Acta Oceanol. Sin., 2021, Vol. 40, No. 7, P. 249–262
https://doi.org/10.1007/s13131-021-1734-1
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Evolving a Bayesian network model with information flow for time series interpolation of multiple ocean variables

Ming Li[1], Ren Zhang[1]*, Kefeng Liu[1]

[1] College of Meteorology and Oceanography, National University of Defense Technology, Nanjing 211101, China

Received 1 June 2020; accepted 21 September 2020

Abstract

Based on Bayesian network (BN) and information flow (IF), a new machine learning-based model named IFBN is put forward to interpolate missing time series of multiple ocean variables. An improved BN structural learning algorithm with IF is designed to mine causal relationships among ocean variables to build network structure. Nondirectional inference mechanism of BN is applied to achieve the synchronous interpolation of multiple missing time series. With the IFBN, all ocean variables are placed in a causal network visually, making full use of information about related variables to fill missing data. More importantly, the synchronous interpolation of multiple variables can avoid model retraining when interpolative objects change. Interpolation experiments show that IFBN has even better interpolation accuracy, effectiveness and stability than existing methods.

Key words: Bayesian network, information flow, time series interpolation, ocean variables

Citation: Li Ming, Zhang Ren, Liu Kefeng. 2021. Evolving a Bayesian network model with information flow for time series interpolation of multiple ocean variables. Acta Oceanologica Sinica, 40(7): 249–262, doi: 10.1007/s13131-021-1734-1

1 Introduction

Time series $x_t(j)$ is a series of observations gathered in chronological order. $x_t(j)$ represents the record of the jth variable at the tth moment. When $j=1$, $x_t(j)$ is a univariate time series; when $j>1$, $x_t(j)$ is a multivariate time series. Time series contain deep knowledge, and data mining of time series has become a hot topic in the field of big data analysis. As we all know, time series are widely existed in oceanology and hydrology, but data loss is inevitable due to the effect of subjective and external factors. Discontinuous time series bring difficulties to studies of regional oceanic and hydrological changes. Therefore, interpolation of missing time series is the premise of oceanography analysis. Tremendous efforts have been made over the last few decades to fill the existing observational data gaps.

Most of classical missing data recovery methods, such as polynomial interpolation, optimal interpolation, auto-regressive moving average model and ensemble Kalman filtering model (Gasca and Sauer, 2000; Kaplan et al., 2000; Zhu, 2016; Zheng et al., 2019), have been widely applied to meteorological and oceanographic interpolation and have obtained many good results. Sheng (2009) applied the data interpolating empirical orthogonal function (DINEOF) model for interpolation of missing sea surface temperature (SST) data, on the basis of decomposition of EOF. Huang et al. (2014) defined the segmented matched interpolation method and developed a linear fitting equation corresponding to different time periods, to reconstruct the ground temperature time series. Liu et al. (2018) adopted Bayesian linear regression to establish mathematical relationship between the data missing stations and their adjacent stations, to interpolate the monthly precipitation time series. In addition, some novel statistical methods have been introduced. Bai et al. (2014) proposed a new information diffusion model optimized by genetic algorithm to fill missing time series of monthly river discharge.

The above interpolation models in geoscience can be sum up as the regression-based methods, which establish a mapping, such as regression equation and interpolation function, and take the predicted value of the mapping model as interpolation result. Besides, the regression-based methods mainly aim at univariate time series and effectively utilize the change rule of interpolative variables, which is better at dealing with linear relationship. However, these methods neglect information of other related variables. In other word, the above time series interpolation algorithms can only handle single variable, and take no considerations of multivariate interaction.

In recent years, with the development of machine learning (ML) and deep learning (DL), many new interpolation algorithms for missing time series have been born in the field of computer information. K-nearest neighbor (KNN) regression, support vector machine (SVM), artificial neural network (ANN) and convolutional neural network (CNN) have been initially used for missing data interpolation in geoscience (Yao, 2006; Zhang, 2013; Bu et al., 2014; Xu et al., 2018; Zheng et al., 2020). Liu et al. (2015) proposed an interpolation method for estimating sea surface salinity using random forest (RF) algorithm. Barth et al. (2020) applied a CNN with error estimates to reconstruct SST satellite observations. These ML-based interpolation methods are appropriate for nonlinear time series. Based on information of related variables, data interpolation of both univariate time series and multivariate time series is achieved, which significantly improve interpolation accuracy.

However, the above ML models have to be retrained when the interpolative object changes because their input and output are fixed in the model, so data interpolation of multiple variables is inefficient. More importantly, some studies have pointed out

Foundation item: The National Natural Science Foundation of China under contract Nos 41875061 and 41976188; the "Double First-Class" Research Program of National University of Defense Technology under contract No. xslw05.
*Corresponding author, E-mail: zrpaper@163.com

that data sets of variables such as radiation, temperature and evapotranspiration reconstructed by different interpolation methods often generate new errors when substituted into the energy balance equation (Jiang et al., 2011). In other word, data reconstructed by separate interpolation without considering influence of related variables may be ineffective in practical application. Aiming at the problems existing in the regression-based and ML-based interpolation methods, the state-of-the-art Bayesian network (BN) is introduced.

BN, an artificial intelligence algorithm, is a combination of probability theory and graph theory. It is good at mining and expressing causal relationships from data. At present, some scholars have applied BN to missing data interpolation (Li, 2006; Gong and Dong, 2010). BN can learn and express causal relationships to construct a reliability network containing multiple variables (Li, 2018a). Missing data can be filled by probabilistic reasoning with making full use of information about related variables. It is worth noting that BN can provide a comprehensive description of network nodes with probability distribution and achieve probabilistic reasoning in any direction. Therefore, it is possible to fill multivariate missing data synchronously with BN. There is no need to retrain the model when interpolative objects change.

BN training is the core of BN modeling, including structural learning and parameter learning. Structural learning is construction of a causal network describing relationship of multiple variables, which is the foundation of BN. Interactions among ocean variables are grossly complicated, so effective identification of relationships from time series and network structure construction are the guarantee of high interpolation accuracy. The key to structural learning is determining arcs and their direction between network nodes. Search-and-score algorithm is the most widely used method, which searches for the optimal network structure according to scoring function in a network space. K2 algorithm, K3 algorithm and genetic algorithm are common algorithms in structural learning (Cooper and Herskovits, 1992; Bouckaert, 1994; Liu et al., 2001; Wang and Yang, 2010; Li and Liu, 2019). However, these algorithms are easy to fall into local optimum and structure arcs have great uncertainty. Actually, structural learning of BN is causal analysis. For the first time, we introduce information flow (IF) (Liang, 2008), an emerging causal analysis theory, for structural learning. We conduct the causal analysis of ocean variables based on IF and optimize network learning, to propose a new time series interpolation model for multiple ocean variables, namely IFBN.

In this paper, we will elaborate that how IFBN learns the relationships among ocean variables by mining the causality from time series, and quantitatively expresses them in the form of probability distribution. After constructing a complete BN with ocean variables, data interpolation of multiple variables can be realized with nondirectional inference mechanism. Time series of all variables in the network can be interpolated by only once model training. IFBN effectively utilizes information about related variables, and realizes simultaneous interpolation of missing multivariate time series.

The remainder of the paper is organized as follows: Section 2 presents the basic theory and techniques. Section 3 introduces the specific modeling techniques of IFBN. The obtained results and analysis are shown in Section 4. Section 5 concludes the paper and discusses the main advan tages of the proposed model.

2 Basic theory

2.1 *Bayesian network*

Bayesian network (BN), also known as Bayesian reliability network, is not only a graphical expression of causal relationship among variables, but also a probabilistic reasoning technique (Pearl, 1998). It can be represented by a binary $B=<G,\theta>$:

(1) $G=(V,E)$ represents a directed acyclic graph. V is a set of nodes where each node represents a variable in the problem domain. E is a set of arcs, and a directed arc represents the causal dependency between variables.

(2) θ is the network parameter, that is the probability distribution of nodes. θ expresses the degree of mutual influence between nodes and presents quantitative characteristics in the knowledge domain.

Assume a set of variables $V=(v_1,\cdots,v_n)$. The mathematical basis of BN is Bayes Theorem showed by Eq. (1), which is also the core of Bayesian inference.

$$P\left(v_i|v_j\right)=\frac{P(v_i,v_j)}{P\left(v_j\right)}=\frac{P\left(v_i\right)\cdot P\left(v_j|v_i\right)}{P\left(v_j\right)}, \tag{1}$$

where $P\left(v_i\right)$ is the prior probability, $P\left(v_j|v_i\right)$ is the conditional probability and $P\left(v_i|\,v_j\right)$ is the posterior probability. Based on $P\left(v_i\right)$, $P\left(v_i|v_j\right)$ could be derived by Bayes Theorem under the relevant conditions $P\left(v_j|v_i\right)$.

After determination of structure and parameter, the joint probability distribution for Bayesian inference can be derived from Eq. (1) under the assumption of conditional independence (Shi, 2012).

$$P\left(v_1,v_2,\cdots,v_n\right)=\prod_{i=1}^{n}P\left(v_i\,|\mathrm{Pa}(\,v_i)\right), \tag{2}$$

where v_i is network node; $\mathrm{Pa}\left(v_i\right)$ is parent node of v_i. Bayesian inference is the calculation of probability distribution of a set of query variables according to evidences of input variables through Eq. (2).

BN provides a reasoning technique based on probability distribution. It does not distinguish forward reasoning or backward reasoning (Li et al., 2008). Each node in the network can be taken as input and output flexibly, so the simultaneous interpolation of missing time series of multiple ocean variables can be achieved with the nondirectional inference mechanism.

The training of BN includes structural learning and parameter learning. The BN structure is the basis of parameter learning and probabilistic reasoning. In order to improve the structural learning, information flow (IF) will be introduced and its theory will be explained in the next subsection.

2.2 *Information flow*

Information flow (IF) is a real physical notion recently rigorized by Liang (2014, 2015) to express causality between two variables in a quantitative way, where causality is measured by the information transfer rate from one variable's time series to another. IF can realize the formalization and quantification in causal analysis.

Following Liang (2014, 2015), given two time series X_1 and X_2, the maximum likelihood estimator of the rate of the IF from X_2 to X_1 is:

$$T_{2\to1}=\frac{C_{11}C_{12}C_{2,\mathrm{d}_1}-C_{12}^2C_{1,\mathrm{d}_1}}{C_{11}^2C_{22}-C_{11}C_{12}^2}, \tag{3}$$

where $\to$ represents the assignment operator, C_{ij} denotes the covariance between X_i and X_j, and C_{i,d_j} is determined as follows. Let

$\dot{X}_j$ be the finite-difference approximation of $\mathrm{d}X_j/\mathrm{d}t$ using the Euler forward scheme:

$$\dot{X}_{j,n} = \frac{X_{j,n+k} - X_{j,n}}{k\Delta t}, \tag{4}$$

with $k=1$ or $k=2$ (the details about how to determine k are referred to Liang (2014) and Δt being the time step). C_{i,d_j} in Eq. (3) is the covariance between X_i and $\dot{X}_j$.

In order to quantify the relative importance of a detected causality, Liang (2015) developed an approach to normalizing the IF:

$$\begin{cases} Z_{2\to1} \equiv |T_{2\to1}| + \left|\dfrac{\mathrm{d}H_1^*}{\mathrm{d}t}\right| + \left|\dfrac{\mathrm{d}H_1^{\mathrm{noise}}}{\mathrm{d}t}\right|, \\ \tau_{2\to1} = \dfrac{T_{2\to1}}{Z_{2\to1}}, \end{cases} \tag{5}$$

where $\equiv$ is the identity sign, H_1^* represents the phase space expansion along the X_1 direction, and H_1^{noise} represents the random effect.

The normalized IF calculated with Eq. (5) can be zero or nonzero. Ideally, if $\tau_{2\to1}=0$, then X_2 does not cause X_1; otherwise there is a causal link between X_1 and X_2. Further, if $\tau_{2\to1}>0$, then X_2 makes X_1 unstable. On the contrary, $\tau_{2\to1}<0$ indicates that X_2 makes X_1 stable. In particular, when the significance level is 0.1, $|\tau_{2\to1}|>1\%$ indicates that the causal relationship is significant.

All in all, IF is very suitable for structural learning of BN. The value of IF reflects the strength of causal dependency between network nodes. The symbol of IF represents the direction of structural arcs (Li and Liu, 2020). With causal analysis of ocean variables based on IF, a more reasonable network structure of ocean variables can be obtained, which is helpful for time series interpolation.

3 Construction of IFBN

In this section, we will propose a new time series interpolation model (IFBN) for multiple ocean variables and explain the model designation in detail. The great advantage of IFBN is synchronous interpolation of missing multivariate time series. The specific technical process is shown in Fig. 1.

First, time series of ocean variables are processed to generate discrete time series for network training. Then, combined with causal analysis by IF, the search-and-score algorithm is used to mine causal relationships among ocean variables and obtain the topological structure. On the basis of this network structure, EM algorithm is employed to learn probability distribution of network nodes, and the complete BN containing ocean variables has been trained. Finally, missing multivariate time series are input for probabilistic reasoning and missing data of different ocean variables are interpolated synchronously. The technical process of IFBN is described in detail below.

3.1 *Time series preprocessing*

As BN is better at processing discrete data, time series of ocean variables are required to be discretized in order to train the BN. The equal interval method is adopted to discretize time series (Li, 2018). Historical data over a period are analyzed to choose suitable intervals according to the maximum and minimum. Consequently, discrete state space of each node is divided with constant interval scale. Different states of network nodes (ocean variables) are represented by discrete numbers.

3.2 *Structural learning*

In order to identify the relationships among ocean variables, we introduce IF to design a new structural learning algorithm, whose process consists of two steps: construction of the initial structure and search for the optimal structure. We first conduct causal analysis of ocean variables based on IF and construct the unconstrained 0/1 optimization problem to get the initial network structure. Then the search-and-score algorithm is adopted to learn the optimal network structure. Details are as follows.

Step 1. Construction and solution of unconstrained 0/1 optimization problem

We calculate IF between different ocean variables, and make a significant analysis of causal relationship to construct the unconstrained 0/1 optimization problem.

BN structure can be represented by an adjacency matrix $\boldsymbol{X} = (x_{ij})$. A network with n nodes can be represented as follows:

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}, \tag{6}$$

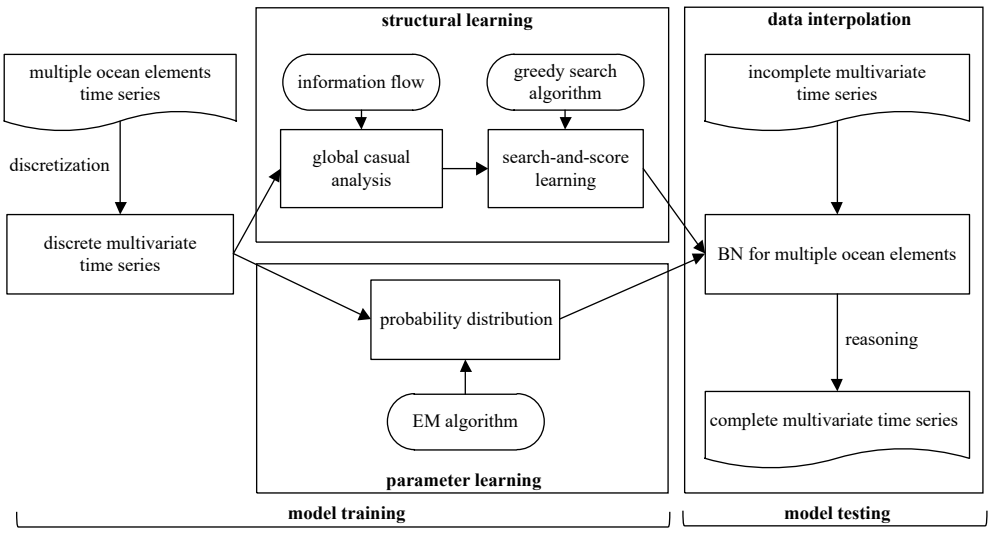


Fig. 1. Technique flowcharts of IFBN.

where $x_{ij}=1$ represents there is an arc between v_i and v_j, while $x_{ij}=0$ represents there is no arc between v_i and v_j.

Definition 1: Construct a mathematical variable to measure the causation of a network based on $X=(x_{ij})$ and τ_{ij}.

$$C(X,\alpha)=\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\tau_{ij}x_{ij}. \tag{7}$$

The larger the value of $C(X,\alpha)$, the more significant the causation of a network. With Definition 1, the unconstrained 0/1 optimization problem can be defined as following:

$$\max C(X,\alpha)=\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\tau_{ij}x_{ij}. \tag{8}$$

Solve the above 0/1 optimization problem to get the optimal adjacency matrix, corresponding to a directed topology, that is the initial network structure used in the next step.

Step 2. Get the optimal structure with search-and-score algorithm

The search-and-score algorithm has two important components: measurement mechanism and search process (Chickering et al., 1995). The measurement mechanism is a structure scoring function, used for evaluating the quality of a network, such as information entropy, Bayesian information criterion (BIC) and minimum description length (MDL). The search process refers to searching the highest rated network in a structure space according to the measure. In our algorithm, structure space is generated from the initial structure based on IF and we adopt BIC and greedy search (GS) algorithm to search for the optimal structure.

The expression of BIC is as follows:

$$\text{BIC\_score}=\sum_{i=1}^{n}\log_{10}P(G|D)-\frac{1}{2}\log_{10}N\cdot\text{Dim}(G), \tag{9}$$

where $\text{Dim}(G)$ represents the dimensions of BN, G represents network structure, and D represents data set.

The basic thought of GS algorithm: start from an initial structure, perform arc addition, arc reduction and arc rotation on the initial structure and score it after each operation. If the score is increased, the operation is retained. The process is iterated until the network score is optimal (Chickering, 2003). It should be pointed out that as IF is introduced, arc rotation is omitted in GS algorithm, and the arc direction is determined according to the IF symbol. The specific algorithm flow is shown in Algorithm 1.

Algorithm 1	**GS algorithm**
Input	V is variable set, D is complete data set of V, G_0 is an initial structure
Output	G is the optimal structure
Step 1	Score the initial network structure $G_0 \rightarrow oldscore$;
Step 2	Perform arc addition, arc reduction and determine arc direction by IF, score the new network structure $G' \rightarrow tempscore$;
	if $tempscore > oldscore$
	$newscore \equiv tempscore$ and keep the corresponding arc operation;
	else
	$newscore \equiv oldscore$ and discard the corresponding arc operation;
	end if
Step 3	If $newsore \rightarrow max$
	return $G \equiv G \rightarrow'$

3.3 *Parameter learning*

In the BN, the relationships among ocean variables are expressed quantitatively by probability distribution, and EM algorithm (Zhou, 2016) is used to learn probability distribution, including prior probability and conditional probability. First, initialize the probability distribution of each node. Then, modify the initial probability distribution according to training data to find its maximum likelihood estimate. In our paper, EM algorithm is achieved by ***BNT*** toolbox (https://download.csdn.net/download/b08514/6942975), so its specific algorithm principle is omitted.

3.4 *Reasoning and interpolation*

After structural learning and parameter learning, the BN with multiple ocean variables is completed. We input missing time series of variables to calculate the posterior probability of interpolative objects by probabilistic reasoning. Missing data of all variables can be filled according to posterior probability distribution. Bayesian reasoning algorithm includes exact algorithm and approximate algorithm (Li, 2018b). Approximate algorithm is usually applied to large-scale network structure to deal with excessive computation. Considering the scale of network in our research, we apply the exact algorithm, joint tree inference algorithm (Liu and Zhang, 2006), to accurate reasoning.

4 Time series interpolation experiment of multiple ocean variables

In this section, we use the proposed IFBN to interpolate missing time series of multiple ocean variables. To further illustrate the validity of IFBN, we also apply the Cubic Spline Interpolation (CSI) algorithm, Back Propagation (BP) neural network and classic BN (CBN) without IF to conduct comparative experiments and error analysis of the results. All experiments are carried out with MATLAB.

4.1 *Data introduction*

We use daily Tropical Atmosphere Ocean (TAO) data released by National Oceanic and Atmospheric Administration (NOAA) for interpolation experiments. The selected ocean variables are wind speed, sea surface temperature, sea level pressure, salinity, relative humidity and density, respectively denoted as WD, SST, SLP, SAL, RH and DEN. We download six complete time series of above variables from January 1, 2017 to December 31, 2018 as experimental data (a total of 718 d). The latitude-longitude coordinate is (0°, 165°E). In our experiments, the first 600 d are taken as training data and the last 118 d are test data. As we all know, these time series are generally asymmetrical and non-normal.

4.2 *IFBN training*

Based on the training data of ocean variables, we train the IFBN according to the modeling process elaborated in Section 3.

(1) Time series preprocessing

After analyzing historical records, we select reasonable division intervals for six ocean variables and denote states with consecutive numbers. The discretization standard is shown in Table 1.

Then we discretize six variables with equal interval to obtain discrete training time series for BN learning as shown in Table 2.

(2) Global causal analysis

Based on the original training time series in Section 4.1, we calculate the normalized IF between different ocean variables according to Eqs (3)–(5), as shown in Table 3. It should be noted that the IF direction is from the row to the column.

Now we take significance analysis of IF in the table: for example, $\tau_{WD\to SST}=0.133\ 6>0$, $\tau_{SST\to WD}=-0.000\ 3<0$, and the former is greater than 1%, so we could judge WD is the cause for SST, that is there may be an arc "WD→SST" in the BN; $\tau_{WD\to SLP}=-0.002\ 5<0$, $\tau_{SLP\to WD}=0.011\ 9>0$, and the latter is greater than 1%, so we could judge that there may be an arc "SLP→WD"; for another example, $\tau_{WD\to RH}=0.023\ 8>0$, $\tau_{RH\to WD}=0.039\ 4>0$, and both are greater than 1%, so causal relationship is inapparent. But $\tau_{WD\to RH}<\tau_{RH\to WD}$, we could make a preliminary judgment that RH is the cause for WD. A preliminary analysis of the causal relationship between ocean variables is carried out by analyzing the value and symbol of IF.

(3) Structural learning and parameter learning

According to Section 3.2, on the basis of the above causal analysis, we design and solve the unconstrained 0/1 optimization problem to get the optimal initial network structure. The adjacency matrix describing the relationship of ocean variables is shown in Fig. 2a, and the corresponding initial network structure is shown in Fig. 2b.

Then we generate the search space based on the initial network structure and GS algorithm is implemented to search for the optimal structure. The iterative convergence curve is shown in Fig. 3a. The optimal structure describing ocean variables is shown in Fig. 3b.

Table 1. Discretization standard of ocean variables

	Ocean variable					
	WD	SST	SLP	SAL	RH	DEN
Interval	1 m/s	0.1°C	1 Pa	0.1‰	1%	0.1 kg/m^3
State label	1–10	1–26	1–9	1–17	1–24	1–19

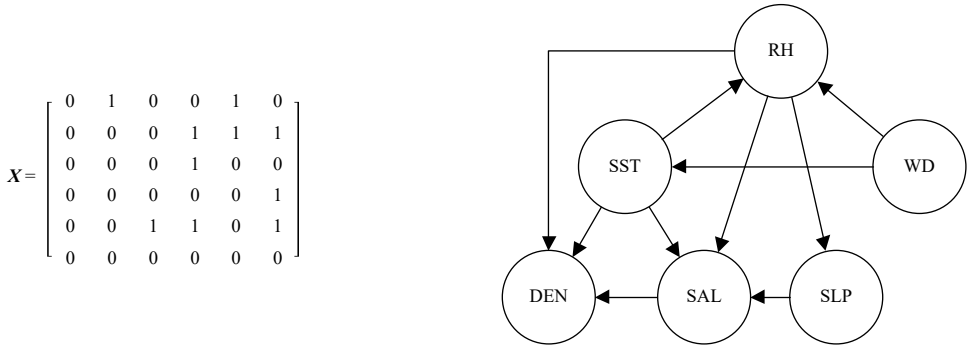
Table 2. Discrete training time series

Variable	Training data							
	1 d	2 d	3 d	4 d	5 d	…	500 d	600 d
WD/(m·s^{-1})	3	2	3	3	5	…	2	5
SST/°C	25	24	24	22	24	…	19	18
SLP/Pa	2	3	2	3	3	…	6	7
SAL/‰	15	15	15	10	8	…	15	15
RH/%	8	12	14	17	10	…	9	7
DEN/(kg·m^{-3})	10	11	11	7	6	…	12	12

Table 3. Standardized IF matrix

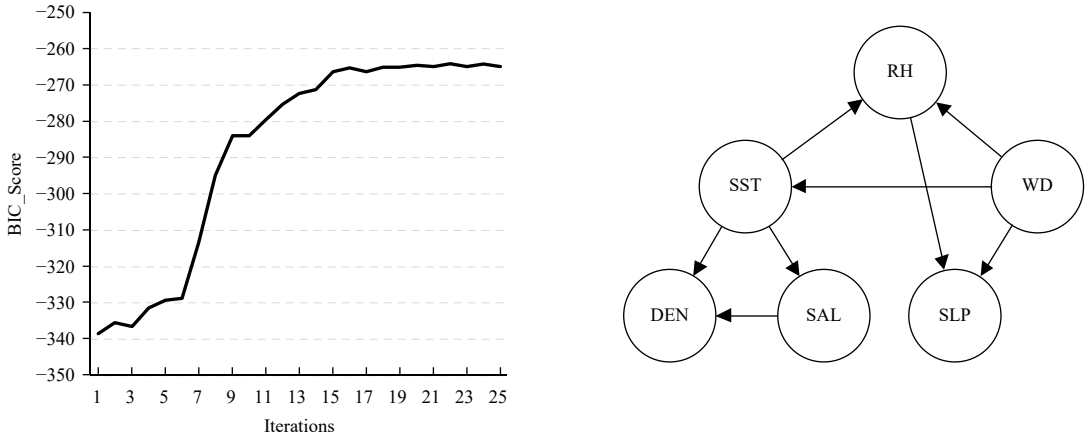
Variable	WD	SST	SLP	SAL	RH	DEN
WD	\	0.133 6	–0.002 5	–0.012 6	0.023 8	0.087 0
SST	–0.000 3	\	0.005 2	0.194 7	0.109 9	0.239 1
SLP	0.011 9	–0.011 8	\	0.027 2	0.006 0	–0.023 9
SAL	0.003 0	0.119 3	0.037 1	\	0.017 1	0.088 6
RH	0.039 4	0.096 2	0.006 7	–0.020 2	\	–0.018 6
DEN	–0.005 2	0.259 3	0.028 1	0.347 4	0.066 9	\

Note: \ means it cannot be calculated.



a. optimal adjacency matrix b. initial network structure

Fig. 2. 0/1 optimization solution of network.



a. iterative convergence curve b. optimal network structure

Fig. 3. Optimization result of GS algorithm.

Based on the network structure, EM algorithm is adopted to calculate the prior probability distribution and conditional probability distribution, which is achieved by BNT toolbox. As an example, Table 4 shows the conditional probability distribution of node SST.

4.3 *Interpolation results of time series*

In Section 4.2, the training of BN with ocean variables has been completed. In the network, causal relationships of ocean variables are expressed quantitatively and intuitively. Then we fill missing data of time series by network reasoning. For the purpose of testing the validity of IFBN, we perform multiple sets of interpolation experiments.

(1) Experiment I

In order to test the interpolation accuracy in different data missing situations, we adopt the missing completely at random (MCAR) method (Bai et al., 2014) to process test time series (a total of 118 d) of six ocean variables, and randomly delete 40%, 50% and 60% of data for each variable time series. To avoid chance, CSI, BP neural network, CBN and IFBN are used to perform interpolation experiments with ten different random samples for each missing rate. The average of ten experiments is

Table 4. Conditional probability distribution of node SST

Condition	P(SST| WD)									
	WD=1	WD=2	WD=3	WD=4	WD=5	WD=6	WD=7	WD=8	WD=9	WD=10
SST=1	0	0	0	0	0	0	0.032 3	0	0	0
SST=2	0	0	0	0	0	0.029 0	0.064 5	0.076 9	0	0
SST=3	0	0	0	0	0.014 3	0.014 5	0	0	0	0
SST=4	0	0	0	0	0.028 6	0.058 0	0.032 3	0	0	0
SST=5	0	0	0	0.012 2	0.014 3	0.014 5	0.032 3	0	0	0
SST=6	0	0	0.021 3	0.024 4	0.028 6	0.029 0	0.032 3	0.153 8	0	0
SST=7	0	0.062 5	0	0.048 8	0.042 9	0.058 0	0.096 8	0.076 9	0	0
SST=8	0	0.031 3	0.021 3	0.012 2	0.042 9	0.043 5	0.032 3	0.076 9	0	0
SST=9	0	0.031 3	0.042 6	0.109 8	0.014 3	0.014 5	0	0.153 8	0	0
SST=10	0.1	0.031 3	0.042 6	0.024 4	0.028 6	0.014 5	0	0.076 9	0	0
SST=11	0	0.031 3	0.021 3	0.012 2	0.028 6	0.072 5	0.032 3	0	0	0
SST=12	0	0	0.063 8	0.048 8	0.071 4	0.087 0	0	0.076 9	0	0
SST=13	0.1	0	0.085 1	0.048 8	0.028 6	0.043 5	0.032 3	0	0	0
SST=14	0.1	0	0	0.036 6	0.057 1	0.058 0	0.096 8	0.076 9	0	0
SST=15	0	0.031 3	0.063 8	0.048 8	0.057 1	0.014 5	0.064 5	0.076 9	0.333 3	0
SST=16	0	0.031 3	0.042 6	0.048 8	0.014 3	0.014 5	0.064 5	0	0	0
SST=17	0	0	0.042 6	0.012 2	0.028 6	0.043 5	0	0	0	0
SST=18	0.2	0.031 3	0.085 1	0.073 2	0.114 3	0.014 5	0.064 0	0	0.333 3	0
SST=19	0.1	0.156 3	0.021 3	0.122 0	0.128 6	0.072 5	0	0.076 9	0	0
SST=20	0	0.062 5	0.063 8	0.085 4	0.014 3	0.058 0	0.096 8	0	0	0
SST=21	0.2	0.187 5	0.148 9	0.134 1	0.071 4	0.058 0	0.032 3	0.076 9	0	1
SST=22	0	0.031 3	0.042 6	0.036 6	0.028 6	0.058 0	0.064 5	0	0	0
SST=23	0.1	0.125 0	0.063 8	0.012 2	0.028 6	0.072 5	0.032 3	0	0	0
SST=24	0	0.093 8	0.021 3	0.048 8	0.085 7	0.029 0	0.096 8	0	0.333 3	0
SST=25	0.1	0.031 3	0.085 1	0	0.028 6	0.029 0	0	0	0	0
SST=26	0	0.031 3	0.021 3	0	0	0	0	0	0	0

Note: WD=1, 2, ···, 10, and SST=1, 2, ···, 26 indicate WD and SST take different discrete states.

taken as the final interpolation.

CSI is a segmental interpolation method, which constructs a cubic function in each segment. In modeling with BP neural network, the interpolation object is output neuron and the five remaining ocean variables are input neurons. The number of network layers is three. Given space limitations, we are not going to repeat detailed steps of CSI and BP as the two methods are mature. It should be pointed out that CSI and BP need to be retrained each time the interpolation object changes. Besides, the difference between CBN and IFBN is the structural learning. There is no causal analysis with IF in structural learning of CBN.

We discretize and input missing time series of six ocean variables into IFBN for probabilistic reasoning, and the model synchronously output the posterior probability distribution of each variable at the missing point with the nondirectional inference mechanism. Take the median of the state interval with maximum probability as the interpolation value. In order to intuitively display the interpolation, Fig. 4 shows a comparison between measured data and interpolation result of once experiment with a data missing rate of 40%, and the rest of interpolation results are shown in Figs A1 and A2.

We choose mean relative error (MRE), THEIL inequality coefficient (TIC) and correlation coefficient (R), as shown by Eqs (10)–(12), to measure the accuracy of interpolation. The smaller the first two, the more accurate the interpolation, while R is on the contrary. The average results of ten interpolation experiments are analyzed as shown in Fig. 5 and Tables 5 and 6.

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y-\hat{y}|}{y}, \tag{10}$$

$$\mathrm{TIC} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y-\hat{y})^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}y^2}+\sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{y}^2}}, \tag{11}$$

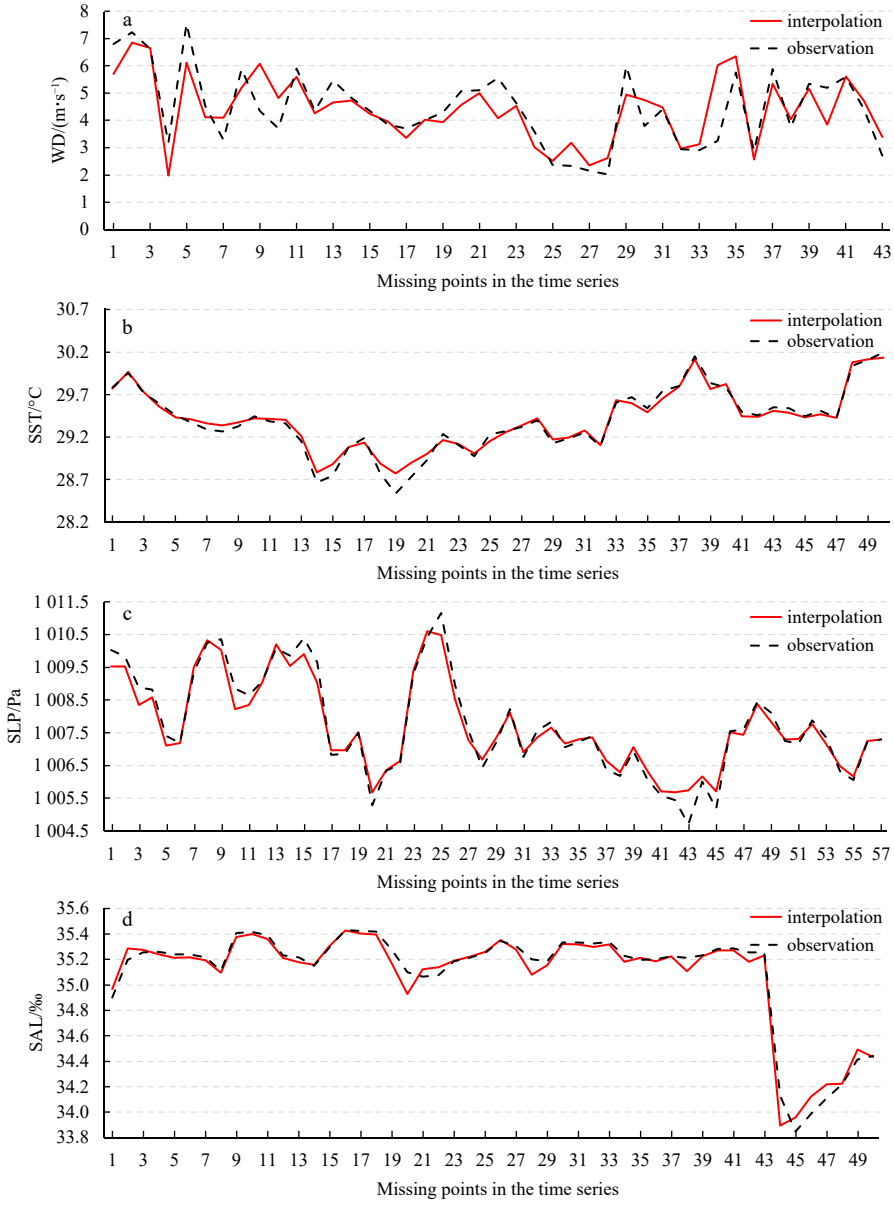


Fig. 4.

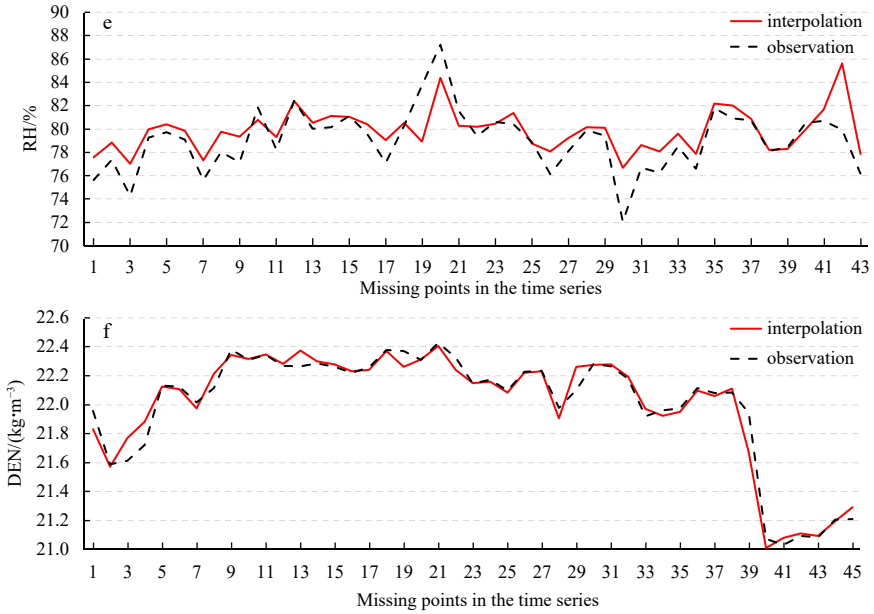


Fig. 4. Interpolation results with the data missing rate of 40%: WD (a), SST (b), SLP (c), SAL (d), RH (e), DEN (f).

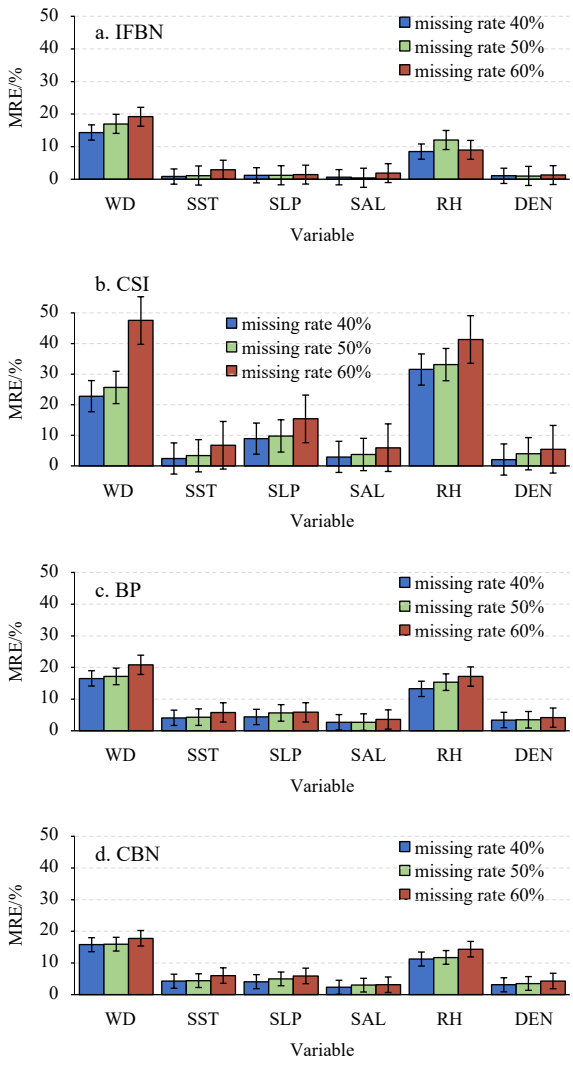


Fig. 5. MRE of each model under different missing rate.

$$R = \frac{\mathrm{Cov}\,(y, \hat{y})}{\sqrt{\mathrm{Var}\,(y) \cdot \mathrm{Var}\,(\hat{y})}}, \tag{12}$$

where y is the observation and $\hat{y}$ is the interpolation. $\mathrm{Cov}\,(y, \hat{y})$ is the covariance of y and $\hat{y}$. $\mathrm{Var}\,(y)$ is the variance of y and $\mathrm{Var}\,(\hat{y})$ is the variance of $\hat{y}$.

As seen from Fig. 4, the interpolation of six ocean variables with IFBN is close to the overall trend of the observation, and the accuracy is even higher. The peak value and low value are correspond to the measured value, and the change trend of time series is accurately expressed. Figure 5 and Tables 5 and 6 show that IFBN has smaller mean MRE (5.284%), smaller mean TIC (0.021 3) and larger mean R (0.887) than CIS (MRE: 15.144%, TIC: 0.044 3, R: 0.783), BP (MRE: 8.117%, TIC: 0.028 1, R: 0.811) and CBN (MRE: 7.546%, TIC: 0.027 6, R: 0.819), indicating that the interpolation of IFBN is more accurate. Besides, CBN has similar performance with BP, worse than IFBN, demonstrating that causal IF can improve the interpolation accuracy.

In addition, when data missing rate is small (40%), the interpolation effect of IFBN is better than the other three interpolation methods. As data missing rate increases (60%), IFBN can still maintain a higher accuracy because it can utilize action information of related variables based on the causal network, less affected by the spatial distribution pattern of missing data of individual variable. By contrast, performances of CSI and BP degrade sharply. IFBN shows great interpolation stability in the face of different degree of data loss, while the interpolation precision of CSI and BP is susceptible to the degree of data loss. Though CBN has similar stability with IFBN, its accuracy is obviously lower than IFBN.

Most important, when CSI and BP are used to interpolate multivariate time series, it is necessary to interpolate each ocean variable separately and perform repeated six interpolation operations. However, IFBN only needs once network training to realize synchronous interpolation of multiple variables.

(2) Experiment II

The length of consecutive missing data has an impact on the interpolation accuracy. In order to discuss the sensitivity of our

Table 5. TIC of each model under different missing rate

Missing rate	Interpolation model	WD	SST	SLP	SAL	RH	DEN
40%	**IFBN**	**0.091 5**	**0.001 1**	**0.001 4**	**0.000 9**	**0.011 5**	**0.001 7**
	CSI	0.182 9	0.005 3	0.002 7	0.002 2	0.022 1	0.004 3
	BP	0.095 6	0.003 3	0.002 5	0.001 6	0.014 1	0.003 5
	CBN	0.089 7	0.003 6	0.002 2	0.001 8	0.013 9	0.003 3
50%	**IFBN**	**0.117 1**	**0.001 4**	**0.001 5**	**0.001 1**	**0.017 5**	**0.001 3**
	CSI	0.177 6	0.013 2	0.006 1	0.005 6	0.039 5	0.009 8
	BP	0.119 4	0.005 8	0.004 3	0.003 4	0.025 2	0.014 3
	CBN	0.112 1	0.005 6	0.004 5	0.003 7	0.031 1	0.012 7
60%	**IFBN**	**0.106 5**	**0.003 7**	**0.002 4**	**0.002 2**	**0.018 9**	**0.001 8**
	CSI	0.215 8	0.015 4	0.010 9	0.009 3	0.062 9	0.013 1
	BP	0.138 2	0.008 3	0.007 4	0.006 5	0.040 7	0.012 1
	CBN	0.141 2	0.007 9	0.007 5	0.006 1	0.039 6	0.011 9

Note: The bold numbers are the results obtained by proposed model.

Table 6. *R* of each model under different missing rate

Missing rate	Interpolation model	WD	SST	SLP	SAL	RH	DEN
40%	**IFBN**	**0.811 5**	**0.989 5**	**0.987 2**	**0.987 1**	**0.788 1**	**0.992 7**
	CSI	0.731 5	0.926 9	0.941 2	0.975 9	0.719 9	0.968 1
	BP	0.636 6	0.915 7	0.935 3	0.985 9	0.720 5	0.973 6
	CBN	0.701 1	0.914 8	0.939 8	0.984 6	0.731 6	0.975 4
50%	**IFBN**	**0.678 6**	**0.988 8**	**0.972 3**	**0.993 5**	**0.690 8**	**0.990 4**
	CSI	0.651 0	0.807 2	0.838 1	0.867 0	0.578 4	0.839 9
	BP	0.512 7	0.943 8	0.905 7	0.975 5	0.596 1	0.816 5
	CBN	0.611 4	0.915 2	0.902 1	0.969 8	0.601 3	0.824 1
60%	**IFBN**	**0.667 8**	**0.934 2**	**0.923 1**	**0.929 1**	**0.654 7**	**0.982 7**
	CSI	0.585 9	0.716 4	0.793 6	0.860 2	0.517 0	0.768 0
	BP	0.618 3	0.883 2	0.826 5	0.973 8	0.574 4	0.798 3
	CBN	0.609 7	0.892 1	0.799 8	0.976 1	0.584 2	0.801 1

Note: The bold numbers are the results obtained by proposed model.

Table 7. The Maximum and averaged length of the consecutive missing data for each variable at different missing rate

Missing rate	Statistics length	WD	SST	SLP	SAL	RH	DEN
40%	maximum/d	4.00	5.00	4.00	4.00	4.00	3.00
	average/d	1.59	1.92	1.78	1.78	1.48	1.32
50%	maximum/d	6.00	5.00	7.00	7.00	6.00	7.00
	average/d	2.32	2.46	3.56	3.91	3.64	2.98
60%	maximum/d	8.00	8.00	7.00	9.00	8.00	9.00
	average/d	3.78	5.23	4.16	5.35	4.62	5.01

proposed IFBN to the length of consecutive missing data, some contrast experiments are conducted. We first do a statistics on the length of consecutive missing data, as shown in Table 7. Then, instead of MCAR method, we set different length (5 d, 10 d and 15 d) of consecutive missing data for each time series and each time series are set three groups of missing data. Three models (IFBN, CSI and BP) are used for interpolation and results are showed in Figs 6 and 7.

When the length of consecutive missing data is 5 d, IFBN has higher interpolation accuracy (mean *R*: 0.848, mean MRE: 4.325%) than BP (mean *R*: 0.769, mean MRE: 5.622%) and CSI (mean *R*: 0.706, mean MRE: 6.234%). As the length increases to 10 days, IFBN still maintains good mean *R* (0.733) and mean MRE (6.267%). Mean *R* of BP decreases slightly (0.617) while that of CSI decreases sharply (0.514). Therefore, IFBN has a lower sensitivity to the length of consecutive missing data than BP and CSI. However, when the length reaches 15 d, the performance of IFBN also decreases dramatically (mean *R*: 0.462, mean MRE: 10.538%), slightly better than BP (mean *R*: 0.427, mean MRE: 11.719%) and CSI (mean *R*: 0.314, mean MRE: 13.365%).

(3) Experiment III

In order to further test the generalization ability of IFBN, we select time series with different date and different position (A, B, and C) are shown in Table 8.

For each position, we divide the time series into training data and test data, and randomly delete 50% of test data. Then we apply three models (IFBN, CSI and BP) according to the above processes to interpolation. The comparison results are shown in Tables 9–11. For six ocean variables time series, IFBN has smaller MRE, smaller TIC, and larger *R* than the other two models, showing that our proposed model is suitable for many different situations and interpolation results are more stable. Besides, IFBN does not have to be retrained if the interpolation object (output) changes, so data interpolation for multiple variables is efficient.

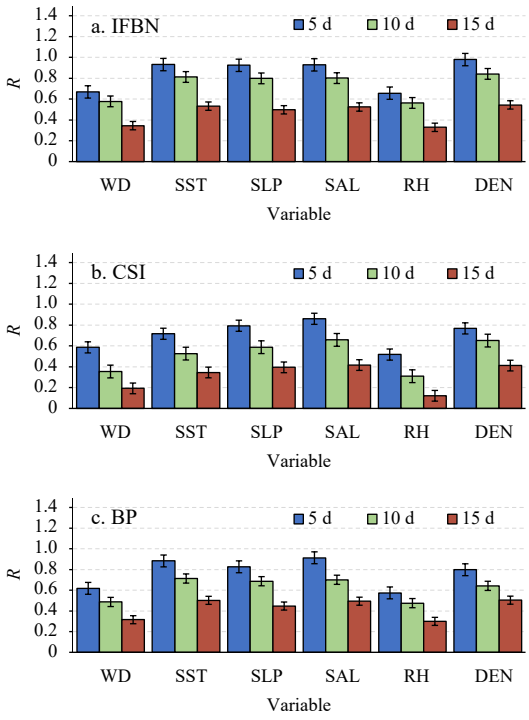


Fig. 6. *R* of each model with different length of consecutive missing data.

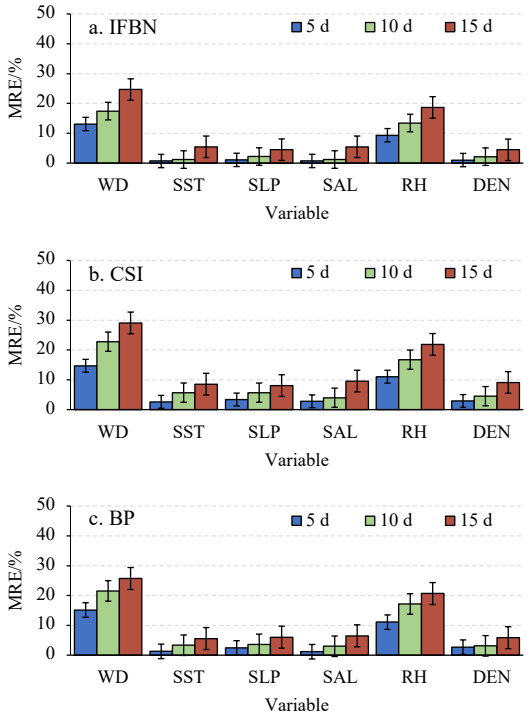


Fig. 7. MRE of each model with different length of consecutive missing data.

Table 8. The imformation of new experiment data

Position name	Peried	Latitude-longitude coordinate
A	Jan. 1, 2013 to Jun. 30, 2015	5°N, 110°W
B	Apr. 1, 2009 to Jul. 31, 2012	12°N, 70°W
C	May 1, 2010 to Jan. 31, 2013	15°S, 90°E

Table 9. Comparative analysis of interpolation results with different model in Position A

Evaluation indicator	Interpolation model	WD	SST	SLP	SAL	RH	DEN
MRE	**IFBN**	**0.167 7**	**0.036 5**	**0.037 9**	**0.027 1**	**0.134 4**	**0.028 7**
	CSI	0.245 8	0.065 4	0.093 9	0.042 3	0.229 1	0.062 1
	BP	0.191 5	0.056 4	0.041 2	0.035 9	0.179 9	0.038 1
TIC	**IFBN**	**0.140 9**	**0.002 6**	**0.002 2**	**0.000 5**	**0.023 9**	**0.002 7**
	CSI	0.179 4	0.005 8	0.004 3	0.004 4	0.075 2	0.014 3
	BP	0.147 6	0.003 2	0.003 1	0.002 6	0.049 5	0.009 8
R	**IFBN**	**0.618 6**	**0.968 8**	**0.960 2**	**0.950 8**	**0.633 3**	**0.910 3**
	CSI	0.582 7	0.923 8	0.835 7	0.865 5	0.514 4	0.736 5
	BP	0.601 2	0.946 9	0.908 1	0.917 1	0.577 1	0.819 9

Note: The bold numbers are the results obtained by proposed model.

Table 10. Comparative analysis of interpolation results with different model in Position B

Evaluation indicator	Interpolation model	WD	SST	SLP	SAL	RH	DEN
MRE	**IFBN**	**0.169 8**	**0.031 8**	**0.040 0**	**0.025 5**	**0.134 9**	**0.027 2**
	CSI	0.241 1	0.064 8	0.096 4	0.043 2	0.225 5	0.059 1
	BP	0.189 3	0.055 2	0.039 0	0.033 1	0.176 4	0.035 6
TIC	**IFBN**	**0.136 4**	**0.005 3**	**0.002 1**	**0.000 3**	**0.021 5**	**0.003 9**
	CSI	0.175 4	0.008 8	0.005 9	0.002 0	0.078 6	0.014 0
	BP	0.150 8	0.003 1	0.002 9	0.002 7	0.047 0	0.008 3
R	**IFBN**	**0.620 5**	**0.968 7**	**0.956 4**	**0.952 8**	**0.636 4**	**0.913 6**
	CSI	0.580 9	0.923 3	0.835 7	0.869 4	0.511 8	0.737 4
	BP	0.605 7	0.948 4	0.912 7	0.921 7	0.581 4	0.820 4

Note: The bold numbers are the results obtained by proposed model.

Table 11. Comparative analysis of interpolation results with different model in Position C

Evaluation indicator	Interpolation model	WD	SST	SLP	SAL	RH	DEN
MRE	**IFBN**	**0.171 9**	**0.039 3**	**0.036 0**	**0.022 9**	**0.130 5**	**0.031 7**
	CSI	0.243 7	0.069 7	0.094 2	0.039 6	0.233 7	0.061 4
	BP	0.194 1	0.052 7	0.037 9	0.040 0	0.174 9	0.042 2
TIC	**IFBN**	**0.143 4**	**0.003 3**	**0.003 2**	**0.000 4**	**0.026 6**	**0.003 5**
	CSI	0.178 2	0.005 5	0.001 9	0.007 7	0.078 4	0.011 9
	BP	0.148 3	0.003 7	0.004 6	0.003 1	0.053 2	0.006 3
R	**IFBN**	**0.614 4**	**0.967 2**	**0.962 1**	**0.955 8**	**0.629 1**	**0.906 7**
	CSI	0.578 2	0.920 4	0.838 2	0.861 3	0.513 4	0.740 2
	BP	0.601 5	0.949 8	0.907 6	0.916 5	0.574 7	0.820 7

Note: The bold numbers are the results obtained by proposed model.

5 Discussion and conclusions

The above interpolation experiments show that IFBN has great interpolation accuracy and stability. Compared with the classical regression-based interpolation method represented by CSI, IFBN takes full advantage of correlations among related variables. It places all related variables in a causal network for interpolation and information of all variables are complementary. Therefore, the performance of IFBN is less affected by the missing rate of data loss. Compared with the ML-based interpolation method represented by BP neural network, IFBN can clearly express causal relationships among variables in the form of probability distribution, rather than establishing an opaque mapping between input and output (so-called "black box") in BP. The causal relationship from data mining ensures the high accuracy of data interpolation.

Most importantly, whether regression-based interpolation models or ML algorithms, only single variable's time series can be interpolated with once model application. It is necessary to process six time series one by one and synchronous interpolation of multivariate time series is impossible. However, the problem will be handled in our proposed IFBN. Missing data of all ocean variables in the BN can be interpolated with only once model training. The nondirectional inference mechanism of BN avoids model retraining though the interpolation variable is different. The innovation of our proposed interpolation model for multiple ocean variables refers to: (1) mine interrelationships among variables with a visual network instead of a "black box", and quantitatively express the relationships with the probability distribution; (2) achieve the synchronous interpolation of multivariate time series efficiently through network reasoning, which avoids retraining model with the change of interpolative objects.

It can be seen from our interpolation experiments that interpolating effect of six ocean variables are different. The interpolating results of SST, SLP, SAL and DEN are better, while the interpolation result of RH is poor and WD is the worst. That may be caused by structural learning and parameter learning of BN. The structure and parameters of IFBN will be optimized in subsequent studies. In addition, IFBN is lacking in use of information about time series itself, and a dynamic Bayesian network will be introduced for time series interpolation.

References

Bai Chengzu, Hong Mei, Wang Dong, et al. 2014. Evolving an information diffusion model using a genetic algorithm for monthly river discharge time series interpolation and forecasting. Journal of Hydrometeorology, 15(6): 2236–2249, doi: 10.1175/JHM-D-13-0184.1

Barth A, Alvera-Azcárate A, Licer M, et al. 2020. DINCAE 1.0: a convolutional neural network with error estimates to reconstruct sea surface temperature satellite observations. Geoscientific Model Development, 13(3): 1609–1622, doi: 10.5194/gmd-13-1609-2020

Bouckaert R R. 1994. A stratified simulation scheme for inference in Bayesian belief networks. In: Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence. Seattle, WA: Morgan Kaufmann Publishers Inc, 110–117

Bu Fanyu, Chen Zhikui, Zhang Qingchen. 2014. Incomplete big data imputation algorithm based on deep learning. Microelectronics & Computer (in Chinese), 31(12): 173–176

Chickering D M. 2003. Optimal structure identification with greedy search. The Journal of Machine Learning Research, 3(3): 507–554

Chickering M, Geiger D, Heckerman D. 1995. Learning Bayesian networks: search methods and experimental results. In: Proceedings of Fifth Conference on Artificial Intelligence and Statistics. Lauderdale, FL: Society for Artificial Intelligence in Statistics

Cooper G F, Herskovits E. 1992. A Bayesian method for the induction of probabilistic networks from data. Machine Learning, 9(4): 309–347

Gasca M, Sauer T. 2000. Polynomial interpolation in several variables. Advances in Computational Mathematics, 12(4): 377, doi: 10.1023/A:1018981505752

Gong Yi, Dong Chen. 2010. Data patching method based on Bayesian network. Journal of Shenyang University of Technology (in Chinese), 32(1): 79–83

Huang Rong, Hu Zeyong, Guan Ting, et al. 2014. Interpolation of temperature data in northern Qinghai-Xizang Plateau and preliminary analysis on its recent variation. Plateau Meteorology (in Chinese), 33(3): 637–646

Jiang Dong, Fu Jingying, Huang Yaohuan, et al. 2011. Reconstruction of time series data of environmental parameters: methods and application. Journal of Geo-Information Science (in Chinese), 13(4): 439–446, doi: 10.3724/SP.J.1047.2011.00439

Kaplan A, Kushnir Y, Cane M A. 2000. Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. Journal of Climate, 13(16): 2987–3002, doi: 10.1175/1520-0442(2000)013<2987:RSOIOH>2.0.CO;2

Li H. 2006. Lost data filling algorithm based on EM and Bayesian network. Computer Engineering and Applications, 46(5): 123–125

Li Ming, Hong Mei, Zhang Ren. 2018a. Improved Bayesian network-based risk model and its application in disaster risk assessment. International Journal of Disaster Risk Science, 9(2): 237–248, doi: 10.1007/s13753-018-0171-z

Li Haitao, Jin Guang, Zhou Jinglun, et al. 2008. Survey of Bayesian network inference algorithms. Systems Engineering and Electronics (in Chinese), 30(5): 935–939

Li Ming, Liu Kefeng. 2018. Application of intelligent dynamic Bayesian network with wavelet analysis for probabilistic prediction of storm track intensity index. Atmosphere, 9(6): 224, doi: 10.3390/atmos9060224

Li Ming, Liu Kefeng. 2019. Causality-based attribute weighting via information flow and genetic algorithm for naive Bayes classifier. IEEE Access, 7: 150630–150641, doi: 10.1109/ACCESS.2019.2947568

Li Ming, Liu Kefeng. 2020. Probabilistic prediction of significant wave

height using dynamic Bayesian network and information flow. Water, 12(8): 2075, doi: 10.3390/w12082075

Li Ming, Zhang Ren, Hong Mei, et al. 2018b. Improved structure learning algorithm of Bayesian network based on information flow. Systems Engineering and Electronics (in Chinese), 40(6): 1385–1390

Liang Xiangsan. 2008. Information flow within stochastic dynamical systems. Physical Review: E, Statistical, Nonlinear, and Soft Matter Physics, 78(3): 031113

Liang Xiangsan. 2014. Unraveling the cause-effect relation between time series. Physical Review: E, Statistical, Nonlinear, and Soft Matter Physics, 90(5–1): 052150

Liang Xiangsan. 2015. Normalizing the causality between time series. Physical Review: E, Statistical, Nonlinear, and Soft Matter Physics, 92(2): 022126, doi: 10.1103/PhysRevE.92.022126

Liu Meiling, Liu Xiangnan, Liu Da, et al. 2015. Multivariable integration method for estimating sea surface salinity in coastal waters from in situ data and remotely sensed data using random forest algorithm. Computers & Geosciences, 75: 44–56

Liu Dayou, Wang Fei, Lu Yinan, et al. 2001. Research on learning Bayesian network structure based on genetic algorithms. Journal of Computer Research & Development (in Chinese), 38(8): 916–922

Liu Tian, Yang Kun, Qin Jun, et al. 2018. Construction and applications of time series of monthly precipitation at weather stations in the central and eastern Qinghai-Tibetan Plateau. Plateau Meteorology (in Chinese), 37(6): 1449–1457

Liu Junna, Zhang Yousheng. 2006. An adaptive joint tree algorithm. In: System Simulation Technology and Its Application Academic Exchange Conference Proceedings. Hefei: China System Simulation Society

Pearl J. 1998. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Berlin: Elsevier Inc

Sheng Zheng, Shi Hanqing, Ding Youzhuan. 2009. Using DINEOF method to reconstruct missing satellite remote sensing sea temperature data. Advances in Marine Science (in Chinese), 27(2): 243–249

Shi Zhifu. 2012. Bayesian Network Theory and its Application in Military System (in Chinese). Beijing: Defense Industry Press

Wang Tong, Yang Jie. 2010. A heuristic method for learning Bayesian networks using discrete particle swarm optimization. Knowledge and Information Systems, 24(2): 269–281, doi: 10.1007/s10115-009-0239-6

Xu Zilong, Xing Zuoxia, Ma Shichang. 2018. Wind power data missing data processing based on adaptive BP neural network. In: Proceedings of the 15th Shenyang Scientific Academic Annual Meeting. Shenyang: Shenyang Science and Technology Association

Yao Zizhen. 2006. A Regression-based K nearest neighbor algorithm for gene function prediction from heterogeneous data. BMC Bioinformatics, 7(1): S11, doi: 10.1186/1471-2105-7-11

Zhang Chan. 2013. A support vector machine-based missing values filling algorithm. Computer Applications and Software (in Chinese), 30(5): 226–228

Zheng Chongwei, Chen Yunge, Zhan Chao, et al. 2019. Source tracing of the swell energy: A case study of the Pacific Ocean. IEEE Access, 7: 139264–139275, doi: 10.1109/ACCESS.2019.2943903

Zheng Chongwei, Liang Bingchen, Chen Xuan, et al. 2020. Diffusion characteristics of swells in the North Indian Ocean. Journal of Ocean University of China, 19(3): 479–488, doi: 10.1007/s11802-020-4282-y

Zhou Zhihua. 2016. Machine Learning (in Chinese). Beijing: Tsinghua University Press

Zhu Ke. 2016. Bootstrapping the portmanteau tests in weak auto-regressive moving average models. Journal of the Royal Statistical Society: Series B, 78(2): 463–485, doi: 10.1111/rssb.12112

Appendix

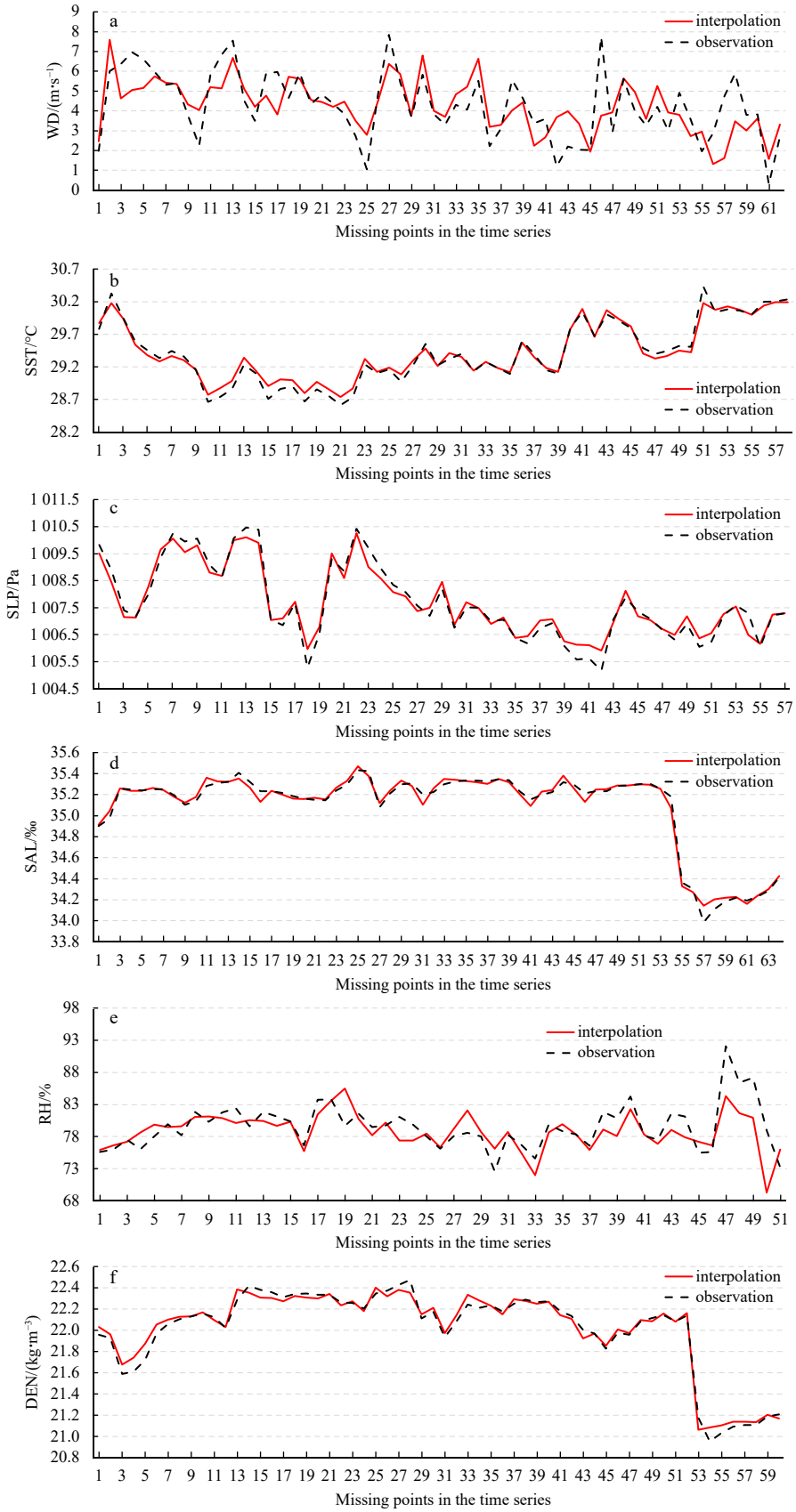


Fig. A1. Interpolation results with the data missing rate of 50%: WD (a), SST (b), SLP (c), SAL (d), RH (e), DEN (f).

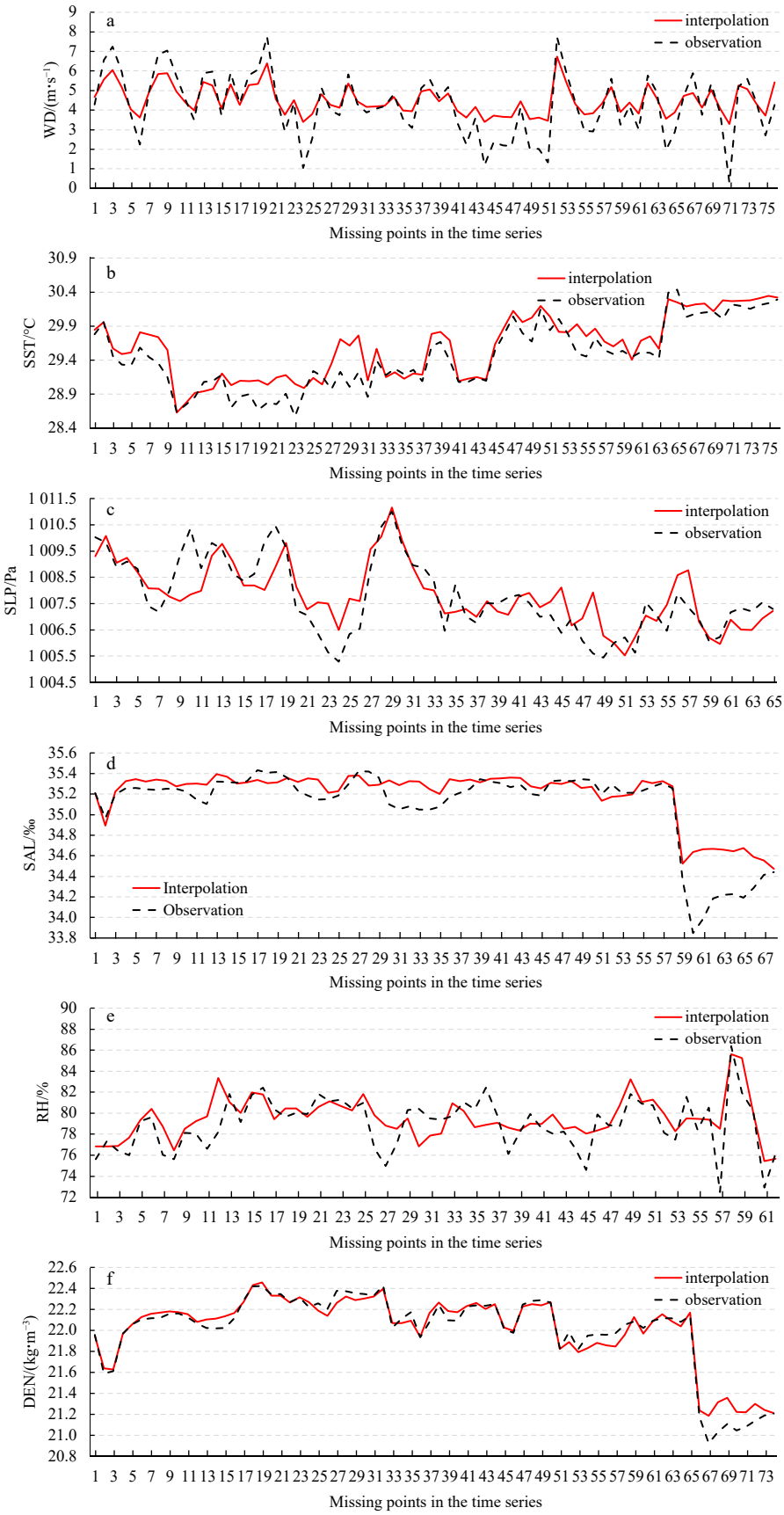


Fig. A2. Interpolation results with the data missing rate of 60%: WD (a), SST (b), SLP (c), SAL (d), RH (e), DEN (f).