https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

An improved method of the Globally Resolved Energy Balance Model by the Bayes network

Zhenxia Liu[1], Zengjie Wang[1], Jian Wang[1], Zhengfang Zhang[1], Dongshuang Li[4,5], Zhaoyuan Yu[1,2,3], Linwang Yuan[1,2,3], and Wen Luo[1,2,3]

[1]School of Geography, Nanjing Normal University, Nanjing, 210023, China
[2]Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China
[3]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, 210023, China
[4]Jiangsu Key Laboratory of Crop Genetics and Physiology/Jiangsu Key Laboratory of Crop Cultivation and Physiology, Agricultural College of Yangzhou University, Yangzhou, China
[5]Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, Yangzhou University, Yangzhou, China,

Correspondence: Wen Luo (09415@njnu.edu.cn)

Abstract. This study introduces an improved method of the Globally Resolved Energy Balance Model(GREB) by the Bayes network. Starting from climate elements relationship included in the GREB model, we reconstruct the model by Bayes network to solve the problem of low model accuracy due to over-reliance on boundary conditions and initial conditions and inability to use observed data for dynamic correction of model parameters. The improved model is applied to the simulation of surface average temperature and atmospheric average temperature based on the $3.75° \times 3.75°$ global data sets by Environmental Prediction (NCEP)/ National Center for Atmospheric Research(NCAR) from 1985 to 2014. The results illustrate that the improved model has higher average accuracy and lower spatial differentiation than the original GREB model. And the improved method provides a strong support for other dynamic model improvements.

1 Introduction

As the global warming progresses, extreme events and meteorological disasters occur frequently(Grant, 2017). Thus, the simulation and prediction of climate have become an important topic in current scientific research(Huang et al., 2019). Climate model is a model that describes the change law of climate system, such as the change law of climate, ocean, atmosphere, ice, etc.(Berrocal et al., 2012), which can be solved by supercomputer and is an important tool for simulating and predicting future climate change(Kay, 2020).

Generally, climate models mainly include two categories, dynamic model and statistical model. Dynamic model can well understand and express the dynamic process of climate by modeling various complex climate processes or interactions, but it still faces two major problems: i) The simulation process overly relies on initial conditions and boundary conditions(Alley et al., 2019; Zhang et al., 2019; Ludescher et al., 2021); ii) The climate model is too complicated, and its internal characteristics

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

cannot be fully expressed(Fan et al., 2021; Zou et al., 2019; Feng et al., 2020). The Globally Resolved Energy Balance Model (GREB) is a simple but representative dynamic model, which is based on energy balance theory(Dommenget and Flöter, 2011). Compared with other dynamic models, the GREB model is a relatively fast tool because it computes about one model year per second on a standard personal computer, which allows conducting sensitivity studies to external forcing within minutes to hours(Dommenget and Flöter, 2011; Dommenget, 2016; Stassen et al., 2019). However, in addition to the two main problems of dynamic models, the GREB model also faces the problem that the model does not respond well to anomalous climate change because the parameters of the GREB model are predetermined and the observed data can hardly be used to dynamically correct the model parameters(Dommenget and Flöter, 2011; Dommenget, 2016). How to solve these problems is an important research topic to improve the GREB model and further extend it to other dynamic models.

On the contrary, statistical model, as another type of climate models, can make good use of historical observation data to dynamically modify the models from data(Feng et al., 2020), and solve the problem that dynamic climate models rely too much on initial and boundary conditions and underutilize full observation data. Therefore, it provides a possible way to solve those defects of the dynamical model by combining that with the statistical model.

Bayes network is a statistical method which combines graph theory and probability (Cai et al., 2013, 2019; Jansen et al., 2003). The method uses graph to express the structure relation of the variables related to the model and has the characteristics of structuring and quantifying the object relation through the causal relation among the parts of the probability computing system(Pearl, 1986), variable logic reasoning and predictive simulation can be realized, and it can use a large amount of historical observation data. As described, it is a possible way to improve the GREB model by the Bayes network.

This paper introduces a method for improving the GREB model by the Bayes network. The aim of method is to solve the problem of low model accuracy due to over-reliance on boundary conditions and initial conditions and inability to use observed data for dynamic correction of model parameters. The following section presents the original GREB model and the method. Section 3 presents the study case and data sets to test the new improved model. Finally, we give a discussion and conclusion of the results.

2 Methods

The improved method is developed according to the following procedure. First, the climate elements contained in different climate processes are selected as nodes by the GREB model. And the structural relationships between different nodes are determined to establish an abstract model of the components and structural relationships of climate processes. Then, the improved climate model is reconstructed based on the structural relationships of the climate processes through Bayes network to achieve the simulation of the target climate elements.

2.1 Climate elements structural relationship

Based on the law of conservation of energy, the GREB model can simulate the main characteristics and global average states of global warming, including seven climatic processes (solar radiation, thermal radiation, hydrological cycle, sensible heat

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

and atmospheric, atmospheric circulation, sea ice and deep ocean) and four main climatic average states (surface average temperature, atmospheric average temperature, average temperature of ocean and atmospheric average humidity). In order to show the inherent mechanism of the evolution of climate process, the surface average temperature and the atmospheric average temperature are selected as the output of the GREB model. The simulation of these two average state variables includes most of the climate processes of the GREB and can reflect the complex coupling process and climate change characteristics of the GREB model.

The scene of simulating the surface average temperature includes solar radiation, thermal radiation, sensible heat and atmospheric, and deep ocean. The main heat source of the surface temperature is solar radiation, some of which is absorbed by the surface temperature, the other part is reflected by the surface temperature, and part of the heat on the surface temperature is transferred in the atmosphere, and some of it is transferred to the ocean below the surface. Each climate elements in this scene can be expressed by a highly simplified equation, which follows the surface average temperature trend equation as follows:

$$\gamma_{surf}\frac{dT_{surf}}{dt} = F_{solor} + F_{thermal} + F_{sense} + F_{latent} + F_{ocean} \tag{1}$$

Where T_{surf} is surface temperature; γ_{surf} is surface specific heat capacity; F_{solor} is the net solar radiation reaching the surface; $F_{thermal}$ is thermal radiation in the atmosphere; F_{sensel} is turbulent heat exchange in the atmosphere; F_{latent} is heat exchange produced by water during phase transitiont; F_{ocean} is turbulent heat exchange in the ocean.

In the scene of simulating the atmospheric average temperature by the GREB model, atmospheric temperature is not only related to the thermal radiation reflected from the surface, but also related to the sensible heat in atmospheric transport and latent heat exchange of water vapor in the atmosphere. Each climate elements in this scene can be expressed by a highly simplified equation, which follows the atmospheric average temperature trend equation as follows:

$$\gamma_{atmos}\frac{dT_{atmos}}{dt} = -F_{sense} + Fa_{thermal} + Q_{latent} + \gamma_{atmos}\left(k\cdot\nabla^2 T_{atmos} - \vec{u}\cdot\nabla T_{atmos}\right) \tag{2}$$

Where T_{atmos} is atmospheric temperature; γ_{atmos} is atmospheric specific heat capacity; F_{sensel} is turbulent heat exchange in the atmosphere; $Fa_{thermal}$ is net thermal radiation of the atmosphere; Q_{latent} is latent heat released by condensation of atmospheric water vapor.

For different climate average state scenes, the climate processes and relationship structures are different. Therefore, the selection of nodes in each climate average state scenes will also be different. Not only the selection of appropriate variables as nodes is very important, but also the number of nodes will directly affect the simulation of the final climate average state. In order to simplify the complex climate evolution process and facilitate calculation, 4-6 climate elements are selected as key nodes in each climate average state scenes, and the attribute states of climate average state elements in each scene are simulated by these nodes.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

Through the trend equation (Eq.1) in the scene of surface average temperature, the relation equation of climate elements can be simplified:

$$\begin{cases} T_s = f(S,O,W,C) \\ O = f(S) \\ W = f(S) \end{cases} \tag{3}$$

Where T_s is surface temperature; S is solar radiation; O is ocean temperature; W is water vapor content; C is cloud cover. That is, surface temperature, solar radiation, cloud cover and water vapor content can be selected as the key nodes of the surface average temperature scene.

Through the trend equation (Eq.2) in the scene of atmospheric average temperature, the relation equation of climate elements can be simplified:

$$\begin{cases} T_a = f(V,W,C) \\ W = f(V) \end{cases} \tag{4}$$

Where T_a is atmospheric temperature; V is wind speed; C is cloud cover; W is water vapor content. That is, atmospheric temperature, wind speed, cloud cover and water vapor content can be selected as the key nodes of the atmospheric average temperature scene.

According to Eq.3, in the scene of surface average temperature, surface temperature (T_s) is controlled by solar radiation (S), cloud cover (C), water vapor (W) and ocean temperature (O). Ocean temperature (O) and water vapor (W) are controlled by solar radiation (S). For the above relationship, the bayes network structure in the surface average temperature scene($Fig1.(a)$) can be constructed. According to Eq.4, in the scene of atmospheric average temperature, atmospheric temperature (T_a) is controlled by cloud cover (C), water vapor (W) and wind speed (V). And water vapor (W) is controlled by wind speed (V). For the above relationship, the Bayes network structure in the atmospheric average temperature scene($Fig1.(b)$) can be constructed.

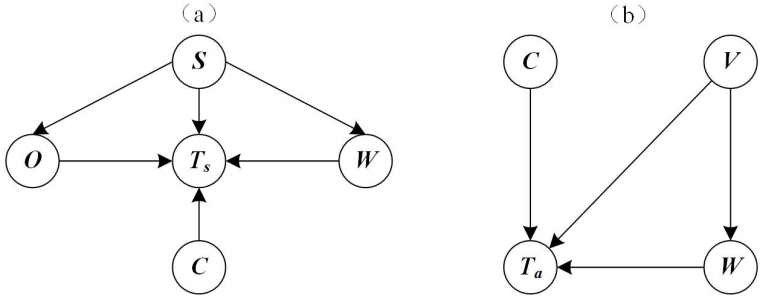

Fig 1. (a) climate elements structural relationship in the surface average temperature scene; (b) climate elements structural relationship in the atmospheric average temperature scene.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

2.2 Bayes network

Bayes network is a probabilistic model that simulates the human reasoning process, which is a combination of graph theory and probability theory, and its network topology is a directed acyclic graph. Where variables are nodes and correlations or causal relationships between variables are directed edges.The dynamic evolution of Bayes network node probabilities is controlled by conditional probabilities, and each node covers a probability distribution table under the joint distribution of the parent nodes, indicating the strength of the relationship between the nodes.(Sahin et al., 2019). When the Bayes network is constructed, given the state of any node, the probability distribution of the states of the remaining nodes can be calculated.

In the Bayes network, the probability of a node can be calculated in the form of probability using prior knowledge and statistical data, namely the Bayes probability (Maher, 2010). Observed sample are defined as: $G = \{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$, where X is event, x is event value or state. When θ is the prior probability of event $X = x$, ζ is prior knowledge, $P(\theta|\zeta)$ is probability density function, then the probability $P(X_{n+1} = x_{n+1}|\theta,\zeta)$ of the $n+1$ event $X_{n+1} = x_{n+1}$ can be obtained from the prior probability density $P(\theta|\zeta)$ and the sample G through the Bayes probability. It can be calculated by total probability formula:

$$\begin{aligned} & P(X_{n+1} = x_{n+1}|\theta,\zeta) \\ = & \int P(X_{n+1} = x_{n+1}|\theta,\mathrm{G},\zeta) P(\theta|\mathrm{G},\zeta)\, d\theta \\ = & \int \theta P(\theta|\mathrm{G},\zeta)\, d\theta \end{aligned} \tag{5}$$

Based on Bayes equation, The posterior probability $P(\theta|\mathrm{G},\zeta)$ is denated as:

$$P(\theta|\mathrm{G},\zeta) = \frac{P(\theta|\zeta)\, P(\mathrm{G}|\theta,\zeta)}{P(\mathrm{G}|\zeta)} \tag{6}$$

Where G is given sample, ζ is priori probability of G,

2.3 Climate evolution process based on Bayes network

In a climatic process composed of several climatic elements, there is an association relationship between climatic elements. These climatic elements are regarded as network nodes, and the association relations between climatic elements are taken as directed edges. The association relationship between nodes is represented by graph model, and the action intensity of association relationship is described quantitatively by conditional probability table. Using the characteristics of Bayes network, the attribute feature state of nodes is inferred by probability. To realize the expression and simulation of the attribute feature state of geographical elements.

A climate process $M_t = \{X(m_1, m_2, \ldots, m_i)|m_{1t}, m_{2t}, \ldots, m_{it}\}$ is composed with i climate elements, $m_1, m_2, \ldots, m_i$, and $X(m_1, m_2, \ldots, m_i)$ is the structural relationship among the elements. Suppose that the climate element m_i has j states, then the sataes set of m_i is $\{W_{m_{i1}}, W_{m_{i2}}, \ldots, W_{m_{ij}}\}$. The climate process is described by a Bayes network $B = (S, X)$. where S is a directed acyclic graph composed of nodes; X is the nodes set of graph, that is climate elements $m_1, m_2, \ldots, m_i$. nodes are connected by directed edges to represent the relationship between climate elements. Each node has an independent

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

conditional probability table, which represents the probability distribution under the joint distribution of its parent nodes. Assume that a climate m_i has one or more parent nodes $m_1, m_2, \ldots, m_e\,(e \leq i-1)$ and states $d_1, d_2, \ldots d_e$, it can be denoted as :$m_1, m_2, \ldots, m_e \rightarrow m_i$. Under the parent node of all possible states, the conditional probability table composed of the set of state probabilities of mdeteci is follow:

$$B_{m_i}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}} = \left\{ (W_{m_{i1}}, P_{W_{m_{i1}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}), (W_{m_{i2}}, P_{W_{m_{i2}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}), \ldots, (W_{m_{ij}}, P_{W_{m_{ij}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}) \right\} \quad (r1 = 1, 2, \ldots, d_1), (r2 = 1, 2, \ldots, d_2), \ldots, (re = 1, 2, \ldots, d_e) \tag{7}$$

Where $B_{m_i}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}$ is a conditional probability table of climate elements m_i; $W_{m_{ij}}$ is jth characteristic state of climate elements m_i;$P_{W_{m_{ij}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}$ is the probability of climate element m_i corresponds to the jth state under the $r1, r2, \ldots re$ characteristic state corresponding to the parent node $m_1, m_2, \ldots, m_e$ expression set. The probability set of climate element m_i at t moment can be denoted as $C_{m_{it}}$:

$$C_{m_{it}} = \left\{ (W_{m_{i1}}, P_{W_{m_{i1}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}), (W_{m_{i2}}, P_{W_{m_{i2}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}), \ldots, (W_{m_{ij}}, P_{W_{m_{ij}}}^{W_{m_{1r1}}, W_{m_{2r2}}, \ldots, W_{m_{ere}}}) \right\} \tag{8}$$

The conditional probability table of each node can be calculated by Eq.5 using training data.

3 Case study

3.1 Data description

In this paper, data produced by National Centers for Environmental Prediction (NCEP)/ National Center for Atmospheric Research(NCAR) is used as the experimental data to evaluate the improved model(IMPM). The data sets include surface temperature(T_s), atmospheric temperature(T_a), solar radiation(S), cloud cover(C), water vapor (W), ocean temperature(O), and wind speed(V) stored as a 3.75° *3.75°(latitude * longitude) grid NC data from 1985 to 2014. In order to facilitate calculation and comparative analysis, all climate data is preprocessed. Firstly, the downloaded climate data is removed from the outliers so that the data are calculated to avoid too large or too small results; secondly, the grid data is resampled and the resampling method is bilinear interpolation. The bilinear interpolation method is used to interpolate the climate data, which not only fills the null values, but also unifies the scale size of the data. Finally, the climate data from 1985 to 2014 were processed as quarterly averages because the changes of climate elements are usually related to seasons.

Based on the theory of climate sensitivity(Annan and Hargreaves, 2006; Dommenget, 2016), climate state data calculated by specific numerical classification of climate data, can be used to replace the specific numerical data to simulate climate change and characterize the long-term trend of climate change and extreme weather conditions. In order to further reduce the amount of calculation and reduce the simulation response time, we use the Natural Breaks method to classify the original data to get

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

the climate average state as the simulation data. Accordingly, the climate elements data is divided into five, seven and nine different classification states to test the improved model and verify the effect of the classification number of climate elements data on the simulation results. Detailed classifications are shown in appendix Table A1, A2, A3.

3.2 State simulation

From Sect.2.1, we choose surface temperature and atmospheric temperature as the simulation objects, and other elements as known objects, and uses the historical data to calculate the conditional probability tables of each nodes through the Bayes network structure with Eq 7. Among them, the training data is the 10-year historical data from 1985 to 1994.

In each climate average state scenes, there are two training methods for the simulated object. The first is to train a conditional probability table using the data in all the grids, and then use the conditional probability table to simulate the states of all grids. The conditional probability table obtained by this training method can reflect the numerical characteristic relationship between climate elements in the whole region. However, it can not show the distribution law of the characteristics of the simulated state in space. The second is to train the data in each grid separately. Because the state grading data in each grids is different, the conditional probability table of the simulated object trained in each grids is also different, and a total of 96 ×48 conditional probability tables are obtained. The conditional probability tables obtained by this training method can accurately reflect the different numerical characteristic relationship between the simulated object and the known object in different regions. However, due to the training of more conditional probability tables, the running time of this training method will be a little longer.

Considering the great differences in the law of climate evolution in different regions, this paper uses the second data training method in state simulation, which first divides the whole world into 96×48 grids, and then uses the data in each grids to train the conditional probability tables of the grid. After the training is completed, the data of the known climate elements will be used to simulate the unknown climate elements from 1995 to 2014. The simulation results are shown in Fig 2,3.

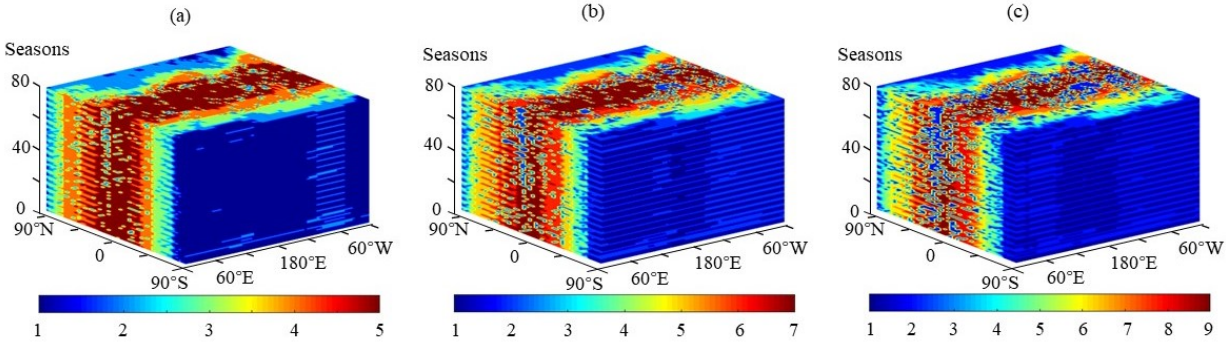

Fig 2. Simulation results of the surface average temperature state. (a) 5 classification; (b) 7 classification; (c) 9 classification.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

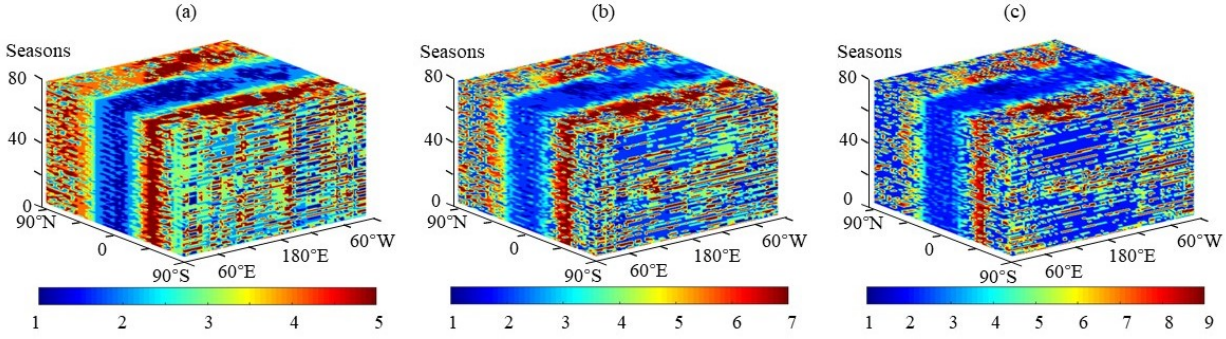

Fig 3. Simulation results of the atmospheric average temperature state. (a) 5 classification; (b) 7 classification; (c) 9 classification.

From Fig 2,it can be clearly observed that the distribution of global surface average temperature has latitude variations. The temperature starts from the equator and decreases with the increase of latitude, so the temperature in the North and South Pole is the lowest. The distribution of surface temperature is basically in line with the real world. Different from the simulation map of surface average temperature, the trend of atmospheric average temperature rises from the equator and increases with the increase of latitude in Fig.7, which is also basically in line with the real world. The tropospheric height of the poles is lower and the tropospheric height of the equator is higher, and which phenomenon leads to the result that temperature of the troposphere at the same height is higher in the poles. Meanwhile, from the simulations in the two scenes, it can be found that with the increase of the state classification, the color of the simulated effect map becomes more complex, which means that the simulated resolution will be higher. It shows that this method can finely simulate the state of surface average temperature and the atmospheric average temperature in the complex environment by increasing the state classification. Given the above, if it takes the actual numerical value, not state as calculating data, it will get a higher resolution, and less training data of each case.

3.3 Comparison with the GREB model

In order to verify the improved method, accuracy (the equal proportion of the state level of the simulated value and the real value) was used as an evaluation index. The state simulation accuracy of each grid is calculated when the number of state classification of climate elements is 5, 7, and 9. And the results are compared with that of the GREB model. The GREB model uses the model code of the GREB in the Monash Simple Climate Model (MSCM) laboratory repository and runs the code in Fortran language.

The all average accuracy of different situations from 1985 to 1994 is shown in Fig 4(IMPM-T: average accuracy of surface temperature by IMPM; GREB-T: average accuracy of surface temperature by GREB; IMPM-A: average accuracy of atmospheric temperature by IMPM; GREB-A: average accuracy of atmospheric temperature by GREB), and the globe accuracy of the surface temperature and the atmospheric temperature are shown in Fig 5, 6. Overall, the comparison results in the two scenes (Fig 4) show that the improved model(IMPM) has a higher simulation accuracy. Since the total number of data remains

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

unchanged, as the number of state classifications increases, the number of training data per state classifications decreases, it results in a decrease in the accuracy of the simulations of the two methods. This implies that the accuracy of the simulation predictions can be stabilized at a high level when there is enough training data in the long-period simulation. It also shows that the improved model can compensate well for the shortcoming of the GREB model to make full use of the observed data. On the other hand, in the improved model, more climate data are involved in the simulation process. It means that more observational data are involved in the model modification, indicating that the improved model can have a more obvious response to climate anomalies. From the analysis above, the comparison results(Fig 4) means that the improved method in this paper is effective.

Meanwhile, it is obvious in the comparison results(Fig 5, 6) that the accuracy of the IMPM is relatively uniform in spatial distribution and has no obvious spatial characteristics. However, the accuracy of the GREB model has obvious characteristics of latitude differentiation. Based on it, the average accuracy variation trend chart along with latitude(Fig 7) is made. The variances of the IMPM in six cases are 0.016(IMPM-T-5), 0.014(IMPM-T-7), 0.014(IMPM-T-7), 0.017(IMPM-A-5), 0.008(IMPM-A-7), 0.004(IMPM-A-9), and the variances of the GREB in six cases are 0.089(GREB-T-5), 0.070(GREB-T-7), 0.060(GREB-T-7), 0.077(GREB-A-5), 0.054(GREB-A-7), 0.036(GREB-A-9). As the chart shown, the fluctuation range of the accuracy of the IMPM is much smaller than that of the GREB along the latitude direction. Especially, the average accuracy of atmospheric temperature is 0 between 30°N and 30° S, and it means that the GREB is not suitable for atmospheric temperature simulation of the region. After considering the relationship of climate elements in the GREB model to build the network, the whole simulation process is based on observation data, so it is not affected by local anomalies, and the accuracy of the improved model is not affected by spatial location. Consequently, the use of the IMPM is not restricted by spatial, and there is no significant difference in the global simulation effect. As described, the IMPM not only has a higher average accuracy, but also has a wider range of application.

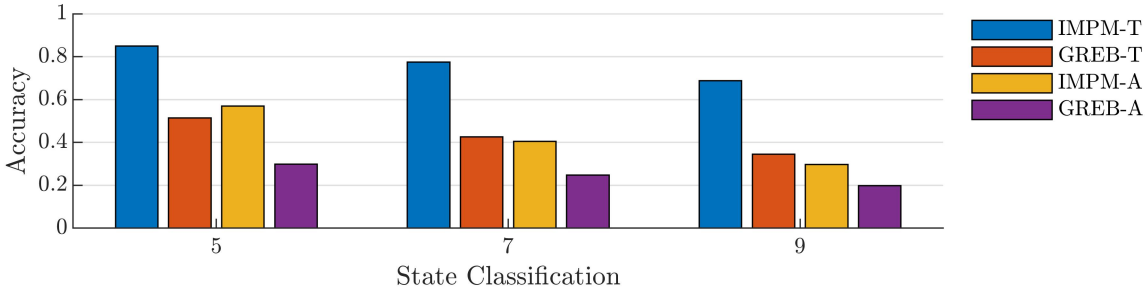

Fig 4. Comparison result of the average accuracy of different methods.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

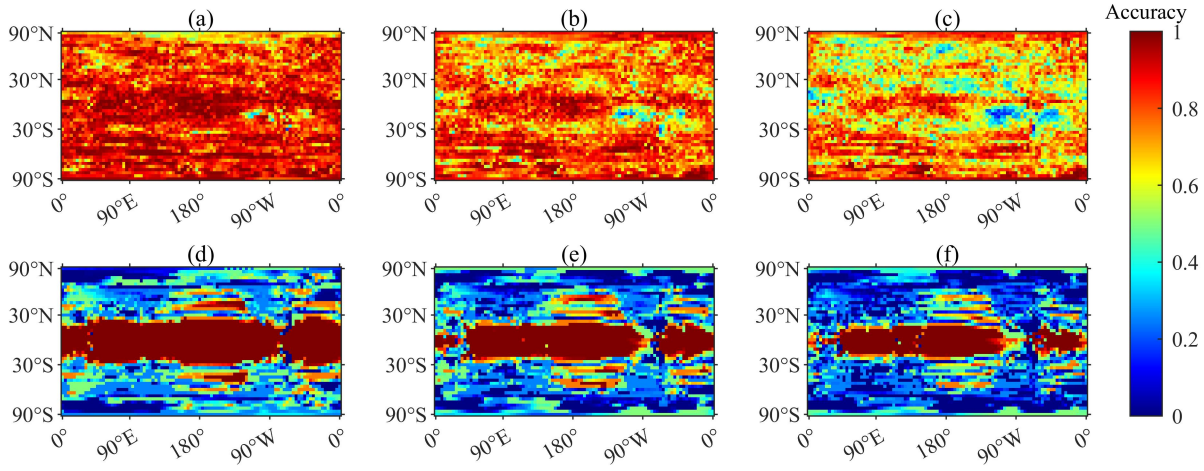

Fig 5. Comparison results of the accuracy of simulation results of surface average temperature. (a) IMPM under 5 classification; (b) IMPM under 7 classification; (c) IMPM under 9 classification; (d) GREB under 5 classification; (e) IMPM under 7 classification; (f) IMPM under 9 classification.

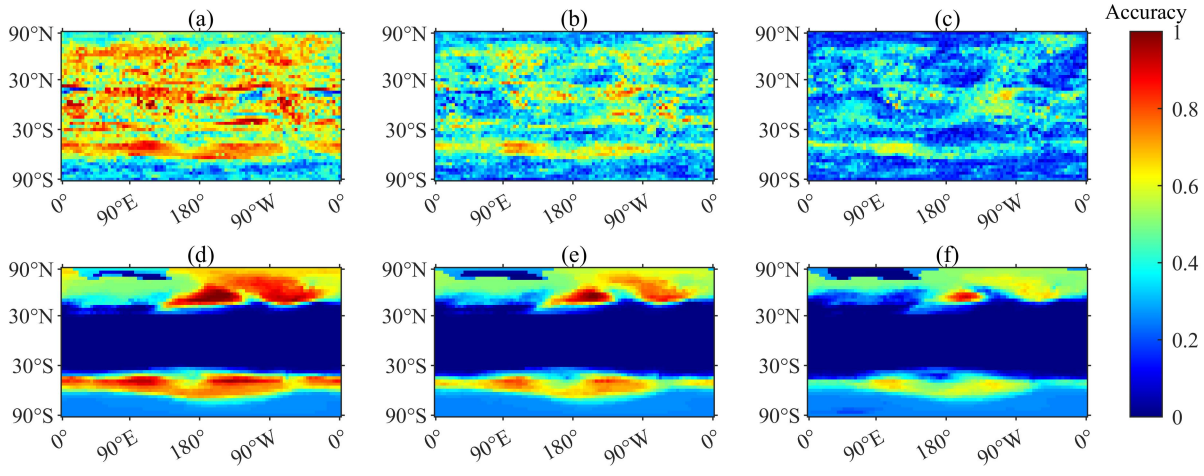

Fig 6. Comparison results of the accuracy of simulation results of atmospheric average temperature. (a) IMPM under 5 classification; (b) IMPM under 7 classification; (c) IMPM under 9 classification; (d) GREB under 5 classification; (e) IMPM under 7 classification; (f) IMPM under 9 classification.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

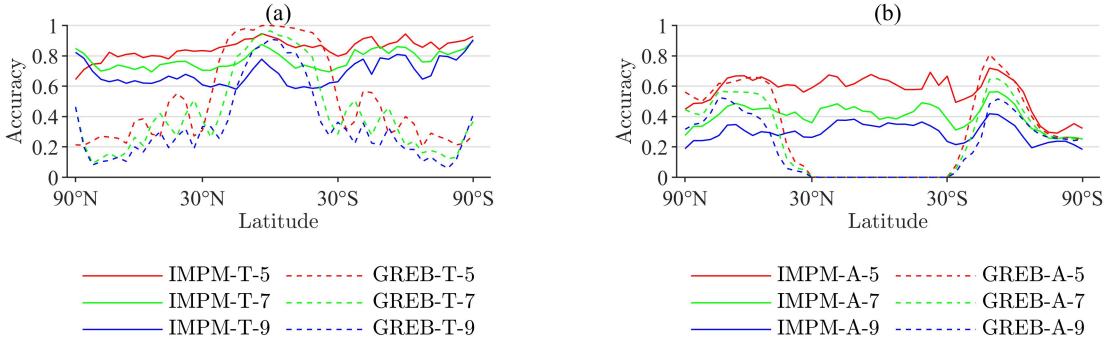

Fig 7. Average accuracy variation trend results of latitude. (a) surface average temperature; (b) atmospheric average temperature.

4 Conclusions and discussions

In this study, we introduced an improved method of the GREB model by Bayes network. It starts from the structural relationship of climate elements within the GREB model and re-models the climate process by Bayes network, which can be well integrated with the observed data to make dynamic corrections with spatial position change to the model simulation process to improve the model simulation accuracy.

The improved model was verified by two cases: surface temperature and atmospheric temperature. The simulation results of the improved model show that the improved model has higher average accuracy and lower spatial variability compared to the GREB model. This means that the improved model has better applicability and stability on a global scale. Although the simulation results of the improved model do not yield specific numerical solutions, but only climate states, they can still be applied well in studies on climate sensitivity(Dommenget, 2016; Kutzbach et al., 2013), extreme weather(Bellprat and Doblas-Reyes, 2016; Chen et al., 2018), and climate threshold(Mahlstein et al., 2015; Vogel et al., 2020), where the climate anomalies characterized by climate states can be a good indicator of climate change trends.

The results of the two study cases not only demonstrate that the improved method can also be used for the simulation of other climate elements within the GREB model. It also reveals that improving dynamical models through statistical methods such as Bayes networks is potential means to improve climate simulation and prediction. This improved approach can overcome the shortcomings of a single dynamical model that cannot accurately describe many nonlinear processes in the climate system and can be applied to other dynamical models. In terms of development, the improved methods for improving climate dynamical models by statistical methods show great possibilities for improving the accuracy of climate predictions.

However, this improved method still has some shortcomings.1) The scientific problem of state classification of climate element attribute features. In this paper, the state of each node is classified by the Natural Breaks method according to the data characteristics and statistical laws of the climate model, but this classification method changes with the data, and the data-based

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

classification model may not be consistent with the actual climate evolution law. Therefore, the following studies can discuss related issues and choose the appropriate feature classification criteria to achieve a balance between different simulation.2) Balance of accuracy and resolution. If the actual values rather than states are used as the calculation parameters, a higher resolution will be obtained, and of course the training data for each case will be reduced, which leads to the loss of accuracy. How to achieve the balance of accuracy and resolution will be an important issue. 3) Applicability of climate evolution models based on Bayesian networks. Stable conditional probability tables can be trained with historical climate data to simulate climate states, but conditional probability tables cannot change over time and cannot be adapted to time-sensitive climate models. The following study can extend the applicability of the method by dynamically training Bayes networks on climate data.

Code availability. The improved method in this paper was conducted in MATLAB R2021a. The code of the improved method used in this paper is archived on Zenodo(https://doi.org/10.5281/zenodo.6472781). The original GREB model uses the model code from the Monash Simple Climate Model (MSCM) laboratory repository for the GREB model and runs the code using the Fortran language. The model code is available from https://doi.org/10.5281/zenodo.2232282

Data availability. The data used in this paper is is archived on Zenodo(https://doi.org/10.5281/zenodo.6472781). The data used for the analysis in this paper have been pre-processed and the original data can be gotten from Environmental Predic-tion(NCEP)/ National Center for Atmospheric Research(NCAR), download from https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.pressure.html

Appendix A: Tables

Table A1. Five levels of climate elements state classification

State	$T(K)$	$A(K)$	S(W/m^2)	O(°C)	$W(kg/kg)$	$C(kg/kg)$	$V(m/s)$
1	<242.73	<198.30	<131.69	<3.44	<0.003	<33.06	<-3.44
2	242.73-264.81	198.29-205.11	131.69-160.62	3.44-10.82	0.003-0.007	33.06-45.51	-3.44 - -0.9
3	264.81-279.03	205.11-212.03	160.62-193.67	10.82-17.86	0.007-0.011	45.51-56.24	-0.90 - 1.51
4	279.03-291.24	212.03-217	193.67-224.48	17.86-24.17	0.011-0.016	56.24-66.1	1.51 - 4.55
5	>291.24	>217.00	>224.48	>24.17	>0.016	>66.10	>4.55

Author contributions. ZL, ZZ, and WL designed the paper's ideas and methods. ZL, ZW, JW, implemented the method of the paper with code. ZL, DL, and WL wrote the paper with considerable input from ZY and LY. ZW revised and checked the language of the paper.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

Table A2. Seven levels of climate elements state classification

State	$T(K)$	$A(K)$	S(W/m^2)	O(°C)	$W(kg/kg)$	$C(kg/kg)$	$V(m/s)$
1	<236.92	<196.02	<122.28	<1.97	<0.003	<28.24	<-4.94
2	236.92-252.70	196.02-200.22	122.28-139.30	1.97-6.54	0.003-0.005	28.24-38.05	-4.94 - -2.56
3	252.70-264.68	200.22-205.53	139.30-162.10	6.54-11.51	0.005-0.008	38.05-46.85	-2.56 - -0.53
4	264.68-275.57	205.53-210.36	162.10-188.67	11.51-16.79	0.008-0.011	46.85-54.51	-0.53-1.14
5	275.57-285.52	>210.36-214.34	188.67-212.89	>16.79-21.61	0.011-0.014	>54.51-61.71	1.14-3.16
6	285.52-294.63	>210.34-217.89	212.89-237.12	>21.61-25.76	0.014-0.017	>61.71-69.05	3.16-5.66
7	>294.63	>217.89	>237.12	>25.76	>0.017	>69.05	>5.66

Table A3. Nine levels of climate elements state classification

State	$T(K)$	$A(K)$	S(W/m^2)	O(°C)	$W(kg/kg)$	$C(kg/kg)$	$V(m/s)$
1	<235.71	<195.10	<121.03	<-0.67	<0.001	<25.31	<-5.37
2	235.71-250.23	195.10-198.07	121.03-136.52	-0.67-2.24	0.001-0.00	25.31-33.31	-5.37- -3.41
3	250.23-259.75	198.07-201.69	136.52-154.71	2.24-6.58	0.003-0.005	33.31-40.68	-3.41- -1.65
4	259.75-268.10	201.69-206.06	154.71-174.83	6.58-11.23	0.005-0.008	40.68-47.40	-1.65 - -0.26
5	268.10-276.19	>206.06-209.96	174.83-193.40	11.23-15.91	0.008-0.01	47.40-53.36	-0.26-0.95
6	276.19-283.38	>209.96-213.38	193.40-210.80	15.91-19.99	0.01-0.013	53.36-59.13	0.95-2.38
7	283.38-290.23	>213.38-216.46	210.80-227.90	19.99-23.60	0.013-0.015	59.13-64.92	2.38-4.18
8	290.23-296.32	>216.46-218.86	227.90-248.00	23.60-26.67	0.015-0.018	64.92-70.91	4.18-6.26
9	>296.32	>218.86	>248.00	>26.67	>0.018	>70.91	>6.26

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research has been supported by the National Natural Science Foundation of China (No.41976186, and 42130103) and the Postdoctoral Science Foundation of China (No. 2021M702757)

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

References

Alley, R. B., Emanuel, K. A., and Zhang, F.: Advances in weather prediction, SCIENCE, 363, 342-344, 10.1126/science.aav7274, 2019.

Annan, J. D. and Hargreaves, J. C.: Using multiple observationally-based constraints to estimate climate sensitivity, GEOPHYSICAL RESEARCH LETTERS, 33, 10.1029/2005GL025259, 2006.

Bellprat, O. and Doblas-Reyes, F.: Attribution of extreme weather and climate events overestimated by unreliable climate simulations, GEOPHYSICAL RESEARCH LETTERS, 43, 2158-2164, 10.1002/2015GL067189, 2016.

Berrocal, V. J., Craigmile, P. F., and Guttorp, P.: Regional climate model assessment using statistical upscaling and downscaling techniques, ENVIRONMETRICS, 23, 482-492, 10.1002/env.2145, 2012.

Cai, B., Liu, Y., Liu, Z., Tian, X., Zhang, Y., and Ji, R.: Application of Bayesian Networks in Quantitative Risk Assessment of Subsea Blowout Preventer Operations, RISK ANALYSIS, 33, 1293-1311, 10.1111/j.1539-6924.2012.01918.x, 2013.

Cai, B., Kong, X., Liu, Y., Lin, J., Yuan, X., Xu, H., and Ji, R.: Application of Bayesian Networks in Reliability Evaluation, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, 15, 2146-2157, 10.1109/TII.2018.2858281, 2019.

Chen, L., Zhang, H., Wu, Q., and Terzija, V.: A Numerical Approach for Hybrid Simulation of Power System Dynamics Considering Extreme Icing Events, IEEE TRANSACTIONS ON SMART GRID, 9, 5038-5046, 10.1109/TSG.2017.2679109, 2018.

Chou, J.: Why do the dynamical models and statistical methods need to be combined?—Also on how to combine, Plateau Meteorol, 5, 77-82, 1986.

Chou, J.: Short term climatic prediction: Present condition, problems and way out, Bimon Xinjiang Meteorol, 26: 1, 4, 2003.

Dommenget, D.: A simple model perturbed physics study of the simulated climate sensitivity uncertainty and its relation to control climate biases, CLIMATE DYNAMICS, 46, 427-447, 10.1007/s00382-015-2591-4, 2016.

Dommenget, D. and Flöter, J.: Conceptual understanding of climate change with a globally resolved energy balance model, Climate Dynamics, 37, 2143-2165, 10.1007/s00382-011-1026-0, 2011.

Fan, J., Meng, J., Ludescher, J., Chen, X., Ashkenazy, Y., Kurths, J., Havlin, S., and Schellnhuber, H. J.: Statistical physics approaches to the complex Earth system, PHYSICS REPORTS-REVIEW SECTION OF PHYSICS LETTERS, 896, 1-84, 10.1016/j.physrep.2020.09.005, 2021.

Feng, G. L., Yang, J., Zhi, R., Zhao, J. H., and Sun, G. Q.: Improved prediction model for flood-season rainfall based on a nonlinear dynamics-statistic combined method, Chaos Solitons & Fractals, 140, 110160, 10.1016/j.chaos.2020.110160, 2020.

Grant, P. R.: Evolution, climate change, and extreme events, SCIENCE, 357, 451-452, 10.1126/science.aao2067, 2017.

Huang, J., Chen, W., Wen, Z., Zhang, G., Li, Z., Zuo, Z., and Zhao, Q.: Review of Chinese atmospheric science research over the past 70 years: Climate and climate change, SCIENCE CHINA-EARTH SCIENCES, 62, 1514-1550, 10.1007/s11430-019-9483-5, 2019.

Jansen, R., Yu, H. Y., Greenbaum, D., Kluger, Y., Krogan, N. J., Chung, S. B., Emili, A., Snyder, M., Greenblatt, J. F., and Gerstein, M.: A Bayesian networks approach for predicting protein-protein interactions from genomic data, SCIENCE, 302, 449-453, 10.1126/science.1087361, 2003.

Kay, J. E.: Early climate models successfully predicted global warming, Nature, 578, 45-46, 10.1038/d41586-020-00243-w, 2020.

Kutzbach, J. E., He, F., Vavrus, S. J., and Ruddiman, W. F.: The dependence of equilibrium climate sensitivity on climate state: Applications to studies of climates colder than present, GEOPHYSICAL RESEARCH LETTERS, 40, 3721-3726, 10.1002/grl.50724, 2013.

https://doi.org/10.5194/egusphere-2022-138
Preprint. Discussion started: 2 August 2022

Ludescher, J., Martin, M., Boers, N., Bunde, A., Ciemer, C., Fan, J., Havlin, S., Kretschmer, M., Kurths, J., Runge, J., Stolbova, V., Surovyatkina, E., and Schellnhuber, H. J.: Network-based forecasting of climate phenomena, PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 118, 10.1073/pnas.1922872118, 2021.

Maher, P.: Bayesian probability, SYNTHESE, 172, 119-127, 10.1007/s11229-009-9471-6, 2010.

Mahlstein, I., Spirig, C., Liniger, M. A., and Appenzeller, C.: Estimating daily climatologies for climate indices derived from climate model data and observations, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 120, 2808-2818, 10.1002/2014JD022327, 2015.

Pearl, J.: Fusion, propagation, and structuring in belief networks, Artificial intelligence, 29, 241-288, 10.1016/0004-3702(86)90072-x, 1986.

Sahin, O., Stewart, R. A., Faivre, G., Ware, D., Tomlinson, R., and Mackey, B.: Spatial Bayesian Network for predicting sea level rise induced coastal erosion in a small Pacific Island, JOURNAL OF ENVIRONMENTAL MANAGEMENT, 238, 341-351, 10.1016/j.jenvman.2019.03.008, 2019.

Stassen, C., Dommenget, D., and Loveday, N.: A hydrological cycle model for the Globally Resolved Energy Balance (GREB) model v1.0, Geoscientific Model Development, 12, 425-440, 10.5194/gmd-12-425-2019, 2019.

Vogel, M. M., Zscheischler, J., Fischer, E. M., and Seneviratne, S. I.: Development of Future Heatwaves for Different Hazard Thresholds, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 125, 10.1029/2019JD032070, 2020.

Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H., and Emanuel, K.: What Is the Predictability Limit of Midlatitude Weather?, JOURNAL OF THE ATMOSPHERIC SCIENCES, 76, 1077-1091, 10.1175/JAS-D-18-0269.1, 2019.

Zou, Y., Donner, R. V., Marwan, N., Donges, J. F., and Kurths, J.: Complex network approaches to nonlinear time series analysis, PHYSICS REPORTS-REVIEW SECTION OF PHYSICS LETTERS, 787, 1-97, 10.1016/j.physrep.2018.10.005, 2019.