous areas is necessary and important for human activities. In contrast to plain areas, the topography of a mountainous area is very complex, which can cause the air temperature to vary within small terrain units. For example, in a 1-km grid, the difference in altitude between the slope top and foot can reach as much as 500 m, and the temperature difference as high as 3°C, assuming that the slope of the grid is 0.5. It is a difficult task for air temperature forecasting and modeling in complex geographical zones because the temperature is influenced by several nonlinear processes, such as the interaction between large-scale air mass circulation and local airflow, airflow-topography interactions, and the interplay between radiation and topographic shading.

Currently, many numerical weather prediction (NWP) models are available for air temperature prediction, for example, the European Centre for Medium-Range Weather Forecasts (ECMWF) model, the fifth-generation mesoscale model (MM5), and the Weather Research and Forecasting (WRF) model. All these NWP models allow day-ahead air temperature prediction, which is usually adopted in practice. However, uncertainties in model

inputs, model parameter estimation, and model structure are unavoidable in NWP modeling. Owing to these uncertainties, the accuracy of NWP models attained in air temperature prediction hardly meets the increasing needs of grid systems (Xu et al., 2021). Moreover, due to their coarse spatial resolution, mostly lower than 1 km (Caldwell et al., 2009; Yan et al., 2020), NWP models have mainly been applied to develop climate change scenarios and perform large-scale studies. These models suffer in their ability to resolve fine-scale air temperature (sub-km scale) at the timescales relevant to biota living in narrow mountain valleys (e.g., 10–30 m; Holden et al., 2011), in microclimates or in areas of complex terrain (Alessi & DeGaetano, 2021; Le Roux et al., 2018; Zhang et al., 2013).

As a result, a variety of techniques, such as downscaling methods, have been developed to bridge the gap between the scale at which data are available and the scale at which it is needed for assessment purposes. Downscaling methods can be divided into two broad groups: dynamical downscaling (DD) and statistical downscaling (SD) methods. A high-resolution regional climate model (RCM) is used for DD method to obtain the fine-scale climate information, and then the large-scale atmospheric state obtained from a general circulation model (GCM) is used for the RCM as the boundary condition (Dickinson et al., 1989; Giorgi & Bates, 1989). It is easier for the physical interpretation of the downscaled results, since a physical principle is used to construct a model (Adachi & Tomita, 2020). However, the DD method has some drawbacks. Although more physically defensible, this approach is much more computationally expensive than the SD method (Harding et al., 2013) and is itself subject to error due to imperfect parameterizations and numerics (Caldwell et al., 2009). The RCM performance is also strongly dependent on the GCM/reanalysis boundary forcing data (Fowler et al., 2007). The SD method deduces fine-scale information using some statistical relations between large- and fine-scale climate information (Tang et al., 2016; Wilby et al., 2004). Statistical methods range from multiple linear regression (e.g., Ke et al., 2011; Schoof & Pryor, 2001) to more complex techniques such as artificial neural networks (e.g., Schoof & Pryor, 2001), and the method of analogs (e.g., Viggiani et al., 2013; Zorita & Storch, 1999). It has been revealed that SD methods attain a comparable accuracy to that of DD methods (Wilby et al., 2004). Compared to DD methods, the SD method is relatively easier and computationally undemanding to implement and provides station-scale climate information from NWP model output (Bechler et al., 2015; Cheng et al., 2008).

There have been many studies on statistical downscaling focused on air temperature. For example, studies have examined the downscaling of the daily mean air temperature (Hofer et al., 2015; Wang et al., 2020), daily minimum air temperature (Holden et al., 2011) and maximum air temperature (Viggiani et al., 2019; Wang et al., 2020). To better extend climate change impact analysis to estimates of future synoptic weather types and meteorological variables, not only daily but also hourly climate scenarios are necessary. There have been studies on hourly air temperature forecasting in urban areas (Yi et al., 2018), but few studies have been conducted in mountainous areas. Despite the above uncertainty, it has been reported that Multiple Linear Regression (MLR) can be used to downscale air temperature (Schoof & Pryor, 2001), and the explanatory variables for the MLR calculations were mainly topographical parameters (Brunetti et al., 2014; Casellas et al., 2020). Therefore, to achieve the purpose of downscaling, we can apply high-resolution terrain data (Bechler et al., 2015) to improve the spatial resolution of WRF data by constructing the relationship between the air temperature and terrain.

To obtain the hourly temperature, the air diurnal temperature cycle (DTC) model is an effective tool. Maximum and minimum air temperatures based on other parameters, such as the hourly air temperature, can be estimated by applying a DTC model as a parametric method (Duan et al., 2012; Gholamnia et al., 2017; Hu et al., 2020). The DTC model can be adopted to improve the inputs required for numerical weather models or assimilated data (Gholamnia et al., 2019). The shape of diurnal temperature curves has been modeled in a variety of ways that range from simple curve-fitting models based on sine curves to more sophisticated techniques utilizing Fourier analysis and complex energy balance models (Gholamnia et al., 2019; Reicosky et al., 1989). Therefore, a DTC model can be constructed based on observation data of weather stations, and a MLR model can then be constructed considering DTC model coefficients and local parameters, such as the altitude, slope, and aspect ratio. Through this process, high-spatial resolution hourly forecasts can be obtained by expanding/contracting the established DTC model. Ground observation of air temperature on a virtually continuous spatial scale is required to construct the downscaling model. In order to reconstruct observation data for the entire surface, interpolation techniques are mandatory (Goovaerts, 1999; Thornton et al., 1997). Spatial interpolation methodologies, such as the inverse distance weighted (IDW) method and kriging technique, are now widely used. However, large uncertainties arise when applying interpolation methods (such as inverse distance weighted method and kriging technique) to areas

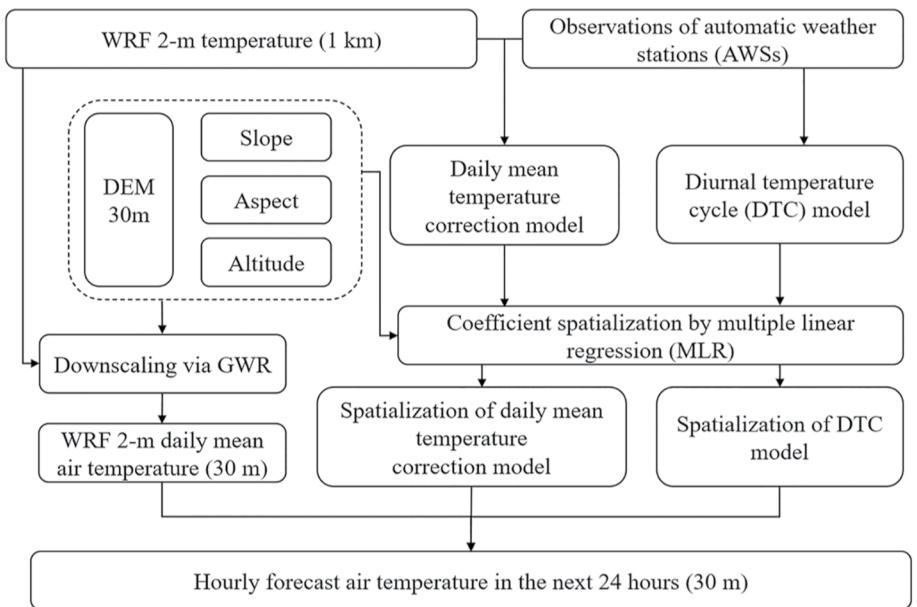


Figure 1. Flow chart of this study.

with a complex terrain (Jiang et al., 2021), while the geographically weighted regression (GWR) method is more effective when analyzing nonstationary spatial parameters and has been widely employed (Zhou et al., 2019). GWR allows the relationships between dependent and explanatory variables to vary over space and directly deals with non-stationarity. For our study, GWR was used to obtain observation data and WRF simulation results with continuous spatial scale, and these data were used by our downscaling model as the input.

A statistical downscaling scheme was designed in this study to forecast the hourly air temperature, with a 30-m spatial resolution, up to 24 hr in advance in a mountainous area from WRF simulations. This study developed a multistep downscaling and correction model to improve the WRF performance in air temperature prediction. The steps are as follows: (a) the WRF air temperature was downscaled from 1-km to a 30-m spatial resolution via the geographically weighted regression (GWR) method, in which the WRF-simulated air temperature was the dependent variable and the aspect, slope, altitude as the independent variables; (b) a DTC model is built with observation data retrieved from automatic weather stations (AWSs) and it is determined how to use the DTC model in forecasting; (c) the daily mean air temperature in WRF downscaling results is corrected based on ground observation data; (d) multiple linear regression is applied to achieve spatialization of the daily mean air temperature correction model and DTC model; (e) combined with the downscaling results obtained via the GWR method, the spatialized multiplicative correction model and the spatialized DTC model are employed to predict the air temperature in the study area. Figure 1 shows a summary of the overall procedure followed in this study. Each step is described in detail in the section of Methods.

2. Study Area and Data

2.1. Study Area

The study area is located in Chongli District, Zhangjiakou city, Hebei Province, China (Figure 2), and it covers ~2,300 km². The bounding coordinates of the study area are 114.8°–115.6°E, 40.8°–41.3°N. This area belongs to the semiarid monsoon climate zone. The complex terrain of Chongli District exhibits an elevation range from approximately 812 m to ~2,169 m, with a mean elevation of 1,485 m. The study area is one of the venues of the 2022 Beijing Winter Olympic Games. In the future, the region will become a hot spot of tourism and sports activities. Therefore, providing correct air temperature forecasts at high spatial and temporal resolutions is necessary in this area.

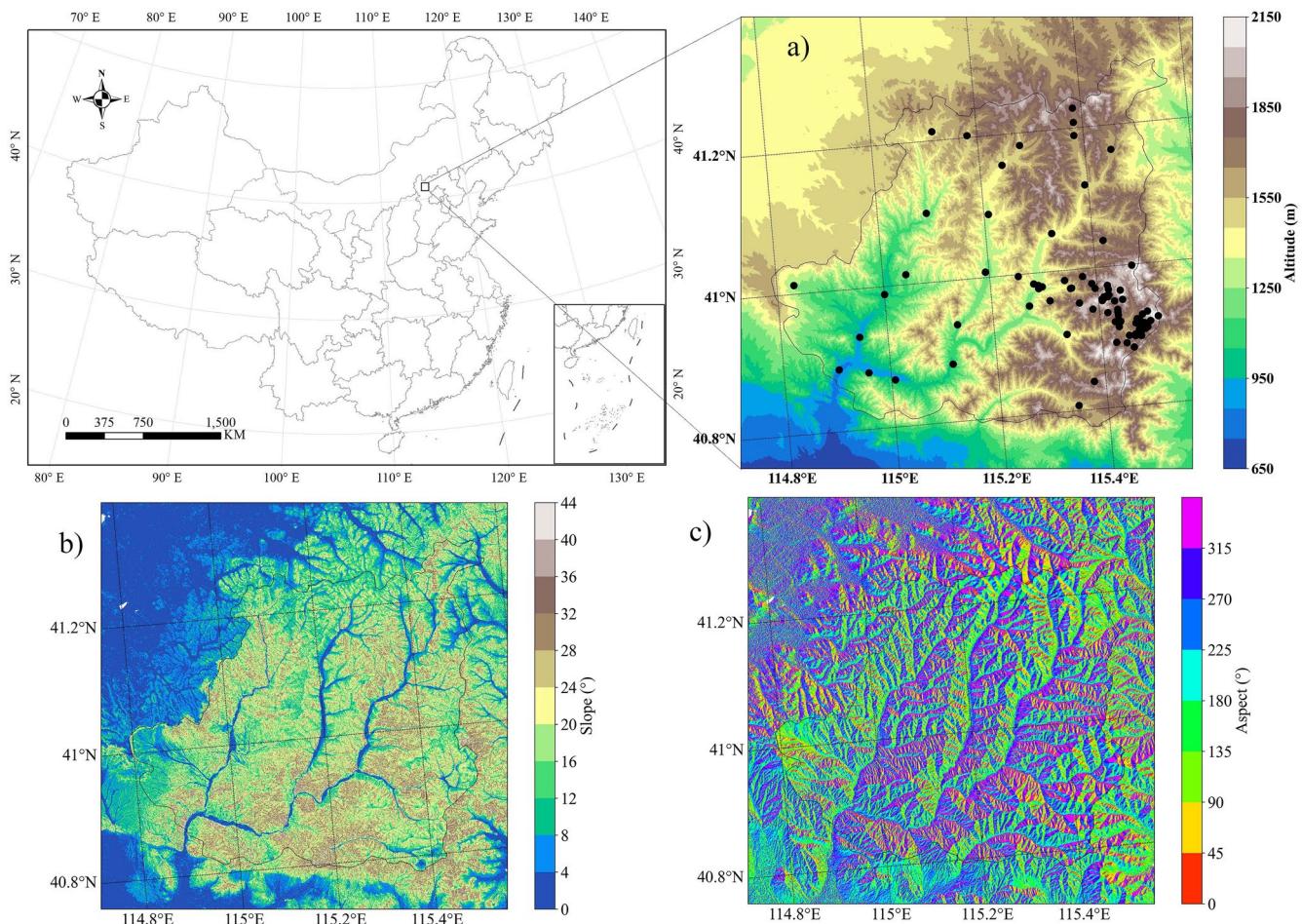


Figure 2. Elevation (a), slope (b), and aspect (c) diagrams of the simulation area. The distribution of 74 automatic weather stations (AWSs) is shown with black dots.

2.2. ASTER DEM

In this study, ASTER DEM was adopted to obtain 30-m resolution altitude data across the whole study area and to calculate the slope and aspect. The ASTER DEM version 2 is a global 1-arc-second resolution elevation data set that was released in October 2011 by the METI, Japan, and NASA. ASTER DEM was generated using optical imagery with a resolution of 15-m collected in space with the METI ASTER sensor mounted on the NASA Terra satellite (Abrams et al., 2010). The sensor has three spectral bands in the visible infrared spectrum, six bands in the shortwave infrared spectrum and six bands in the thermal infrared spectrum. First, the 1-arc-second (~30-m) GDEM1 was released in June 2009 and covered the globe from 83°N latitude to 83°S latitude. GDEM1 was derived from ~1.2 million images and exhibited a vertical accuracy of 20 m at the 95% confidence level (ASTER GDEM Validation Team, 2009).

2.3. WRF Simulations

The WRF model is a numerical weather prediction model for mesoscale weather systems for simulation and real-time forecasting purposes. It was developed collaboratively by NCAR and the National Centers for Environmental Prediction (NCEP; see <http://www.mmm.ucar.edu/wrf/users/>). It employs scale analysis to resolve fluid dynamics and thermodynamic equations that express atmospheric motion in future atmospheric circulation and weather prediction (Xu et al., 2021).

In this study, WRF version 3.7.1 with its center located at (108°E, 38°N) was configured with three domains at 9, 3, and 1-km horizontal resolutions, and then the simulation results of 1-km horizontal resolution were used in

this study. The model vertical resolution was discretized with 50 full terrain-following σ levels with the model top located at 50 hPa in all domains. There were 550 grid points along the east-west direction and 484 grid points along the north-south direction in domain 03 (d03). The National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) Final (FNL) operational global analysis data with a resolution of 0.5° was used to determine the initial conditions under ingestion at 6-hr intervals. Each episode was initialized at 12:00 UTC and was independently run during periods of 36 hr. The first 6 hours were considered the spin-up time. The various physical parameterization options utilized in the present study were as follows: the Thompson parametrization scheme was chosen for the microphysics, the Yonsei University (YSU) parametrization scheme was chosen for the planetary boundary layer (PBL), the RRTM longwave scheme and Dudhia shortwave scheme were employed for radiation, and finally, the Noah land-surface model was adopted as the surface layer scheme. The time range of the simulation data was from 7 January 2020 (D7), to 29 January 29 (D29) 2020, for a total of 23 days (In the following, D7 represents 7 January 2020, D8 represents 8 January 2020, and so on).

2.4. Ground Observation Data

Observed hourly 2-m air temperature data at 74 automatic weather stations (AWSs; as shown in Figure 2) in Chongli District, Zhangjiakou city, Hebei Province, China, were used in this study. These observations are part of records retrieved from gradually evolving AWSs that are owned and maintained by the Hebei Meteorological Administration. These stations are located at an average altitude of $\sim 1,500$ m. These data were reviewed through a quality control procedure to identify any invalid and missing data, likewise in Feng et al. (2004). The time range of the ground observation data was the same as WRF simulation data.

3. Methods

3.1. Geographically Weighted Regression Method

In this study, the geographically weighted regression (GWR) method was used in two steps: one was to downscale the resolution of WRF from 1 km to 30 m (WRF-simulated), and the other was to obtain observation data spatial distributions (GWR-interpolation). WRF-simulated with 30-m resolution was the input for subsequent model establishment. The GWR-interpolation results were used as a benchmark to compare the spatial bias in the WRF simulation results (WRF-simulated) and combined model results (WRF-downscaled). The GWR is a local modeling technique appropriate for spatial data with some degree of spatial dependence. The aim of GWR is to examine the existence of spatial non-stationarity in the relationship between a dependent variable and a set of independent variables (Georganos et al., 2017). The GWR method calculates regression coefficients at each individual location toward the ordinary least squares (OLS) model that calculates the coefficients across the whole study area (Fotheringham et al., 1997). The GWR model is expressed as:

$$y(m) = \beta_0 + \sum_{k=1}^p \beta_k x_k(m) + \varepsilon(m) \quad (1)$$

where $y(m)$ is the value of the response variable y at location m , x_k is the value of the k th independent variable, β_0 is the regression constant, β_k is the correlation coefficient for the independent predictor variable x_k , and $\varepsilon(m)$ is the random error term associated with location m (Fotheringham & Oshan, 2016). In this study, the response variable was the air temperature (WRF output or observation data), and the independent variables were the slope, aspect, and altitude, which were derived from the DEM, with 30-m resolution.

3.2. Air Diurnal Temperature Cycle Model

According to the Fourier series theory, a complex periodic function or periodic sequence is formed via the superposition of sine waves of different amplitudes and phases. The air diurnal temperature cycle (DTC) model is based on the method of extracting known periodic changes in elements and analyzing the law of sequence changes. The DTC model has been widely applied to capture key diurnal characteristics, for example, the daily maximum, with a high degree of accuracy, and other applications, such as cloud screening, emissivity estimation, and capturing the diurnal patterns of urban heat islands during a short period (Duan et al., 2014; Zhou et al., 2013).

The DTC model has often been adopted to fit and repair historical data but has rarely been applied to temperature forecasting. In this paper, we use this model to forecast the air temperature.

We applied the Fourier transform method to convert the diurnal temperature variation into a set of harmonics (Wang et al., 2018). The periodic air temperature variation can be transformed into the sum of harmonics with different frequencies as follows:

$$T(t) = \bar{T} + \sum_{k=1}^M \Delta \tilde{T}_k \sin(\omega_k t + \varphi_k) \quad (2)$$

where $T(t)$ is the air temperature at time t , t is the local solar time during a day and its value ranges from 0 to 23, \bar{T} is the daily mean temperature, M is the order of the Fourier function, $\Delta \tilde{T}$ is the amplitude, φ is the phase and $\omega = 2\pi/P$ is the fundamental angular frequency, with the period P equal to 24 hr in regard to the daily cycle.

Regarding the daily air temperature cycle, daily and semi-daily harmonics represent the fundamental periodic variation components. A strong daily harmonics signal reflects the primary forcing mechanism, that is, solar radiation. The semi-daily cycle is mainly imposed by the abrupt nighttime zeroing of solar radiation, in addition to heat storage in the soil and atmosphere. The observations during a day can be treated as the sum of the above two periodic variation components and other random variation components, as follows:

$$T(t) = \bar{T} + \Delta \tilde{T}_{d1} \sin\left(\frac{2\pi}{24}t + \varphi_{d1}\right) + \Delta \tilde{T}_{d2} \sin\left(\frac{2\pi}{12}t + \varphi_{d2}\right) \quad (3)$$

The major parameters include the daily mean temperature \bar{T} , amplitude $\Delta \tilde{T}_{di}$ and phase φ_{di} of the daily harmonics (24-hr harmonics, $i = 1$) and the semi-daily cycle (12-hr harmonics, $i = 2$). φ can be calculated based on the peak/valley temperature occurrence time. According to statistics of the observation data, the above harmonics perform the best when setting the peak temperature in the study area during a daily cycle to occur at 14:00 and the valley temperature to occur at 7:00. All the harmonics functions of the 74 AWSs on the previous day were established via curve fitting.

To apply these harmonics to air temperature forecasting, the next day's mean temperature and amplitude $\Delta \tilde{T}_{di}$ are the key parameters. Excluding the daily mean temperature from the above harmonics, the remaining harmonics represent the variation in the temperature difference between the daily mean temperature and the temperature at each local solar time. Through experiments, the temperature difference on the next day can be better predicted by multiplying the harmonics of the previous day with a scaling factor. The scaling factor can be adjusted as follows:

$$r = \frac{\bar{T}_f}{\bar{T}_o} \quad (4)$$

where \bar{T} is the mean air temperature, f is the forecast day, and o is the observation of the previous day.

However, the daily mean air temperature on the forecast day is unknown when forecasting the air temperature. In this study, we applied the daily mean temperature determined via WRF simulations as the daily mean air temperature on the forecast day. This produces a large error in the result if the daily mean air temperature of the WRF simulations is directly used, because there is a random error in WRF simulations. Therefore, we first revised the WRF daily mean air temperature using the following correction method.

3.3. Multiplicative Correction of WRF Simulations

The daily mean temperature on the forecast day is the key variable when using the DTC model to forecast the air temperature. WRF simulated air temperatures have significant biases (Wang & Kotamarthi, 2014), hence bias correction of WRF simulation results is a useful method to obtain the daily mean temperature on the forecast day. Error correction models have been developed to improve the accuracy of numerical output by training the relationship between prediction errors and related variables (Xu et al., 2021). Application of bias correction techniques has become increasingly popular to perform local correction of deterministic model outputs with empirical equations. These empirical equations include additive correction equations (mean bias subtraction), multiplicative correction equations, hybrid models, model output statistics (MOS), and Kalman filter-based correction equations (Mok et al., 2017; Pierce et al., 2015). Additive correction adjusts the forecast at a particular

monitoring station by adding the temporal average of the measured bias on previous days to the forecast of a deterministic model (Wilczak et al., 2006), while multiplicative correction is performed by multiplying the forecast with a ratio (Borrego et al., 2011). This ratio is obtained by dividing the sum of the observations at the station over the past few days by the sum of the forecasts obtained with a deterministic model during the same period. The hybrid forecast is a special case of additive correction, which only considers the bias on the previous day to perform correction (Neal et al., 2014; Silibello et al., 2015). Regarding MOS, a correction model based on linear regression is generated between the measurements and forecasts or a set of independent variables such as meteorological measurements (with their respective parameters; Konovalov et al., 2009), which is applied to posterior processing after training with historical data. Kalman filter-based bias correction and additive and multiplicative model biases are adaptively estimated at each time step based on the weighted combination of prior estimates and the contribution of new measurements (Borrego et al., 2011).

Multiplicative correction, also referred to as linear correction, is a simple bias-correction technique. It considers a scaling factor between the observations and simulations during the calibration period to reduce the future bias (Ines & Hansen, 2006). In this method, assuming that the ratios between field measurements (observations) and WRF simulations on the previous day and forecast day remain the same, the air temperature on the next day can then be corrected by the ratio on the previous day. The forecasted air temperature corrected from WRF simulations was adjusted as follows:

$$\tilde{T} = \frac{\sum_{i=1}^n T_o}{\sum_{i=1}^n T_m} \times T_{m-f} \quad (5)$$

where T is the air temperature based on either observations (o) or modeling (m) during a historic training period, f denotes the forecast day, and n is the number of days of historical data.

In this paper, we used Equation 5 to correct the WRF mean air temperature. Therefore, T_o was replaced with the mean air temperature according to historical AWS data, T_m was replaced with the mean air temperature based on historical WRF data, and T_{m-f} was replaced with the mean air temperature of WRF forecast data. All the WRF values used here were downscaling results obtained via the GWR method. In this study, the nearest grid point of the WRF value is adopted as the forecast value at the given AWSs.

The WRF daily mean air temperature calculated via Equation 5 is the daily mean temperature on the forecast day. Combining Equations 3–5, the forecast equation can be expressed as follows:

$$T(t) = \frac{\sum_{i=1}^n \tilde{T}_o}{\sum_{i=1}^n \tilde{T}_m} \times \bar{T}_{m-f} \left\{ 1 + \frac{1}{\tilde{T}_o} \times \left(\Delta \tilde{T}_{d1} \sin \left(\frac{2\pi}{24} t + \varphi_{d1} \right) + \Delta \tilde{T}_{d2} \sin \left(\frac{2\pi}{12} t + \varphi_{d2} \right) \right) \right\} \quad (6)$$

where t is the local solar time, ranging from 0 to 23, and T is the forecast temperature.

3.4. Spatialization of the DTC Model and Daily Mean Air Temperature Correction Model

All the above models can only forecast the temperature at AWSs. It is not enough to correct and downscale the temperature at the considered AWSs, and the goal is to produce spatial predictions across the entire study area. Therefore, the DTC model and daily mean air temperature correction model require spatialization. In regard to spatialization, spatial interpolation methodologies are widely used, such as the inverse distance weighted (IDW) method and kriging technique. However, large uncertainties arise, and most WRF grids are not properly considered when applying them to areas with complex terrains. Although there is bias in WRF simulations, the interpolation result is not better than that of the WRF model. Therefore, all grid WRF data must be corrected before forecasting in the entire study area. In addition, the DTC model based on AWS data must also be spatialized. Then, the air temperature at each hour in the entire study area can be predicted on the forecast day.

As expressed in Equation 5, the equation can be divided into two parts. The left part is a scaling factor calculated by the WRF model and the observation data on the previous day. The scaling factor is the correction coefficient of the WRF simulations for the next day. We established a MLR equation for the scaling factor, in which the scaling factor is the dependent variable and the altitude, slope, and aspect as the independent variables. Through the MLR equation for the scaling factor and via substitution of the independent variables of all WRF grids into this equation, the corrected air temperature in all grids can be calculated. The corrected WRF data was used as the

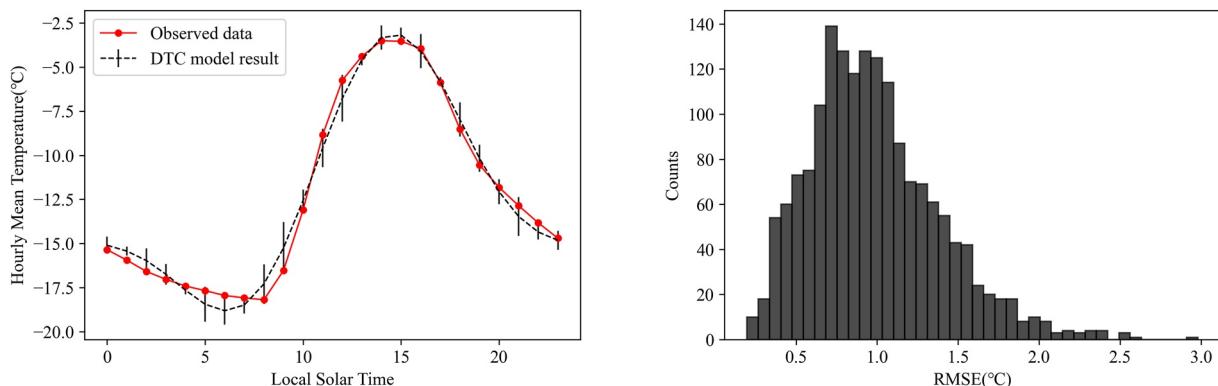


Figure 3. The diurnal temperature cycle (DTC) model and observed air temperature curves of the hourly average value from an automatic weather station (AWS) (left; each point represents a mean temperature of local solar time during all 23 days. The AWS number is 54,304, located at 40.9542°N and 115.2697°E, with an elevation of 1,485 m.), and a histogram of the root-mean-square error (RMSE) between the DTC model results and observed data from all 74 AWSs during the 23 days (right).

basic data for the DTC model. A MLR model for $\Delta\tilde{T}_{di}$ was also built, where $\Delta\tilde{T}_{di}$ is the dependent variable and the altitude, slope, and aspect are independent variables. Then, the air temperature spatial distribution in the study area at each local solar time was forecasted by combining the above two MLR equations and the scaling factor.

4. Results and Discussion

4.1. DTC Fitting Performance at the AWSs

The DTC model can represent most of the daily variations, as shown in Figure 3 (left). The root-mean-square error (RMSE) between the DTC model results and the observed data is 0.64°C throughout the daily cycle at all 74 AWSs. The RMSE histogram of all 74 AWSs shows that most RMSE values (1,520 records out of a total of 1,702) are less than 1.50°C, and 973 records are less than 1.0°C, with a mean value of 0.98°C at all stations during the 23 days (Figure 3 right).

Excluding the daily mean temperature retrieved from the DTC model, the DTC model further represents the difference between the daily mean temperature and temperature at each local solar time. A scatter plot of the simulated hourly temperature difference during the total 23 days for all AWSs against the measured value at all AWSs is shown in Figure 4 (left). The results indicate a very high reconstruction accuracy ($R^2 = 0.945$) for the

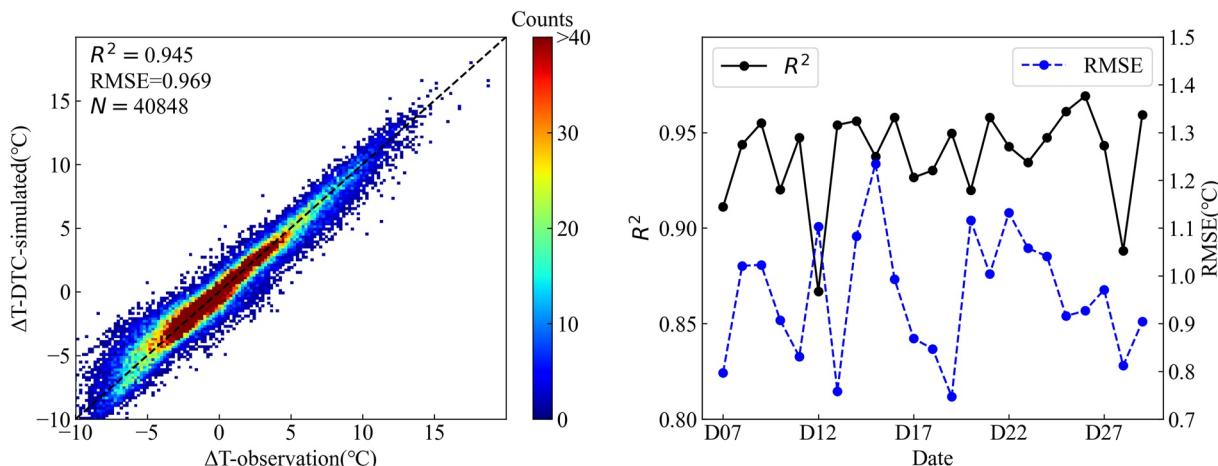


Figure 4. The relationship between ΔT -observation and ΔT -diurnal temperature cycle (DTC)-simulated (each point represents 1 hour at one AWS station) during the total 23 days (left), and the scatter plot of daily mean R^2 and root-mean-square error (RMSE) between ΔT -observation and ΔT -DTC-simulated during the total 23 days (right). ΔT -observation is the observed hourly difference from the daily mean temperature. ΔT -DTC-simulated is the DTC model-simulated hourly difference from the daily mean temperature.

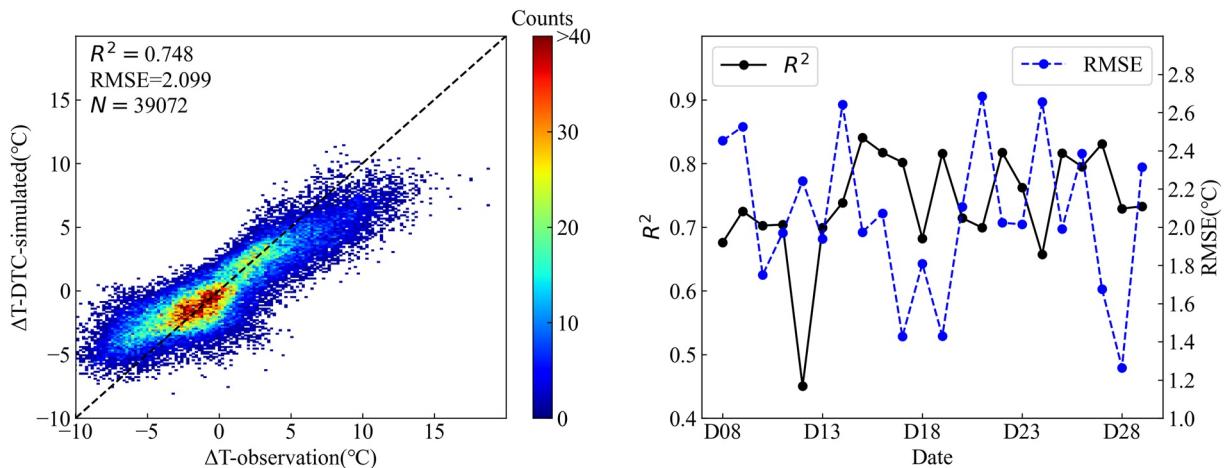


Figure 5. As in Figure 4, but for the model application to the forecast days (D8–D29, 22 days in total).

air temperature differences. The distribution of DTC model-simulated results fitted the one-to-one line well. During the 23-day period, the majority of R^2 values were greater than 0.9, and all RMSE values were less than 1.3° C (Figure 4 right). The DTC model achieves a very good performance in fitting the diurnal air temperature differences.

Then the DTC model was applied on the forecast day, and the performance is shown in Figure 5 (left). The average R^2 value was 0.748 during the 22 forecast days for all AWSs, and the average RMSE was 2.10° C. The distribution of the DTC model-simulated results also fits the one-to-one line well. The mean values of the daily R^2 and RMSE between the DTC model and observation were 0.74° and 2.06° C, respectively, indicating a very good performance in fitting the diurnal air temperature on the forecast day (Figure 5 right). Therefore, the DTC model can be used to forecast and correct the air temperature differences on the forecast day.

However, it was found that DTC performed poorly on D14, D21, and D24 (Figure 5 right). Take D14 as an example. It exhibited warm bias at cold temperatures and cold bias at warm temperatures (Figure 6 right), which mainly comes from WRF simulation, as shown in Figure 6. The downscaling model in this study can reduce these biases to some extent, but not totally. There will be continuous improvement in the follow-up research to reduce this kind of bias.

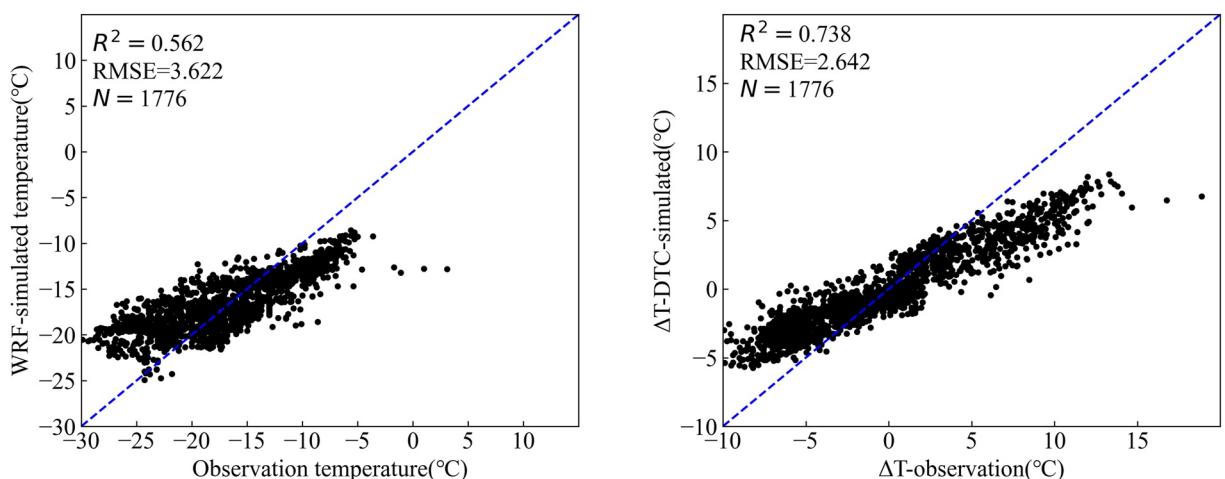


Figure 6. Scatter plot of observation temperature versus weather research and forecasting (WRF)-simulated temperature on D14 (left), and ΔT -observation versus ΔT -diurnal temperature cycle (DTC)-simulated on D14 (right; each point represents 1 hour at one automatic weather station [AWS] station).

Table 1

Weather Research and Forecasting (WRF) Simulation (WRF-Simulated) and Multiplicative Correction Method (WRF-Corrected) Root-Mean-Square Error (RMSE) ($^{\circ}\text{C}$) (n Denotes the Number of Days of Historical Data)

Date	WRF-simulated	WRF-corrected			
		$n = 1$	$n = 2$	$n = 3$	$n = 4$
D8	1.87	2.43	--	--	--
D9	2.43	1.16	2.59	--	--
D10	1.22	2.51	2.96	2.09	--
D11	1.25	1.51	1.40	1.37	1.26
D12	0.97	1.46	2.98	1.17	1.34
D13	2.31	2.30	2.95	2.12	2.29
D14	1.95	2.18	2.26	1.38	1.32
D15	2.28	1.29	4.20	1.89	1.89
D16	2.83	1.31	1.92	1.85	2.00
D17	1.39	2.24	1.69	1.40	1.33
D18	1.23	0.73	1.73	1.64	1.60
D19	2.76	2.03	1.65	2.61	2.67
D20	1.89	1.31	0.98	0.83	1.08
D21	3.15	2.10	4.76	2.16	2.13
D22	3.40	0.86	1.88	1.70	1.89
D23	2.72	1.35	1.71	0.63	0.63
D24	2.76	1.08	1.24	0.98	0.86
D25	2.96	1.03	0.91	0.50	0.59
D26	2.67	1.35	1.71	1.15	1.21
D27	1.51	1.01	1.13	1.07	1.13
D28	1.92	1.72	1.73	1.15	1.10
D29	2.13	1.42	1.37	0.87	1.02
Mean	2.16	1.56	2.08	1.43	1.44

4.2. Performance of the Multiplicative Correction Method at the AWSs

In this paper, the daily mean air temperature, which was regarded as the main meteorological variable, was corrected. The correct daily mean air temperature over the next 24 hr at all AWSs was calculated with Equation 5. The accuracy of the corrected results was verified against the observation data. The RMSE was calculated to assess the performance of the method. The smaller the RMSE value is, the more accurate the corrected results.

The original RMSE value of the WRF simulation results (WRF-simulated) and RMSE of the multiplicative correction results (WRF-corrected) were calculated. The results are listed in Table 1. With increasing n , the result does not improve. Therefore, n was set to 3 in the following research. Moreover, the daily mean temperature used to predict the next day's temperature was also determined.

4.3. Combined Downscaling Model Performance at the AWSs

Hourly difference temperature from the daily mean values has been downscaled and corrected by the DTC model. To get the final downscaling air temperature, the only unknown variable is the daily mean temperature on the forecast day when forecasting the temperature on the next day at the AWSs. The daily mean temperature on the forecast day is obtained from the WRF downscaling results via the GWR method. The GWR-derived results, multiplicative correction model, and DTC model were combined to predict the air temperature at the AWSs.

The predictive power of the combined model was evaluated against the observation data set extracted from the AWSs, but the data retrieved from all 74 stations was not involved in model building. All the hourly values at all

Table 2

Performance Statistics of the Weather Research and Forecasting (WRF) Simulation Results (WRF-Simulated) and Combined Model Results (WRF-Downscaled) Against the Ground Station Observation Data Over the Mean Value at All Automatic Weather Stations (AWSs)

Date	WRF-simulated			WRF-downscaled		
	R ²	RMSE	MAE	R ²	RMSE	MAE
D10	0.47	2.61	2.06	0.40	2.76	2.15
D11	0.60	2.40	1.88	0.61	2.36	2.39
D12	0.55	2.26	1.77	0.45	2.50	2.41
D13	0.35	3.14	2.32	0.46	2.85	2.43
D14	0.56	3.62	2.94	0.73	2.86	1.60
D15	0.53	3.62	3.08	0.73	2.73	2.0
D16	0.43	4.09	3.47	0.71	2.90	2.34
D17	0.38	2.73	2.22	0.66	2.00	1.75
D18	0.46	2.53	1.95	0.50	2.45	2.93
D19	0.26	3.35	2.93	0.42	2.98	2.43
D20	0.46	3.09	2.43	0.72	2.22	1.60
D21	0.36	4.19	3.62	0.58	3.40	2.11
D22	0.21	4.44	3.92	0.69	2.80	1.59
D23	0.27	3.78	3.21	0.76	2.14	2.12
D24	0.28	4.04	3.53	0.65	2.81	1.64
D25	0.30	4.08	3.48	0.83	2.02	1.40
D26	0.48	4.02	3.45	0.77	2.68	2.02
D27	0.52	2.94	2.41	0.75	2.12	2.05
D28	0.26	3.23	2.62	0.65	1.70	2.15
D29	0.49	3.39	2.85	0.73	2.48	1.79
Total	0.60	3.44	2.81	0.78	2.57	2.10

stations on the 20 days were compared (D10-D29). The raw WRF simulation results (WRF-simulated), combined model results (WRF-downscaled) and measured values at the AWSs were compared. Three different statistical scores, which are the most adopted accuracy measures, were calculated to assess the model performance, including the coefficient of determination (R^2), RMSE and mean absolute error (MAE). The statistical results are provided in Table 2.

We observed a notable increase in the performance metrics of the combined model results over the WRF simulation results as expected. The R^2 value is higher than WRF simulation results. The RMSE and MAE values decrease by 0.87°C and 0.71°C, respectively. The combined model achieves a good performance at the AWSs regarding the correction and downscaling of the air temperature forecasts. As shown in Figure 7, WRF-downscaled results fitted one-to-one line well, and the combined model reduced the warm bias at cold temperatures and cold bias at warm temperatures.

4.4. Downscaling Model Spatial Uncertainty Assessment

After verifying the correlation between the AWSs data, WRF simulation results and combined model results, we can investigate the spatial correlation between the WRF simulation results and combined model results at locations where ground stations do not occur. GWR interpolation was used to obtain observation data spatial distributions (GWR-interpolation). In this section, the spatial distributions of the WRF simulation results (WRF-simulated), combined model results (WRF-downscaled), GWR-interpolation results of air temperature observation, and their spatial bias are compared.

Spatial forecasting data was calculated with the combined model, of which the coefficients had already been spatialized. The independent variables of the GWR downscaling model were obtained with a 30-m resolution DEM, so we could obtain combined model results at a 30-m spatial resolution. Regarding spatial forecasting, 50 AWSs (~2/3 of the total number) were selected to build the combined model, and the other 24 stations (~1/3 of the total number) were used to validate the combined model. The RMSE values of WRF-simulated and WRF-downscaled model results on the 20 forecast days are shown in Figure 8. The combined model attains a better performance than does the raw WRF simulation model.

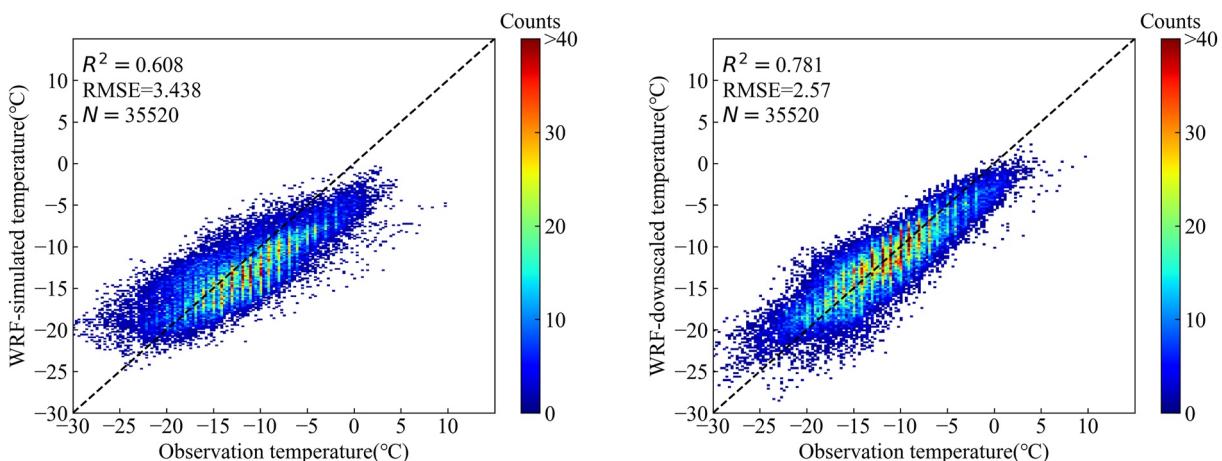


Figure 7. The relationship between observation temperature and weather research and forecasting (WRF)-simulated temperature (each point represents one hour at one automatic weather station [AWS] station) during the 20 forecast days (left), and the relationship between observation temperature and WRF-downscaled temperature during the 20 forecast days (right).

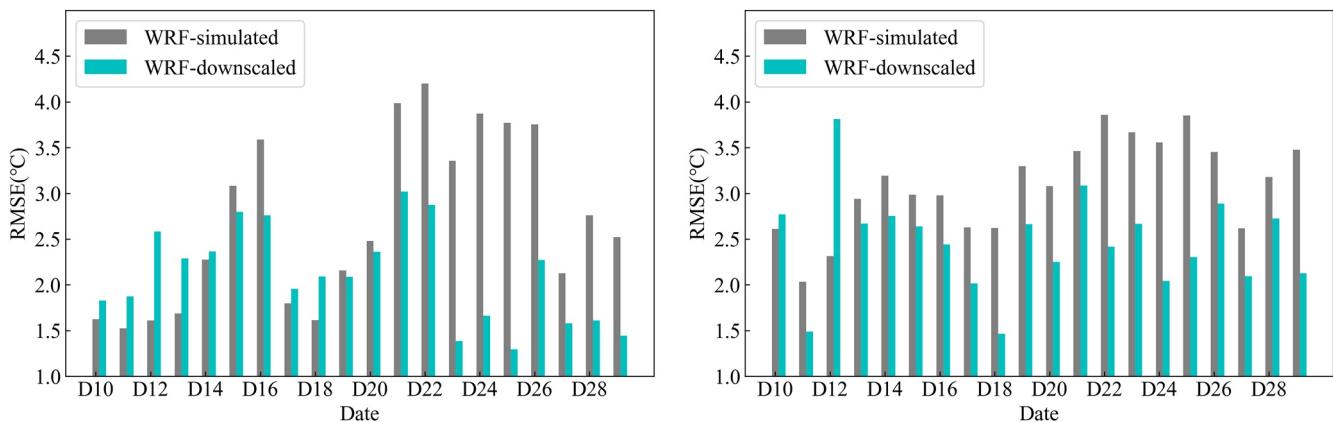


Figure 8. Root-mean-square error (RMSE) of the combined model results (weather research and forecasting [WRF]-downscaled) and WRF simulation results (WRF-simulated) during the 20 forecast days (left; daily mean value of 24 verification automatic weather stations [AWSs]), and RMSE of the combined model results (WRF-downscaled) and WRF simulation results (WRF-simulated) on the 20 forecast days (right; daily mean value of 2,000 random points over the domain).

The mean RMSE value of WRF simulated during the 20 days is 2.69°C , and that of the combined model is 2.11°C , which decreased by 0.58°C on average. During the total 20 forecast days, 13 days' results produced from the combined model were lower than WRF simulated, with a mean value decreased by 1.12°C . The other 7 days' results showed a slightly increasing trend, with a mean value of 0.41°C . The same pattern was observed for MAE, in which the mean value decreased from 2.30°C of the WRF simulation to 1.74°C of the combined model.

Then, 2,000 random points were generated, and WRF-simulated, WRF-downscaled and observation values were extracted from their spatialization results by these random points over the domain. The RMSE values of WRF-simulated and WRF-downscaled model results on the 20 forecast days showed the same pattern with the above 24 verification AWSs (Figure 6). In most cases, compared to WRF-simulated, the RMSE of WRF-downscaled greatly decreased by 0.62°C on average (from 3.09°C to 2.47°C), and the mean MAE decreased by 0.50°C (from 2.51°C to 2.01°C).

As shown in Figure 8, the RMSE of WRF-downscaled was lower than WRF-simulation most days, but the result of D12 was always obviously higher. The mean temperature calculated in Equation 5 was higher than the real value when there was a sudden drop in the temperature, since the mean value was an average state if n was set to 3 (n is the number of days of historical data in Equation 6). During D11 to D13, there was a cold front passage in this study (Figure 9 left). The higher mean data led to DTC model results significantly higher than the real value

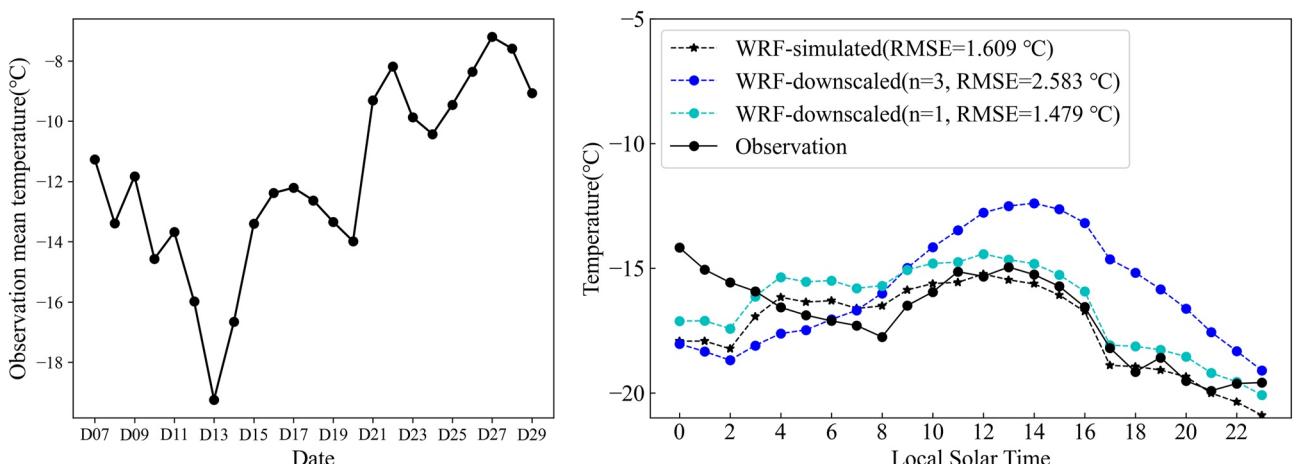


Figure 9. Dynamic pattern of observation mean temperature during the total of 23 days (left; each point represents the daily mean temperature of all 74 automatic weather stations [AWSs]), and comparison among values of the combined model downscaling results (weather research and forecasting [WRF]-downscaled; n is the number of days of historical data in Equation 6), the WRF simulation results (WRF-simulated) and observation data on D12 (right; each point represents the hourly mean temperature of 24 verification AWSs).

during the day, and significantly lower at night (Figure 9 right). It was found that this error could be well corrected by setting the coefficient n of Equations 6–1. As shown in Figure 9 (right), the RMSE of WRF-downscaled decreased from 2.58°C ($n = 3$) to 1.48°C ($n = 1$).

A scatter plot of the WRF simulation results (WRF-simulated) and combined model downscaling results (WRF-downscaled) on D17 against the GWR-interpolated air temperature values at the four times is shown in Figure 10. It is observed that the distribution of the WRF-downscaled results fits the one-to-one line well, but the distribution of the WRF-simulated results does not fit the line very well. Moreover, R^2 of the WRF-downscaled results is greater than that of the WRF-simulated results. These results indicate that the downscaled WRF model attains a very good performance in correcting and downscaling the forecasted air temperature in this mountainous area.

Although there is a bias between the interpolation results (GWR-interpolation) and the real spatial distribution, the result of GWR-interpolation can approximately express the real spatial distribution of air temperature. Here, we use the GWR interpolation results as a benchmark to compare the spatial bias in the WRF-simulated and WRF-downscaled model results. The spatial distributions of the WRF simulation, WRF downscaling, and GWR interpolation results during four consecutive days at 2 and 14 hr are shown in Figures 8 and 9, respectively. Figure 11 shows that the WRF simulation results are generally higher than the GWR interpolation results at 2 hr local solar time, the spatial distribution of the WRF downscaling results is closer to that of the GWR interpolation results, and WRF downscaling can display more detailed air temperature changes attributed to terrain fluctuations. In contrast to the results at 2 hr, the results at 14 hr (Figure 12) show that the WRF simulation values are generally lower than the GWR interpolation values. The WRF-downscaled results are similar to those at 2 hr, that is, closer to the observation interpolation results, and can display detailed changes caused by terrain variations. Through these spatial distributions, it can be found that the spatial distribution of the combined model prediction results is closer to that of the interpolation results based on observation data determined via GWR. In general, the WRF-simulated temperature values were higher than the observation data at the ground stations at night and lower during the daytime.

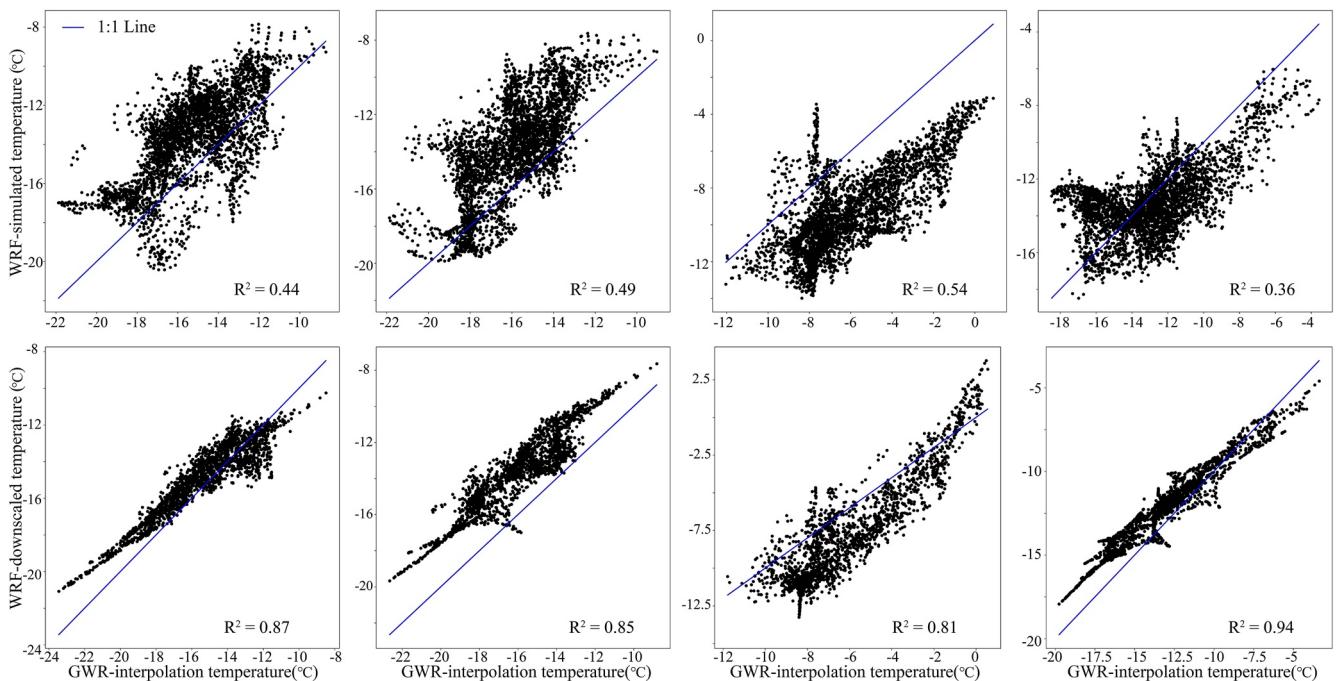


Figure 10. Scatter plot of the weather research and forecasting (WRF) simulation results versus observation interpolation results (top; WRF-simulated) and the combined model results versus observation interpolation results (bottom; WRF-downscaled) at four local solar times on D17 (the results at 2, 8, 14, and 20 hr were shown from left to right).

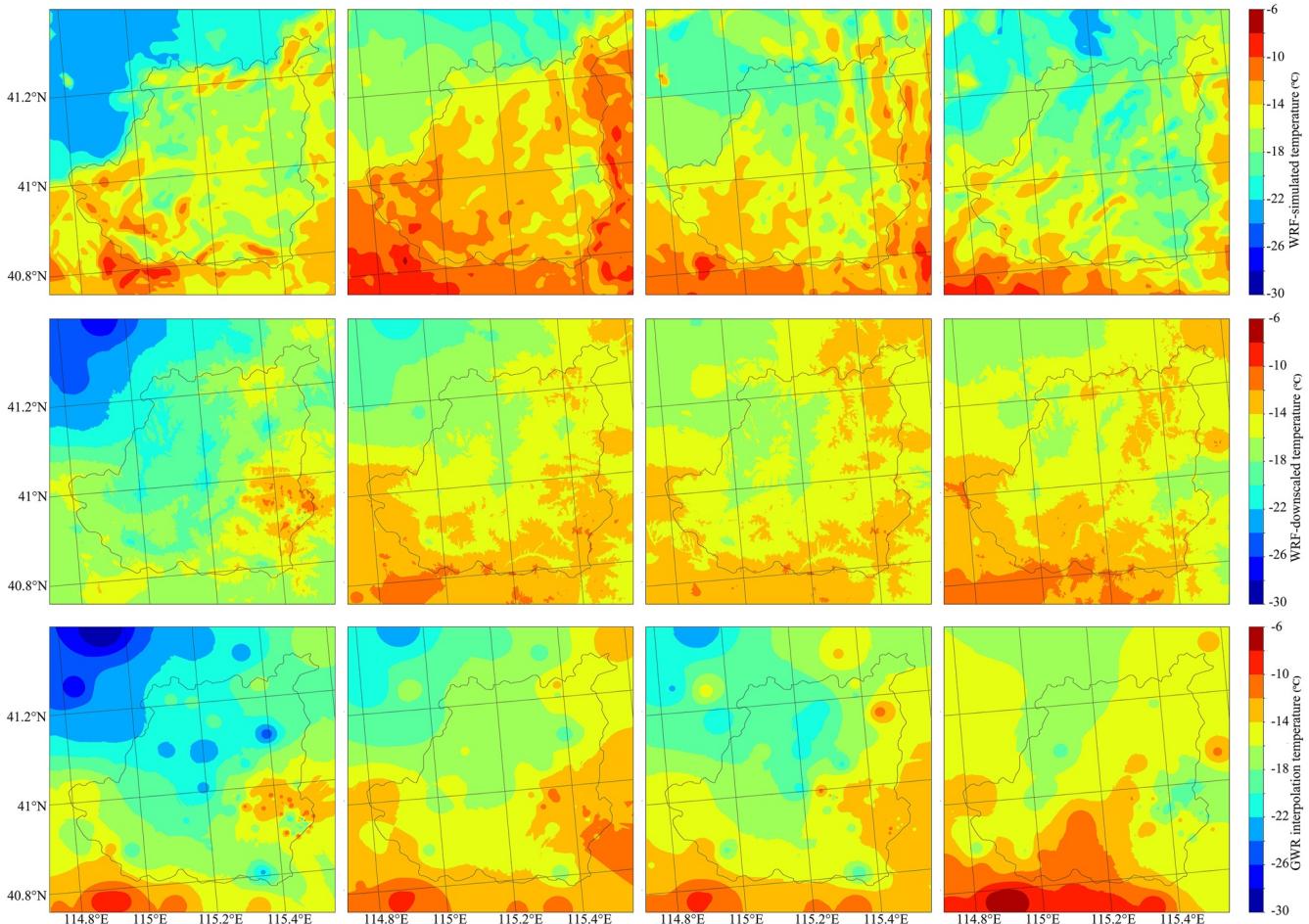


Figure 11. Air temperature spatial distribution of the weather research and forecasting (WRF) simulation results (WRF-simulated), combined model downscaling results (WRF-downscaled) and observation data-based interpolation results (geographically weighted regression [GWR] interpolation) at 2 hr (the results on D16, D17, D18 and D19 were shown from left to right; the WRF-simulated, WRF-downscaled and GWR-interpolation results were shown from top to bottom).

The spatial distribution of the bias between the combined model downscaling results (WRF-downscaled) and the observation data-based GWR interpolation results (GWR-interpolation) at 2 and 14 hr are shown in Figure 13. At 2 hr, the bias in most areas is negative, suggesting that the WRF downscaling results are higher than the GWR interpolation results. Combined with Figure 11, the WRF simulation results are generally higher than the GWR interpolation results at 2 hr local solar time. The generally negative bias is the cause for this phenomenon. The minimum-bias areas are located in the northeastern and southeastern parts of the study area on D16, D17, and D18, and these areas are located at higher altitudes than those of the other areas. In contrast to these three days, the minimum-bias areas were mainly located in the southwestern part on D19, and the southwest part has a lower altitude throughout the whole study area. We found that the bias at 2 hr was mainly caused by a temperature inversion, which commonly occurs at night. As shown in Figure 14, the observation temperature at the AWSs increases with increasing altitude on D16, D17, and D18 (under normal conditions, the air temperature usually decreases with the altitude). However, the WRF simulations fail to predict the temperature inversion phenomenon. The WRF simulation temperature at the AWSs decreases with increasing altitude. This leads to a bias increase with increasing altitude when a temperature inversion occurs. Although a temperature inversion occurs on D17, the top of the inversion layer is located at ~1,600 m. For this reason, the absolute bias on D17 is lower than that on D16 and D18 in high-altitude areas. In contrast to these three days, inversion did not occur on D19. The spatial pattern on D19 is different from that on the other three days, and the maximum absolute bias occurs in the low-altitude area located in the southwestern part of the study area. In conclusion, temperature inversion affects the accuracy of temperature forecasts at night. Prediction of inversion and its intensity may help to improve the accuracy of air temperature predictions in mountainous areas at night.

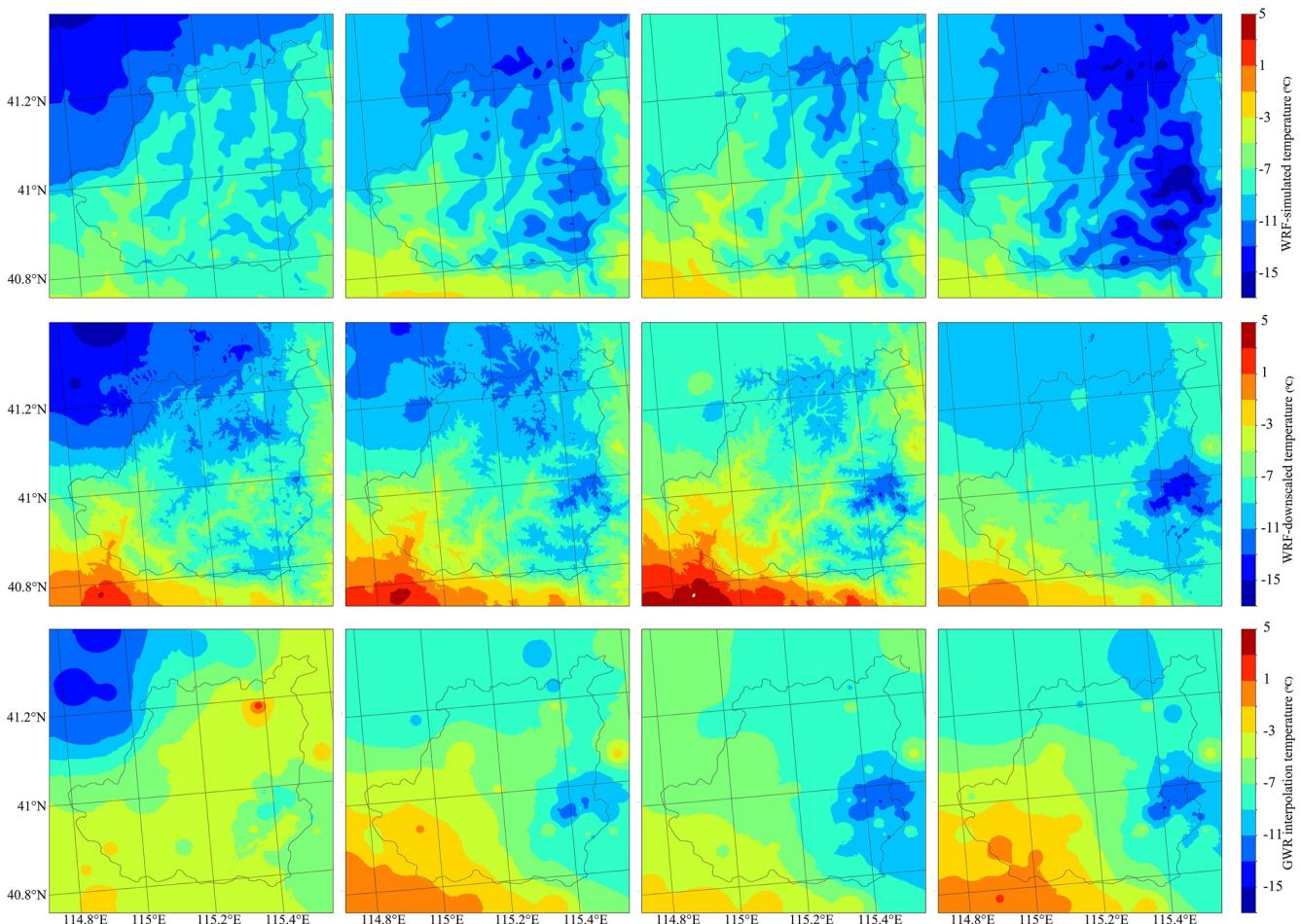


Figure 12. The same as in Figure 11, but for 14 hr.

At 14 hr, the bias in most areas is positive, suggesting that the WRF downscaling results are lower than the GWR interpolation results. The generally positive bias is caused by the overall lower WRF simulation values than the GWR interpolation results at 14 hr. The MAE values of the WRF-simulated temperatures at 14 hr are 3.48°C, 3.39°C, 1.52°C, and 4.74°C on D16, D17, D18 and D19, respectively. The MAE value at 14 hr was much higher than that at the other times. After downscaling, the MAE values of the WRF downscaling results are 2.90°C, 1.95°C, 1.56°C, and 1.85°C on D16, D17, D18, and D19, respectively. There is a notable decrease in MAE on all forecast days at 14 hr. However, MAE remains higher at 14 hr than that at the other local solar times because the raw bias in the WRF simulations is very high at 14 hr. This leads to an overall high bias in the spatial distribution of the WRF downscaling results at 14 hr. The combined downscaling model performed well when downscaling the air temperature in this mountainous area. However, its accuracy depends on the accuracy of the WRF simulation model.

5. Conclusions

This study evaluated the use of a combination of the GWR method, bias correction, DTC model and MLR in WRF downscaling and correction in a mountainous area to reduce the bias and improve the 2-m air temperature accuracy and spatial resolution of WRF simulations. The 2-m air temperature series predicted with the WRF model was first downscaled to 30 m from a 1-km grid resolution via GWR. The DTC model was used to fit diurnal variation curves of the air temperature during 1 day, and these curves were applied to predict future 24-hr hourly air temperatures. A multiplicative correction was adopted to correct the GWR downscaling results and to improve the accuracy of the daily mean air temperature determined via WRF simulations. MLR was employed in

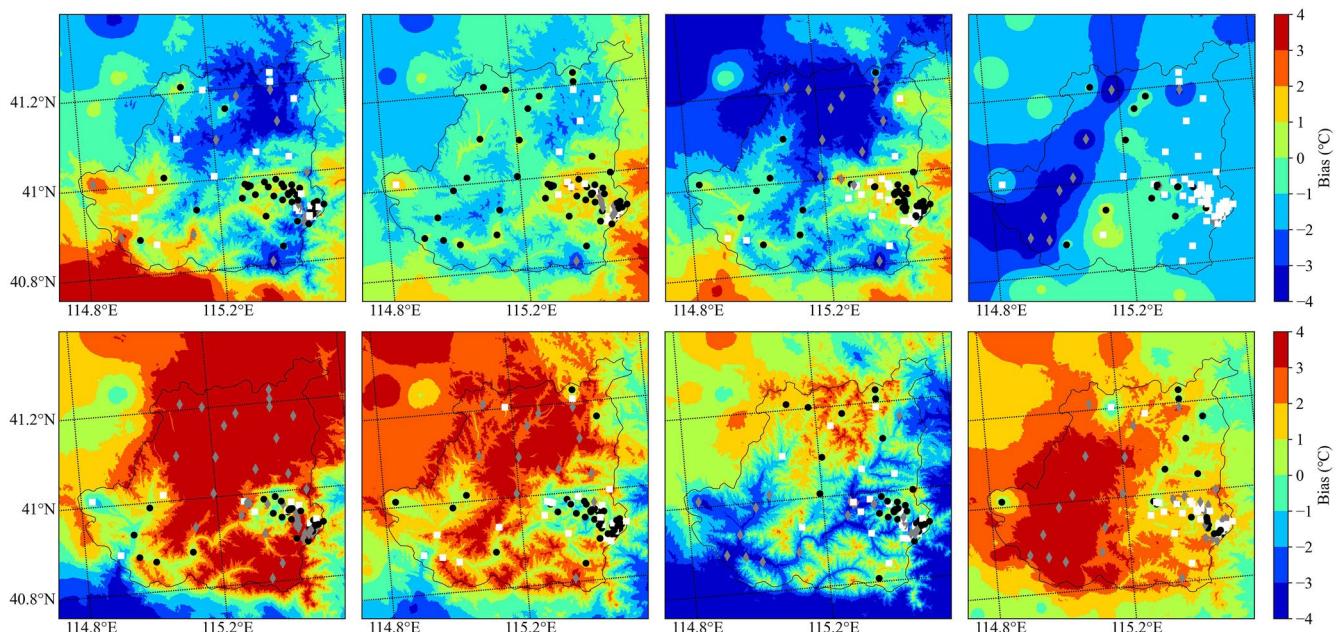


Figure 13. The spatial distribution of the air temperature bias between the geographically weighted regression (GWR) interpolation and weather research and forecasting (WRF) downscaling results (the results on D16, D17, D18, and D19 were shown from left to right). The top row shows the results at 2 hr, and the bottom row shows the results at 14 hr. The black dots indicate mean absolute error (MAE) biases less than 1°C, the white square dots indicate MAE biases between 1°C and 2°C, and the gray diamond-shaped dots represent MAE biases greater than 2°C.

the spatialization of all correction and DTC coefficients to train the relationship between these coefficients and the slope, aspect, and altitude of the study area. We tested this approach against ground observation data provided by the Hebei Meteorological Administration.

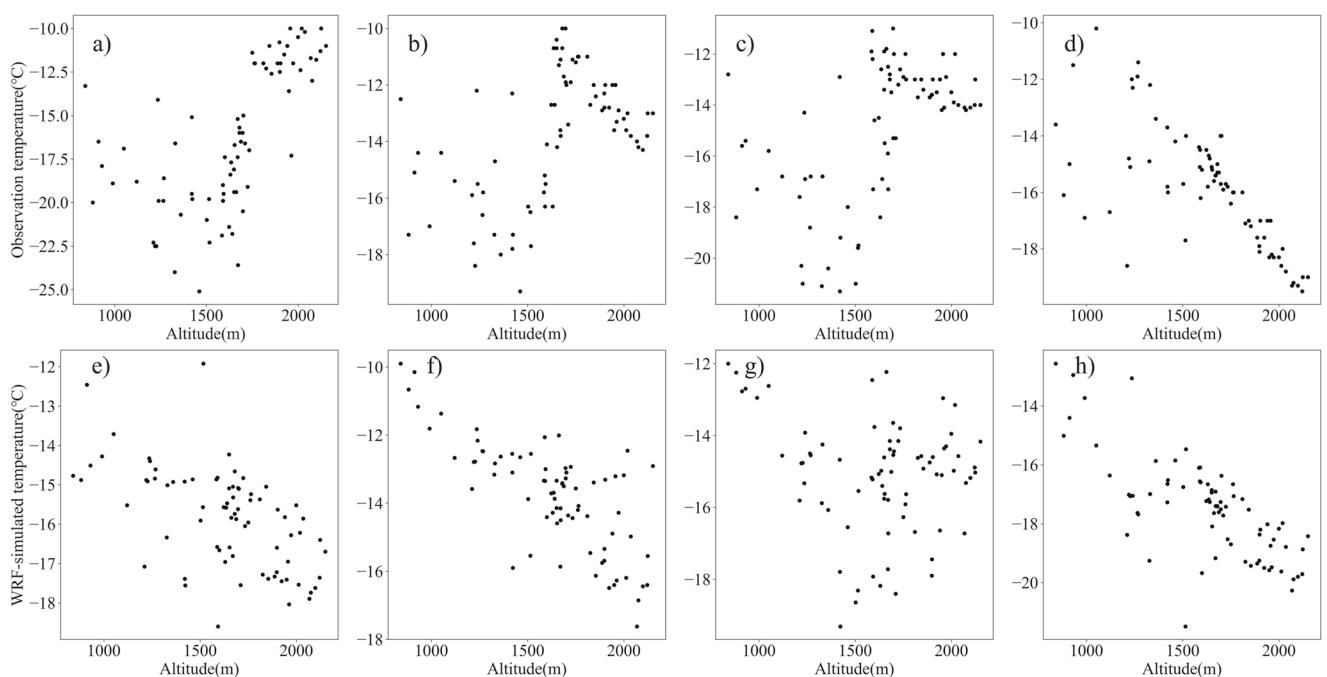


Figure 14. Scatter plots of the raw weather research and forecasting (WRF) simulation results (WRF-simulated) versus observed temperature and altitude of all automatic weather stations (AWSs) at 2 hr (the top row shows the observed temperature, and the bottom row shows the WRF simulation results; the results on D16, D17, D18, and D19 were shown from left to right).

The combined downscaling model results (WRF-downscaled) can better express the detailed features of the air temperature than the WRF simulation results can. Moreover, the influence of the topography is reduced, and the forecasting accuracy is improved. The scheme was designed to operate at hourly time scales, and hourly values are generally highly variable and more notably affected by weather conditions than are daily or monthly means. The research presented here shows that the air temperature determined in WRF simulations can be downscaled and corrected by this method, and the accuracy and spatial resolution of temperature forecasts can be improved in this area. Compared to the WRF simulation results, RMSE of the WRF-downscaled results decreased 0.87°C at the AWS level and by 0.62°C over the domain, and MAE decreased by 0.71°C and 0.51°C, respectively, at these two levels. Moreover, the correlation between the WRF-downscaled and observation data (or observation data-based GWR interpolation results) matched the one-to-one line better than did the WRF simulation results. The spatial distribution of the WRF downscaling results was closer to that of the observation data-based GWR interpolation results than that of the WRF simulation results.

There are three major findings. First, the multiplicative correction method is useful to correct the WRF simulation air temperature, and the accuracy of this method is higher when correcting the daily mean air temperature than when correcting the hourly air temperature. Moreover, for short-term forecasting (in this study, daily forecasting), better results can be obtained by setting n to 3. The coefficient of the combined model should be adjusted when there was a sudden change in the temperature, such as a cold passage. Second, the DTC model can represent most of the daily variations, and it performs well in forecasting. Third, the proposed 2-m air temperature prediction model attains a high prediction accuracy and stability, provides high-resolution day-ahead predictions and is suitable for actual temperature prediction in this mountainous area.

As the combined downscaling model in this paper does not require a long computing time and many computational resources for implementation and data sources are easy to obtain, this methodology can readily be extended to other variables, regions, and numerical models. With additional data, such as the slope, aspect, and land cover, it should be possible to improve the spatialization effect of the models and the overall model performance. It is anticipated to further develop high-spatial and high-temporal resolution temperature products across a large region of mountainous areas in North China.

Despite the good performance of the combined downscaling model, there are certain disadvantages, for example, the smooth curve of the DTC model hardly fits any sudden change in the temperature. This kind of sudden change in temperature is not only the result of random errors but actually occurs in the diurnal temperature cycle. Moreover, temperature inversion imposes a great influence on air temperature forecasts at night. To forecast sudden changes, temperature inversion prediction and elimination are required, which are not addressed in this paper, and should be studied in detail in follow-up research.

Data Availability Statement

The observation data are not publicly available due to the relevant national regulations on the management of meteorological data. Readers can access them through contact with the Hebei Meteorological Administration. The DEM data were obtained from the Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (ASTGTM) Version 2 (<https://lpdaac.usgs.gov/products/astgtmv002/>).

References

- Abdel-Aal, R. E. (2004). Hourly temperature forecasting using abductive networks. *Engineering Applications of Artificial Intelligence*, 17(5), 543–556. <https://doi.org/10.1016/j.engappai.2004.04.002>
- Abrams, M., Bailey, B., Tsu, H., & Hato, M. (2010). The aster global DEM. *Photogrammetric Engineering & Remote Sensing*, 76(4), 344–348.
- Adachi, S. A., & Tomita, H. (2020). Methodology of the constraint condition in dynamical downscaling for regional climate evaluation: A review. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD032166. <https://doi.org/10.1029/2019JD032166>
- Alessi, M. J., & DeGaetano, A. T. (2021). A comparison of statistical and dynamical downscaling methods for short-term weather forecasts in the US Northeast. *Meteorological Applications*, 28, e1976. <https://doi.org/10.1002/met.1976>
- Bechler, A., Vrac, M., & Bel, L. (2015). A spatial hybrid approach for downscaling of extreme precipitation fields. *Journal of Geophysical Research: Atmospheres*, 120, 4534–4550. <https://doi.org/10.1002/2014JD022558>
- Ben-Nakhi, A. E., & Mahmoud, M. A. (2004). Cooling load prediction for buildings using general regression neural networks. *Energy Conversion and Management*, 45, 2127–2141. <https://doi.org/10.1016/j.enconman.2003.10.009>
- Borrego, C., Monteiro, A., Pay, M. T., Ribeiro, I., Miranda, A. I., Basart, S., & Baldasano, J. M. (2011). How bias-correction can improve air quality forecasts over Portugal. *Atmospheric Environment*, 45, 6629–6641. <https://doi.org/10.1016/j.atmosenv.2011.09.006>

- Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., & Spinoni, J. (2014). High-resolution temperature climatology for Italy: Interpolation method intercomparison. *International Journal of Climatology*, 34, 1278–1296. <https://doi.org/10.1002/joc.3764>
- Caldwell, P., Chin, H. N. S., Bader, D. C., & Bala, G. (2009). Evaluation of a WRF dynamical downscaling simulation over California. *Climatic Change*, 95, 499–521. <https://doi.org/10.1007/s10584-009-9583-5>
- Casellas, E., Bech, J., Veciana, R., Miro, J. R., Sairouni, A., & Pineda, N. (2020). A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, 246, 105103. <https://doi.org/10.1016/j.atmosres.2020.105103>
- Cheng, C. S., Li, G., Li, Q., & Auld, H. (2008). Statistical downscaling of hourly and daily climate scenarios for various meteorological variables in South-central Canada. *Theoretical and Applied Climatology*, 91, 129–147. <https://doi.org/10.1007/s00704-007-0302-8>
- Chung, U., Seo, H. H., Hwang, K. H., Hwang, B. S., Choi, J., Lee, J. T., & Yun, J. I. (2006). Minimum temperature mapping over complex terrain by estimating cold air accumulation potential. *Agricultural and Forest Meteorology*, 137(1–2), 15–24. <https://doi.org/10.1016/j.agrformet.2005.12.011>
- Cifuentes, J., Marulanda, G., Bello, A., & Reneses, J. (2020). Air temperature forecasting using machine learning techniques: A review. *Energies*, 13, 4215. <https://doi.org/10.3390/en13164215>
- Dickinson, R. E., Errico, R. M., Giorgi, F., & Bates, G. T. (1989). A regional climate model for the western United States. *Climatic Change*, 15(3), 383–422. <https://doi.org/10.1007/BF00240465>
- Dombayci, Ö. A., & Gölcü, M. (2009). Daily means ambient temperature prediction using artificial neural network method: A case study of Turkey. *Renewable Energy*, 34, 1158–1161.
- Duan, S. B., Li, Z. L., Tang, B. H., Wu, H., Tang, R., Bi, Y., & Zhou, G. (2014). Estimation of diurnal cycle of land surface temperature at high temporal and spatial resolution from clear-sky MODIS data. *Remote Sensing*, 6, 3247–3262. <https://doi.org/10.3390/rs6043247>
- Duan, S. B., Li, Z. L., Wang, N., Wu, H., & Tang, B. H. (2012). Evaluation of six land-surface diurnal temperature cycle models using clear-sky in situ and satellite data. *Remote Sensing of Environment*, 124, 15–25. <https://doi.org/10.1016/j.rse.2012.04.016>
- Feng, S., Hu, Q., & Qian, W. (2004). Quality control of daily meteorological data in China, 1951–2000: A new dataset. *International Journal of Climatology*, 24(7), 853–870. <https://doi.org/10.1002/joc.1047>
- Fotheringham, A. S., Charlton, M., & Brunsdon, C. (1997). Measuring spatial variations in relationships with geographically weighted regression. *Measuring spatial variations in relationships with geographically weighted regression* (pp. 60–82). Springer. https://doi.org/10.1007/978-3-662-03499-6_4
- Fotheringham, A. S., & Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: Dispelling the myth. *Journal of Geographical Systems*, 18(4), 303–329. <https://doi.org/10.1007/s10109-016-0239-5>
- Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27, 1547–1578. <https://doi.org/10.1002/joc.1556>
- Georganos, S., Abdi, A. M., Tenenbaum, D. E., & Kalogirou, S. (2017). Examining the NDVI-rainfall relationship in the semi-arid Sahel using geographically weighted regression. *Journal of Arid Environments*, 146, 64–74. <https://doi.org/10.1016/j.jaridenv.2017.06.004>
- Gholamnia, M., Alavipanah, S. K., Boloorani, A. D., Hamzeh, S., & Kiavarz, M. (2017). Diurnal air temperature modeling based on the land surface temperature. *Remote Sensing*, 9, 915. <https://doi.org/10.3390/rs9090915>
- Gholamnia, M., Alavipanah, S. K., Boloorani, A. D., Hamzeh, S., & Kiavarz, M. (2019). A new method to model diurnal air temperature cycle. *Theoretical and Applied Climatology*, 137, 229–238. <https://doi.org/10.1007/s00704-018-2587-1>
- Giorgi, F., & Bates, G. T. (1989). The climatological skill of a regional model over complex terrain. *Monthly Weather Review*, 117, 2325–2347. [https://doi.org/10.1175/1520-0493\(1989\)117<2325:tcsoar>2.0.co;2](https://doi.org/10.1175/1520-0493(1989)117<2325:tcsoar>2.0.co;2)
- Goovaerts, P. (1999). Geostatistics in soil science: State-of-the-art and perspectives. *Geoderma*, 89, 1–45. [https://doi.org/10.1016/s0016-7061\(98\)00078-0](https://doi.org/10.1016/s0016-7061(98)00078-0)
- Harding, K. J., Snyder, P. K., & Liess, S. (2013). Use of dynamical downscaling to improve the simulation of Central U.S. warm season precipitation in CMIP5 models. *Journal of Geophysical Research*, 118, 12522–12536. <https://doi.org/10.1002/2013JD019994>
- Hofer, M., Marzeion, B., & Mölg, T. (2015). A statistical downscaling method for daily air temperature in data-sparse, glaciated mountain environments. *Geoscientific Model Development*, 8, 579–593. <https://doi.org/10.5194/gmd-8-579-2015>
- Holden, Z. A., Abatzoglou, J. T., Luce, C. H., & Baggett, L. S. (2011). Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agricultural and Forest Meteorology*, 151, 1066–1073. <https://doi.org/10.1016/j.agrformet.2011.03.011>
- Hu, L., Sun, Y., Collins, G., & Fu, P. (2020). Improved estimates of monthly land surface temperature from MODIS using a diurnal temperature cycle (DTC) model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 168, 131–140. <https://doi.org/10.1016/j.isprsjprs.2020.08.007>
- Ines, A. V. M., & Hansen, J. W. (2006). Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138, 44–53. <https://doi.org/10.1016/j.agrformet.2006.03.009>
- Jiang, Y., Yang, K., Shao, C., Zhou, X., Zhao, L., Chen, Y., & Wu, H. (2021). A downscaling approach for constructing high-resolution precipitation dataset over the Tibetan Plateau from ERA5 reanalysis. *Atmospheric Research*, 256, 105574. <https://doi.org/10.1016/j.atmosres.2021.105574>
- Ke, Z. J., Zhang, P. Q., Chen, L. J., & Du, L. M. (2011). An experiment of a statistical downscaling forecast model for summer precipitation over China. *Atmospheric and Oceanic Science Letters*, 4, 270–275. <https://doi.org/10.1080/16742834.2011.11446941>
- Konovalov, I. B., Beekmann, M., Meleux, F., Dutot, A., & Foret, G. (2009). Combining deterministic and statistical approaches for PM10 forecasting in Europe. *Atmospheric Environment*, 43, 6425–6434. <https://doi.org/10.1016/j.atmosenv.2009.06.039>
- Le Roux, R., Katurji, M., Zawar-Reza, P., Quénol, H., & Sturman, A. (2018). Comparison of statistical and dynamical downscaling results from the WRF model. *Environmental Modelling & Software*, 100, 67–73. <https://doi.org/10.1016/j.envsoft.2017.11.002>
- Li, S., Goel, L., & Wang, P. (2016). An ensemble approach for short-term load forecasting by extreme learning machine. *Applied Energy*, 170, 22–29. <https://doi.org/10.1016/j.apenergy.2016.02.114>
- Mok, K. M., Miranda, A. I., Yuen, K. V., Hoi, K. I., Monteiro, A., & Ribeiro, I. (2017). Selection of bias correction models for improving the daily PM10 forecasts of WRF-EURAD in Porto, Portugal. *Atmospheric Pollution Research*, 8, 628–639. <https://doi.org/10.1016/j.apr.2016.12.010>
- Neal, L. S., Agnew, P., Moseley, S., Ordóñez, C., Savage, N. H., & Tilbee, M. (2014). Application of a statistical post-processing technique to a gridded, operational, air quality forecast. *Atmospheric Environment*, 98, 385–393. <https://doi.org/10.1016/j.atmosenv.2014.09.004>
- Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, 16, 2421–2442. <https://doi.org/10.1175/jhm-d-14-0236.1>
- Reicosky, D. C., Winkelman, L. J., Baker, J. M., & Baker, D. G. (1989). Accuracy of hourly air temperatures calculated from daily minima and maxima. *Agricultural and Forest Meteorology*, 46, 193–209. [https://doi.org/10.1016/0168-1923\(89\)90064-6](https://doi.org/10.1016/0168-1923(89)90064-6)
- Schoof, J. T., & Pryor, S. C. (2001). Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *International Journal of Climatology*, 21, 773–790. <https://doi.org/10.1002/joc.655>
- Silibello, C., D'Allura, A., Finardi, S., Bolignano, A., & Sozzi, R. (2015). Application of bias adjustment techniques to improve air quality forecasts. *Atmospheric Pollution Research*, 6, 928–938. <https://doi.org/10.1016/j.apr.2015.04.002>

- Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., & Murphy, J. M. (2007). Improved surface temperature prediction for the coming decade from a global climate model. *Science*, 317, 796–799. <https://doi.org/10.1126/science.1139540>
- Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., & Wu, J. (2016). Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *Journal of Geophysical Research: Atmospheres*, 121, 2110–2129. <https://doi.org/10.1002/2015JD023977>
- Thornton, P. E., Running, S. W., & White, M. A. (1997). Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, 190, 214–251. [https://doi.org/10.1016/S0022-1694\(96\)03128-9](https://doi.org/10.1016/S0022-1694(96)03128-9)
- Vigaud, N., Vrac, M., & Caballero, Y. (2013). Probabilistic downscaling of GCM scenarios over southern India. *International Journal of Climatology*, 33, 1248–1263. <https://doi.org/10.1002/joc.3509>
- Viggiani, M., Busetto, L., Cimini, D., Di Paola, F., Gheraldi, E., Ranghetti, L., et al. (2019). A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements. *Agricultural and Forest Meteorology*, 276–277, 107590. <https://doi.org/10.1016/j.agrformet.2019.05.021>
- Wang, J., Fonseca, R. M., Rutledge, K., Martín-Torres, J., & Yu, J. (2020). A hybrid statistical-dynamical downscaling of air temperature over Scandinavia using the WRF model. *Advances in Atmospheric Sciences*, 37, 57–74. <https://doi.org/10.1007/s00376-019-9091-0>
- Wang, J., & Kotamarthi, V. R. (2014). Downscaling with a nested regional climate model in near-surface fields over the contiguous United States. *Journal of Geophysical Research: Atmospheres*, 119, 8778–8797. <https://doi.org/10.1002/2014JD021696>
- Wang, K., Li, Y., Luo, Z., Yin, S., & Chan, P. K. (2018). Harmonic analysis of 130-year hourly air temperature in Hong Kong: Detecting urban warming from the perspective of annual and daily cycles. *Climate Dynamics*, 51, 613–625. <https://doi.org/10.1007/s00382-017-3944-y>
- Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., & Mearns, L. O. (2004). Guidelines for use of climate scenarios developed from statistical downscaling methods. *Supporting material of the intergovernmental panel on climate change* (p. 27). DDC of IPCC TGCIA.
- Wilczak, J., McKeen, S., Djalalova, I., Grell, G., Peckham, S., Gong, W., et al. (2006). Bias-corrected ensemble and probabilistic forecasts of surface ozone over eastern North America during the summer of 2004. *Journal of Geophysical Research*, 111. <https://doi.org/10.1029/2006JD007598>
- Xu, W., Liu, P., Cheng, L., Zhou, Y., Xia, Q., Gong, Y., & Liu, Y. (2021). Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renewable Energy*, 163, 772–782. <https://doi.org/10.1016/j.renene.2020.09.032>
- Yan, D., Liu, T., Dong, W., Liao, X., Luo, S., Wu, K., et al. (2020). Integrating remote sensing data with WRF model for improved 2-m temperature and humidity simulations in China. *Dynamics of Atmospheres and Oceans*, 89, 101127. <https://doi.org/10.1016/j.dynatmoce.2019.101127>
- Yi, C., Shin, Y., & Roh, J. W. (2018). Development of an urban high-resolution air temperature forecast system for local weather information services based on statistical downscaling. *Atmosphere*, 9, 164. <https://doi.org/10.3390/atmos9050164>
- Zhang, H., Pu, Z., & Zhang, X. (2013). Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain. *Weather and Forecasting*, 28(3), 893–914. <https://doi.org/10.1175/WAF-D-12-00109.1>
- Zhou, J., Chen, Y., Zhang, X., & Zhan, W. (2013). Modelling the diurnal variations of urban heat islands with multi-source satellite data. *International Journal of Remote Sensing*, 34(21), 7568–7588. <https://doi.org/10.1080/01431161.2013.821576>
- Zhou, Q., Wang, C., & Fang, S. (2019). Application of geographically weighted regression (GWR) in the analysis of the cause of haze pollution in China. *Atmospheric Pollution Research*, 10, 835–846. <https://doi.org/10.1016/j.apr.2018.12.012>
- Zorita, E., & Storch, H. (1999). The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, 12, 2474–2489. [https://doi.org/10.1175/1520-0442\(1999\)012<2474:tamaas>2.0.co;2](https://doi.org/10.1175/1520-0442(1999)012<2474:tamaas>2.0.co;2)