CLIGEN Parameter Regionalization for Mainland China

Wenting Wang[1,2], Shuiqing Yin[2], Bofu Yu[3], Shaodong Wang[2]

[1] Zhuhai Branch of State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University at Zhuhai, Zhuhai 519087, People's Republic of China

[2] State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, People's Republic of China

[3] Australian Rivers Institute, School of Engineering and Built Environment, Griffith University, Nathan, Qld 4111, Australia

Correspondence to: Shuiqing Yin (yinshuiqing@bnu.edu.cn)

Abstract. Stochastic weather generator CLIGEN can simulate long-term weather sequences as input to WEPP for erosion predictions. Its use, however, has been somewhat restricted by limited observations at high spatial-temporal resolutions. Long-term daily temperature, daily and hourly precipitation data from 2405 stations and daily solar radiation from 130 stations distributed across mainland China were collected to develop the most critical set of site-specific parameter values for CLIGEN. Ordinary Kriging (OK) and Universal Kriging (UK) with auxiliary covariables, i.e. longitude, latitude, elevation, and the mean annual rainfall were~~as~~ used to interpolate parameter values into a 10 km × 10 km grid and ~~parameter~~the interpolation accuracy was evaluated based on the leave-one-out cross-validation. ~~The r~~Results showed that ~~demonstrated~~ UK generally outperformed OK. The -root mean square error between UK-interpolated and observed temperature related parameters was < 1.63℃ (2.94℉). ~~and that~~ The Nash-Sutcliffe efficiency coefficient~~s (NSEs) between UK-interpolated and observed parameters~~ for precipitation and solar radiation related parameters w~~ere~~as ≥ ~~greater than~~ 0.87, apart from that for the skewness coefficient ~~of daily precipitation~~, which was 0.78. ~~tTvalues Nash-Sutcliffe efficiency coefficients (NSEs) between UK interpolated and observed parameters werewere greater than 0.85 for all parameters apart from the standard deviation of solar radiation, skewness coefficient of daily precipitation,NSEs for precipitation and solar radiation related parameters were greater 0.87, apart from skewness coefficient of daily precipitation, which was 0.78 and cumulative distribution of relative time to peak intensity, with relatively lower interpolation accuracy (NSE > 0.66).~~ In addition, CLIGEN~~-~~-simulated daily weather sequences using UK-interpolated and observed parameters ~~inputs~~ showed consistent statistics and frequency distributions. The mean absolute discrepancy between the two sequences ~~in~~ for ~~the average and standard deviation of the~~ temperature ~~was~~ was < ~~less than~~ 0.51℃~~.~~, and ~~The~~ the mean absolute relative discrepancy ~~for the same statistics~~ for ~~the~~ solar radiation, precipitation amount, duration and maximum 30-min intensity was < ~~I_{30} were less than~~ 5% in terms of the mean and standard deviation. These CLIGEN parameter values at ~~the~~ 10 km resolution would meet the minimum ~~WEPP climate~~data requirements for WEPP application throughout ~~in~~ mainland China. The dataset is available~~availability~~ at - -http://clicia.bnu.edu.cn/data/cligen.html and http://doi.org/10.12275/bnu.clicia.CLIGEN.CN.gridinput.001 (Wang et al., 2020).

Keywords: CLIGEN, input parameters, database, China, storm pattern

1 Introduction

Weather generators (WGs) are stochastic models that can generate arbitrarily long sequences of weather variables with statistical properties that are similar to observations for a specific location or area (Yin and Chen, 2020). Early WGs were originally developed to provide surrogate climate series for hydrological, soil erosion, and agricultural models when the observed data could not satisfy the application requirements due to missing data, limited record length or spatial coverage (Wilks and Wilby, 1999). Since the 1990s, WGs have received increased attention as a statistical downscaling tool for the assessment of climate change impact (Katz and Parlange, 1996; Maraun et al., 2010). While global climate models (GCMs) / regional climate models (RCMs) have been used for climate projections, outputs from these models were often too coarse to meet the requirements of earth surface process models in terms of spatial-temporal resolutions and were biased compared with observations. Statistical downscaling methods, mainly including perfect prognosis (PP), model output statistics (MOS) and WGs, can be used to downscale and bias-correct the output from GCM/RCMs prior to earth surface model applications (Maraun and Widmann, 2018; Yin and Chen, 2020).

CLIGEN is a stochastic WG developed based on the generators used in the EPIC and SWRRB models (Williams et al., 1985; Williams et al., 1984) and was released in 1995 initially accompanying the process-based Water Erosion Prediction Project (WEPP) model from the United ~~accompanied by the process-based soil erosion model Water Erosion Prediction Project (WEPP) by the United~~ States Department of Agriculture (Nicks et al., 1995). CLIGEN can simulate a series of long-term climate data in daily scale, including maximum and minimum temperatures, precipitation, solar radiation, dew point, wind velocity and direction. In addition, CLIGEN can generate three inter-storm variables in sub-daily scale, including storm duration, time to peak intensity (t_p) and the ratio of the peak intensity to the average intensity (i_p), from which an unlimited length of high-resolution breakpoint data can be generated (Flanagan et al., 2001; Nicks et al., 1995; Yu, 2003).

Of the ten CLIGEN-simulated weather elements, seven, namely daily maximum and minimum temperature, daily precipitation, duration, t_p, i_p, and daily solar radiation, are all that are required for predicting hydrological processes, soil erosion, and bio-production (Arnold et al., 1998; Flanagan et al., 2001; USDA-ARS, 2013~~Foster, 2005~~; Wallis and Griffiths, 1995). These seven climate elements are considered to meet the minimum data requirements for WEPP ~~if modeling wind-induced snow drift is not needed (Flanagan and Livingston, 1995)~~. As CLIGEN is independent of WEPP, it can be used to provide simulated climate series for other surface process models as well (Flanagan et al., 2014; Yu, 2002).

Table 1

Thirteen groups of input parameters related to temperature, solar radiation and precipitation as listed in Table 1 are all parameters needed by CLIGEN to generate the aforementioned seven climate elements. As a site-specific weather generator, input parameters for CLIGEN can be directly prepared for ~~site~~stations~~s~~ with observed data. CLIGEN was initially released in the United States with a set of 2600 weather station parameter files (Flanagan et al., 2001). Parameters for the daily temperature and daily precipitation were calculated directly based on the observations of temperature and precipitation from each station.

Parameters for daily solar radiation and storm pattern were based on 142 weather stations with daily solar radiation and sub-daily rainfall observations first, and then extended to other 2000 more stations using the triangulation interpolation method (Scheele and Hall, 2000).

Parameter regionalization, which extends model parameter values from stations with observations to areas/regions without observations, is required when the model is going to be used in these areas/regions. Commonly used parameter regionalization methods can be categorized as follows: (1) the parametric transplantation method, where a reference area that is spatially near or has similar climate characteristics to the target area is first selected, then the parameters of the reference area are extended to the target area (Cheng et al., 2016); (2) spatial interpolation method such as Thiessen polygon, inverse distance weighted, or ordinary Kriging, that interpolate parameter values based on spatial correlations of parameters among multiple ~~sites~~ stations (Hutchinson, 1995); (3) parameter transfer as a function of regional properties such as multiple regression, based on correlations between parameters and regional characteristics (Cowpertwait et al., 1996); (4) regionalization considering both the spatial correlation of parameters and the correlation between parameters and regional characteristics, including external drift Kriging~~,~~ and universal Kriging~~,~~ that can be treated as combination methods to take advantage of method (2) and (3) (Haberlandt, 1998; Semenov and Brooks, 1999).

Accuracy of parameter regionalization is known to be influenced by several factors. Firstly, regionalization of climate variables with lower or regular spatial variability generally performs better than highly heterogeneous and discontinuous variables.~~ Xu et al. (2018) attempted to regionalize monthly temperature and precipitation in the Kangdian region of China and noted that the accuracy of interpolation for the temperature was higher than that for the precipitationthe *root mean square error* (RMSE) of the temperature was less than that of the precipitation.~~ Secondly, for the same climate variable, temporal resolution plays an important role. The climate variable at a monthly or annual scale tends to perform better than variables at a daily or hourly scale because data with finer resolutions possess greater spatial variability. Thirdly, adopted approaches affect the efficiency of regionalization. For example, Wilks (2008) compared and evaluated the interpolation accuracy of four spatial interpolation methods for parameters of WGEN (Weather GENerator), a weather generator developed by Richardson and Wright (1984), ~~parameters and~~ and results showed that locally weighted regressions outperformed Thiessen polygons and domain-wide ('global') regressions. The accuracy of interpolation can be improved by adopting auxiliary covariables that are correlated with the regionalized climate variables into the regionalization process (Hengl et al., 2007). For example, elevation is frequently used as an auxiliary covariable and has been found to improve the interpolation of temperature and precipitation (Carrera-Hernández and Gaskin, 2007; Ly et al., 2013; Verworn and Haberlandt, 2011), especially in mountainous regions with complex terrains (Xu et al., 2018).

Several studies have been attempted at regionalization of CLIGEN input parameters. Regionalization of CLIGEN input parameters for WEPP have combined the parametric transplantation~~parameter transport~~ and spatial interpolation. When CLIGEN was developed in the U.S. to provide climate input to WEPP, parameter values for 2600 stations were regionalized based on inverse distance weighting (IDW). In the WEPP application, users identify the targeted location, for which daily weather sequences using parameters from the nearest stations will be automatically generated directly or by interpolation from surrounding stations (up to 20 stations within a distance of one degree of latitude/longitude). The parameter files and the internally installed interpolation in the WEPP application has facilitated application of CLIGEN/WEPP in the US. However, the accuracy of regionalized parameters has not been evaluated and the effect on generated weather sequences using the interpolated parameters are largely unknown.

Chen (2008) explored four spatial interpolation methods, inverse distance weighting (IDW), ordinary Kriging (OK), global polynomial interpolation (GPI), and local polynomial interpolation (LPI), to regionalize the daily temperature and precipitation related input parameters of CLIGEN for 12 stations in the Loess Plateau of China. Paired t-tests showed that the temperature and precipitation series generated using interpolated input parameters we~~a~~re not significantly different from those generated using input parameters computed using observations for the 12 ~~sites~~ stations considered (Chen, 2008). However, solar radiation and storm pattern-related parameters used to generate daily solar radiation and storm characteristics were not considered in Chen's study (Chen, 2008). Input parameters for simulating the 7 weather variables mentioned above, listed in Table 1, meet the minimum data requirements for WEPP at a specific ~~site~~station. Without ~~temperature,~~ solar radiation and storm pattern-related parameter values, CLIGEN cannot be used to generate the required weather sequences for WEPP.

The overall aim of this study was to enable widespread use of CLIGEN to generate daily ~~precipitation,~~ temperature, ~~and~~ solar radiation, precipitation and sub-daily precipitation variables anywhere in mainland China and to gain better understanding of the performance of various spatial interpretation techniques. Specific objectives of this study were to (1) assemble CLIGEN input parameter values for 2405 ~~sites~~ stations in mainland China based on meteorological observations; (2) evaluate spatial interpolation techniques for regionalizing CLIGEN parameters; (3) produce grid-based CLIGEN temperature, solar radiation and precipitation parameter values at 10 km resolution for mainland China.

2 Data and methods

2.1 Data collection

Four datasets consisting of daily temperature, daily rainfall, and hourly rainfall from 2405 meteorological stations, and solar radiation data from 130 stations distributed across mainland China were collected (Fig. 1) from the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA) and had~~ve~~ been quality controlled by NMIC. Data lengths were different for these four datasets (Table 2). Daily temperature and daily rainfall data were characterized by longer periods of observation for most stations compared with hourly rainfall data, especially for stations located in the northwest arid area and the Qinghai-Tibet plateau where gauges for observing hourly rainfall for some stations were installed very late (Zhao, 1983; Wang and Zuo, 2009). Based on these four data sets, a total of 156 parameter values were calculated for each station. It should be noted that the 12th value of TimePk is equal to 1 by definition and 155 parameters were involved in the calculation and interpolation. The siphon rain gauges used to record hourly rainfall were stopped in winter to avoid freezing failures; therefore, hourly rainfall was only available for the warm rainy season for some northern and western stations. Nine stations distributed in ~~the~~ North China (Miyun, Zhengzhou, Ha'erbin), Northwest China (Lanzhou, Wulumuqi), the Tibet Plateau (Lasa), and South China (Fuzhou, Changsha, Haikou) were selected to further display the regional differences and monthly variability of input parameters (Fig. 1).

Fig. 1.

Table 2.

2.2 Site-based input parameters and simulation

CLIGEN requires 13 groups of input parameters and 12 values for each group to stochastically simulate temperature, solar radiation and precipitation (Table 1). Temperature-related input parameters, TMAX AV, SD TMAX, TMIN AV, and SD TMIN are used to simulate the daily maximum and minimum temperature for each simulated day and to decide whether the simulated precipitation occurred as snowfall or rainfall (Table 1). These four values can be calculated using daily maximum and minimum temperature data for each month directly. Solar radiation related inputs SOL.RAD and SD SOL are used to generate daily solar radiation and can be directly obtained from observed daily solar radiation.

The wet-following-wet and wet-following-dry day transition probabilities, P(W|D) and P(W|W) are used to determine the occurrence of rainy days with a first-order two-states Markov chain prepared as follows:

$$P(W|W) = \frac{N_{ww}}{N_{wd}+N_{ww}}, \tag{1}$$

$$P(W|D) = \frac{N_{dw}}{N_{dw}+N_{dd}}, \tag{2}$$

in which, N_{ww}, N_{wd}, N_{dw}, N_{dd} represent the number of days in a month that a wet day followed a wet day, a wet day followed a dry day, a dry day followed a wet day, and a dry day followed a dry day, respectively. For each simulated wet day, MEAN P, S DEV P, and SKEW P are used to simulate the daily precipitation amount using a skewness normal distribution. These three parameters can be computed directly from daily precipitation month by month. As CLIGEN assumes there is only one storm occurring on a wet day, daily precipitation ~~amount~~ depths in CLIGEN are equal to storm precipitation amount.

MX.5P and TimePk are used to simulate inter-storm variables, including storm duration (D, h) and two normalized dimensionless variables, the ratio of peak intensity to average intensity (i_p), and the ratio of time to the peak intensity to storm duration (t_p) (Nicks et al., 1995; Yu, 2002; Yu, 2003; Zhang and Garbrecht, 2003). MX.5P represents the average maximum 30-min intensity for each month. The maximum 30-min intensity for a wet day is denoted as I_{30}. If a month has *n* wet days, the maximum I_{30} among *n* wet days can be denoted by maxI_{30}; and for a specific month in a data series of *k* years, the MX.5P is given by:

$$\mathrm{MX.5P} = \frac{1}{k}\sum maxI_{30} \tag{3}$$

~~Theoretically~~Ideally, MX.5P values are ~~expected to~~should be prepared using rainfall data with ~~an~~ a resolution of ~~observed i~~ 30 min or less. Depending on the temporal resolution, I_{30} can be calculated directly from moving averages of the original data over successive 30 -minutes. ~~Considering~~ Given the limited availability of ~~aforementioned~~ high-resolution rainfall observations for this study, MX.5P was ~~calculated~~ estimated ~~in this study~~ using hourly data described in detail elsewhere~~in reference to methods developed by~~ (Wang et al.~~(~~, 2018~~b~~)~~)~~.

In CLIGEN (Nicks et al., 1995), as in Arnold and Williams (1989), it is assumed that the magnitude of precipitation intensity decreases exponentially from the maximum rate when time distribution of precipitation intensities is discarded. ~~Rainfall intensity is basically assumed to be ranked from high to low in CLIGEN (Nicks et al., 1995);~~ T~~t~~herefore, the precipitation depth $P_{\Delta t}$ in any given interval Δt can be described by:

$$P_{\Delta t} = i_p \int_0^{\Delta t} e^{-t/\tau} dt = \tau i_p \left(1 - e^{-\Delta t/\tau}\right) \tag{4}$$

For hourly data, the interval Δt = 1 h, and the maximum 1 h precipitation P_{1h} and maximum 2 h precipitation P_{2h} were known:

$$\frac{P_{1h}}{P_{2h}} = \frac{1-e^{-1/\tau}}{1-e^{-2/\tau}} \tag{5}$$

where τ can be solved and then i_p can be readily obtained as,

$$i_p = \frac{P_{1h}}{\tau(1-e^{-\frac{1}{\tau}})} \tag{6}$$

Once τ and i_p are known, the maximum 30-min precipitation $P_{0.5}$ can be determined as,

$$P_{0.5h} = \tau\, i_p (1 - e^{-\frac{1}{2\tau}}) \tag{7}$$

The maximum 30-min rainfall intensity is given simply as,

$$I_{30min} = 2P_{0.5h} \tag{8}$$

In reference to Wang et al. (2018~~b~~), ~~TimePk~~ MX.5P can be directly prepared using hourly rainfall data.

There are 12 discrete values of TimePk for each ~~site~~station, describing an empirical cumulative probability distribution of time to peak (Nicks et al., 1995). The observed interval is Δt and the storm duration, D, consists of n intervals. If the peak intensity occurs in the ith interval, time to peak intensity, T_p is estimated as,

$$T_p = (i - \tfrac{1}{2})\Delta t \tag{9}$$

and time to peak as a fraction of duration is,

$$t_p = \frac{T_p}{D} = \frac{(i-0.5)}{n} \tag{10}$$

If $Ntp(i)$ is the number of wet days from all data records with $t_p \leq$ i/12 for i = 1, 2, …12, then

$$TimePk(i) = \frac{Ntp(i)}{Ntp(12)} \tag{11}$$

TimePk computed using 1-min rainfall data and hourly rainfall data differs slightly, and it has some small influence on CLIGEN~~-~~-simulated intensity and duration (Wang et al., 2018). Therefore, ~~In reference to Wang et al. (2018b),~~ TimePk was prepared directly using hourly data in this study for consistence.~~ as well as MX.5P.~~ Given the time increment (Δt) of 1 hour, and known storm duration (D) for each wet day, TimePk can be computed using equations (9) to (11). It is worth noting that the 12th parameter value of TimePk for all stations equals to 1 (equation 11).

2.3 Spatial interpolation by Kriging

Kriging interpolation is a spatial interpolation method that gives the best linear unbiased prediction of intermediate values, assuming a Gaussian process governed by prior covariance. For a research region with n samples at spatial locations $\boldsymbol{x}_i$ (i = 1, 2, …, n), $Z(\boldsymbol{x}_i)$ are the sample values at $\boldsymbol{x}_i$. At an unknown target point $\boldsymbol{x}_0$, the estimated value $\hat{Z}(\boldsymbol{x}_0)$ can be expressed as a weighted average of the known observations $Z(\boldsymbol{x}_i)$ (Wackernagel, 2013):

$$\hat{Z}(\boldsymbol{x}_0) = \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{x}_i) \tag{12}$$

where λ_i are the weighting coefficients of the known sample values $Z(\boldsymbol{x}_i)$, which depend on the spatial autocorrelation structure of the sample values and should minimize the prediction error variance. Assuming the variable value $Z(\boldsymbol{x})$ can be modeled as a combination of a deterministic trend $\mu(\boldsymbol{x})$ and an auto-correlated random error $\varepsilon(\boldsymbol{x})$, $Z(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \varepsilon(\boldsymbol{x})$, then the best linear unbiased prediction requires $E[\hat{Z}(\boldsymbol{x}_0) - Z(\boldsymbol{x}_0)] = 0$ and $Var[\hat{Z}(\boldsymbol{x}_0) - Z(\boldsymbol{x}_0)]$ is minimized. *Ordinary Kriging* (*OK*) assumes that the trend is constant but unknown, $\mu(\boldsymbol{x}) = m$, while in *universal Kriging* (UK), the trend is assumed to be a linear combination of some known covariables f_l, $\mu(\boldsymbol{x}) = \sum_{l=1}^{k} \beta_l f_l$. *Universal Kriging* (UK) ~~takes into account~~considers the relationship between the target variable and the auxiliary covariables. Soil, elevation, temperature, and remote sensing images ~~were~~ are commonly used auxiliary covariables (Haberlandt, 1998; Li et al., 2014; McKenzie and Ryan, 1999; Semenov and Brooks, 1999).

Both OK and UK were ~~adopted~~ used to interpolate the CLIGEN input parameters in this study. Stepwise regression was conducted to select appropriate covariables for UK. The longitude, latitude, elevation, and annual rainfall amount were found correlated with the ~~twelve groups of parameters CLIGEN~~ parameters one for each month for CLIGEN with the exception of the SKEW P (Table 1), ~~and were selected as auxiliary covariables for these twelve groups of parameters~~therefore, all these four variables were adopted as auxiliary covariables when UK was conducted to interpolate these twelve groups of parameters.~~.~~ SKEW P had low correlations with all four of these covariates but good correlation with parameters MEAN P and SDEV P. Therefore, MEAN P and SDEV P were selected as covariables during the interpolation of SKEW P.

2.4 Assessment of interpolation accuracy

A leave-one-out cross-validation method was ~~applied~~ used to evaluate the interpolation accuracy of OK and UK. First, one of the 2405 stations was excluded from data analysis and treated as unknown, data for the remaining 2404 stations were then used to predict parameter values for the excluded station using OK or UK. This leave-one-out procedure was repeated for 155 parameters for each of the 2405 stations (13 groups × 12 input parameters -1, as the value of 12th parameter of TimePk is always 1, Table 1). Denoting CLIGEN parameters based on observations as P_O and the corresponding predicted CLIGEN parameters obtained using OK or UK as P_K, three indicators, *root mean square error* (RMSE), *Nash-Sutcliffe efficiency coefficient* (NSE), and *percent bias* (PBIAS) were selected to evaluate and compare the performances of OK and UK as follows (Yin et al., 2019): ~~The input parameters prepared using observation were denoted as~~ $\cancel{P_{ij}^{obs}}$ ~~(~~$\cancel{i = 1, 2, \ldots, 2405}$ ~~stations;~~ $\cancel{j = 1, 2, \ldots 131}$ ~~input parameter values), and the corresponding inputs interpolated using OK (UK) as~~ $\cancel{P_{ij}^{OK}}$ $\cancel{(P_{ij}^{UK})}$~~. For a specific parameter value~~ $\cancel{j_{th}}$~~, assumed the value for the~~ $\cancel{i_{th}}$ ~~station was unknown and removed~~ $\cancel{P_{ij}^{obs}}$ ~~from all stations. Use the remaining stations to predict~~ $\cancel{P_{ij}^{OK}}$ $\cancel{(P_{ij}^{UK})}$ ~~of~~ $\cancel{x_i}$ ~~using OK (UK), respectively. Following this procedure, two sets of input parameters for 2405 stations predicted by OK and UK were obtained and compared with parameters determined from observations to evaluate two interpolation methods.~~

~~Four indicators,~~ ~~*Nash-Sutcliffe efficiency coefficient*~~ ~~(NSE),~~ ~~*percent bias*~~ ~~(PBIAS),~~ ~~*root mean square error*~~ ~~(RMSE), and~~ ~~*RMSE-observations standard deviation ratio*~~ ~~(RSR), were selected to evaluate and compare the performances of OK and UK as follows (Yin et al., 2019):~~

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_n (P_O - P_K)^2 \cancel{\sum_{i=1}^{n}(P_{ij}^{obs} - P_{ij}^{K})^2}}, \qquad (13)$$

$$\mathrm{NSE} = 1 - \frac{\sum_n (P_O - P_K)^2 \cancel{\sum_{i=1}^{n}(P_{ij}^{obs} - P_{ij}^{K})^2}}{\sum_n (P_O - \overline{P_O})^2 \cancel{\sum_{i=1}^{n}(P_{ij}^{obs} - \overline{P^{obs}})^2}}, \qquad (\cancel{13}14)$$

$$\mathrm{PBIAS} = \frac{\sum_n (P_O - P_K) \cancel{\sum_{i=1}^{n}(P_{ij}^{obs} - P_{ij}^{K})}}{\sum_n P_O \cancel{\sum_{i=1}^{n} P_{ij}^{obs}}} * 100, \qquad (\cancel{14}15)$$

$$\cancel{\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{ij}^{obs} - P_{ij}^{K})^2}}, \qquad \cancel{(15)}$$

As

$$\cancel{\mathrm{RSR} = \frac{RMSE}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{ij}^{obs} - \bar{O})^2}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{ij}^{obs} - P_{ij}^{K})^2}}{\sqrt{}}}. \qquad \cancel{(16)}$$

NSE and PBIAS are inappropriate for temperature-related parameters which are in interval scales, and the same is true of probabilities. NSE and PBIAS were computed for parameters in ratio scales only, i.e. MEAN P, S DEV P, SKEW P, SOL.RAD, and SD SOL. By calculating ~~of~~ the above ~~four~~ three indicators~~, for each input parameter values,~~ the better of the two interpolation techniques, OK and UK, was determined and applied to calculate the regionalization of CLIGEN input parameters for mainland China. A two-dimensional grid database was established at a spatial resolution of 10 km × 10 km based on the 155 sets of ~~156 parameter layers~~interpolated parameters ~~in total~~.

Input parameters based on observed data and interpolated data using the better interpolation technique were input into CLIGEN to evaluate the influence of regionalized parameters on the simulation. For each station, 100 years of continuous climate series were generated using the default CLIGEN stochastic seed without interpolation between months, and the simulated data predicted ~~by P^{obs} and P^{K} were denoted as G^{obs} and G^{K},~~ by P_O and P_K were denoted as G_O and G_K, respectively. The maximum and minimum temperature (°C), daily solar radiation (Langley), daily rainfall amount (mm), storm duration (h), i_p and t_p of each simulation day were derived from ~~G_i^{obs} and G_i^{k}~~ G_O and G_K ~~f~~for each station, and the maximum 30-min intensity (I_{30}, mm/h) was calculated based on an assumed bi-exponential storm pattern (Yu, 2002). CLIGEN input parameter values are required to have US customary unit as shown in Table 1, while CLIGEN output is produced in SI as input to WEPP.

Three basic statistics, the average, standard deviation and skewness coefficient were calculated for each CLIGEN-generated variable. The ~~*Absolute*~~ *absolute error* (AE) and *mean absolute error~~s~~* (MAE) were calculated to examine the differences between the two sets of statistics for generated temperatures. *Relative error* (RE) and *mean absolute relative error~~s~~* (MARE) were calculated to examine the differences between the two sets of statistics for generated daily solar radiation, daily precipitation and sub-daily storm pattern:

$$|AE| = |G_O - G_K| \qquad (\cancel{17}16)$$

$$\mathrm{MAE} = \frac{1}{2405}\sum|(G_O - G_K)| \qquad (\cancel{18}17)$$

$$|RE| = 100\%\,|(G_O - G_K)/G_O| \qquad (\cancel{19}18)$$

$$\mathrm{MARE} = \frac{100\%}{2405}\sum|(G_O - G_K)/G_O| \qquad (\cancel{20}19)$$

3 Results

3.1 Spatial-temporal distribution of CLIGEN input parameters

Thirteen groups of CLIGEN temperature and precipitation parameters from 2405 stations and solar radiation parameters from 130 stations were plotted to ~~exhibit~~ examine the inter-annual variation and the differences among parameters (Fig. 2). The average max-temperature and min-temperature, TMAX AV and TMIN AV (in unit of °F ~~°F~~, 1°F = 1°C/1.8+32) ~~°F = 1°C/1.8 + 32)~~, and the average and standard deviation of solar radiation, SOL.RAD and SD SOL (in unit of Langley, 1Ly = 4.184*10^{-2}MJ/m^2) ~~1 Ly = 4.184 * 10^{-2} MJ/m²)~~ showed strong seasonality and the spatial variance ~~value~~ became smaller ~~convergent~~ from the cold season to the warm one (Fig. 2a, 2c, 2e-f). The spatial distribution of CLIGEN temperatures and solar radiation related inputs in August based on the UK-interpolated results were depicted as examples (Fig. 3), from which we can find a differentiation rule for latitude and vertical zonality for TMAX AV, TMIN AV (Fig. 3a-b). SD TMAX and SD

TMIN varied with season with a similar pattern and with generally higher values in spring and autumn (Fig. 3c-d), because these two seasons are transitional periods between warm and cold seasons when temperature fluctuations are larger.

Fig. 2.

Fig. 3.

The average and standard deviation of daily precipitation, MEAN P, S DEV P (in unit of inch, 1 inch = 25.4 mm), and the average monthly maximum 30-min intensity, MX.5P (in unit of inch/h, 1 inch/h = 25.4 mm/h), showed a similar seasonal pattern with the parameter values becoming gradually higher from the cold season to the warm (Fig. 2g-h). Precipitation in China is influenced by the East Asian summer monsoon and the location relative to land and sea. From the spatial distribution of daily precipitation in August we found a general decreasing trend from southeast to southwest (Fig. 4a-b). The August rain belt is located in North and Northeast China, while the South China region is controlled by the subtropical high-pressure belt and experiences a summer drought. Therefore, MEAN P and MX.5P in North China ~~was~~ were apparently greater than in South China. In comparison, skewness of daily precipitation, SKEW P, showed imperceptible differences among months and no apparent latitudinal or longitudinal zonality (Fig. 4c). This may be one of the reasons leading to the low spatial interpolation accuracy of SKEW P.

Fig. 4.

The wet-following-dry transition probability P(W~~/~~|D) showed a clear inter-annual variability in that the probability increased from cold season to warm (Fig. 2j), while the wet-following-wet transition probability P(W|~~/~~W) was characterized by greater regional differences but smaller monthly variability for most stations compared with P(W|~~/~~D) (Fig. 2k). The spatial-temporal variation in these two transition probabilities revealed the stepwise northward progress of the East Asian monsoon and the North-South advance of the Frontal cyclone (Liao et al., 2004). Due to the pre-monsoon rainy season before June, strong convection in summer, and the retreating monsoon rain belt after August, the southern region was characterized by a longer rainy season than North China (Yu and Zhou, 2007). Therefore, P(W|~~/~~W) of the southern region was generally higher than other regions and its seasonal variations were relatively insignificant (Fig. 5b).

Fig. 5.

MX.5P of nine example stations showed the regional differences more clearly in that the parameters of southern stations were relatively higher (Fig. 5c). Differences among southern and northern stations became gradually smaller in the warm season. It should be noted that the narrower range of MX.5P in winter was partially related to the limited availability of hourly data. Due to the restriction of low temperatures on siphon rain gauge observations, MX.5P in cold seasons were available for fewer stations than in warm seasons.

TimePk consists of 12 discrete values representing the cumulative distribution of time to peak intensity ranging from 0 to 1 for a specific location. The sixth value for TimePk represents the cumulative ratio of storms with peak intensity occurring before 1/2 duration, and related ratios for 2405 stations ranging from 60% to 80% (Fig. 2m). TimePk for nine example stations shows the cumulative ratio of time to peak intensity in different regions, consistently indicating that most ~~storms~~ peak intensities tend to occur earlier during the storms, with no obvious regional differences found for this parameter (Fig. 5d).

3.2 Evaluation of interpolated parameters using OK and UK

3.2.1 Parameters at the~~a~~ daily scale

The leave-one-out cross-validation showed that four groups of temperature parameters, TMAX AV, SD TMAX, TMIN AV, SD TMIN, two groups of solar radiation, SOL.RAD, SD SOL, and four groups of precipitation parameters at daily scale, MEAN P, S DEV P, P(W~~/~~|D) and P(W~~/~~|W), were well predicted by *ordinary Kriging* (OK) and *universal Kriging* (UK). RMSE for all these parameters were relatively low compared with the average of observed inputs (Table 3). For all these four groups of temperature related parameters, RMSE between the UK-interpolated and observed were less than 2.94°F (1.63℃). ~~The average~~ NSE ~~over 12 months~~ were~~as~~ greater than 0.~~88~~ 87 for ~~all these 8~~parameters of MEAN P, S DEV P, SOL.RAD, and SD SOL ~~groups of parameters~~ in ratio scales. The PBIAS were all smaller than 1%, suggesting that parameters based on observation and interpolation have a very close average trend and showed no obvious bias. In contrast, the interpolated accuracy of ~~two groups of solar radiation parameters, SOL.RAD, SD SOL, and~~ the skewness coefficient of daily precipitation, SKEW P, were not very satisfactory ~~(Table 3)~~, with NSE being ~~0.46-0.80~~0.48 using OK and 0.~~0.66-0.85~~78 using UK. ~~The relatively lower interpolation accuracy of solar radiation related parameters was partially related to the sparsity of stations involved in the interpolation.~~ Parameters related to daily average (TMAX AV, TMIN AV, SOL.RAD and MEAN P) were generally better predicted than corresponding parameters related to standard deviation (SD TMAX, SD TMIN, SD SOL and S DEV P), and the skewness coefficient was the least accurately simulated. ~~In addition, the interpolation accuracy tended to be lower in the warm season (May to Sept.) compared with the yearly rest period (Fig. 6a-f).~~

Table 3.

In comparison with OK, the overall and monthly predicted accuracy using UK with auxiliary covariables obviously improved TMAX AV~~ and~~, TMIN AV ~~in the warm season~~, SOL.RAD, MEAN P, ~~SOL.RAD~~SKEW P, P(W|W), and P(W|D) ~~in the cold season and SD SOL in March~~ (Fig. 6). ~~. The predicted quality for SD TMAX, MEAN P, S DEV P, P(W|W), and P(W|D) was somewhat improved by UK, as these groups of parameters already had high accuracy when using OK to interpolate, resulting in a small range of improvement.~~ The predicted accuracy for ~~the minimum temperature (~~SD ~~TMIN~~TMAX and ~~)~~ S DEV P using the two techniques showed no evident difference. For SD TMIN and SD SOL, ~~using the two techniques showed no evident difference~~the predicted accuracies were approximate ~~,~~ except for July, when the ~~NSE~~ RMSE of UK ~~was~~ were obviously ~~lower~~ larger than OK and the reason was unclear. Although the prediction of SKEW P using UK was not as good as other parameters at a daily scale, the improvement compared with OK was quite obvious, as the ~~average~~ NSE over 12 months increased from 0.~~458~~ 48 for OK to 0.~~769~~ 78 for UK, and the RMSE decreased from 0.73 mm to 0.47 mm (Table 3).

Predicted inputs using OK and UK versus inputs based on observations from August were plotted to show the difference between two methods as examples (Fig. 7a-7k).

Fig. 6.

3.2.2 Parameters at ~~a~~ the sub-daily scale

Cross-validation results showed that the interpolation ~~accuracies~~ of the two ~~storm pattern related~~ parameters related to storm patterns, i.e. MX.5P and TimePk ~~were~~ performed ~~not as good as precipitation related parameters on a daily scale~~well. Three~~Four~~ cross-validation statistics for these two parameters using two methods were numerically ~~close~~ similar ~~(Table 3) for both parameters~~(Table 3). NSE over 12 months for MX.5P interpolated with OK and UK were both equal to 0.95. ~~After taking auxiliary covariates for interpolation using UK, the prediction improved only slightly.~~ The seasonal ~~annual~~ varia~~nce~~tion ~~of~~ in ~~NSE~~ RMSE based on OK and UK ~~varied~~ follows ~~in~~ a similar pattern ~~within the year~~ (Fig. 6l-m). For~~the parameter of~~ TimePk, the ~~NSE~~ RMSE ~~of~~ using OK were slightly ~~higher~~ lower than those~~at~~ ~~from~~ using UK ~~from~~for- the 3^{th}, 4^{th}, and 5^{th}~~Jane to May~~ parameters, ~~,~~ but ~~reversed~~ slightly higher ~~during the~~ for the others~~rest period~~. ~~In comparison, MX.5P performed better than TimePk. The interpolation accuracy of TimePk was the lowest among all 13 groups of input parameters (Table 3).~~

Fig. 7.

Interpolation accuracy has been adequately estimated through cross-validation, and these results ~~agreed~~ indicated that the accuracy of interpolation results based on UK was generally higher than those based on OK. Therefore, two sets of CLIGEN-simulated climate series using observed inputs and UK-interpolated inputs were generated and compared to further evaluate the regionalized parameters using UK for the simulation of CLIGEN.

3.3 Assessment of parameters' regionalization on the CLIGEN outputs

3.3.1 Simulated climate elements at a daily scale

CLIGEN-~~-~~simulated daily temperature and solar radiation based on UK-interpolated input parameters agreed well with those simulated based on observed parameters. The average, standard deviation and skewness coefficient of generated daily maximum temperature, minimum temperature, solar radiation and daily precipitation generated using observed and interpolated input parameters were calculated for each station, and the simulated accuracy of the average and standard deviation were found to be better than that of the skewness coefficient. The RMSE~~s~~ ~~The NSE~~ of the ~~average~~ mean and standard deviation were all ~~greater than~~less than 0.79°C, 18 Ly/day (0.75 MJ/day), ~~0.97 for~~ 0.71 mm, respectively, for ~~generated~~ daily ~~climate elements~~temperatures, solar radiation and precipitation~~ at a daily scale~~ (Table 4 & Table 5). The NSE of the skewness coefficient for ~~temperature and~~ solar radiation was ~~ranged from~~ 0.~~94~~56~~-0.95~~, obviously ~~slightly~~ lower than that for the mean ~~corresponding average~~ and standard deviation (Table 4). ~~By contrast~~Meanwhile, the NSE of the skewness coefficient of daily precipitation was low ~~as low as 0.48~~ (Table 5~~).~~), indicating a relatively low ~~This may be attributed to the lower~~ interpolation accuracy of SKEW P~~,~~. In fact, the~~with the~~ ~~lowest~~ accuracy of SKEW P was the lowest among all input parameters (Table 3).

Table 4.

The *absolute error* (AE) of the average, standard deviation and skewness coefficient between the simulated daily temperature of G_O $\cancel{G^{obs}}$ and G_K $\cancel{G^{UK}}$ were statistically similar (Table 4). The *mean absolute error~~s~~* (MAE~~s~~) over 2405 stations were all lower than 0.51℃ ~~℃~~. For daily solar radiation, the *relative errors* (REs) for ~~three statistics~~the mean and standard deviation were ~~all~~ lower than 10~~2~~% for more than 90% stations, and the *mean absolute relative error* (MARE) were lower than 4~~0.1~~%.

Table 5.

For generated daily precipitation, 94.1% and 91.4% of stations yielded RE~~s~~ of the average and standard deviation below 10%, and the MARE for 2405 stations were 3.72~~-~~% and 4.56%, respectively. Bias between annual rainy days of G_O and G_K ~~G^{UK} and G^{obs}~~ was small as well. RE~~s~~ of 92.9% of stations were lesser than 10%. The frequency distribution of daily precipitation generated using two sets of inputs were well matched for most stations. Fig. 8a depicted the frequency distributions of simulated daily precipitation for Fuzhou station as an example, with RE slightly higher than MARE over 2405 stations. Meanwhile, some stations do not satisfactorily simulate the frequency distribution. The frequency distribution of Tuokexun, whose simulation quality was approximately the worst among 2405 stations was offered as an example (Fig. 8d). It showed that the frequency of daily precipitation ranging from 0-1 mm was under~~-~~estimated, whereas that for values greater than 1 mm was over~~-~~estimated (Fig. 8d).

Fig. 8.

3.3.2 Simulated storm pattern related variables

The average and standard deviation of storm duration and the maximum 30-min intensity (I_{30}) generated using observed and UK-interpolated input parameters possessed a generally small bias. The NSE of the average and standard deviation for both duration and I_{30} were above 0.87. Compared with the average and standard deviation, the accuracy of skewness was the worst, with the NSE being 0.26 for the duration and 0.66 for the peak intensity index. Comparison of the frequency distribution of the duration and I_{30} for Fuzhou station showed that the frequency of simulated storm patterns was~~ere~~ well preserved using data employing UK-interpolated parameters (Fig. 8b-c). The frequency distribution of the duration and I_{30} for Tuokexun station showed that interpolated parameters seemed to underestimate low values and overestimate high values (Fig. 8e-f).

4 Discussion

Both AE and RE indexes were adopted to evaluate the simulated results in this study. The RE index was applied for solar radiation and precipitation related outputs, while the AE index was applied for the assessment of temperature-related outputs~~.~~, as~~This is because we find that~~ RE was not an appropriate ~~index~~ indicator to evaluate the temperature which was in interval scale. ~~for~~ S~~some s~~tations located in high latitude or high altitude areas where the mean annual temperature may be close to zero resulting in an extremely high derived RE. For example, the mean maximum temperature of Qian'an station (Fig. 1) using observed inputs was -0.01°C~~°C~~ and that using interpolated inputs was -0.33°C, ~~°C,~~ resulting in a~~n~~ RE between the two values of ~~was~~ 2912.7%, which was an extremely large error. However, the mean maximum temperature simulated using the two data sets were very similar, with an AE of 0.32°C ~~°C~~. ~~We've checked more than 100 stations with extremely high REs for maximum temperature, and all were in similar situation (Fig. 9).~~ If RE was used to evaluate the simulated temperature, the actual simulation quality may be strongly underestimated. Therefore, AE were used to demonstrate errors between generated temperature based on observed and interpolated inputs.

~~**Fig. 9.**~~

The frequency distributions of CLIGEN~~-~~simulated daily precipitation, duration and peak intensity at Tuokexun station using observed inputs were all not well preserved by those simulated using UK-interpolated inputs (Fig. 8). The simulation quality for Tuokexun was almost the worst among 2405 stations, as RE~~s~~ for all these three precipitation related variables were greater than 99% of stations. This may be explained partially because Tuokexun is located in the northwest arid area of China (Fig. 1), with a station density of $0.97/10^4 \cdot km^2$, much lower than that in the Eastern Monsoon Area (Table ~~7~~6). Stations involved in the interpolation were separated by far distances, with a negative influence on the interpolation accuracy (Oliver and Webster, 2014). Other stations with extremely low simulated quality similar to Tuokexun are almost located in the northwest arid area or Qinghai-Tibet Plateau where the station density is lower. The MAE and MARE for generated temperature and precipitation in the eastern monsoon area were the lowest among three physical-geographical regions of China (Table 6). The standard error of the interpolation results for the two parameters, i.e. TMAX AV and MEAN P in August are shown as an ~~an~~ example~~s~~ (Fig. 9). It can be seen ~~from the figures that~~ that the errors are relatively high in the western part of China, especially in the north-western ~~south-western~~ part of Qinghai-Tibet Plateau, where there is a large area without stations and characterized with the highest standard errors for both parameters (Fig. 1 and Fig. 9). ~~The MAE for generated temperature and the MARE for generated precipitation related variables in the eastern monsoon area were the lowest among three physical-geographical regions of China (Table 7).~~

~~**Table 7.**~~**Fig. 9.**

The number and density of weather stations for solar radiation were considerably less than for those for temperature and precipitation (Table 6~~7~~). However, the ~~simulated~~ mean and standard deviation of daily solar radiation using the UK-interpolated parameters was in good agreement with that simulated using observation-based parameter values (Table 4), and MARE of solar radiation was similar to that of daily precipitation.~~. MARE for solar radiation across all stations was the lowest among all simulated weathe r elements. MAREs were similar for the three geographical regions with the difference among them varying from 0.08% to 0.13%. S~~Solar radiation is characterized with much lower spatial variability in comparison to~~-~~ that for the temperature and precipitation. As a result, solar radiation-related parameters were easier to regionalize and parameter values could readily be interpolated for regions with~~out~~ limited observations.

~~Fig. 10.~~Table 6.

CLIGEN-input parameters in the US ~~is~~ are regionalized from 2600 stations using the *inverse distance weighted method* (IDW), which was employed in the initial attempt to regionalize CLIGEN input parameters. In this study, UK was adopted to interpolate CLIGEN parameters for mainland China. Interpolated parameter values using IDW and UK were compared for four selected parameters in August as shown in Fig. 10. It can be seen that UK performed better than IDW for all four parameters selected. UK-interpolated parameter values were concentrated mostly along the 1:1 line. The ~~NSEs~~ RMSE of all four groups of parameters interpolated using UK were lowe~~ large~~r than those predicted using IDW. Noticeable improvement was noted for SKEW P, with the ~~NSE~~ RMSE improving~~ed~~ from 0.~~27~~84 to 0.~~74~~49 using UK instead of IDW. Therefore, UK appears to be consistently superior to IDW when regionalizing CLIGEN input parameters based on the limited comparison for selected parameters.

Fig. 10.

5 Data availability

The g~~i~~ridded CLIGEN input parameter dataset of China at 10 km resolution is available~~ility~~ at the homepage of CLI~~li~~mate Change I~~i~~mpact A~~a~~ssessment (CLICIA) group ~~—~~at http://clicia.bnu.edu.cn/data/cligen.html. Additional materials including the data manual and grid information are also available~~ility~~ at the same website and can be downloaded.

6 Summary and Conclusion

The widely used stochastic weather generator CLIGEN can simulate long-term climate data to drive hydrological, soil erosion, and crop-yield models. Limitations in high spatial-temporal observations, especially at the sub-daily scale, have partially restricted its application. Daily temperature, daily precipitation, and hourly precipitation data for 2405 stations and daily solar radiation for 130 stations distributed across mainland China were collected to establish the CLIGEN input parameter files and to explore an appropriate method for regionalizing these parameters from stations to the entire region. The predicted quality using two interpolation techniques, OK and UK, were compared and fully assessed, yielding the following results:

1) UK generally performed better than OK when interpolating CLIGEN parameters. Compared with OK the

interpolation accuracy was markedly improved for parameters TMAX AV, TMIN AV, SOL.RAD, ~~SD SOL~~MEAN P, SKEW P, P(W|~~/~~D) and P(W~~/~~|W)~~, and slightly improved for parameters SD TMAX, MEAN P and S DEV P~~. For rest parameters, ~~The~~ the comparative interpolation accuracies were numerically approximate between the two techniques.

2) UK can accurately predict ~~the~~ temperature, solar radiation and precipitation input parameters for CLIGEN. RMSE in UK-interpolated parameter values for temperature were less than 1.63°C (2.94°F), and ~~The *Nash-Sutcliffe efficiency coefficient* (NSE)~~ values ~~obtained using the observed parameters and UK-~~interpolated ~~predicted parameters were all greater than 0.85 for most parameters expect for SD SOL, SKWE P and Time Pk. The interpolation accuracies for these final three parameters were relatively lower, with NSEs greater than 0.66.~~~~*s*~~NSEs for precipitation and solar radiation parameters were all greater than 0.87, except for the skewness coefficient (SKEW P) with a relatively lower interpolation accuracy (NSE = 0.78).

3) Basic statistics and frequency distributions for CLIGEN-simulated climate elements using UK-interpolated parameters agreed well with those simulated using observations. The *mean absolute error~~s~~* (MAE~~s~~) for the average, standard deviation and skewness coefficient for the two simulated series of temperature across 2405 stations were all less than 0.5°C~~1°C~~. The *mean absolute relative error~~s~~* (MARE~~s~~) for same statistics for simulated solar radiation were less than 0.1%. MARE~~s~~ for the average and standard deviation for precipitation amount, duration and I_{30} ~~are~~ were less than 5.0%, while errors for skewness coefficient for these three groups of parameters were less than 10.1%.

The developed gridded input parameter database can be applied using CLIGEN, with an established and reliable simulation quality, to the stochastic simulation of temperature, solar radiation and precipitation at a daily scale and to precipitation at a sub-daily scale for any single point in China. CLIGEN can simulate the dew point and wind as well, not regionalized in this study. As a site-based weather generator, simulated climate series using CLIGEN are independent of each other and are lack of spatial correlations among stations. Further research might focus on the rebuilding of correlations among climate elements and between nearby ~~sites~~stations.

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the National Key R&D Program (No.2018YFC0507006), the National Natural Science Foundation of China (No. 41877068) and China Postdoctoral Science Foundation (No. 2020M680433). We also would like to thank the high-performance computing support from the Center for Geodata and Analysis, Faculty of Geographical Science, Beijing Normal University [https://gda.bnu.edu.cn/].

Author Contributions

Wenting Wang calculated the input parameters, developed the programming code, and wrote the original draft; Shuiqing Yin provided the main conceptualization, supervised the project, and reviewed the draft; Bofu Yu ~~gave~~ provided ~~advises~~ advices about the methodology and reviewed the draft; Shaodong Wang reviewed the draft.

References

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R.: Large area hydrologic modeling and assessment Part 1: Model development, J. Am. Water Resour. As., 34, 73-89, https://doi.org/10.1111/j.1752-1688.1998.tb05961.x, 1998.

Arnold, J. G., Williams, J. R.: Stochastic generation of internal storm structure at a point, Trans. ASAE, 32: 161-167, https://doi.org/10.13031/2013.30976, 1989.

Carrera-Hernández, J. J., and Gaskin, S.J: Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico, J. Hydrol., 336, 231-249, https://doi.org/10.1016/j.jhydrol.2006.12.021, 2007.

Chen J: Applicability assessment and improvement of CLIGEN for the Loess Plateau of China, Thesis for Master Degree, Northwest A&F University, Xi'an, 2008. (in Chinese with English abstract).

Cheng, Y., Ao, T., Li, X., and Wu, B.: Runoff simulation by SWAT model based on parameters transfer method in ungauged catchments of middle reaches of Jialing River, Transactions of the Chinese Society of Agricultural Engineering, 32, 81-86, https://doi.org/10.11975/j.issn.1002-6819.2016.13.012, 2016. (in Chinese with English abstract)

Cowpertwait, P. S. P., O'Connell, P. E., Metcalfe, A. V., and Mawdsley, J. A.: Stochastic point process modelling of rainfall. II. Regionalisation and disaggregation, J. Hydrology, 175, 47-65, https://doi.org/10.1016/S0022-1694(96)80005-9, 1996.

Flanagan, D. C., Meyer, C. R., Yu, B., and Scheele, D. L.: Evaluation and enhancement of the CLIGEN weather generator: Soil Erosion Research for the 21st Century: Proc, International Symposium, American Society of Agricultural Engineers, St. Joseph, Michigan, USA, pp. 107-110, 2001.

Flanagan, D. C., Trotochaud, J., Wallace, C. W., and Engel, B. A.: Tool for obtaining projected future climate inputs for the WEPP and SWAT models: 21st Century Watershed Technology Conference and Workshop Improving Water Quality and the Environment Conference Proceedings, University of Waikato, New Zealand, American Society of Agricultural and Biological Engineers, St. Joseph, Michigan, USA, 8 pp., 2014.

USDA-ARS: Science Documentation, Revised Universal Soil Loss Equation, Version 2 (RUSLE2), USDA-Agricultural Research Service, Washington, D.C., USA, 2013.

Haberlandt, U.: Stochastic rainfall synthesis using regionalized model parameters, J. Hydrol. Eng., 3, 160-168, https://doi.org/10.1061/(ASCE)1084-0699(1998)3:3(160), 1998.

Hengl, T., Heuvelink, G.B., and Rossiter, D.G.: About regression-kriging: From equations to case studies, Comput. & Geosci., 33, 1301-1315, https://doi.org/10.1016/j.cageo.2007.05.001, 2007.

Hutchinson, M. F.: Interpolating mean rainfall using thin plate smoothing splines, Int. J. Geogr. Inf. Syst., 9, 385-403, https://doi.org/10.1080/02693799508902045, 1995.

Katz, R. W., and Parlange, M. B.: Mixtures of stochastic processes: application to statistical downscaling, Clim. Res., 7, 185- 193, https://doi.org/10.3354/cr007185, 1996.

Li, L., Shi, R., Zhang, L., Zhang, J., and Gao, W.: The data fusion of aerosol optical thickness using universal kriging and stepwise regression in East China, Remote Sens. Model. Ecosyst. Sustain. XI, 9221, 922112, https://doi.org/10.1117/12.2061764, 2014.

Liao, Y., Zhang, Q., and Chen, D.: Precipitation simulation in China with a weather generator, Acta Geographica Sinica, 59, 698-707, 2004. (in Chinese with English abstract).

Ly, S., Charles, C., and Degré, A.: Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review, Biotechnologie, agronomie, société et environnement, 17, 392-406, http://hdl.handle.net/2268/136084, 2013.

Maraun, D., and Widmann, M.: Statistical downscaling and bias correction for climate research, Cambridge University Press, Cambridge, United Kingdom, https://doi.org/10.1017/9781107588783, 2018.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., and Themeßl , M.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, https://doi.org/10.1029/2009RG000314, 2010.

McKenzie, N. J., and Ryan, P. J.: Spatial prediction of soil properties using environmental correlation, Geoderma, 89, 67-94, https://doi.org/10.1016/S0016-7061(98)00137-2, 1999.

Nicks, A. D., Lane, L. J., and Gander, G. A.: Weather generator, Chapter 2, US Department of Agriculture (USDA) Water Erosion Prediction Project, Technical Documentation, National Soil Erosion Research Laboratory (NSERL) Report, 2.1-2.22, 1995.

Oliver, M. A., and Webster, R.: A tutorial guide to geostatistics: Computing and modelling variograms and kriging, Catena, 113, 56-69, https://doi.org/10.1016/j.catena.2013.09.006, 2014.

Richardson, C.W., Wright, D.A.: WGEN: A model for generating daily weather variables, 1984.

Scheele, D.L., and Hall, D.E.: Corrections and improvements to the CLIGEN climate database, Moscow, Idaho, USDA Forest Service, Rocky Mountain Research Station, 2000.

Semenov, M. A., and R. J. Brooks: Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain, Clim. Res., 11, 137-148, https://doi.org/10.3354/cr011137, 1999.

Verworn, A., and Haberlandt, U.: Spatial interpolation of hourly rainfall-effect of additional information, variogram inference and storm properties, Hydrol. Earth Syst. Sc., 15, 569-584, http://dx.doi.org/10.5194/hess-15-569-2011, 2011.

Wackernagel, H.: Multivariate geostatistics: an introduction with applications, Springer, Berlin, Heidelberg, Science & Business Media, https://doi.org/10.1007/978-3-662-05294-5, 2013.

Wang, J. A., and Zuo, W: Geographic atlas of China. Beijing, China, China Atlas Press, 2009.

Wang, W., Yin, S., Flanagan, D. C., Yu, B.: Comparing CLIGEN-generated storm patterns with 1-minute and hourly precipitation data from China. Journal of Applied Meteorology and Climatology, 57, 2005-2017, https://doi.org/10.1175/JAMC-D-18-0079.1, 2018.

Wang, W., Yin, S., Yu, B., and Wang, S.: CLIGEN parameter regionalization for mainland China, http://doi.org/10.12275/bnu.clicia.CLIGEN.CN.gridinput.001, 2020.

Wallis, T. W. R., and Griffiths, J. F.: An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC), Agr. Forest Meteorol., 73, 115-133, https://doi.org/10.1016/0168-1923(94)02172-G, 1995.

Wilks, D. S.: High-resolution spatial interpolation of weather generator parameters using local weighted regressions, Agr. Forest. Meteorol., 148: 111-120, https://doi.org/10.1016/j.agrformet.2007.09.005, 2008.

Wilks, D. S., and Wilby, R. L.: The weather generation game: a review of stochastic weather models, Prog. Phys. Geog., 23, 329-357, https://doi.org/10.1177/030913339902300302, 1999.

Williams, J. R., Nicks, A. D., and Arnold, J. G.: Simulator for water resources in rural basins. J. Hydraul. Eng., 111, 970-986, https://doi.org/10.1061/(ASCE)0733-9429(1985)111:6(970), 1985.

Williams, J. R., Jones, C. A., and Dyke, P.: A modeling approach to determining the relationship between erosion and soil productivity, Trans. ASAE, 27, 129-0144, https://doi.org/10.13031/2013.32748, 1984.

Xu, X., Xu, Y., Sun, Q., Xie, T., and Zhang, H: Comparison study on meteorological spatial interpolation approaches in Kangdian region of China, Journal of Central China Normal University, 52, 122-129, https://doi.org/10.19603/j.cnki.1000-1190.2018.01.020, 2018. (in Chinese with English Abstract).

Yu, B.: Using CLIGEN to generate RUSLE climate inputs, Trans. ASAE, 45, 993-1001, https://doi.org/10.13031/2013.9952, 2002.

Yu, B.: An assessment of uncalibrated CLIGEN in Australia, Agr. Forest Meteorol., 119, 131-148, https://doi.org/10.1016/S0168-1923(03)00141-2, 2003.

Yu, R., and Zhou, T.: Seasonality and three-dimensional structure of interdecadal change in the East Asian Monsoon, J. Climate, 20, 5344-5355, https://doi.org/10.1175/2007JCLI1559.1, 2007.

Zhang, X. C., and Garbrecht, J. D.: Evaluation of CLIGEN precipitation parameters and their implication on WEPP runoff and erosion prediction. Trans. ASAE, 46, 311-320, https://doi.org/10.13031/2013.12982, 2003.

Yin, S., and Chen, D.: Weather generators, In Oxford Research Encyclopedia of Climate Science, Oxford University Press, https://doi.org/10.1093/acrefore/9780190228620.013.768, 2020.

Yin, S., Xue, X., Yue, T., Xie Y., Gao G: Spatiotemporal distribution and return period of rainfall erosivity in China. Transactions of the Chinese Society of Agricultural Engineering, 35, 105-113, https://doi.org/10.11975/j.issn.1002-6819.2019.09.013, 2019. (in Chinese with English abstract)

Zhao, S.: A new scheme for comprehensive physical regionalization in China. Acta Geographica Sinica, 38, 1-10, 1983. (in Chinese with English abstract).

~~Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams: Large area hydrologic modeling and assessment Part 1: Model development. Journal of American Water Resources Association, 34(1): 73-89, 1998.~~

~~Carrera-Hernández, J.J. and Gaskin, S.J: Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. Journal of Hydrology, 336(3-4): 231-249, 2007.~~

~~Chen J: Applicability Assessment and Improvement of CLIGEN for the Loess Plateau of China. Thesis for Master Degree, Northwest A&F University, Xi'an. (in Chinese with English abstract), 2008.~~

~~Cheng, Y., T. Ao, X. Li, and B. Wu: Runoff simulation by SWAT model based on parameters transfer method in ungauged catchments of middle reaches of Jialing River. Transactions of the Chinese Society of Agricultural Engineering, 32(13): 81-86, 2016.~~

~~Cowpertwait, P. S. P., P. E. O'Connell, A. V. Metcalfe, and J. A. Mawdsley: Stochastic point process modelling of rainfall. II. Regionalisation and disaggregation. Journal of Hydrology, 175(1-4): 47-65, 1996.~~

~~Flanagan, D.C. and Livingston, S.J.: WEPP user summary. NSERL report, v. 11, 1995.~~

~~Flanagan, D. C., C. R. Meyer, B. Yu, and D. L. Scheele: Evaluation and enhancement of the CLIGEN weather generator: Soil Erosion Research for the 21st Century: Proc. International Symposium, 107-110, 2001.~~

~~Flanagan, D. C., J. Trotochaud, C. W. Wallace, and B. A. Engel: Tool for obtaining projected future climate inputs for the WEPP and SWAT models: 21st Century Watershed Technology Conference and Workshop Improving Water Quality and the Environment Conference Proceedings, 3-6 November 2014, University of Waikato, New Zealand, 1-9, 2014.~~

~~Foster, G. P.: Science documentation: Revised universal soil loss equation, Version 2 (RUSLE 2). USDA-Agricultural Research Service, Washington, DC, 2005.~~

~~Haberlandt, U.: Stochastic rainfall synthesis using regionalized model parameters. Journal of hydrologic engineering, 3(3): 160-168, 1998.~~

~~Hengl, T., Heuvelink, G.B. and Rossiter, D.G.: About regression-kriging: From equations to case studies. Computers & geosciences, 33(10): 1301-1315, 2007.~~

~~Hutchinson, M. F.: Interpolating mean rainfall using thin plate smoothing splines. International journal of geographical information systems, 9(4): 385-403, 1995.~~

~~Katz, R. W., & M. B. Parlange: Mixtures of stochastic processes: application to statistical downscaling. Climate Research, 7(2): 185-193, 1996.~~

~~Li, L., R. Shi, L. Zhang, J. Zhang, and W. Gao: The data fusion of aerosol optical thickness using universal kriging and stepwise regression in East China. Remote Sensing and Modeling of Ecosystems for Sustainability XI, 9221:922112, 2014.~~

~~Liao, Y., Q. Zhang, and D. Chen: Precipitation simulation in China with a weather generator. Acta Geographica Sinica, 59(5): 698-707, 2004. (in Chinese with English abstract).~~

~~Ly, S., Charles, C. and Degré, A: Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. Biotechnologie, agronomie, société et environnement, 17(2): 392-406, 2013.~~

~~Maraun, Douglas, & Martin Widmann: Statistical downscaling and bias correction for climate research. Cambridge University Press, 2018.~~

~~Maraun, D., F. Wetterhall, A. M. Ireson, R. E. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. W. Rust, T. Sauter, and M. Themeßl: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Reviews of Geophysics, 48(3), 2010.~~

~~McKenzie, N. J., and P. J. Ryan: Spatial prediction of soil properties using environmental correlation. Geoderma, 89(1-2): 67-94, 1999.~~

~~Nicks, A. D., L. J. Lane, and G. A. Gander: Chapter 2. Weather generator. Pp. 2.1-2.22: US Department of Agriculture (USDA) Water Erosion Prediction Project: Technical Documentation, National Soil Erosion Research Laboratory (NSERL) Report, 1995.~~

~~Oliver, M. A., and R. Webster: A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena, 113: 56-69, 2014.~~

~~Scheele, D.L. and Hall, D.E.: Corrections and improvements to the CLIGEN climate database. Moscow, Idaho: USDA Forest Service, Rocky Mountain Research Station, 2000.~~

~~Semenov, M. A., and R. J. Brooks: Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. Climate Research, 11(2): 137-148, 1999.~~

~~Verworn, A. and Haberlandt, U.: Spatial interpolation of hourly rainfall – effect of additional information, variogram inference and storm properties. Hydrology and Earth System Sciences, 15(2): 569-584, 2011.~~

~~Wackernagel, H. Multivariate geostatistics: an introduction with applications, Springer Science & Business Media, 2013.~~

~~Wallis, T. W. R., and J. F. Griffiths: An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC). Agricultural and Forest Meteorology, 73(1-2): 115-133, 1995.Wang, J. A., & Zuo, W: Geographic atlas of China. Beijing, China: China Atlas Press, 2009.~~

~~Wilks, D. S., and R. L. Wilby: The weather generation game: a review of stochastic weather models. Progress in physical geography, 23(3): 329-357, 1999.~~

~~Williams, J. R., A. D. Nicks, and J. G. Arnold: Simulator for water resources in rural basins. Journal of Hydraulic Engineering, 111(6): 970-986, 1985.~~

~~Williams, J. R., C. A. Jones, and P. Dyke: A modeling approach to determining the relationship between erosion and soil productivity. Transactions of the ASAE, 27(1): 129-0144, 1984.~~

~~Xu, X., Xu, Y., Sun, Q., Xie, Ting, and Zhang, H: Comparison study on meteorological spatial interpolation approaches in Kangdian region of China. Journal of Central China Normal University, 52(1): 122-129, 2018. (in Chinese with English Abstract).~~

~~Yu, B.: Using CLIGEN to generate RUSLE climate inputs, Transactions of the ASAE, 45(4): 993-1001, 2002.~~

~~Yu, B.: An assessment of uncalibrated CLIGEN in Australia: Agricultural and forest meteorology. 119(3-4):131-148, 2003.~~

~~Yu, R., and T. Zhou: Seasonality and Three-Dimensional Structure of Interdecadal Change in the East Asian Monsoon. Journal of Climate, 20(21): 5344-5355, 2007.~~

~~Zhang, X. C., and J. D. Garbrecht: Evaluation of CLIGEN precipitation parameters and their implication on WEPP runoff and erosion prediction. Transactions of the ASAE, 46(2): 311-320, 2003.~~

~~Yin, S., and Chen, D: Weather generators. In Oxford Research Encyclopedia of Climate Science. Oxford University Press. doi: https://doi.org/10.1093/acrefore/9780190228620.013.768, 2020.~~

~~Yin, S., Xue, X., Yue, T., Xie Y., Gao G: Spatiotemporal distribution and return period of rainfall erosivity in China. Transactions of the Chinese Society of Agricultural Engineering, 35(9): 105-113, 2019.~~

~~Zhao, S: A new scheme for comprehensive physical regionalization in China. Acta Geographica Sinica, 38(1): 1-10, 1983. (in Chinese with English abstract).~~

~~Wang, J. A., & Zuo, W: Geographic atlas of China. Beijing, China: China Atlas Press, 2009.~~

Table 1: Summary of CLIGEN input parameters and the data used for the calculation of parameters.

Inputs	Parameter description	Unit[1]~~s~~	Number of parameters	Data used
TMAX AV	Average of daily maximum temperature	°F~~°F~~	Monthly, 12 in total	Daily temperature
SD TMAX	Standard deviation of daily maximum temperature	°F~~°F~~	Monthly, 12 in total	Daily temperature
TMIN AV	Average of daily minimum temperature	°F~~°F~~	Monthly, 12 in total	Daily temperature
SD TMIN	Standard deviation of daily minimum temperature	°F~~°F~~	Monthly, 12 in total	Daily temperature
SOL.RAD	Average of daily solar radiation	Langley	Monthly, 12 in total	Daily solar radiation
SD SOL	Standard deviation of daily solar radiation	Langley	Monthly, 12 in total	Daily solar radiation
MEAN P	Mean precipitation on rainy days	inch	Monthly, 12 in total	Daily precipitation
S DEV P	Standard deviation of precipitation on rainy days	inch	Monthly, 12 in total	Daily precipitation
SKEW P	The skewness coefficient of precipitation on rainy days	~~inch~~	Monthly, 12 in total	Daily precipitation
P(W/D)	The probability ~~to~~ of a wet day ~~from~~ following a dry day		Monthly, 12 in total	Daily precipitation
P(W/W)	The probability ~~to~~ of a wet day ~~from~~ following a wet day		Monthly, 12 in total	Daily precipitation
MX.5P	Maximum rainfall intensity per 30 min (0.5 hour) of a month	inch/h	Monthly, 12 in total	Hourly precipitation
TimePk[2]	Relative time to the peak rainfall intensity		Cumulative frequency, 12 in total	Hourly precipitation

[1]CLIGEN input parameter values are required to have US customary unit.

[2]The 12th parameter of TimePk for all stations is equal to 1.

Table 2: Data lengths for daily temperature, ~~daily solar ration,~~ daily and hourly precipitation and daily solar ration ~~from~~for stations used in this study.

Data length (years)	Daily Temperature (1951-2014)	Daily rainfall (1951-2015)	Hourly rainfall (1951-2012)	Daily solar radiation (1957-2017)
<=10	19	16	215	5
10~20	17	19	34	9
20~30	20	20	94	44
30~50	269	240	1302	16
>50	2080	2110	760	56
Sum	2405	2405	2405	130

Table 3: Comparison of the accuracy of OK and UK using the leave-one-out cross-validation.

CLIGEN inputs	Observations		~~RMSE[2]~~RMSE[3]		~~NSE[3]~~NSE[4]		PBIAS[4] (%)	
	AV[1]	S DEV[2]	OK	UK	OK	UK	OK	UK
TMAX AV (°F)	67.54	18.02	2.94	1.34	-	-	-	-
SD TMAX (°F)	7.58	1.91	0.36	0.35	-	-	-	-
TMIN AV (°F)	48.91	19.84	2.67	1.58	-	-	-	-
SD TMIN (°F)	6.05	1.94	0.45	0.46	-	-	-	-
SOL.RAD (Langley)	347.46	116.18	30.59	27.11	0.93	0.95	0.14	0.24
SD SOL (Langley)	138.70	41.33	14.34	15.14	0.88	0.87	-0.05	0.97
MEAN P (inch)	0.26	0.16	0.03	0.02	0.97	0.98	-0.02	0.07
S DEV P (inch)	0.40	0.27	0.05	0.05	0.96	0.97	-0.06	0.01
SKEW P	3.12	1.01	0.73	0.47	0.48	0.78	0.08	0.09
P(W/D)	0.23	0.12	0.03	0.02	-	-	-	-
P(W/W)	0.53	0.15	0.04	0.03	-	-	-	-
MX.5P (inch/h)	0.93	0.64	0.14	0.14	0.95	0.95	-0.05	0.04
TimePk	0.58	0.32	0.01	0.01	-	-	-	-

[1]Overall average (AV) and [2]standard deviation (S DEV) for all months and station~~ite~~s, and the unit is identical with parameters;

~~[2]The~~ [3]The unit of RMSE is identical with the unit of each group of parameters;

~~[3]NSE~~ [4]NSE and PBIAS were only calculated for parameters in the ratio scale with true zero.

Table 4: Comparison of CLIGEN generated daily temperature and solar radiation based on observ~~at~~ed~~ion~~ input parameters and UK-~~-~~interpolat~~ed~~~~ion~~ ~~parameters~~ones ~~simulation~~.

Estimation indicators ~~Statistics~~	Daily maximum temperature (°C)			Daily maximum temperature (°C)				Daily solar radiation (Langley/day ~~Ly~~)		
	AV[1]	S DEV[2]	SKEW[3]	AV	S DEV	SKEW		AV	S DEV	SKEW
RMSE ~~(°F or Ly)~~~~NSE~~	0.68~~0.98~~	0.25~~0.99~~	0.03~~0.95~~	0.79~~0.99~~	0.35~~0.98~~	0.04~~0.94~~		18.00~~0.99~~	7.24~~0.98~~	0.07~~0.94~~
~~PBIAS~~	~~-0.1~~	~~0.05~~	~~-0.33~~	~~0.01~~	~~0.05~~	~~-0.23~~		~~0.01~~	~~0.05~~	~~-0.23~~
NSE~~RMSE~~	-~~0.68~~	-~~0.25~~	-~~0.03~~	-~~0.79~~	-~~0.35~~	-~~0.04~~		0.87~~0.79~~	0.87~~0.35~~	0.56~~0.04~~
PBIAS (%)	-	-	-	-	-	-		0.39	0.39	-0.14
~~RSR~~	~~0.14~~	~~0.1~~	~~0.22~~	~~0.12~~	~~0.14~~	~~0.25~~		~~0.12~~	~~0.14~~	~~0.25~~
|AE|[4]	(%)[4]	(%)	(%)	(%)	(%)	(%)	|RE|~~[5]~~	(%)[5]~~[5]~~	(%)	(%)
< 1°C~~°C~~	93.7	99	100	86.2	97.5	100	< 10%	9~~9.2~~3.3	91.7~~99.2~~	60.8~~99.2~~
< 2°C~~°C~~	98.5	99.8	100	97.4	99.6	100	< 20%	99.2~~100~~	~~100~~99.2	~~100~~83.3
< 5°C~~°C~~	99.8	100	100	99.9	100	100	< 50%	100	100	93.3
MAE (°C)~~(°C)~~	0.51	0.21	0.02	0.34	0.14	0.02	MARE (%)	3.81~~0.08~~	4.00~~0.05~~	16.75~~0.09~~

[1]The average (AV), [2]the standard deviation (S DEV), and [3]the skewness coefficient (SKEW) of daily maximum/minimum temperature and solar radiation simulated by CLIGEN.

[4]Percent of stations with |AE| in a range.

[5]Percent of stations with |RE| in a range.

Table 5: Comparison of CLIGEN-generated daily rainfall and annual~~yearly~~ rainy days based on observed~~ation~~ input parameters and UK-~~-~~interpolated~~ion~~ ones~~parameters simulation~~.

Estimation indicators	Daily precipitation (mm)			Annual rainy days	Storm duration (h)			I_{30} (mm/h)		
	AV[1]	S DEV[2]	SKEW[3]	AV	AV	S DEV	SKEW	AV	S DEV	SKEW
RMSE~~NSE~~	0.36~~0.98~~	0.71~~0.97~~	0.63~~0.48~~	7.62~~0.97~~	0.21~~0.92~~	0.17~~0.87~~	0.23~~0.26~~	0.28~~0.99~~	0.52~~0.98~~	0.24~~0.66~~
~~PBIAS~~	~~-0.06~~	~~0.27~~	~~0.94~~	~~-0.01~~	~~0.28~~	~~0.73~~	~~0.13~~	~~-0.34~~	~~-0.2~~	~~-0.15~~
NSE~~RMSE~~	0.98~~0.36~~	0.97~~0.71~~	0.48~~0.63~~	0.97~~7.62~~	0.92~~0.21~~	0.87~~0.17~~	0.26~~0.23~~	0.99~~0.28~~	0.98~~0.52~~	0.66~~0.24~~
PBIAS	-0.06	0.27	0.94	-0.01	0.28	0.73	0.13	-0.34	-0.2	-0.15
~~RSR~~	~~0.15~~	~~0.16~~	~~0.72~~	~~0.18~~	~~0.28~~	~~0.36~~	~~0.86~~	~~0.11~~	~~0.12~~	~~0.58~~
|RE|	(%)[4]	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
< 10%	94.1	91.4	61.2	92.9	94.7	90.8	74.1	97.7	96.7	88.6
< 20%	98.6	98.6	87.4	98.4	98.8	97.9	93.5	99.7	99.4	98.3
< 50%	100	99.9	99.6	99.7	99.9	99.8	99.7	100	99.9	100
MARE (%)	3.72	4.56	10.07	4.09	3.47	4.61	7.71	2.36	3.07	5.08

[1]The average (AV), [2]the standard deviation (S DEV), and [3]the skewness coefficient (SKEW) of daily precipitation, annual rainy days, storm duration and I_{30}~~maximum/minimum temperature and solar radiation simulated~~ by CLIGEN.

[4]Percent of stations with |RE| in a range.

Table ~~7~~6: Station density and simulation quality of CLIGEN for~~of~~ three Chinese physical-geographical regions.

	Eastern Monsoon Area	Northwest Arid Area	Qinghai-Tibet Plateau
Temperature and precipitation			
No. of stations	2044	233	128
Density (n/10^4·km^2)	4.57	0.97	0.50
MAE of ~~Min~~ Max ~~Temperature~~ Temperature (°C)~~(°C)~~	0.44	0.90	0.93
MAE of ~~Max~~ Min Temperature (°C) ~~(°C)~~	0.30	0.42	0.82
MARE of Daily ~~precipitation~~ Precipitation (%)	3.13	6.92	7.25
MARE of Duration (%)	2.95	5.93	7.31
MARE of I_{30} (%)	2.00	4.50	4.11
Solar radiation			
No. of stations	92	26	12
Density (n/10^4·km^2)	0.21	0.11	0.05
MARE of Daily ~~solar~~ Solar ~~radiation~~ Radiation (%)	3.92~~0.08~~	~~0.07~~2.87	5.14~~0.13~~

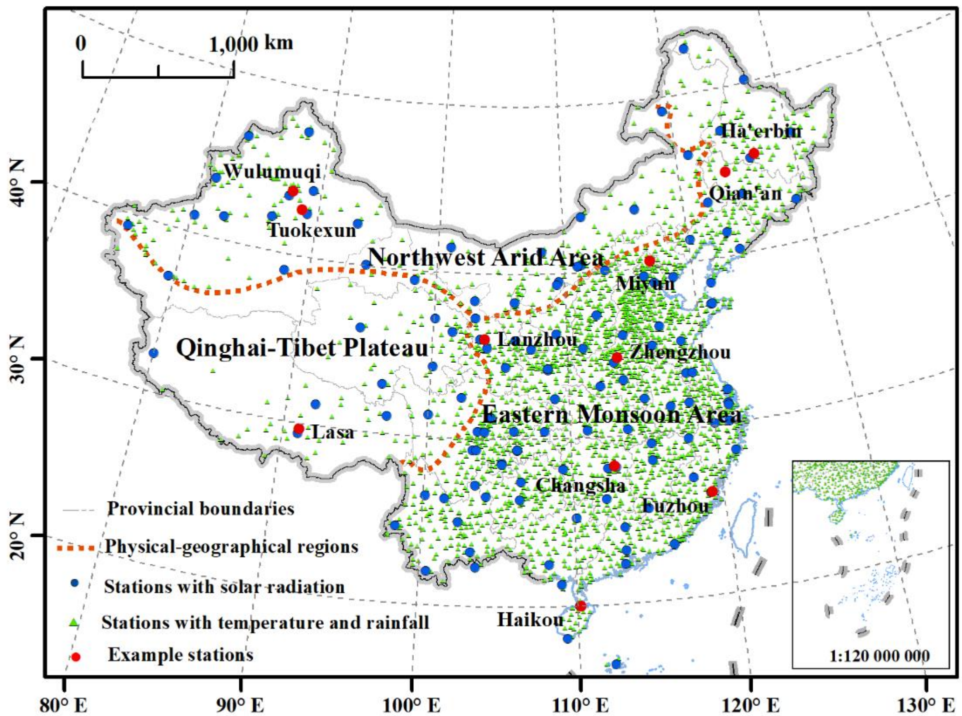


Figure 1: Locations of meteorological stations used in this study.

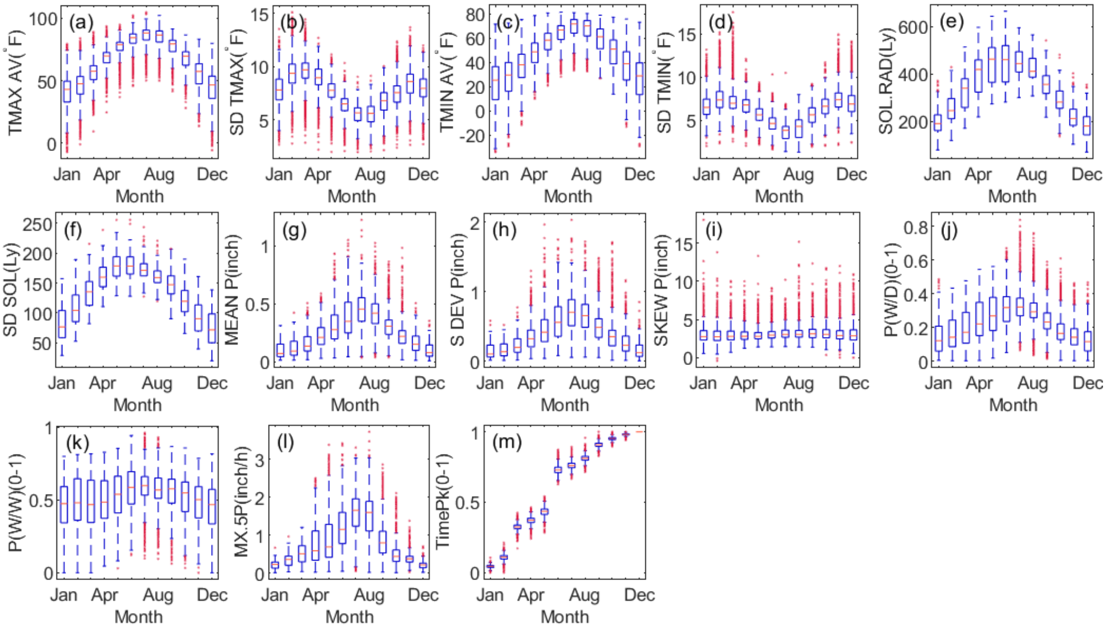


Figure 2: Boxplot of CLIGEN temperature, solar radiation, and precipitation parameters obtained from observations in mainland China.

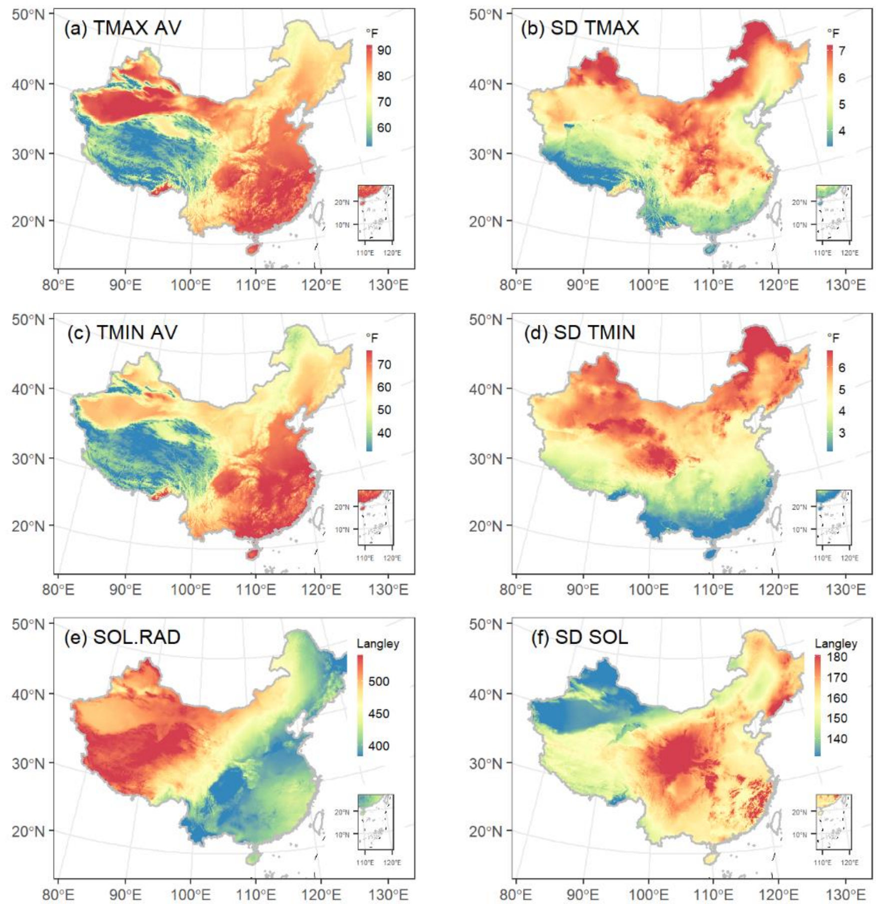


Figure 3: Spatial distribution of CLIGEN temperature-related parameters of mainland China in August. All parameters were regionalized using *universal Kriging*.

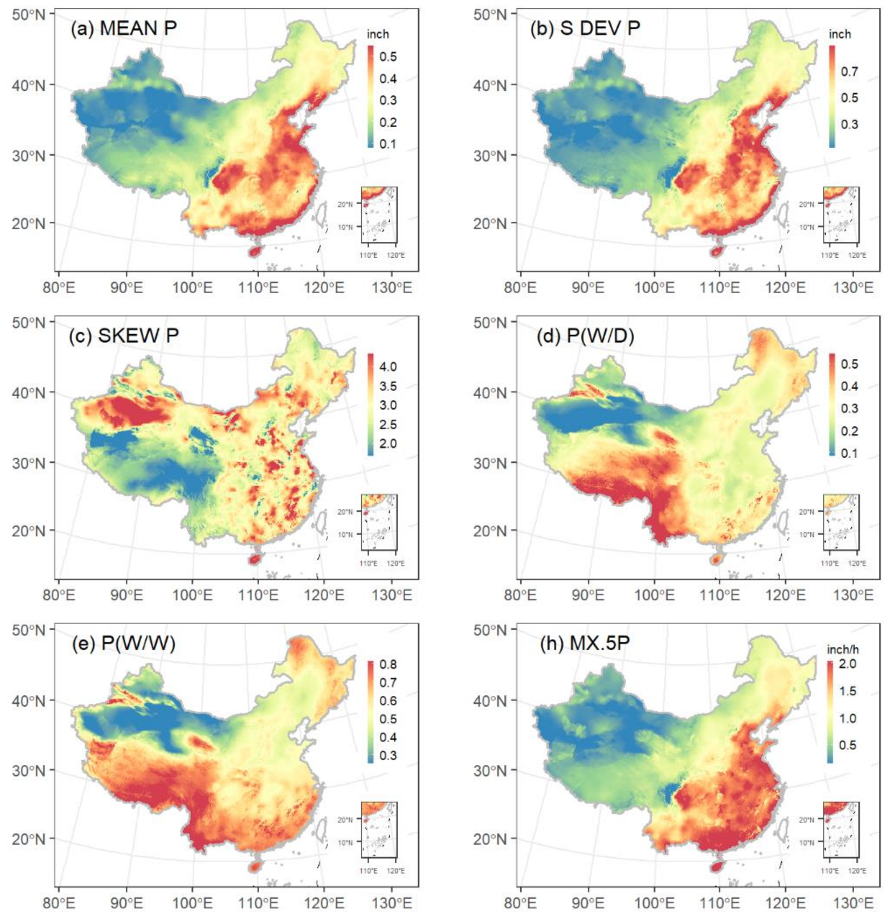


Figure 4: Spatial distribution of CLIGEN precipitation related parameters of mainland China in August. All parameters were regionalized using *universal Kriging*.

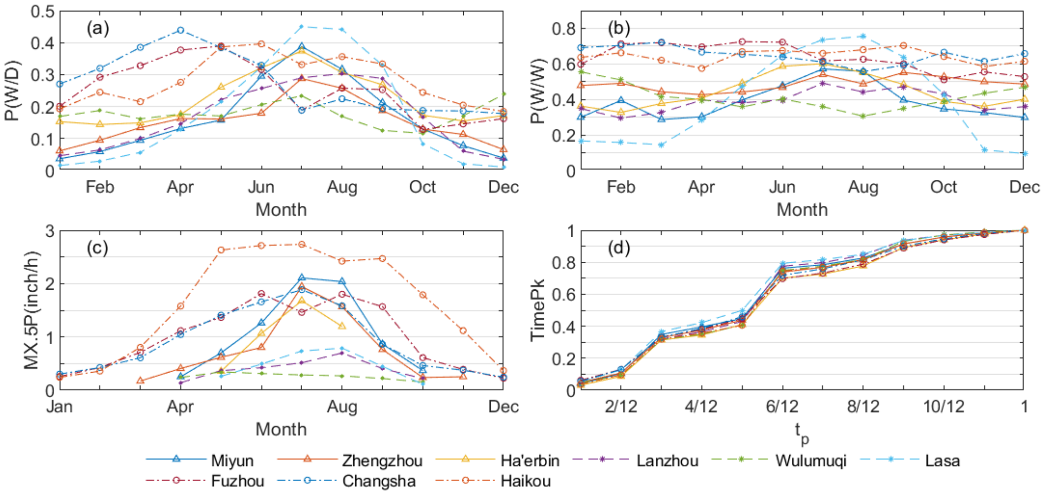
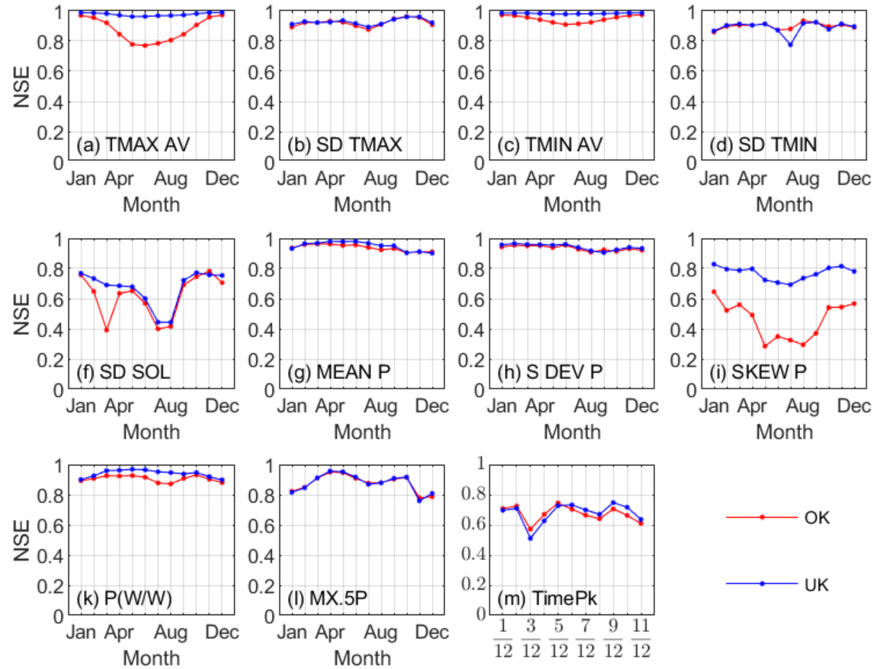


Figure 5: **P(W/D), P(W/W), MX.5P and TimePk of nine stations determined by observed daily precipitation.**

730


(a) TMAX AV
(b) SD TMAX
(c) TMIN AV
(d) SD TMIN
(e) SOL.RAD
(f) SD SOL
(g) MEAN P
(h) S DEV P
(i) SKEW P
(j) P(W/D)
(k) P(W/W)
(l) MX.5P
(m) TimePk
NSE
Month
OK
UK

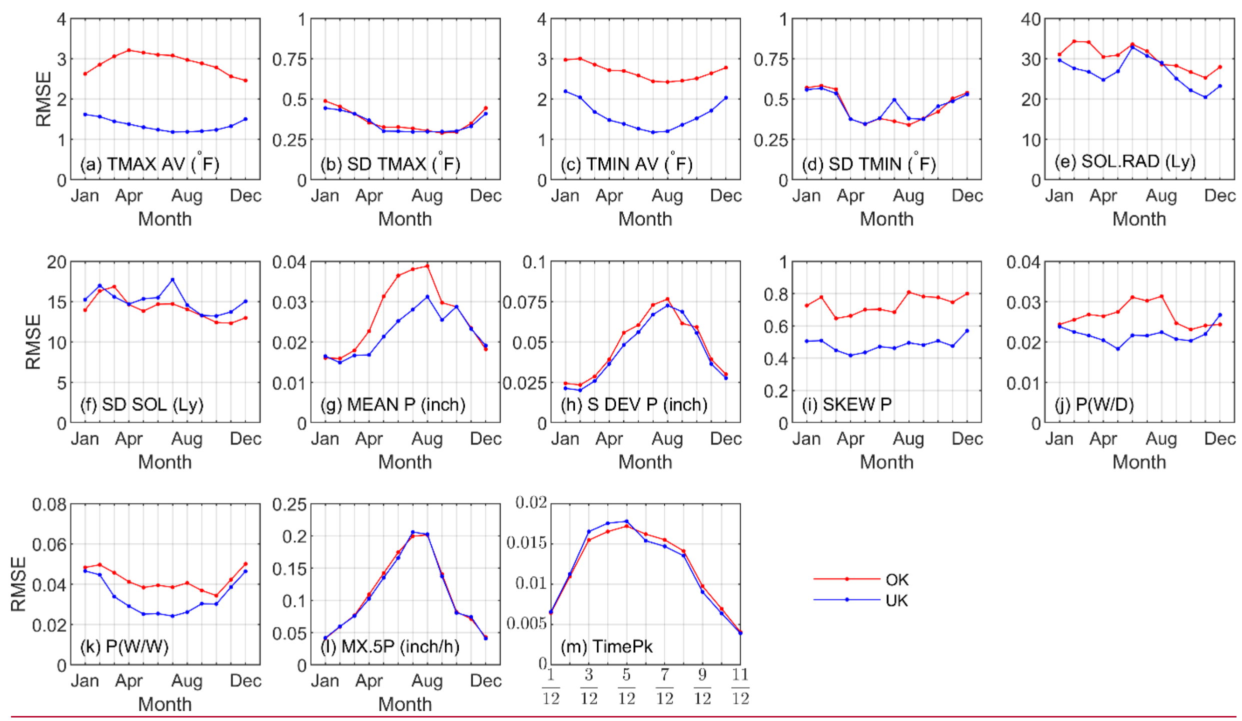

(a) TMAX AV (°F)
(b) SD TMAX (°F)
(c) TMIN AV (°F)
(d) SD TMIN (°F)
(e) SOL.RAD (Ly)
(f) SD SOL (Ly)
(g) MEAN P (inch)
(h) S DEV P (inch)
(i) SKEW P
(j) P(W/D)
(k) P(W/W)
(l) MX.5P (inch/h)
(m) TimePk
RMSE
Month
OK
UK

Figure 6: Comparison of the interpolation quality in terms of the *root mean square error* (RMSE) ~~Nash-Stucliffe coefficient of efficiency (NSE)~~ using *ordinary Kriging* (OK) and *universal Kriging* (UK) for temperature, solar radiation, and precipitation parameters.

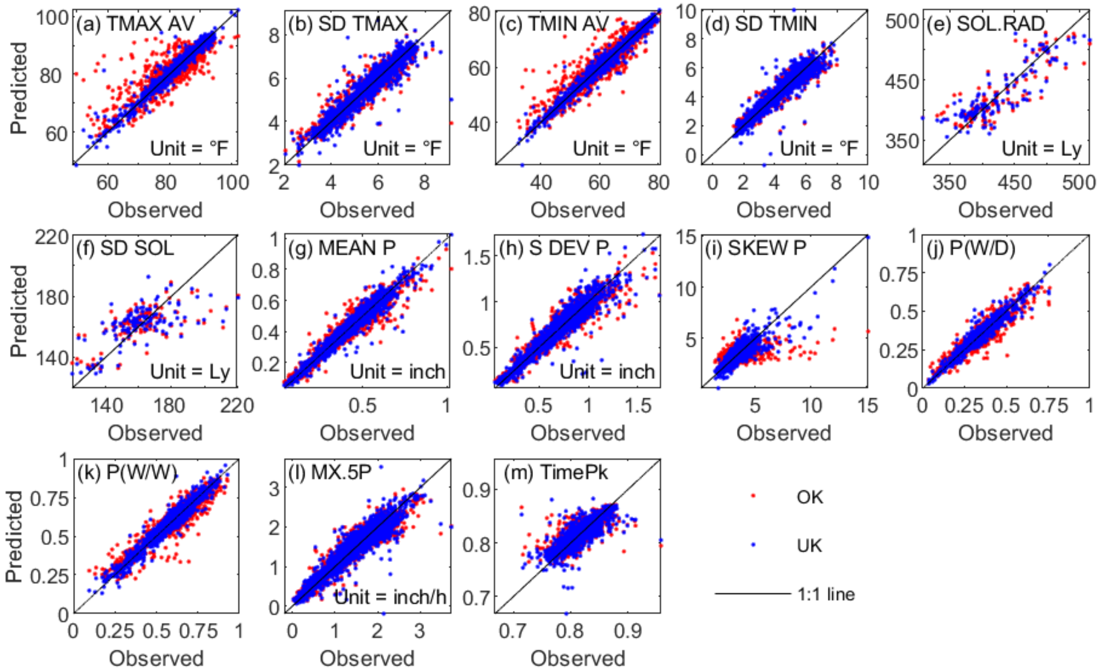


Figure 7: Comparison of the interpolation quality using *ordinary Kriging* (OK) and *universal Kriging* (UK) for CLIGEN temperature, solar radiation, and precipitation parameters in ~~.~~August, and the 8th parameters of TimePk.

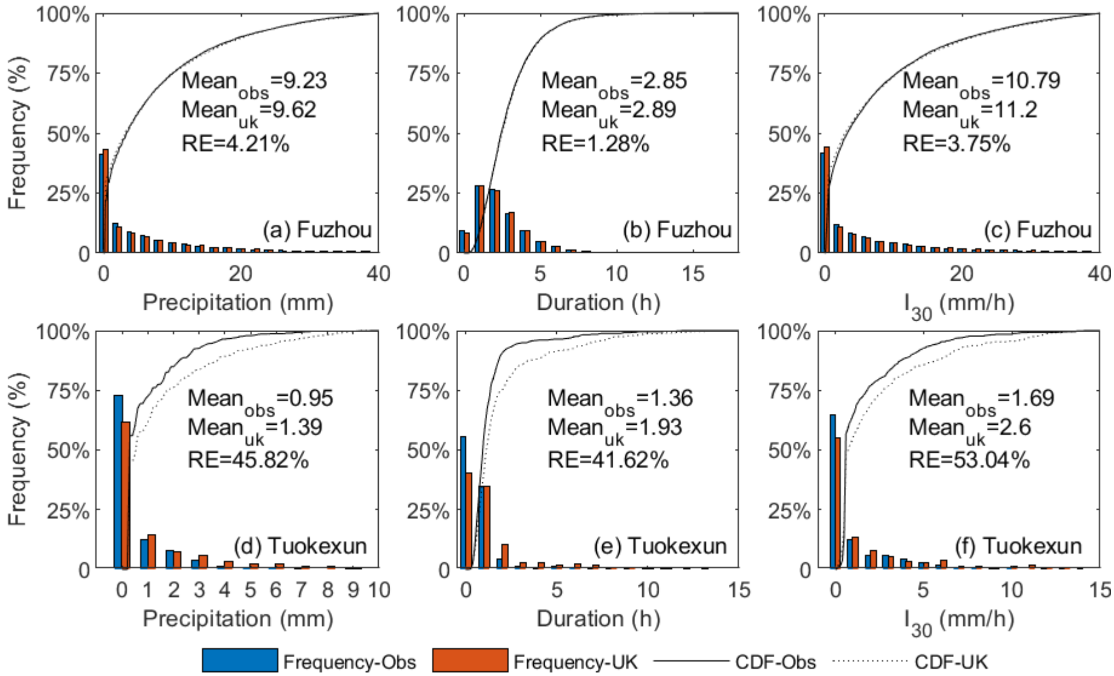
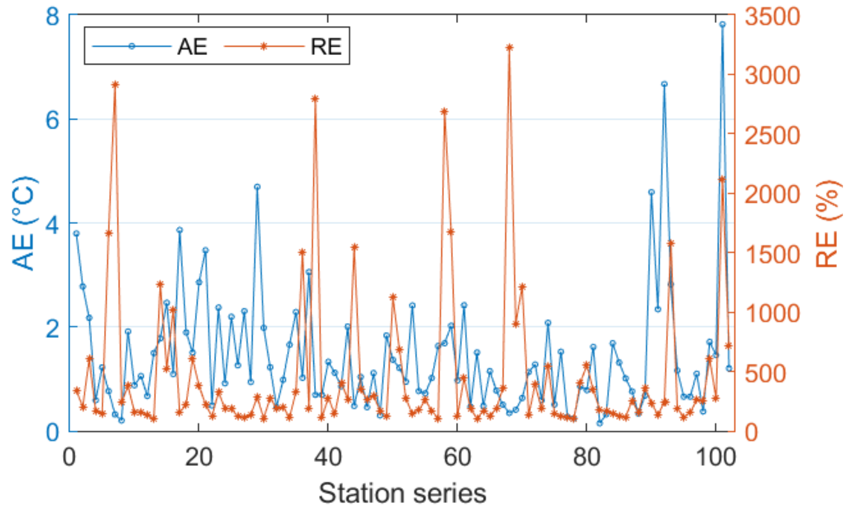
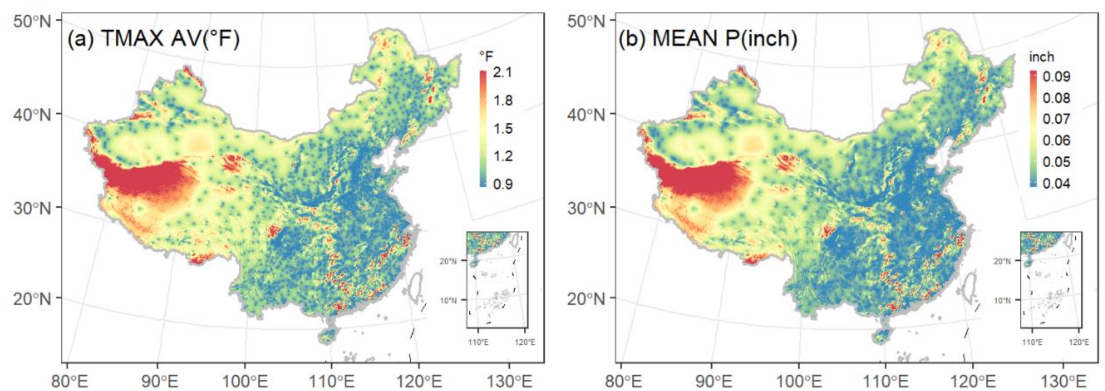


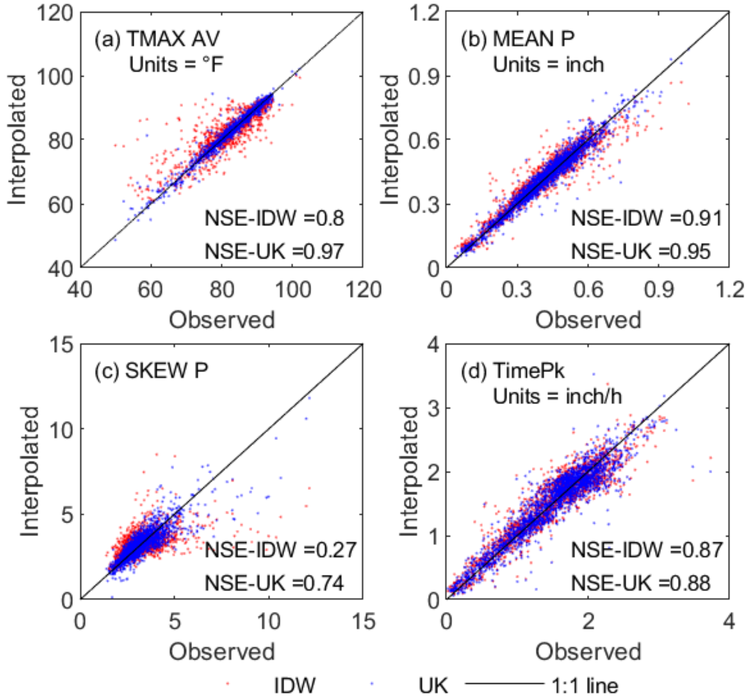
Figure 8. Frequency distribution of daily precipitation, duration, and maximum 30-min intensity (I_{30}) generated by CLIGEN using inputs based on observations and interpolation predicted parameters: Fuzhou station (a-c) and Tuokexun station (d-f) as examples.



~~**Figure 9. Comparison of the absolute error (AE, ℃) and relative error (RE, %) of the simulated average of maximum temperature based on observed and UK-interpolated inputs by CLIGEN for 102 stations with extremely large RE.**~~



<u>Figure 9. Spatial distribution of the standard error for interpolation results of TMAX AV (a) and MEAN P (b) using Universal Kriging.</u>


(a) TMAX AV
Units = °F
NSE-IDW =0.8
NSE-UK =0.97
(b) MEAN P
Units = inch
NSE-IDW =0.91
NSE-UK =0.95
(c) SKEW P
NSE-IDW =0.27
NSE-UK =0.74
(d) TimePk
Units = inch/h
NSE-IDW =0.87
NSE-UK =0.88
Interpolated
Observed
IDW
UK
1:1 line

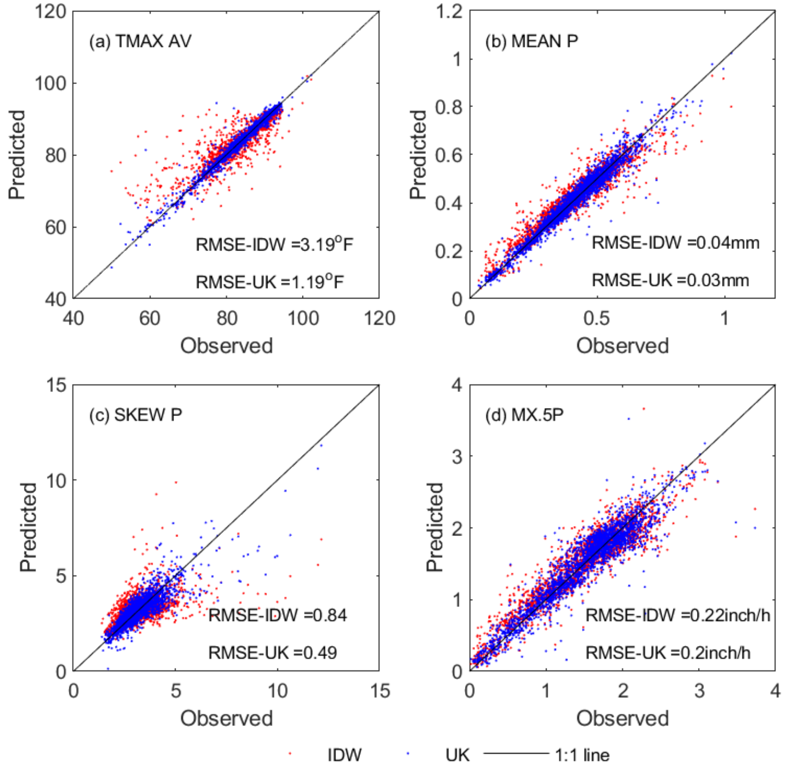


Figure 10: **Comparison of interpolation quality using *universal Kriging* (UK) and the *inverse distance weighted method* (IDW) for CLIGEN temperature and precipitation related parameters for 2405 ~~sites~~ stations in ~~summer (~~August~~)~~.**