Use of Google Earth Engine to Generate a 20-Year 1 Km × 1 Km Monthly Air Temperature Product Over Yellow River Basin

Meiling Gao [ID], Zhenhong Li, *Senior Member, IEEE*, Zhenyu Tan [ID], Huifang Li [ID], and Jianbing Peng

Abstract—Near-surface air temperature (NSAT) is a key parameter in climate changes, environmental ecosystem monitoring, and human settlement issues. As it is difficult for *in situ* observations to capture the spatial distribution characteristics of NSAT in great detail, various methods have been developed to use remotely sensed land surface temperature and other auxiliary variables to estimate the NSAT. Among them, machine learning turns out to be an exhilarating choice due to its superior performance. However, for long-term dataset estimation, when using machine learning methods with abundant related variables, computation loads and data storage cannot be ignored. Fortunately, Google Earth Engine (GEE) provides a wealth of geospatial datasets and powerful parallel computation capability. In this article, GEE has been utilized to demonstrate the feasibility of estimating long-term 1-km NSAT with two machine-learning models (random forest and deep neural network). After testing the effects of the input variables and the model portability, the 20-year monthly mean 1 km × 1 km NSAT over the Yellow River basin was generated and analyzed. Compared with *in situ* observations, the overall RMSE, MAE, R^2, and R of the NSAT product are 0.429, 0.302, 0.998, and 0.999, respectively. Specifically, for each observation station, R and R^2 are greater than 0.997, and RMSE and MAE are smaller than 0.5 and 0.4, respectively. The comparisons with three existing NSAT products show that our product outperforms the other three products. Once again, it is demonstrated in this study that GEE is a powerful platform to generate valuable products with ready-to-use datasets and computation resources.

Index Terms—Air temperature, Google Earth Engine (GEE), machine learning, model portability, variable effectiveness.

Meiling Gao and Zhenhong Li are with the College of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China, with the Key Laboratory of Western China's Mineral Resource and Geological Engineering, Ministry of Education, Xi'an 710054, China, and also with the Big Data Center for Geosciences and Satellites, Xi'an 710054, China (e-mail: gaomeiling@chd.edu.cn; zhenhong.li@chd.edu.cn).

Zhenyu Tan is with the College of Urban and Environmental Sciences, Northwest University, Xi'an 710127, China (e-mail: tanzhenyu@nwu.edu.cn).

Huifang Li is with the School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China (e-mail: huifangli@whu.edu.cn).

Jianbing Peng is with the College of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China, and also with the Key Laboratory of Western China's Mineral Resource and Geological Engineering, Ministry of Education, Xi'an 710054, China (e-mail: dicexy_1@chd.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3116258

I. INTRODUCTION

NEAR-SURFACE air temperature (for simplicity hereafter referred to as air temperature) is a key parameter in climate changes, and air temperature with sufficient spatial and temporal resolutions is critical for people to understand energy balance, environment monitoring, agriculture ecosystems, urban heat island, human health, agriculture, etc. [1]–[4]. The air temperature is a critical factor for natural environments due to the effectivity for plant growth and the interaction effects with soil moisture, precipitation, evapotranspiration, transpiration, etc. [5]. For urban human living environment, the air temperature not only can determine the outdoor and indoor thermal comfort but also affect the air conditioner energy consumption [6]. Hence, the air temperature dataset can be used to study the ecology system circulation and help improve the habitability and comfortableness of the living environment [7]. In recent years, the global warming has become a big concern and one fundamental task is to obtain air temperature datasets [8]. Generally speaking, *in situ* observations are the most popular way to obtain air temperature [9]. Although *in situ* air temperature often has a fine temporal resolution, the sparse distribution of the observation stations makes it difficult (if not possible) to capture the detailed spatial characteristics of air temperature [10]. On the other hand, land surface temperature (LST) retrieved by remotely sensed images covers wide regions with detailed spatial information [11], [12]. Furthermore, it has been proved that air temperature is highly correlated with LST [13], [14], which hence provides an opportunity to estimate air temperature with greater spatial details using LST and other auxiliary information.

In the past two decades, a range of methods have been developed to leverage LST to estimate air temperature (e.g., in [15]–[18]), which can be classified into the following categories.

1) The temperature-vegetation index method [18], [19] is based on the hypothesis that the temperature of vegetation canopy is a proxy for the near-surface air temperature.
2) The energy balance method is based on the theory that the net radiation is equal to the sum of the surface heat fluxes [20], [21].
3) The spatial interpolation methods, such as the Kriging interpolation, spline interpolation, and inversion distance weighting, interpolate sparse *in situ* observations to spatially continuous data [22], [23].
4) Classic regression methods. For example, multiple linear regression method was used to determine the relationship between air temperature and MODIS LST with the

assistance of other environment variables [24], [25]; geographically weighted regression is another popular way to estimate air temperature [23], [26].

5) Machine learning methods. The random forest (RF) is much effective in remote sensing classification and regression [27], [28], and successfully used to estimate near-surface air temperature [29], [30]. Deep learning, the new generation of machine learning, is also proved to perform exceptionally to estimate daily air temperature [16].

Among the above-mentioned five types of LST-based methods, the machine learning approach appears to outperform traditional methods in terms of accuracy [16], [25]. It should be noted that some variables are required as auxiliary information when the machine learning methods are employed to derive air temperature from LST. For example, Xu *et al.* [31] utilized 11 environment variables from MODIS data, digital elevation model (DEM), and topographic index data to estimate monthly air temperature in the Tibetan Plateau. Yao *et al.* [32] used enhanced vegetation index (EVI), location information, solar radiation, topographic index, and DEM to retrieve air temperature. Different variables were utilized in different previous studies, and they can be classified into four types: remotely sensed data, data assimilation or reanalysis data, socioeconomic data, and geographical/temporal information. It has been reported that the best performance could be achieved when all the four types of data were involved in the estimation model [16]. However, the more variables, the more data storage, the more computationally intensive the machine learning methods tend to be than the traditional methods. But the requirements of data storage and computation loads cannot be neglected, especially for long-term and spatially detailed air temperature estimation.

Google Earth Engine (GEE) gathered and archived abundant data resources, including reanalysis datasets, remotely sensed products, and population data. Meanwhile, GEE provides large-scale parallel computation capability for scientific researches powered by Google cloud infrastructure [33]–[36]. Therefore, it is worth exploring whether GEE can be used to estimate long-term spatial detailed air temperature products.

In this article, two machine learning methods, namely RF and deep neural network (DNN), are employed to generate 1 km × 1 km air temperature products over Yellow River basin for 20 years during the period from March 2000 to February 2020 using the GEE platform. Auxiliary information is usually required for the air temperature estimation model; the model accuracy often increases with more input variables, but this can cost higher data storage and computation loads, which may be a problem under the circumstances with the resource-limited hardware. Thus, the tradeoff between the number of input variables and the accuracy of the estimation model should be made. Moreover, it should also be examined whether the air temperature estimation model trained over one region during a specific period can be migrated to another region or period. The objectives of this article are as follows:

1) to explore the feasibility to estimate long-term 1 km × 1 km air temperature with machining learning methods based on GEE;
2) to assess the sensitivity of variables with different datasets archived in GEE in the air temperature estimation model;
3) to examine the portability of the air temperature estimation model; and

TABLE I
VARIABLES USED IN THIS STUDY THAT ARE PROVIDED BY GEE

Variables	Spatial resolution	Temporal resolution	Datasets
LST	1 km	8-day	MODIS(MOD11A2)
EVI	1 km	16-day	MODIS(MOD13A2)
ALBEDO	500 m	16-day	MODIS(MCD43A3)
DEM	30 m	/	SRTM
SLOPE	30 m	/	Generated by DEM
LONGWAVE	0.25° degrees	3-hourly	GLDAS-2.1
SHORTWAVE	0.25° degrees	3-hourly	GLDAS-2.1
HUMI	0.25° degrees	3-hourly	GLDAS-2.1
RAIN	0.25° degrees	3-hourly	GLDAS-2.1
WIND	0.25° degrees	3-hourly	GLDAS-2.1
POPULATION	1 km	5-year	GPWv411
LC	500 m	1-year	MODIS(MCD12Q1)
LONG	1 km	/	/
LAT	1 km	/	/
YEAR	/	/	/
MONTH	/	/	/

4) to investigate the spatial and temporal characteristics of air temperature in the study region during the past 20 years.

The rest of this article is organized as follows. The study region, datasets, and experiment design are presented in Section II. The results and discussion are detailed in Section III. The last section summarizes the conclusions and future work.

II. MATERIALS AND METHODS

A. Study Region

Yellow River, the mother river of China, is the second longest river in China. Yellow River basin (E96°–119°, N32°–42°) is a center of Chinese politics, economy, and culture, which crosses nine provinces (Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Henan, and Shandong) with the east–west extent of about 1900 km and the north–south extent of about 1100 km. The total area of the Yellow River basin is about $795\,000\ \text{km}^2$.

The elevation varies from -8 m in the east to 6241 m in the west over the Yellow River basin. The climatic conditions also show significant spatial and temporal variability. For example, the solar radiation changes profoundly between seasons; there are also big differences in precipitation in different years. The observed air temperature is higher in the south than in the north and it is higher in the eastern region than in the western region. More than one million people live in the Yellow River basin, where the air temperature is a significant parameter for the life and production of the people in this region.

B. Data Sources

Two kinds of datasets were used in this study: one is the data archived in the GEE platform, and the other is the data collected from *in situ* stations. The data used in GEE are shown in Table I, which can be further subclassified as four types.

1) The climatological variables: downward longwave radiation (LONGWAVE), downward shortwave radiation (SHORTWAVE), humidity (HUMI), rain speed (RAIN), and wind speed (WIND) produced by global land data assimilation system (GLDAS, [37]) are involved in our experiments. All of these variables have a spatial resolution of 0.25° and a temporal resolution is 3-hourly.
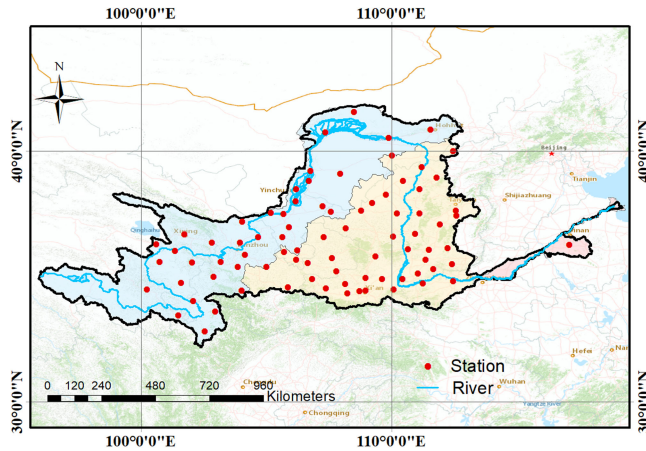
Fig. 1. Map of Yellow River basin. Note that (a) red points represent *in situ* stations and the blue line indicates the Yellow River. (b) Three different colors from left to right represent the upper, middle, and lower reaches.

Fig. 2. Air temperature estimation using GEE with machine-learning methods.

2) The surface-related variables: LST, land cover (LC), ALBEDO, EVI, SRTM DEM, and SLOPE are used in the air temperature estimation model. The LST, LC, ALB-DEO, and EVI are obtained from MODIS products version 6, i.e., MOD11A2 for LST, MOD13A2 for EVI, MCD43A3 for albedo, and MCD12Q1 for LC. The spatial and temporal resolutions of the surface-related variables are listed in Table I.

3) The human-related variables: the Gridded Population of World Version 4 (GPWv4) provides the 1-km grid population every five years, i.e., 2000, 2005, 2010, 2015, and 2020. In this study, one population data is considered to be the actual population in its preceding five-year period. For example, the population data of 2005 in GPWv4 is used as the population during the period of 2001–2005.

4) Other variables: four variables including longitude (LONG), latitude (LAT), year, and month are selected in this study.

The second dataset is the air temperature collected from *in situ* stations. 81 *in situ* stations located in the Yellow River basin are used in this study, as shown in Fig. 1. The temporal resolution of the measured air temperature is daily. The observed 2-m daily mean temperature is checked through the observation quality control (QC) label provided by the metadata and only observations with QC of "0" are used in the study. The observed air temperature can be acquired from the website.[1] Because the daily mean air temperature in the dataset is acquired in 2-m level, the air temperature in the following is also specified to the 2-m high level.

Before estimating the air temperature, all the variables were resampled to a monthly temporal resolution. To be more specific, the monthly mean value was calculated when the original temporal resolution is finer than a month. To the contrary, when the original temporal resolution is larger than a month, the variables kept holding the same values in this period, i.e., LC kept the same in every month within in a year. The scale of the output was set to 1 km in GEE. The data used in this study cover from March 2000 to February 2020.
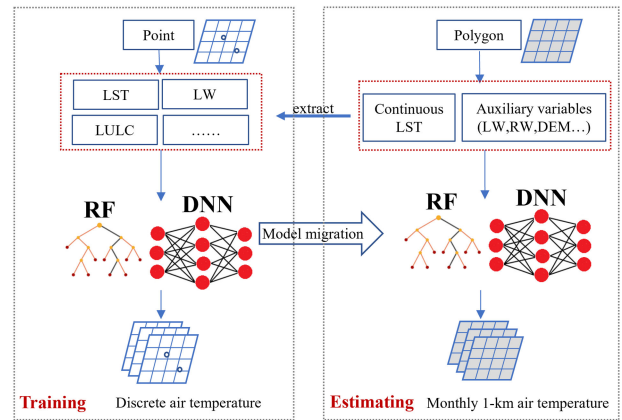
[1][Online]. Available: http://data.cma.cn/

C. Air Temperature Estimation Method

The GEE and machine learning methods were employed to estimate the long-term 1 km × 1 km air temperature product, as shown in Fig. 2. There are following two phases in the air temperature estimation.

1) *Training phase*. For each *in situ* station, the point values were extracted from the archived variable-related dataset in GEE at the given time. The variables in Table I extracted from GEE, and the air temperature observed in each station composes the training dataset. After filtering the dataset, there are 18 900 data records in the training dataset. Then, the two machining learning models (i.e., the RF and DNN) were used for air temperature estimation. The two models were trained and ten-cross validation was used to evaluate the training results.

2) *Estimating phase*. The trained model on station points was migrated to the spatially continuous raster dataset, and the model was run on GEE to produce the long-term 1 km × 1 km air temperature product. Overall, the training phase was performed at the discrete point level, and then the model was migrated to the continuous surface level on GEE.

The RF is a classical machine learning method that fits a number of decision trees and uses the ensemble models to improve the predictive accuracy [38], [39]. RF has been proved to be an outstanding method for remotely sensed image classification and regression. Essentially, the RF is an ensemble of multiple decision trees. The learnable parameters of each decision tree in the forest are usually different from each other, thus, forming different tree models. The core idea of the RF is that each tree might be likely to perform relatively good, but may over fitted for some data. If there are many trees, then the ensemble of these trees may reduce the overfitting. The DNN is a simple deep machine learning model [40] and can be easily implemented on GEE. The DNN can be regarded as the further development of artificial neutral network. A simple DNN is usually constructed with several hidden layers. Each hidden layer consists of one linear transformation and a nonlinear activation function. The commonly used nonlinear activation is the ReLU function. By stacking several hidden layers, the DNN model can learn complex relationship between input(s) and output(s). Generally,

TABLE II
DESIGN OF THE EXPERIMENTS TO ASSESS THE EFFECTIVENESS OF THE
INPUT VARIABLES

Cases	Input variables
Effe1	LST, LONGWAVE, HUMI
Effe2	LST, LONGWAVE, HUMI, SHORTWAVE
Effe3	LST, LONGWAVE, HUMI, SHORTWAVE, EVI
Effe4	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN
Effe5	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM
Effe6	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ABLEDO
Effe7	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG
Effe8	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND
Effe9	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH
Effe10	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH, LC
Effe11	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH, LC, SLOPE
Effe12	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH, LC, SLOPE, POPULATION
Effe13	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH, LC, SLOPE, POPULATION, YEAR
Effe14	LST, LONGWAVE, HUMI, SHORTWAVE, EVI, RAIN, DEM, ALBEDO, LONG, WIND, MONTH, LC, SLOPE, POPULATION, YEAR, LAT

TABLE III
DESIGN OF THE EXPERIMENT TO TEST THE PORTABILITY OF THE MODELS

Cases	Training dataset description
EXP_before	Only the data from March 2000 to February 2010 are selected
EXP_after	Only the data from March 2010 to February 2020 are selected
EXP_mon1	The data in March-August of each year are selected
EXP_mon2	The data in September-February of each year are selected
EXP_west	The dataset is separated by the median longitude of the region, and the western half is selected as the training dataset
EXP_east	The dataset is separated by the median longitude of the region, and the eastern half is selected as the training dataset
EXP_south	The dataset is separated by the median latitude of the region, and the southern half is selected as the training dataset
EXP_north	The dataset is separated by the median latitude of the region, and the northern half is selected as the training dataset
EXP_all	The original dataset is regarded as the training dataset

a deeper model usually can boost the prediction accuracy in some extent, but it may cause the overfitting problem. In this article, the RF and DNN models were trained to establish the complex relationships between the various variables and the air temperature at the discrete point level, then the learned rules were employed in the surface level to generate the 1 km × 1 km air temperature product.

D. Experiment Design

The experiments designed in this study can be separated into two groups. First, we aim to assess the sensitivity of related variables. The only difference among these cases in this group is the number of input variables. To be specific, first of all, we gathered the statistics of the relationship between each variable and the air temperature. The variables were ranked according to their correlation coefficients with the air temperature. The first case comprised of the 3 variables with the greatest absolute correlation coefficient (Effe1 in Table II), the second case included the 4 variables with the greatest absolute correlation coefficient (Effe2 in Table II), the rest can be conducted in the same manner, and the last case (Effe14) involved 16 variables as the inputs. All the quantitative results were validated with the ten-cross test.

The second group was designed to examine whether the model trained within a specific region over a specific time period can be migrated to another region or period. The following cases shown as in Table III were designed. To test the portability of the estimation model over different time periods, the data in the first and last six months every year were selected as the training dataset for cases EXP_mon1 and EXP_mon2, respectively. The first and last ten-year datasets were regarded as the training dataset in cases EXP_before and EXP_after, respectively. To explore whether the estimation model can be migrated spatially, the western, eastern, southern, and northern

parts of the whole datasets were regarded as the training dataset for cases of EXP_west, EXP_east, EXP_south, and EXP_north, respectively. Only half of the dataset was used in training phases, and all the data were used to evaluate the portability of the estimation model.

E. Model Validation Metrics

The following four metrics were used to evaluate the model performance.
1) R is the Pearson correlation coefficient; the greater the value, the higher correlation between the two variables.
2) R^2 is the coefficient of determination, often used in regression models to evaluate the agreement between predicted and actual values. If the value of R^2 is close to 1, this means that a good agreement is obtained.
3) The root-mean-square error (RMSE) is the standard deviation of the residuals (i.e., the difference between the predicted and actual values). It is a measure of how spread out these residuals are; the smaller the RMSE, the better the model.
4) Mean absolute error (MAE) is a measure of the mean absolute error between the predicted and actual values. The lower the MAE, the better the model.

III. RESULTS

A. Effectiveness of Related Variables

Because the correlation coefficients for some variables are positive, while others are negative, the absolute R is used here, as shown in Fig. 3. The first column shows the relationship between each variable and air temperature. The LST, longwave radiation, relative humidity, and shortwave radiation have the highest correlation coefficients with air temperature, and their absolute R values are greater than 0.7, followed by the EVI and rain speed with correlation coefficients of 0.63 and 0.61, respectively. The absolute R values for other variables of air temperature are lower than 0.5. Besides, it should be noted that there are internal correlations between these input variables. For example, the longwave radiation and LST are strongly correlated
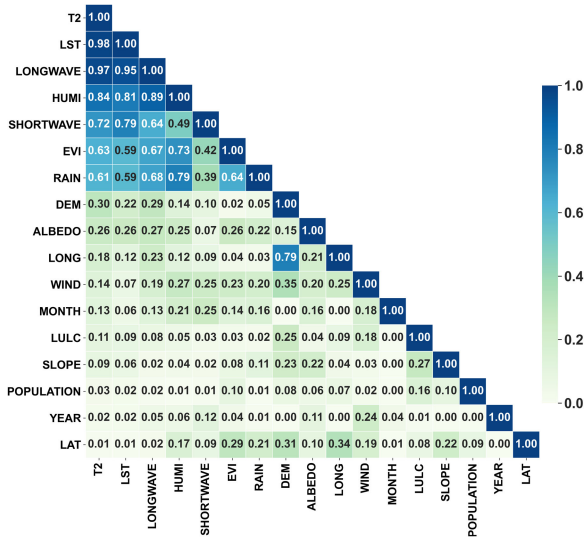
Fig. 3. Absolute Pearson correlation coefficients (R) between variables.
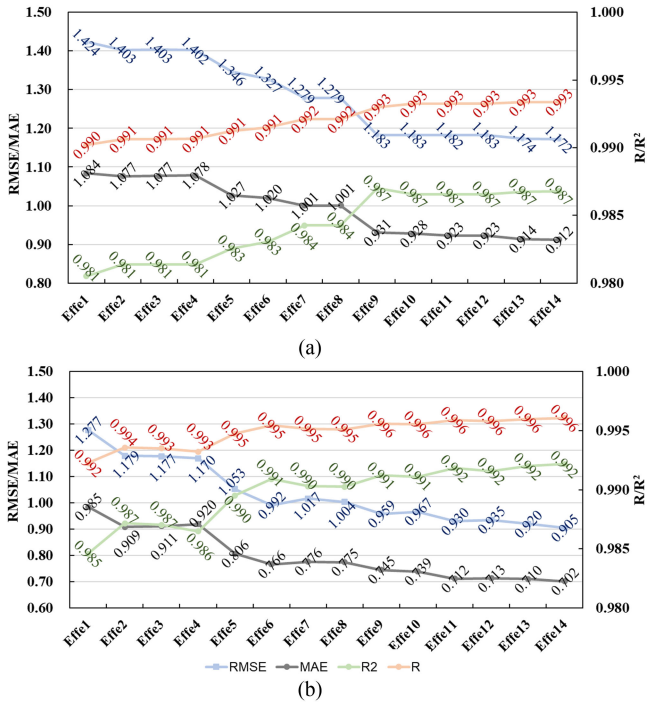

Fig. 4. Performance of the estimation model with different input variables: (a) DNN and (b) RF (Unit: K for RMSE and MAE).

with their absolute R is 0.95. The absolute R between each two of LST, longwave radiation, humidity, shortwave, rain speed, and EVI are larger than 0.39. In other words, the variables those have the higher correlation coefficients with air temperature seem also higher related with each other. For the variables with the absolute R less than 0.5 (compared with air temperature), the absolute R between each two of them are also lower, except that the longitude and DEM is highly correlated with the absolute R is 0.79.

Fig. 4 shows the results of the effectiveness of the input variables. As mentioned earlier, case Effe1 consists of the three

variables with the highest absolute R. The RMSE, MAE, R^2, and R are 1.424, 1.084, 0.981, and 0.990, respectively, for the DNN model [see Fig. 4(a)] and 1.277, 0.985, 0.985 and 0.992, respectively, for the RF model [see Fig. 4(b)]. In case Effe2, the RMSE and MAE obviously decrease while R and R^2 increase for the RF model. But the performance of the DNN is slightly improved. In case Effe5, more variables are included, the performance of the DNN model improves with lower RMSE and MAE values and higher R and R^2 values. For both the DNN and the RF model, the R^2 and R tend to be stable with the increase in the number of input variables, but the RMSE and MAE continue to decrease with more input variables. Both the DNN and RF models gain the best performance while all the relevant variables selected in this study are included.

B. Portability of the Model

In the training phase, with less training data, the training accuracy in both the DNN and RF models does not change much when half of the data points (the cases except EXP_all in Tables IV and V) or all the data points are used to train the model (EXP_all). In the estimating phase, the prediction accuracy of the cases with less training data is not as good as the EXP_all case. The first four cases were designed to test the temporal portability, and the results in Tables IV and V show that the RMSE and MAE in the EXP_before, EXP_after, EXP_month1, and EXP_month2 are greater than those in the EXP_all in the estimating phase. R^2 and R in the EXP_before, EXP_after, EXP_month1, and EXP_month2 are lower than those in the EXP_all. The EXP_west, EXP_east, EXP_south, and EXP_north cases were designed for the spatial portability, and similar results with temporal portability can be observed in Tables IV and V.

It is obvious that the portability in year scale (EXP_before and EXP_after) is slightly better than the portability of other cases. For both the DNN and the RF models, the accuracy of the EXP_month1 is worse than the others. The portability accuracy decreases compared with all the datasets involved in training, but the accuracy is still acceptable with the RMSE less than 1.2 K and MAE less than 0.8 K. R^2 and R are greater than 0.987 in the RF model except for portability in month scale. For DNN model, except for the portability in year scale, the accuracy of other cases seems suffer nonignorable negative effect with less training data.

In summary, the more data involved in the training phase, the better accuracy would be obtained in the estimating phase. However, if there are insufficient data points, the model trained within a specific region or during a specific time period can be migrated to another region or period.

C. Optimized Configurations of the Model

The hyperparameters of the "RF_Plain" in Table VI are set as the above RF. To test the effectiveness of hyperparameter to the model, grid search is employed to find the optimized hyper-parameters for the RF. The "RF_Optimized" in Table VI is the optimized version using grid search. It can be found that the performance of the optimized RF is better than defaults settings regarding both the RMSE and R^2. For the above DNN, there are 4 hidden layers and the neurons in each hidden layer are

TABLE IV
ACCURACY OF THE MODEL WITH DIFFERENT TRAINING DATASETS (DNN)

Experiment	Training accuracy				Estimation accuracy			
	RMSE	MAE	R^2	R	RMSE	MAE	R^2	R
EXP_before	1.130	0.892	0.988	0.994	1.177	0.925	0.987	0.994
EXP_after	1.109	0.873	0.989	0.994	1.186	0.934	0.987	0.993
EXP_month1	1.009	0.794	0.981	0.991	2.031	1.466	0.961	0.982
EXP_month2	1.128	0.886	0.982	0.991	1.470	1.169	0.980	0.993
EXP_west	1.112	0.879	0.987	0.994	1.582	1.231	0.976	0.990
EXP_east	1.043	0.827	0.990	0.995	1.568	1.138	0.977	0.990
EXP_south	1.147	0.914	0.986	0.993	1.987	1.207	0.963	0.982
EXP_north	1.048	0.827	0.990	0.995	1.299	0.967	0.984	0.992
EXP_all	1.172	0.912	0.987	0.993	1.112	0.877	0.988	0.994

TABLE V
ACCURACY OF THE MODEL WITH DIFFERENT TRAINING DATASETS (RF)

Experiment	Training accuracy				Estimation accuracy			
	RMSE	MAE	R^2	R	RMSE	MAE	R^2	R
EXP_before	0.963	0.739	0.991	0.996	0.821	0.570	0.994	0.997
EXP_after	0.913	0.696	0.992	0.996	0.880	0.605	0.993	0.996
EXP_month1	0.857	0.661	0.987	0.993	2.258	1.256	0.952	0.981
EXP_month2	0.940	0.719	0.987	0.994	1.955	1.243	0.964	0.987
EXP_west	0.926	0.707	0.991	0.996	1.155	0.733	0.987	0.994
EXP_east	0.874	0.678	0.993	0.997	1.105	0.712	0.988	0.994
EXP_south	0.935	0.709	0.991	0.995	1.001	0.686	0.991	0.996
EXP_north	0.852	0.656	0.994	0.997	1.130	0.735	0.988	0.994
EXP_all	0.905	0.696	0.992	0.996	0.429	0.302	0.998	0.999

TABLE VI
COMPARISON OF RF PERFORMANCE BEFORE AND AFTER PARAMETERS
OPTIMIZED

Parameters	RF_Plain	RF_Optimzed
n_estimators	100	133
max_features	None	4
min_samples_leaf	1	1
RMSE	0.905	0.826
R^2	0.992	0.994

TABLE VII
COMPARISON OF DNN PERFORMANCE UNDER DIFFERENT SETTINGS

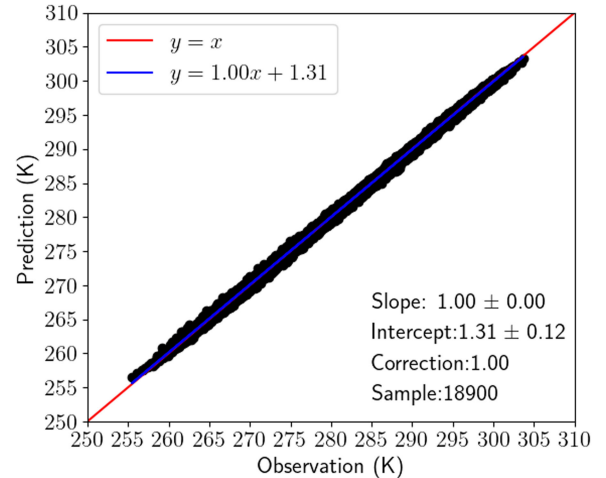| Cases	Layers number	Neurons setup	RMSE	R^2
DNN_case1	3	32,64,1	1.173	0.987
DNN_case2	4	32,64,32,1	1.141	0.988
DNN_case3	4	32.128.64.1	1.131	0.988
DNN_case4	6	32,64,128,64,32,1	1.120	0.988
DNN_case5	8	32,64,128,256,128,64,32,1	1.005	0.990

Fig. 5. Scatter plot of observed and predicted air temperature. Note that the red line represents a perfect fit, whereas the blue line indicates the least squares regression line.

set to 32, 128, 64, and 1, respectively, shown as DNN_case3 in Table VII. The other cases in Table VII are designed by changing the number of hidden layer and neuron nodes in each layer. It can be found that the model can gain a better performance with more neurons by comparing DNN_case2 and DNN_case3. Moreover, from the DNN_case1 and DNN_case2 or DNN_case3 and DNN_case4, it can be concluded that with the same settings of the maximum neurons, more hidden layer also can boost the accuracy. Overall, it can be seen from Fig. 4 that the performance of the RF model is better than the DNN model. The RMSE of the RF model is less than 0.9 and the R^2 is larger than 0.994 with the optimized hyperparameters (see Table VI). The accuracy of DNN is improved when its depth and width are increased (see DNN_case5 in Table VII). However, the training of DNN model is much time-consuming, the RF model can be trained within seconds with the same hardware specifications. Considering the time cost and model accuracy, the RF with the optimized parameters was used in the generation of long-term 1-km monthly air temperature.

D. Accuracy Evaluation for the Estimated Air Temperature

1) Overall Accuracy: Assuming that the relationship between the observed and predicted is linear, i.e., prediction $=$ $a \times$ observation $+ b$. It is obvious in Fig. 5 that the prediction
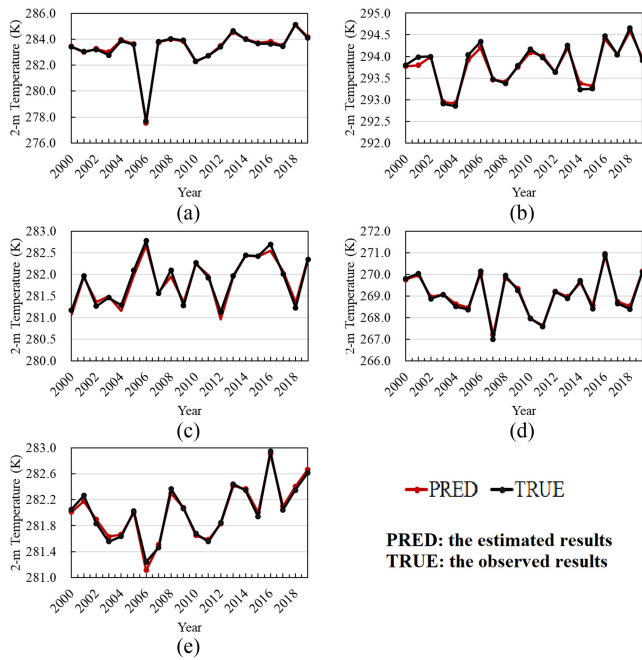
Fig. 6.   Temporal trends of predicted and observed air temperature (RF).

has good consistence with the observed air temperature. The slope of the regression line is 1.00 and the intercept is 1.31, calculated with the 18 900 samples in the experiment. The 18 900 observations are the monthly mean air temperature collected from all the *in situ* stations located in the Yellow River basin observed from March 2000 to February 2020. The predictions are extracted from the air temperature product estimated using the RF method. The standard deviation of the slope and intercept is 0 and 0.12, respectively, with 95% confidence coefficient. Moreover, the correlation coefficient is close to 1, showing a high prediction accuracy.

2) Temporal Accuracy: Fig. 6 presents the seasonal accuracy of the estimated air temperature. Each point represents the mean air temperature in each season over Yellow River basin. In the boreal spring (March–May) and boreal winter (December–February), the predicted and observed results are similar during the period from 2000 to 2020. There is only a quite small bias in the two seasons. The biases of mean air temperature in the boreal summer (June–August) and boreal autumn (September–November) are greater than those in the boreal spring and boreal winter, but the temporal variation trend is still similar between the observed and estimated air temperature. The variation trend of yearly mean air temperature also shows a good consistence between the two datasets. It is clear that the estimation model is able to capture the temporal variation trend.

3) Spatial Accuracy: Fig. 7 shows the spatial distribution of the RMSE, MAE, R^2, and *R*. In most stations, the RMSE is less than 0.4 K, while the MAE is less than 0.3 K. The R^2 and R in most stations are larger than 0.997 and 0.999, respectively. The maximum RMSE and MAE are about 0.5 K and 0.4 K respectively, and the minimum R and R^2 are about 0.999 and 0.996 respectively. Overall, each station's accuracy

stays high regarding all the metrics during the evaluation period from March 2000 to February 2020.

E. Temporal and Spatial Variations of 2-m Temperature Over the Yellow River Basin

1) Temporal Variations of 2-m Temperature: Based on the long-term 1 km × 1 km air temperature product estimated by the RF model, the temporal variations from March 2000 to February 2020 of each season within the upper, middle, and lower reaches are analyzed. The slopes of the linear regression lines in blue in Fig. 8(a)–(c) are 0.0642, 0.043, and 0.0547 respectively, indicating that the air temperature increases significantly in the upper reach in spring, summer, and autumn, but the slope of the winter is less than 0.01, which means that the increase trend is not obvious. For the middle reach, the air temperature increases in each season, but the increase trend is not as high as in the upper reaches. For the lower reach, the air temperature increases similarly in each season with the slope of each regression line in the range of 0.03–0.04. Noteworthy, the hottest year during 2000–2020 varies in different seasons. In spring, the hottest year is 2018 and the coldest year is 2010. The hottest year in summer is 2010 for the upper reaches, but 2018 for middle and lower reaches, meanwhile the coldest year is within 2003–2004. In autumn, the air temperature of 2012 is lower than other years, the hot year is 2015 in the upper reach, but it is 2006 in the other two reaches. In winter, air temperature of 2016 in the upper and middle reaches is the hottest, whereas the air temperature of 2011 in each reach is the lowest. Overall, the extreme air temperature exists in different years for specific region and season, but the hottest year is usually within 2015–2019 and the coldest year is usually within 2003–2004 or 2010–2012.

2) Spatial Distribution of 2-m Temperature: Fig. 9 shows the spatial distribution of monthly mean 2-m temperature in 2016. It can be found that the temperature is hottest in July and coldest in January. The air temperature increases from January to July, then decreases from July to January. From the perspective of spatial distribution, the temperature is lower in the western region than in the east of the Yellow River basin region. Note that the lower reach is always the warmest in each month, the upper reach is the coldest, and the air temperature increases from north to south in the midstream reach.

IV. DISCUSSION

A. Comparison of Air Temperature and LST

Figs. 10 and 11 illustrate the comparisons between the LST and air temperature. The 20-year (March, 2000 to February, 2020) averaged air temperature and LST for each season are used here. The air temperature comes from the above generated product, and the LST obtained from MODIS LST product. The spatial resolution of both are 1 km. High similarity can be observed between these two temperature products in the spatial distribution (see Fig. 10), and the correlation coefficients are 0.96, 0.96, 0.97, and 0.92 for the four seasons, respectively (see Fig. 11). The differences between the LST and air temperature in spring and summer are larger than those in autumn and winter (see Fig. 10), which can also be observed in Fig. 11 with the points in winter and autumn closer to the gray line than in spring
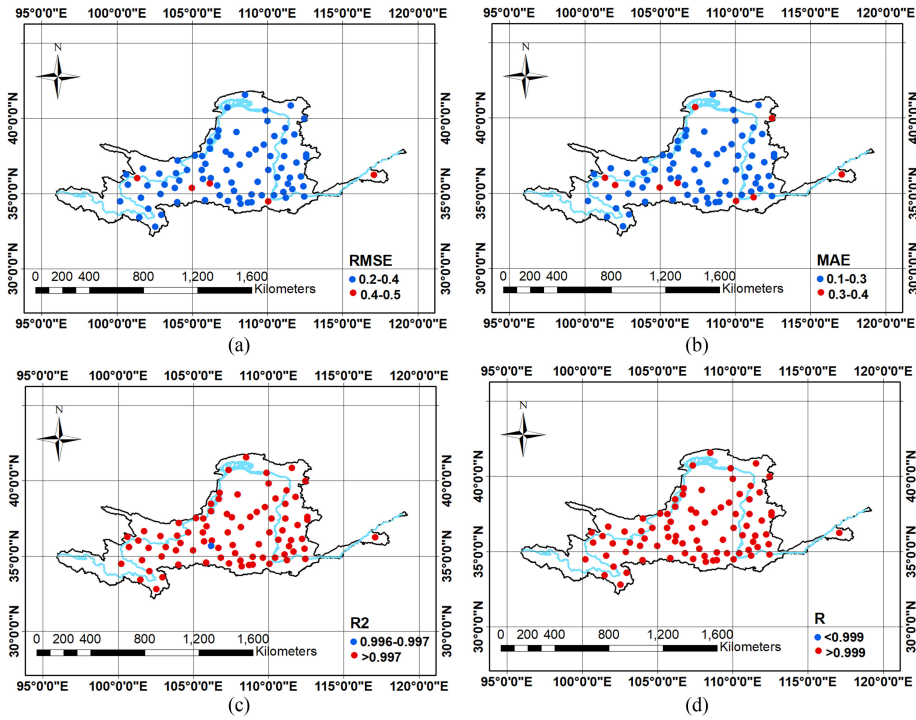
Fig. 7.   Spatial distribution of (a) RMSE, (b) MAE, (c) R^2, and (d) R over each observation station (Unit: K for RMSE and MAE).
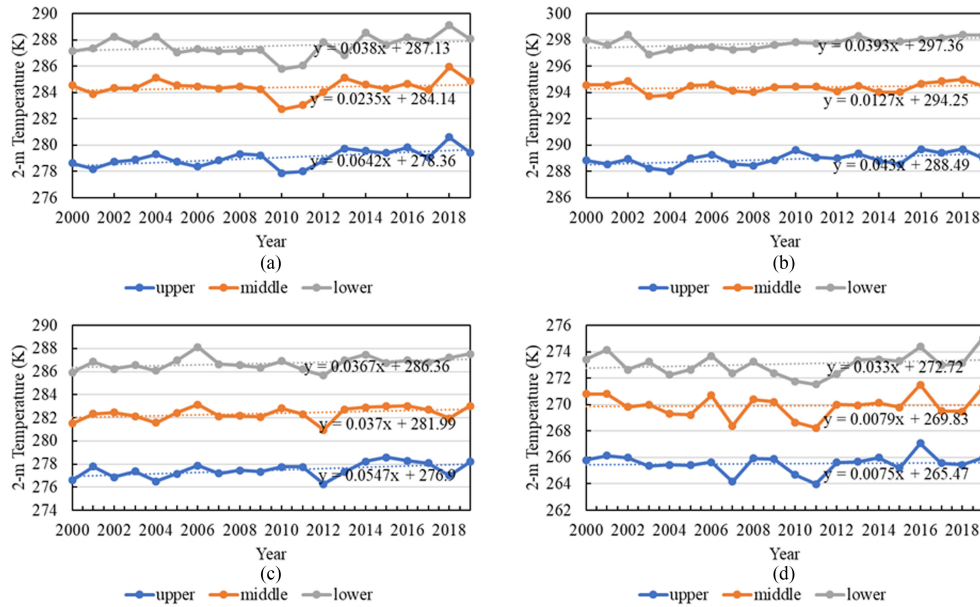

Fig. 8.   Temporal variations of mean 2-m temperature over the Yellow River basin (a: Spring; b: Summer; c: Autumn; d: Winter).

and summer. Most points located in the upper left of the gray line in Fig. 11, indicating that the LST is higher than the air temperature. Besides, the variation range of air temperature is smaller than the LST, for example, the air temperature varies in 260–280 K in winter, whereas the LST varies in 250–280 K. In the 1-km scale, the LST shows somewhat of heterogeneity than the air temperature, which may be due to the fact that LST is more sensitive to land cover [41]. Besides, the location, land cover, vegetation fraction, and elevation are important influence factors that can cause the differences between LST and air temperature [42]. The air temperature and LST are highly correlated, and their differences and driving factors need to be further examined. In our experiment, there are some missing values in the 1 km × 1 km air temperature product (observed in Fig. 10). The reason is that the air temperature is estimated by overall 16 variables, so the missing or low quality of any one variable in a grid can lead to a blank value in the air temperature product. This could be fixed with the spatial deficiency reconstruction method in the future.
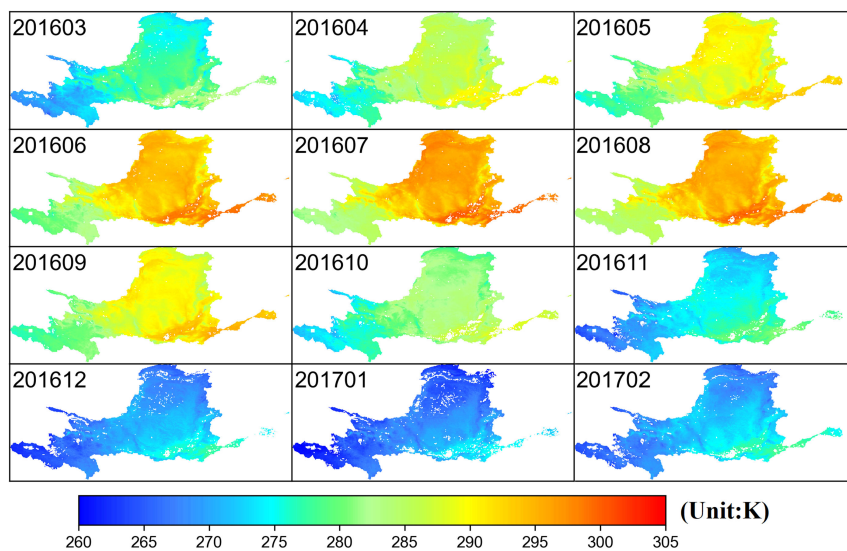
Fig. 9.   Spatial distribution of monthly 2-m temperature in 2016 over Yellow River basin.
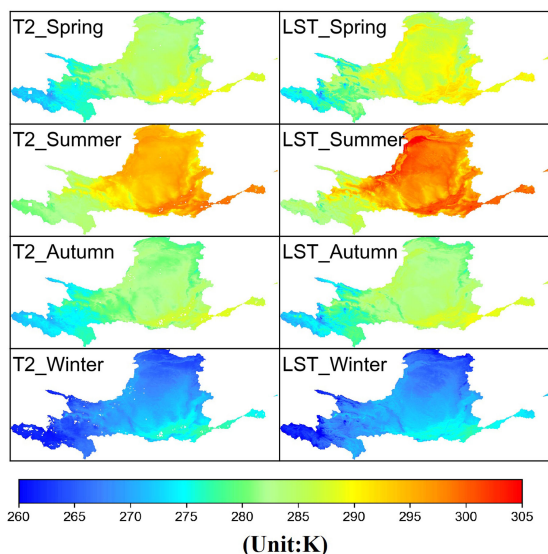

Fig. 10.   Spatial distribution of seasonal air temperature and LST.

B. Comparison of Different Air Temperature Products

The following three and our air temperature products were compared in this study.

1) The monthly air temperature product produced by Peng *et al.* [43], hereafter for brief, called PENG's product. The dataset was spatially downscaled from CRU TS v4.02 with WorldClim datasets based on Delta downscaling method [44] and can be obtained from[2]

2) The monthly air temperature provided in [45], hereafter for brief, called JING's product. This dataset was downscaled from the monthly NCEP/NCAR mean air temperature

product with the supplement of DEM, NDVI, and geographical location by the regression tree model [46]. The dataset can be obtained from[3]

3) The monthly air temperature product of ECMWF ERA5 climate reanalysis dataset from GEE was also involved in the comparisons.

The air temperature products of PENG and JING have the same spatial resolution (1-km) with ours. For convenience, the ECMWF product was resampled to 1-km from about 0.1° on the GEE for comparison. Fig. 12 shows that the ECMWF is significantly coarser than others due to its original coarse resolution. The four products have similar spatial patterns, but are different in details. To evaluate the accuracy of the four products, the observed air temperature from July 2009 to June 2010 were used to calculate the quantitative metrics. When comparing the products with the station data, the value of the grid in which the station located was regarded as the ground truth of the product. The four metrics: RMSE, MAE, *R*, and ME (mean residuals of all the station) are shown in Fig. 13. Fig. 13 suggests that the air temperature generated by this study shows the best performance in terms of all of the metrics. Besides, the maximum residuals of PENG, JING, ECMWF, and our air temperature are 6.96 K, 5.36 K, 7.91 K, and 2.63 K, respectively; the minimum residuals are -2.8 K, -598 K, -10.07 K, and -1.91 K, respectively. Therefore, all the metrics and min–max residuals indicate that the air temperature produced by this study has the highest accuracy. The four products were generated using different methods. The first two are the results of downscaling from reanalysis data, whereas the third is the reanalysis product. The air temperature in this study is a synthetic product with multiple datasets including model simulation, remotely sensed data, population data, *in situ* observed data, etc. There are more variables involved in this study due to the abundant resources on GEE, which may lead to the highest accuracy. Except for the accuracy of our product, another advantage is that our method does not need to download abundant input datasets and then

[2][Online]. Available: https://data.tpdc.ac.cn/

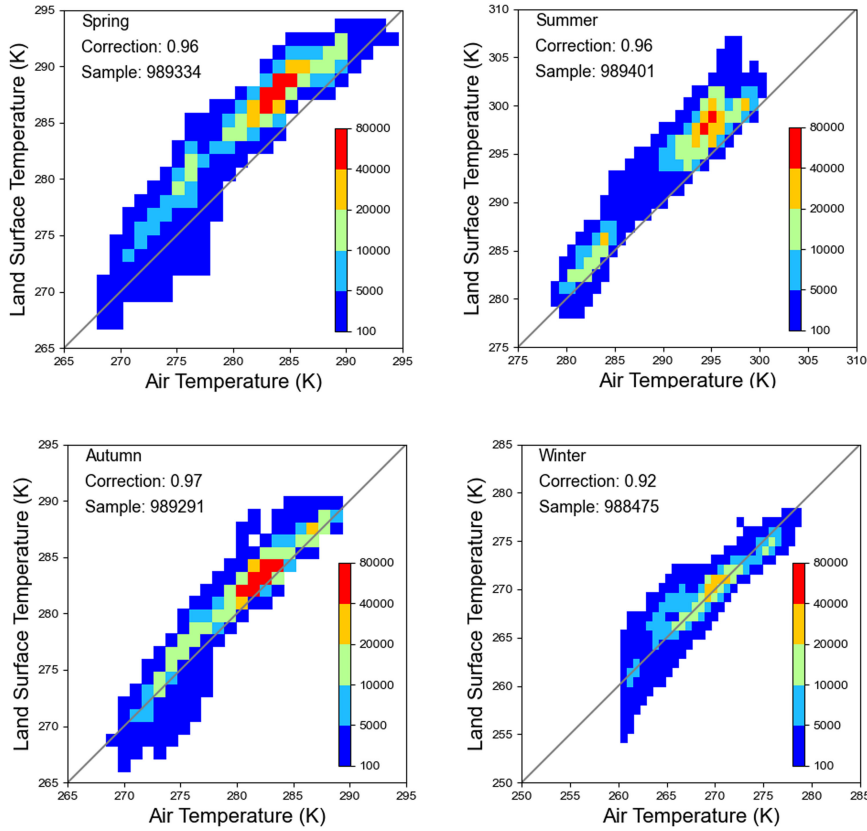[3][Online]. Available: http://geodoi.ac.cn/.

Fig. 11. Comparison between air temperature and LST for each season. Note that the color of the "air temperature"—"LST" points indicate the point density, and the color bar shows the number of points.
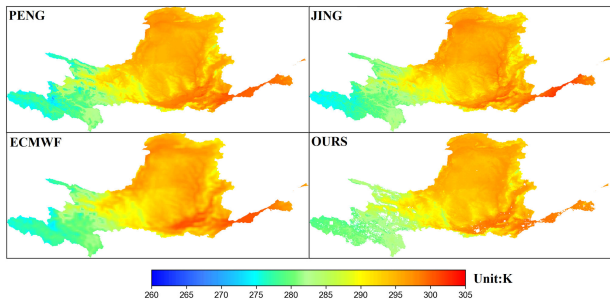

Fig. 12. Spatial distribution of four different air temperature products in June 2010. Note that the ECMWF data are resampled from 0.1° to 1-km here, and the spatial resolution of other three data is also 1-km consistent with their original resolution.
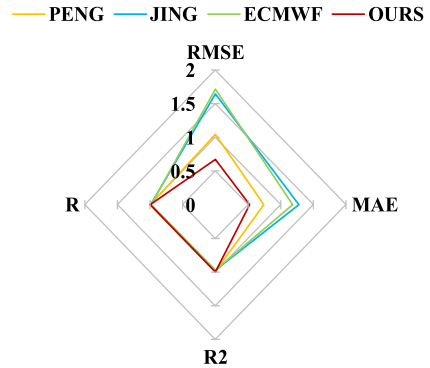

Fig. 13. Quantitative evaluation of the four air temperature products from July 2009 to June 2010.

process them offline, the cloud-based GEE platform takes care of the data storage and computation loads.

V. CONCLUSION

In this article, it is demonstrated that GEE can be utilized to generate long-term 1 km × 1 km air temperature product with machine learning methods. The 20-year monthly mean 1 km × 1 km air temperature product over Yellow River basin is open for public and can be accessed from the cloud disk.[4] The

[4][Online]. Available: https://pan.baidu.com/s/19gdSFx0JSNwF6ZsaDtc4nw

main findings of this study are: (i) The performance of the DNN model and the RF model is related to their hyper-parameters. Considering the time cost and model accuracy, the RF model is better than the DNN model for air temperature estimation in this study; (ii) The air temperature estimation models are shown the best performance under the condition that all the relevant variables selected in this study are involved. Both the RF and DNN can be migrated with a decreased accuracy; (iii) The mean air temperature increases in each reach but the increase speed is different for the specific season and region. The air temperature in the south is higher than in the north, while the eastern region

is hotter than the western region; and (iv) Compared with three other air temperature products, the air temperature generated in this study shows the best performance.

The monthly mean 1-km air temperature are generated with high accuracy from March 2000 to February 2020 in this study, which demonstrates the feasibility to produce the air temperature using GEE. The longer the time span, the more beneficial for research on climate and ecological environment issues; therefore, we will expand the time span by combing multiple remotely-sensed data in the future. For example, Landsat data covers the time range from 1972 to present, and we will utilize the long-time series data to produce air temperature with a longer time span.

Considering the rapid change characteristic of air temperature, the finer temporal resolution is considerably needed. Currently, most input variables from remotely-sensed data used in this study is 8-day or 16-day, but there are also MODIS products with finer temporal resolution (e.g., daily); The products from GLDAS even have a 3-hourly resolution. Hence, it is possible to estimate air temperature with finer temporal resolution, e.g., daily air temperature. It should be noted that the missing data is a big challenge in daily MODIS products. For example, the statistics results show that there are only 5 MYD11A1 (daily LST observed from Aqua) images with clouds less than 10% in the summer of 2013 over Xi'an city. Overall, the MODIS LST can be used to produce daily air temperature, and the GEE platform makes it not that hard, but the missing data keeps to be an unsolved problem. One possible solution is to fix the missing data in daily LST products, then use them in the air temperature estimation, which is one focus of our future study.

REFERENCES

[1] Y. Fan and H. van den Dool, "A global monthly land surface air temperature analysis for 1948–present," *J. Geophys. Res., Atmos.,* vol. 113, no. D1, pp. 1–11, 2008. [Online]. Available: https://doi.org/10.1029/2007JD008470

[2] A. Zanobetti and J. Schwartz, "Temperature and mortality in nine U.S. cities," *Epidemiology,* vol. 19, no. 4, pp. 563–570, 2008, doi: 10.1097/EDE.0b013e31816d652d.

[3] A. Rosenfeld, M. Dorman, J. Schwartz, V. Novack, A. C. Just, and I. Kloog, "Estimating daily minimum, maximum, and mean near surface air temperature using hybrid satellite models across Israel," *Environ. Res.,* vol. 159, pp. 297–312, 2017. [Online]. Available: https://doi.org/10.1016/j.envres.2017.08.017

[4] L. Sheng, X. Tang, H. You, Q. Gu, and H. Hu, "Comparison of the urban heat island intensity quantified by using air temperature and Landsat land surface temperature in Hangzhou, China," *Ecol. Indicators,* vol. 72, pp. 738–746, 2017. [Online]. Available: https://doi.org/10.1016/j.ecolind.2016.09.009

[5] E. Khesali and M. Mobasheri, "A method in near-surface estimation of air temperature (NEAT) in times following the satellite passing time using MODIS images," *Adv. Space Res.,* vol. 65, no. 10, pp. 2339–2347, 2020, doi: 10.1016/j.asr.2020.02.006.

[6] R. S. dos Santos, "Estimating spatio-temporal air temperature in London (U.K.) using machine learning and Earth observation satellite data," *Int. J. Appl. Earth Obs. Geoinf.,* vol. 88, 2020, Art. no. 102066. [Online]. Available: https://doi.org/10.1016/j.jag.2020.102066

[7] J. Peng *et al.*, "Scientific research framework of liveable Yellow River," *J. Eng. Geol,* vol. 38, no. 2, pp. 189–201, 2020.

[8] J. E. Box *et al.*, "Key indicators of Arctic climate change: 1971–2017," *Environ. Res. Lett.,* vol. 14, no. 4, 2019, Art. no. 045010, doi: 10.1088/1748-9326/aafc1b.

[9] C. Bravo *et al.*, "Air temperature characteristics, distribution, and impact on modeled ablation for the South Patagonia Icefield," *J. Geophys. Res., Atmos.,* vol. 124, no. 2, pp. 907–925, 2019, doi: 10.1029/2018jd028857.

[10] S.-B. Duan *et al.*, "Validation of collection 6 MODIS land surface temperature product using in situ measurements," *Remote Sens. Environ.,* vol. 225, pp. 16–29, 2019. [Online]. Available: https://doi.org/10.1016/j.rse.2019.02.020

[11] Z. Wan and Z. L. Li, "Radiance-based validation of the V5 MODIS land-surface temperature product," *Int. J. Remote Sens.,* vol. 29, no. 17/18, pp. 5373–5395, 2008.

[12] Z. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sens. Environ.,* vol. 140, pp. 36–45, 2014.

[13] D. Mutiibwa, S. Strachan, and T. Albright, "Land surface temperature and surface air temperature in complex terrain," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,* vol. 8, no. 10, pp. 4762–4774, Oct. 2015.

[14] N. C. Pepin, E. E. Maeda, and R. Williams, "Use of remotely sensed land surface temperature as a proxy for air temperatures at high elevations: Findings from a 5000 m elevational transect across Kilimanjaro," *J. Geophys. Res., Atmos.,* vol. 121, no. 17, pp. 9998–10015, 2016, doi: 10.1002/2016jd025497.

[15] K. Stahl, R. Moore, J. Floyer, M. Asplin, and I. McKendry, "Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density," *Agric. Forest Meteorol.,* vol. 139, no. 3/4, pp. 224–236, 2006.

[16] H. Shen, Y. Jiang, T. Li, Q. Cheng, C. Zeng, and L. Zhang, "Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data," *Remote Sens. Environ.,* vol. 240, 2020, Art. no. 111692, doi: 10.1016/j.rse.2020.111692.

[17] M. Cobaner, H. Citakoglu, O. Kisi, and T. Haktanir, "Estimation of mean monthly air temperatures in Turkey," *Comput. Electron. Agric.,* vol. 109, pp. 71–79, 2014, doi: 10.1016/j.compag.2014.09.007.

[18] W. Zhu, A. Lǘ, and S. Jia, "Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products," *Remote Sens. Environ.,* vol. 130, pp. 62–73, 2013. [Online]. Available: https://doi.org/10.1016/j.rse.2012.10.034

[19] H. Nieto, I. Sandholt, I. Aguado, E. Chuvieco, and S. Stisen, "Air temperature estimation with MSG-SEVIRI data: Calibration and validation of the TVX algorithm for the Iberian Peninsula," *Remote Sens. Environ.,* vol. 115, no. 1, pp. 107–116, 2011. [Online]. Available: https://doi.org/10.1016/j.rse.2010.08.010

[20] P. Hou, Y. Chen, W. Qiao, G. Cao, W. Jiang, and J. Li, "Near-surface air temperature retrieval from satellite images and influence by wetlands in urban region," *Theor. Appl. Climatol.,* vol. 111, no. 1, pp. 109–118, 2013, doi: 10.1007/s00704-012-0629-7.

[21] R. Zhang, Y. Rong, J. Tian, H. Su, Z.-L. Li, and S. Liu, "A remote sensing method for estimating surface air temperature and surface vapor pressure on a regional scale," *Remote Sens.,* vol. 7, no. 5, pp. 6005–6025, 2015, doi: 10.3390/rs70506005.

[22] K. Stahl, R. D. Moore, J. A. Floyer, M. G. Asplin, and I. G. McKendry, "Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density," *Agric. Forest Meteorol.,* vol. 139, no. 3, pp. 224–236, 2006. [Online]. Available: https://doi.org/10.1016/j.agrformet.2006.07.004

[23] M. Wang *et al.*, "Comparison of spatial interpolation and regression analysis models for an estimation of monthly near surface air temperature in China," *Remote Sens.,* vol. 9, no. 12, pp. 1–16, 2017, doi: 10.3390/rs9121278.

[24] Y. Yang, W. Cai, and J. Yang, "Evaluation of MODIS land surface temperature data to estimate near-surface air temperature in Northeast China," *Remote Sens.,* vol. 9, no. 5, pp. 1–19, 2017, doi: 10.3390/rs9050410.

[25] Y. Xu, A. Knudby, and H. C. Ho, "Estimating daily maximum air temperature from MODIS in British Columbia, Canada," *Int. J. Remote Sens.,* vol. 35, no. 24, pp. 8108–8121, 2014, doi: 10.1080/01431161.2014.978957.

[26] F. Chen, Y. Liu, Q. Liu, and F. Qin, "A statistical method based on remote sensing for the estimation of air temperature in China," *Int. J. Climatol.,* vol. 35, no. 8, pp. 2131–2143, 2015.

[27] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.,* vol. 114, pp. 24–31, 2016. [Online]. Available: https://doi.org/10.1016/j.isprsjprs.2016.01.011

[28] H. Ebrahimy, H. Aghighi, M. Azadbakht, M. Amani, S. Mahdavi, and A. A. Matkan, "Downscaling MODIS land surface temperature product using an adaptive random forest regression method and Google Earth Engine for a 19-years spatiotemporal trend analysis over Iran," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,* vol. 14, pp. 2103–2112, 2021.

[29] T. Appelhans, E. Mwangomo, D. R. Hardy, A. Hemp, and T. Nauss, "Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania," *Spatial Statist.,* vol. 14, pp. 91–113, 2015, doi: 10.1016/j.spasta.2015.05.008.

[30] C. Yoo, J. Im, S. Park, and L. J. Quackenbush, "Estimation of daily maximum and minimum air temperatures in urban landscapes using MODIS time series satellite data," *ISPRS J. Photogramm. Remote Sens.,* vol. 137, pp. 149–162, 2018. [Online]. Available: https://doi.org/10.1016/j.isprsjprs.2018.01.018

[31] Y. Xu, A. Knudby, Y. Shen, and Y. Liu, "Mapping monthly air temperature in the Tibetan Plateau from MODIS data based on machine learning methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,* vol. 11, no. 2, pp. 345–354, Feb. 2018.

[32] R. Yao *et al.*, "Developing a temporally accurate air temperature dataset for Mainland China," *Sci. Total Environ.,* vol. 706, 2020, Art. no. 136037. [Online]. Available: https://doi.org/10.1016/j.scitotenv.2019.136037

[33] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.,* vol. 202, pp. 18–27, 2017, doi: 10.1016/j.rse.2017.06.031.

[34] L. Kumar and O. Mutanga, "Google Earth Engine applications since inception: Usage, trends, and potential," *Remote Sens.,* vol. 10, no. 10, pp. 1–15, 2018, doi: 10.3390/rs10101509.

[35] M. Amani *et al.*, "Google Earth Engine cloud computing platform for remote sensing big data applications: A comprehensive review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,* vol. 13, pp. 5326–5350, 2020.

[36] H. Tamiminia, B. Salehi, M. Mahdianpari, L. Quackenbush, S. Adeli, and B. Brisco, "Google Earth Engine for geo-big data applications: A meta-analysis and systematic review," *ISPRS J. Photogramm. Remote Sens.,* vol. 164, pp. 152–170, 2020, doi: 10.1016/j.isprsjprs.2020.04.001.

[37] M. Rodell *et al.*, "The global land data assimilation system," *Bull. Amer. Meteorol. Soc.,* vol. 85, no. 3, pp. 381–394, 2004.

[38] A. Dey, "Machine learning algorithms: A review," *Int. J. Comput. Sci. Inf. Technol.,* vol. 7, no. 3, pp. 1174–1179, 2016.

[39] S. Ray, "A quick review of machine learning algorithms," in *Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput.*, Feb. 2019, pp. 35–39, doi: 10.1109/COMITCon.2019.8862451.

[40] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[41] T. R. Oke, G. Mills, A. Christen, and J. A. Voogt, *Urban Climates*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[42] E. J. Good, D. J. Ghent, C. E. Bulgin, and J. J. Remedios, "A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series," *J. Geophys. Res., Atmos.,* vol. 122, no. 17, pp. 9185–9210, 2017, doi: 10.1002/2017jd026880.

[43] S. Peng, "1-km monthly mean temperature dataset for China (1901-2017)," National Tibetan Plateau Data Center, 2019, doi: 10.11888/Meteoro.tpdc.270961.

[44] S. Peng, Y. Ding, W. Liu, and Z. Li, "1 km monthly temperature and precipitation dataset for China from 1901 to 2017," *Earth System Sci. Data,* vol. 11, no. 4, pp. 1931–1946, 2019.

[45] W. Jing, Y. Yang, and X. Le, "Historical monthly temperature data of China at 1 km resolution [DB/OL]," Global Change Research Data Publishing & Repository.

[46] W. Jing, Y. Yang, and X. Le, "Monthly temperature dataset of China at 1 km resolution " *J. Glob. Change Data Discov.,* vol. 1, pp. 66–73, 2017.

Zhenhong Li (Senior Member, IEEE) received the B.Eng. degree in geodesy from Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1997, and the Ph.D. degree in GPS, geodesy, and navigation from the University College London, London, U.K., in 2005.

He is currently a Professor of imaging geodesy with the College of Geological Engineering and Geomatics, Chang'an University, Xi'an, China. He has more than 24 years of research experience in geodesy and remote sensing for geohazards (e.g., earthquakes and landslides), infrastructure stability, and environment monitoring. He is also the principal investigator of the Generic Atmospheric Correction Online Service for interferometric synthetic aperture radar.

Dr. Li is a Fellow of the International Association of Geodesy and an Associate Editor for Advances in Space Research and Remote Sensing.

Zhenyu Tan received the B.Sc. degree in geographical information science from Chang'an University, Xi'an, China, in 2013, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2019.

He is currently a Lecturer with the College of Urban and Environmental Sciences, Northwest University, Xi'an, China. His main research interests include remote sensing environment monitoring, data fusion, and deep learning in geosciences.
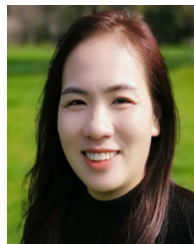
Huifang Li received the B.Sc. degree in geographical information science from China University of Mining and Technology, Xuzhou, China, in 2008, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2013.

She is currently a Professor with the School of Resources and Environmental Science, Wuhan University. Her research interests include radiometric correction of remote sensing images, including cloud correction, shadow correction, and urban thermal environment analysis and alleviation.

Meiling Gao received the B.Sc. degree in geographical information science from Chang'an University, Xi'an, China, in 2014, and the Ph.D. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2019.

She is currently a Lecturer with the School of Geology Engineering and Geomatics, Chang'an University, Xi'an, China. Her main research interests include spatiotemporal data reconstruction and urban thermal environment.

Jianbing Peng received the B.Sc. degree in geology from China University of Geosciences, Wuhan, China, in 1978, and the Ph.D. degree in engineering geology from Chang'an University, Xi'an, China, in 1999.

He is currently a Professor of engineering geology with the College of Geological Engineering and Geomatics, Chang'an University. He has more than 40 years of research experience in engineering geology and geohazards (e.g., landslides, ground fissures, and land subsidence). He is an academician of the Chinese of Academy of Sciences and the President of the Engineering Geology Committee of China.