

Research Article

Design of Deep Belief Networks for Short-Term Prediction of Drought Index Using Data in the Huaihe River Basin

Junfei Chen,^{1,2} Qiongji Jin,^{1,2} and Jing Chao³

¹ State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

² Business School, Hohai University, Nanjing 210098, China

³ State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China

Correspondence should be addressed to Junfei Chen, chenjunfei@yahoo.com.cn

Received 18 January 2012; Revised 20 February 2012; Accepted 21 February 2012

Academic Editor: Ming Li

Copyright © 2012 Junfei Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the global climate change, drought disasters occur frequently. Drought prediction is an important content for drought disaster management, planning and management of water resource systems of a river basin. In this study, a short-term drought prediction model based on deep belief networks (DBNs) is proposed to predict the time series of different time-scale standardized precipitation index (SPI). The DBN model is applied to predict the drought time series in the Huaihe River Basin, China. Compared with BP neural network, the DBN-based drought prediction model has shown better predictive skills than the BP neural network for the different time-scale SPI. This research can improve drought prediction technology and be helpful for water resources managers and decision makers in managing drought disasters.

1. Introduction

With the global environmental degradation and water resource shortages, droughts are becoming increasingly eye catching and have aroused the attention of many countries and regions. Drought is considered the most complex but least understood of all natural hazards, affecting more people than any other disasters [1]. In recent years, drought disasters continuously happened and caused serious impact on production and life in China. The losses caused by drought ranked the first in all natural hazards in China [2]. For example, in the extreme drought in Southwest China during 2009 to 2010, five provinces and cities suffered droughts which have seriously threatened people's life and economic production activities. The Chinese northern region also suffered severe drought in 2011. Long and

severe droughts have direct impacts on industrial production, people's lives, and ecological environment and even lead to desertification and other natural disasters. Droughts have become serious constraints to the sustainable development of Chinese society and economy [3, 4]. Drought prediction is an important content in the planning and management of water resource systems of a river basin. How to effectively monitor and forecast droughts has become the research focus, which can help to take effective strategies and measures to mitigate the damages of droughts.

There are some forecasting methods used in drought prediction fields. Lohani and Loganathan used a nonhomogeneous Markov chain model to characterize the stochastic behavior of drought, and an early-warning system in the form of a decision tree enumerating is proposed for drought management [5]. Jia et al. established a grey-time series combined method (GTCM) to predict annual precipitation of Huangcun Meteorological Observation, Daxing county, Beijing [6]. Yang et al. proposed a chaotic Bayesian method based on multiple criteria decision making to forecast nonlinear hydrological time series, which can be applied in drought forecast [7]. The predictability of drought severity from spatiotemporal varying indices of large-scale climate phenomena was studied by integrating linear and nonlinear statistical data models, and the model was used for the Murray-Darling Basin (MDB) in Australia [8]. Meteorological droughts were characterized using the standardized precipitation index (SPI) developed by McKee et al. [9]. Drought classes based on standardized precipitation index (SPI) values were derived by Markov chain model in Alentejo, Southern Portugal [10]. Peng et al. used weighted Markov chain to predict the future drought index, weighted by the standardized self-coefficients. The drought indexes of Nanjing city from 1959 to 2004 were a specific application with this method and satisfactory results were obtained [11]. SPI is calculated from monthly precipitation data collected from 36 weather stations in Guanzhong plain and Weibei tablelands, and the Markov chain model with weights was applied to predict SPI drought intensity by using standardized self-coefficients as weights [12]. The vegetation temperature condition index (VTCI) based on remote sensing data is used for drought monitoring. The ARIMA models were developed to simulate the VTCI series and be used in Guanzhong Plain in China [13]. The loglinear modeling for three-dimensional contingency tables was used for short-term prediction of drought severity classes. The results show that three-dimensional loglinear modeling of monthly drought class transitions is able to capture the trends for both drought initiation, establishment, and drought dissipation [14]. Mishra and Desai compared linear stochastic model (ARIMA/SARIMA), recursive multistep neural network (RMSNN), and direct multisteps neural network (DMSNN) for drought forecasting by using standardized precipitation index (SPI) series as drought index in the Kansabati River Basin in India [15]. Traditionally, forecasting research and practice has been dominated by conventional statistical methods. Recently, the study of long range or long memory has received many attentions in forecasting. Hurst developed a test for long-range dependence and found significant long-term correlations among fluctuations in the Nile's outflows and described these correlations in terms of power laws [16]. Mathematical models with long-range dependence were first introduced to statistics by Mandelbrot and his workers [17–19]. Long-range dependence is often encountered in practice, not only in hydrology, geophysics, and finance, but also in all fields of statistical applications [20–24]. Pelletier and Turcotte present power spectra of time-series data for tree ring width chronologies, atmospheric temperatures, river discharges, and precipitation averaged over hundreds of stations worldwide. They thought that long-range persistence can have a dramatic effect on the likelihood of severe hydrologic drought and computed recurrence intervals for droughts of different magnitudes,

durations, and coefficients of variation [25]. Radziejewski and Kundzewicz computed fractal dimensions of crossings of Warta flows by using a novel variant of the box-counting method, and spectral properties are compared between the time series of flows [26]. Li et al. computed long-range dependence (LRD) of sea level and thought that sea level is multiscaled and heavy tailed [27].

Recently, deep belief networks (DBNs) are proposed by Hinton. The DBN is a probabilistic generative model, the bottom layer is observable, and the multiple hidden layers are created by stacking multiple restricted Boltzmann machines (RBMs) on top of each other [28]. Hinton et al. derived a way to perform fast, greedy learning of deep belief networks (DBN) one layer at a time, with the top two layers forming an undirected bipartite graph [29]. DBNs and restricted Boltzmann machines (RBMs) have already been applied successfully to solve many problems [30]. Lee et al. present a convolutional deep belief networks and are used to scale the realistic image sizes [31]. A novel text classification approach based on deep belief networks is proposed, and the proposed method outperforms traditional classifier based on support vector machine [32]. Zhou et al. present a discriminative deep belief networks (DDBNs) to address the image classification problem with limited labeled data. Experiments on the artificial dataset and real image datasets show that DDBN outperforms most semisupervised algorithms [33]. Chao et al. proposed a deep belief network (DBN) to forecast the foreign exchange rate. In their experiments, both British pound/US dollar and Indian rupee/US dollar exchange rates are forecasted, and the results show that deep belief networks (DBNs) achieve better performance than feed-forward neural networks [34]. Deep learning techniques have also been shown to perform significantly better than other techniques for problems such as image classification and handwriting analysis [31].

In this paper, we propose a deep belief network (DBN) for short-term prediction of drought index. The aims of this study are to present and evaluate the performance of DBN model as a drought prediction method. This model was applied to forecast drought index using standardized precipitation index (SPI) series in the Huaihe River Basin, China. The results are compared and analyzed with BP neural network for demonstration of the validity of the DBN model. The remainder of the paper is organized as follows. In Section 2, the standardized precipitation index (SPI) and BP neural network are introduced, and the deep belief networks (DBN) model for drought index prediction is proposed. In Section 3, a case is studied, and discussions are arranged. Finally in Section 4, the main conclusions and a discussion for future work are given.

2. Methodology

2.1. Standardized Precipitation Index (SPI)

The SPI was formulated by Mckee et al. of the Colorado Climate Center in 1993. The purpose is to assign a single numeric value to the precipitation which can be compared across regions with markedly different climates [11]. The SPI is an index based on the probability of precipitation for any time scale. Technically, the SPI is the number of standard deviations that the observed value would deviate from the long-term mean, for a normally distributed random variable. The SPI can be computed for different time scales and can provide early warning of drought and help assess drought severity. The SPI is a probability index that considers only precipitation, while Palmer's indices are water balance indices that consider water supply (precipitation), demand (evapotranspiration), and loss (runoff). So, SPI is less

complex than PDSI [35]. Now, the standardized precipitation index (SPI) is widely accepted and used throughout the world [36]. The computing procedure of the SPI value is as follows [37, 38].

Assuming that a precipitation series of some time scale is x , then its probability density function of Γ distribution is expressed as

$$f(x) = \frac{1}{\beta^\gamma \Gamma(\gamma)} x^{\gamma-1} e^{-x/\beta}, \quad x > 0, \quad (2.1)$$

where $\Gamma(\gamma)$ is a gamma function and $\Gamma(\gamma) = \int_0^\infty x^{\gamma-1} e^{-x} dx$. β and γ are the shape parameter and the scale parameter, respectively, and $\beta > 0$, $\gamma > 0$. The precipitation value $x > 0$.

The shape and scale parameters can be estimated by the maximum likelihood method as follows:

$$\hat{\gamma} = \frac{1 + \sqrt{1 + 4A/3}}{4A}, \quad (2.2)$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\gamma}},$$

where $A = \ln(\bar{x}) - (1/n) \sum_{i=1}^n \ln x_i$, n stands for the number of precipitation observations, x_i are the samples of the precipitation data, and \bar{x} is the mean of these samples.

The gamma distribution is not defined for $x = 0$; however, the actual precipitation can be 0. Therefore, cumulative probability of precipitation for a certain time scale can be calculated using the following formula [38, 39]:

$$H(x) = u + (1 - u)F(x), \quad (2.3)$$

where $F(x) = (1/\Gamma(\hat{\gamma})) \int_0^x t^{\hat{\gamma}-1} e^{-t} dt$ and $t = x/\hat{\beta}$. u is the probability of zero precipitation and can be calculated as m/n . m is the total number of precipitation series, and n is the number of zeros in the precipitation series.

The cumulative probability, $H(x)$, is then transformed to the standard normal random variable with mean as zero and variance as one. Following Edwards and Mckee [40] and Hughes and Saunders [41], SPI can be obtained as follows:

$$\text{SPI} = \begin{cases} -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} \right) & \text{for } 0 < H(x) \leq 0.5, \\ t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} & \text{for } 0.5 < H(x) < 1, \end{cases} \quad (2.4)$$

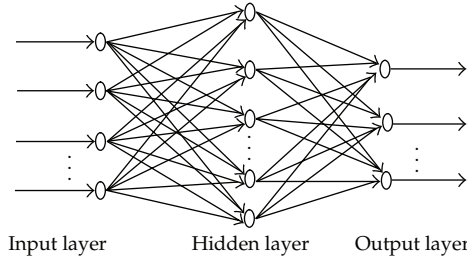


Figure 1: The structure of BP neural network.

where

$$t = \begin{cases} \sqrt{\ln\left(\frac{1}{H^2(x)}\right)} & \text{for } 0 < H(x) \leq 0.5, \\ \sqrt{\ln\left(\frac{1}{(1-H(x))^2}\right)} & \text{for } 0.5 < H(x) < 1. \end{cases} \quad (2.5)$$

In (2.4), the c_i and d_i are parameters during the computing process and $c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, and $d_3 = 0.001308$.

According to SPI, drought can be classified. When the value of SPI is continuously negative, a drought event occurs. The event ends when the SPI becomes positive.

2.2. Backpropagation Neural Network (BPNN)

The BP neural network is a kind of multilayer feed-forward networks with training by error backpropagation algorithm [42]. It is a kind of supervised learning neural network, the principle behind which involves using the steepest gradient descent method to reach any small approximation. A general model of the BP neural network has a structure as described in Figure 1.

In Figure 1, there are three layers contained in BP: input layer, hidden layer, and output layer. Two nodes of each adjacent layer are directly connected, which is called a link. Each link has a weighted value presenting the relational degree between two nodes. The algorithm of BP neural network is to input the training samples from the input layer and then obtain the calculation output through the operation of corresponding thresholds, functions, and connection weights between nodes [42, 43]. The node function has usually selected S-type function as follows:

$$f(x) = \frac{1}{1 + e^{-x/Q}}. \quad (2.6)$$

The Q in the equation is a Sigmoid parameter which is the form of adjusted activation function, and the specific algorithm is introduced in [44]. The output error is obtained by the comparison between the calculation output and the sample output. If the error does not meet the requirements, the network weights and thresholds usually are adjusted along the

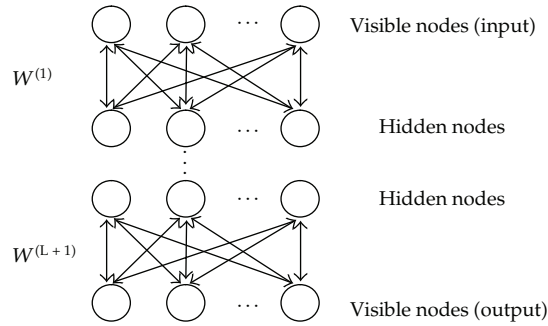


Figure 2: A DBN structure with L hidden layers.

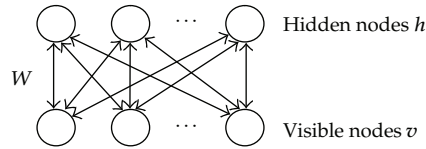


Figure 3: A RBM structure.

negative gradient direction of network error and finally reach the minimum network error [45]. The number of hidden layer nodes is firstly determined by employing an empirical formula in the design stage and finally adjusted by comparing the efficiencies of different numbers of hidden layer nodes in neural network training stage [46].

2.3. Deep Belief Networks

A deep belief network (DBN) is a generative model with an input layer and an output layer, separated by many layers of hidden stochastic units. The multilayer neural network can efficiently be trained by composing RBMs using the feature activations of one layer as the training data for the next. Figure 2 shows an example of a DBN structure [28].

Usually a DBN consists of two kinds of different layers. They are visible layer and hidden layer. Visible layers contain input nodes and output nodes, and hidden layers contain hidden nodes. Hinton et al. proposed a greedy layerwise unsupervised learning algorithm for DBNs which is based on sequence training with restricted Boltzmann machines (RBMs) [28, 34]. A restricted Boltzmann machine (RBM) is composed of two different layers of units, with weighted connection between them. It consists of one layer of visible nodes (neurons) and one layer of hidden units. Figure 3 shows an RBM structure. Nodes in each layer have no connections between them and are connected to all other units in another layer. Connections between nodes are bidirectional and symmetric. Restricted Boltzmann machines (RBMs) have been used as generative models of many different types of data including labeled or unlabeled images windows of mel-cepstral coefficients that represent speech, and so on. Their most important use is as learning modules that are composed to form deep belief nets [28].

Let v_i and h_j represent the states of visible node i and hidden node j , respectively. For binary state nodes, that is, v_i and $h_j \in \{0, 1\}$, the state of h_j is set to 1 with probabilities [47]:

$$p_{h_j} = p(h_j = 1 | v) = \sigma\left(b_j + \sum_i w_{ij} v_i\right), \quad (2.7)$$

where $\sigma(x)$ is the logistic sigmoid function $1/(1 + \exp(-x))$, b_j is the bias of j , and v_i is the binary state. w_{ij} is the weight between v_i and h_j . After binary states have been chosen for the hidden units, then set the state of v_i to be 1 with probability

$$p_{v_i} = p(v_i = 1 | h) = \sigma\left(b_i + \sum_j w_{ij} h_j\right). \quad (2.8)$$

The training process of the RBM is described as follows. Firstly, a training sample is presented to the visible nodes, and the $\{v_i\}$ is obtained. Then the hidden nodes state that $\{h_j\}$ are sampled according to probabilities in (2.7). This process is repeated once more to update the visible and then the hidden nodes, and the one-step "reconstructed" states v'_i and h'_j are obtained. The update in a weight is given as follows:

$$\Delta w_{ij} = \eta \left(\langle v_i h_j \rangle - \langle v'_i h'_j \rangle \right), \quad (2.9)$$

where η is the learning rate, and $\langle \cdot \rangle$ refers to the expectation of the training data.

A continuous restricted Boltzmann machine (CRBM) is considered by Chao et al. [34] and Chen and Murray [48]. Suppose the inputs nodes with state $\{s_i\}$, then the output nodes s_j can be computed as follows:

$$s_j = \varphi_j \left(\sum_i w_{ij} s_i + \sigma \cdot N_j(0, 1) \right), \quad (2.10)$$

where $\varphi_j(x)$ is a sigmoid function with lower and upper asymptotes at θ_L and θ_H , $\varphi_j(x_j) = \theta_L + (\theta_H - \theta_L) \cdot (1 / (1 + e^{-a_j x_j}))$. $N_j(0, 1)$ represents a unit Gaussian. σ is a constant, and parameter a_j is a "noise-control" parameter which controls the slope of the sigmoid function [49]. The update equations for w_{ij} and a_j are

$$\begin{aligned} \Delta w_{ij} &= \eta_w \left(\langle s_i s_j \rangle - \langle s'_i s'_j \rangle \right), \\ \Delta a_j &= \frac{\eta_a}{a_j^2} \left(\langle s_j^2 \rangle - \langle s'^2_j \rangle \right), \end{aligned} \quad (2.11)$$

where η_w and η_a represent the learning rates, s'_j denotes the one-step sampled state of node j , and $\langle \cdot \rangle$ refers to the expectation of the training data. We train sequentially as many RBMs

as the number of hidden layers in the DBN to construct a DBN model. We adopt the learning algorithm according to [28, 34, 50]. The method of stacking CRBMs makes it possible to train many layers of hidden units efficiently and is one of the most common deep learning strategies. As each new layer is added, the overall generative model gets better. This process of learning is continued until a prescribed number of hidden layers in the DBN have been trained. In order to apply DBN model to drought prediction using SPI series, the DBN model with two hidden layers is selected in this paper. The main steps using DBN model for drought index prediction are as follows.

Step 1. Compute the different time-scale SPI series by precipitation data.

The different time-scale SPI series are computed by precipitation data by the description method in Section 2.1, and different time-scale SPI series are obtained.

Step 2. Normalize the SPI series by formula (2.12) as follows:

$$SPI' = \frac{SPI - SPI_{\min}}{SPI_{\max} - SPI_{\min}}, \quad (2.12)$$

where SPI' and SPI represent the normalized and original SPI data, respectively. The SPI_{\min} represents the minimum value of the corresponding SPI series, and SPI_{\max} represents the maximum value of the corresponding SPI series.

Step 3. Determine the optimal network structure by experiments.

Determine the number of input nodes, the numbers of the first hidden and second hidden nodes, and weight coefficients by learning algorithm. The data of SPI series are split into two parts. The first part is used as a training sample, and the rest is used as a testing sample. During the training process, the network structures for different time-scale SPI series are determined according to the criterion of smallest RMSE and MAE.

Step 4. Forecast drought index based on DBN model and results analysis.

3. Case Study

3.1. Experimental Design

We use four data sets of precipitation in the experiments. Four hydrologic stations were considered in this study. They are Bengbu, Fuyang, Xuchang, and Zhumadian in Huaihe River Basin which is located in the eastern part of China. Data sets contain monthly precipitation during 1958–2006. These data are used to calculate four different time scales of standardized precipitation index (SPI), that is, SPI3, SPI6, SPI9, and SPI12. Taking the SPI3 as an example, all of the SPI sets are divided into two parts. The observations during 1958–1999 are as training set, and the remaining observations during 2000–2006 are as testing set.

Our purpose of this research is to explore if the DBN model can be used well in drought prediction by using the monthly rainfall data of four hydrologic stations from January 1958 to 2006 to calculate different time scales of SPI in Huaihe River Basin.

Table 1: The CRBM results of Bengbu SPI3.

Number of input nodes	Number of hidden nodes	RMSE	MAE	Number of input nodes	Number of hidden nodes	RMSE	MAE
2	5	0.1270	0.0980	7	5	0.1107	0.0869
	10	0.1276	0.0983		10	0.1114	0.0877
	15	0.1306	0.1002		15	0.1113	0.0878
	20	0.1276	0.0983		20	0.1113	0.0877
	25	0.1274	0.0982		25	0.1123	0.0885
3	5	0.1270	0.0987	8	5	0.1113	0.0865
	10	0.1293	0.1000		10	0.1120	0.0871
	15	0.1272	0.0988		15	0.1115	0.0866
	20	0.1274	0.0989		20	0.1117	0.0869
	25	0.1273	0.0988		25	0.1097	0.0864
4	5	0.1292	0.0993	9	5	0.1105	0.0863
	10	0.1291	0.0990		10	0.1117	0.0876
	15	0.1283	0.0988		15	0.1119	0.0876
	20	0.1301	0.0995		20	0.1120	0.0876
	25	0.1286	0.0992		25	0.1181	0.0945
5	5	0.1225	0.0929	10	5	0.1112	0.0874
	10	0.1230	0.0927		10	0.1116	0.0875
	15	0.1235	0.0931		15	0.1123	0.0884
	20	0.1232	0.0928		20	0.1115	0.0885
	25	0.1248	0.0939		25	0.1121	0.0886
6	5	0.1235	0.0931		5	0.1235	0.0931
	10	0.1239	0.0930		10	0.1239	0.0930
	15	0.1236	0.0934		15	0.1236	0.0934
	20	0.1235	0.0930		20	0.1235	0.0930
	25	0.1244	0.0942		25	0.1244	0.0942

In this paper, we use two criteria to evaluate the performance of a DBN in drought forecasting. They are root mean square error (RMSE) and mean absolute error (MAE). The formulas of this two predictive accuracy measures are listed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^T (y_i - y'_i)^2}{T}}, \quad (3.1)$$

$$\text{MAE} = \frac{\sum_{i=1}^T |y_i - y'_i|}{T},$$

where y_i is the observations of SPI, y'_i is the predicted SPI values, and T is the total number of predictions.

Table 2: The DBN results of Bengbu SPI3.

Number of input nodes	Number of first hidden nodes	Number of second hidden nodes	RMSE	MAE	Number of input nodes	Number of first hidden nodes	Number of second hidden nodes	RMSE	MAE		
8	5	5	0.6924	0.5497	9	5	5	0.6905	0.5510		
		10	0.6911	0.5567			10	0.6842	0.5425		
		15	0.7103	0.5787			15	0.6858	0.5473		
		20	0.7006	0.5712			20	0.7138	0.5860		
		25	0.8051	0.6858			25	0.7068	0.5769		
	10	5	0.7048	0.5702	20	10	5	0.6881	0.5472		
		10	0.6911	0.5506			10	0.6899	0.5479		
		15	0.6962	0.5616			15	0.6954	0.5548		
		20	0.7147	0.5874			20	0.7090	0.5775		
		25	0.7749	0.6606			25	0.7032	0.5805		
	15	5	5	0.6852	0.5453	25	5	5	0.6930	0.5545	
			10	0.6926	0.5596			10	0.7165	0.5887	
			15	0.6969	0.5636			15	0.7300	0.6106	
			20	0.7265	0.6113			20	0.7218	0.6070	
			25	0.7316	0.6236			25	0.7488	0.6375	
		20	5	0.7224	0.6001		25	10	5	0.6915	0.5516
			10	0.6923	0.5525				10	0.6962	0.5592
			15	0.7152	0.5940				15	0.7407	0.6245
			20	0.6955	0.5560				20	0.7140	0.5885
			25	0.7794	0.6679				25	0.7309	0.6136
25	5	0.6935	0.5570	25	15	5	0.6919	0.5563			
	10	0.6877	0.5496			10	0.7013	0.5637			
	15	0.7037	0.5761			15	0.6949	0.5617			
	20	0.7470	0.6322			20	0.7669	0.6566			
	25	0.7174	0.5868			25	0.7214	0.6104			

We use the learning sample to find an optimal network structure for these four different time-scales SPI. Taking the SPI3 of Bengbu data as an example, we explain how to determine an optimal network structure. In our experiment, the DBN has two hidden layers. The key for our experiment is to determine the numbers of input and hidden nodes. We determine the optimal number of input nodes and two hidden layer nodes by experiments. On one hand, neural networks with too few hidden nodes may not have enough power to model the data. On the other hand, neural networks with too many hidden nodes may lead to overfitting problems and finally result in poor forecasting performance [30]. In our experiment, the number of input nodes and hidden nodes of the DBN network structures is selected by experimentation. The number of input nodes ranges from 2 to 10. Because the forecasting performance of neural networks is not as sensitive to the number of hidden nodes as to the number of input nodes, so the number of hidden nodes is selected by five levels, that is, 5, 10, 15, 20, and 25. We did the experiment for 45 times to find the optimal structure of DBN. We compared the RMSE and MAE, and we determined the number of every layer node. The results are shown in Table 1.

Table 3: The optimal network structures of DBN.

SPI series	Station	Number of input nodes	Number of first hidden nodes	Number of second hidden nodes
SPI3	Bengbu	9	5	10
	Fuyang	8	20	15
	Xuchang	8	20	15
	Zhumadian	8	20	15
SPI6	Bengbu	10	5	10
	Fuyang	7	5	5
	Xuchang	7	5	5
	Zhumadian	8	5	5
SPI9	Bengbu	10	5	5
	Fuyang	7	5	5
	Xuchang	6	5	5
	Zhumadian	10	5	15
SPI12	Bengbu	8	5	5
	Fuyang	10	5	5
	Xuchang	9	5	5
	Zhumadian	10	5	5

In Table 1, we find when the CRBM structure is 8-25, the RMSE is the smallest, and when the CRBM structure is 9-5, the MAE is the smallest. We can find that the most optimal structure is most likely to appear when the number of input nodes is 8 or 9. Then we do the next step. The results of the next step have just been shown in Table 2. We can find that the best DBN structure is 9-5-10-1. That is, the DBN has 9 input nodes, 5 nodes in the first hidden layer, 10 nodes in the second hidden layer, and 1 output node, and the RMSE and MAE are the smallest of all.

According to above processes, we can determine the optimal structures of DBN for the four stations and different time-scale SPI series. We try nine levels of input nodes from 2 to 10 in combination with five hidden nodes (5, 10, 15, 20, and 25) for CRBM training. We can find the optimal network structure in a similar way for all of the SPI series. The optimal network structures of DBN for the different four stations and different time-scale SPI series are shown in Table 3.

3.2. Results and Discussion

In this paper, the DBN and BP neural network model are used for forecasting the different time-scale SPI series, and the results of their prediction are compared. The quantitative performance evaluations of DBN and BP neural network are carried out by using RMSE and MAE. The results are shown in Table 4.

We can find that the prediction errors of the DBN are smaller than the prediction errors of BP neural network in Table 4. The errors results demonstrate that DBN model is suitable for the drought prediction in the Huaihe River Basin. DBN model can obtain smaller RMSE and MAE compared with BP neural network. With the change of the time scale of SPI from little to large, the RMSE and MAE become smaller. That is, the fitting

Table 4: The comparison of RMSE and MAE between BP and DBN.

Station	Model	Errors	SPI3	SPI6	SPI9	SPI12
Bengbu	DBN	RMSE	0.6842	0.6592	0.5355	0.4797
		MAE	0.5425	0.5274	0.3959	0.3553
	BP neural network	RMSE	0.9897	0.6987	0.5899	0.5809
		MAE	0.7564	0.5523	0.4157	0.4532
Fuyang	DBN	RMSE	0.8112	0.6634	0.5590	0.5620
		MAE	0.6527	0.4812	0.3923	0.4282
	BP neural network	RMSE	1.0876	0.8022	0.8032	0.5773
		MAE	0.8202	0.5867	0.5509	0.4080
Xuchang	DBN	RMSE	0.7258	0.5764	0.5262	0.4236
		MAE	0.5714	0.4342	0.3880	0.2976
	BP neural network	RMSE	0.8223	0.6938	0.6783	0.4454
		MAE	0.6786	0.5411	0.4725	0.3268
Zhumadian	DBN	RMSE	0.7794	0.6239	0.5686	0.4990
		MAE	0.6276	0.4792	0.3811	0.3504
	BP neural network	RMSE	1.0780	0.7956	0.7996	0.4474
		MAE	0.8147	0.6336	0.5365	0.3144

results of SPI12 are better than SPI9, SPI9 is better than SPI6, and SPI6 is better than SPI3. In a word, DBN has a higher precision in drought prediction based on SPI than BP neural network.

Figure 4 shows the test results of SPI3, SPI6, SPI9, and SPI12 of Bengbu station. It is obvious that the prediction values of different time-scale SPI series are very close to the actual ones. The comparison results between observations and predicted data of Fuyang station, Xuchang station, and Zhumadian station are shown in Figures 5, 6, and 7 using DBN and BP neural network for SPI6 series.

From Figures 5, 6, and 7, the predicted data of SPI based on DBN model agreed with observations very well. The majority of DBN outputs are nearer to the real SPI values than those of BP neural network. The results show that the DBN model is appropriate for short term of drought index and can obtain higher precision.

4. Conclusion

In this paper, we proposed a deep belief network (DBN) for short-time drought index prediction. The forecasting model based on DBN is used to forecast different time-scale SPI series of four stations in Huaihe River Basin, China. Compared with the BP neural network, the DBN-based model is more reliable and efficient for short-term prediction of drought index. The errors results show that the DBN model outperforms the BP neural network. This study shows that the DBN model is a useful tool for drought prediction. Due to the complexity of the formation mechanism of the drought disasters and the long memory of hydrological data, some new method which can deal with long-range dependence will be

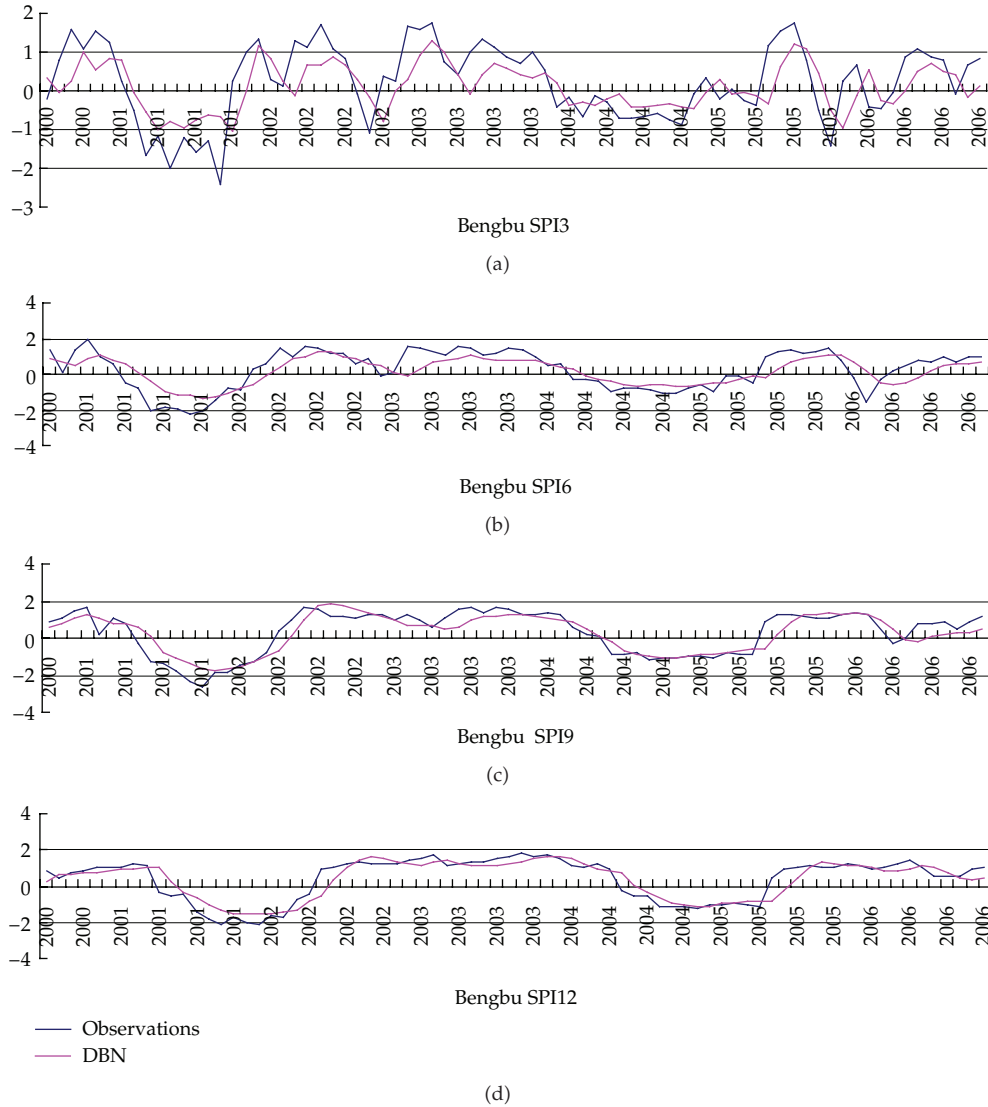


Figure 4: Results comparison between observations and predicted data using DBN for different time-scale SPI series of Bengbu station.

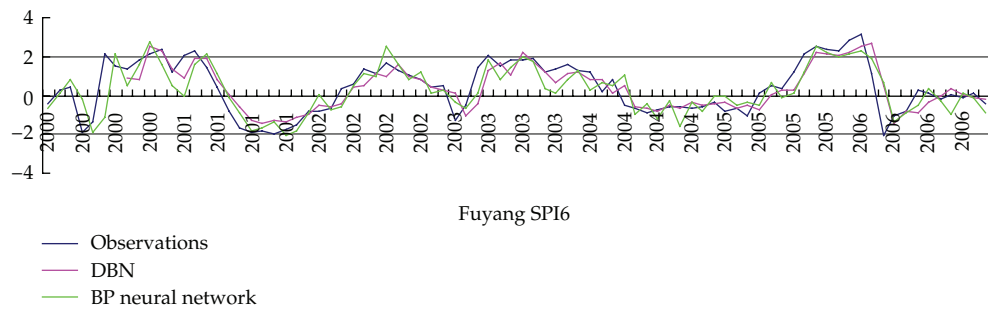


Figure 5: A comparison of DBN and BP neural network for SPI6 series of Fuyang station.

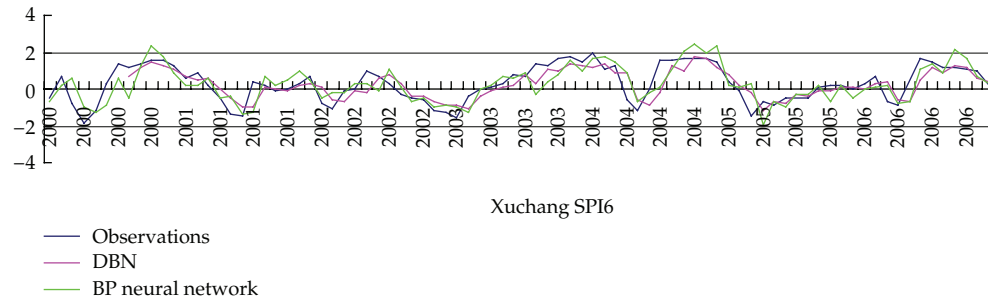


Figure 6: A comparison of DBN and BP neural network for SPI6 series of Xuchang station.

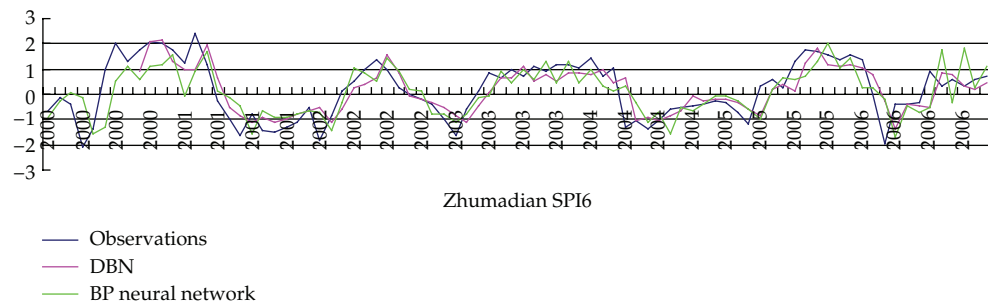


Figure 7: A Comparison of DBN and BP neural network for SPI6 series of Zhumadian station.

thought about, and further studies are needed to deal with more complex situations for drought prediction.

Acknowledgments

This work was supported partially by the National Society Science Fund of China (09CJY020), the National Nature Science Foundation of China (90924027), the Special Fund of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, China (2011585312), Public-Interest Industry Project of Chinese Ministry of Water Resources (200801027), the Science and Technology Projects of Yunnan Province (2010CA013), the Fundamental Research Funds for the Central Universities of Hohai University, and 2010 Jiangsu Province Qing Lan Project.

References

- [1] D. A. Wilhite, "Drought as a natural hazard: concepts and definitions," in *Drought: A Global Assessment*, D. A. Wilhite, Ed., Routledge Hazards and Disasters Series, Routledge, New York, NY, USA, 2000.
- [2] G. Hagman, *Prevention Better than Cure: Report on Human and Environmental Disasters in the Third World*, Swedish Red Cross, Stockholm, Sweden, 1984.
- [3] L. C. Song, Z. Y. Deng, and A. X. Dong, *Drought Disaster*, China Meteorological Press, Beijing, China, 2003.
- [4] R. H. Huang, C. Y. Li, and S. W. Wang, *Climatic Disasters*, China Meteorological Press, Beijing, China, 2003.

- [5] V. K. Lohani and G. V. Loganathan, "An early warning system for drought management using the Palmer Drought Index," *Journal of the American Water Resources Association*, vol. 33, no. 6, pp. 1375–1386, 1997.
- [6] H. F. Jia, Y. Q. Zheng, Y. Y. Ding, and B. Cao, "Grey-time series combined forecasting model and its application in annual precipitation," *Systems Engineering-Theory & Practice*, vol. 8, pp. 122–126, 1998.
- [7] X. H. Yang, D. X. She, Z. F. Yang, Q. H. Tang, and J. Q. Li, "Chaotic bayesian method based on multiple criteria decision making (MCDM) for forecasting nonlinear hydrological time series," *International Journal of Nonlinear Sciences and Numerical Simulation*, vol. 10, no. 11-12, pp. 1595–1610, 2009.
- [8] A. P. Barros and G. J. Bowden, "Toward long-lead operational forecasts of drought: an experimental study in the Murray-Darling River Basin," *Journal of Hydrology*, vol. 357, no. 3-4, pp. 349–367, 2008.
- [9] T. B. McKee, N. J. Doeskin,, and J. Kleist, "The relationship of drought frequency and duration to time scales," in *Proceedings of the 8th Conference on Applied Climatology*, pp. 179–184, American Meteorological Society, Boston, Mass, USA, 1993.
- [10] A. A. Paulo, E. Ferreira, C. Coelho, and L. S. Pereira, "Drought class transition analysis through Markov and Loglinear models, an approach to early warning," *Agricultural Water Management*, vol. 77, no. 1–3, pp. 59–81, 2005.
- [11] S. Z. Peng, Z. Wei, C. Y. Dou, and J. Z. Xu, "Model for evaluating the regional drought index with the weighted Markov chain and its application," *Xitong Gongcheng Lilun yu Shijian/System Engineering Theory and Practice*, vol. 29, no. 9, pp. 173–178, 2009.
- [12] Y. J. Wang, J. M. Liu, P. X. Wang et al., "Prediction of drought occurrence based on the standardized precipitation index and the Markov chain model with weights," *Agricultural Research in the Arid Areas*, vol. 5, pp. 198–203, 2007.
- [13] P. Han, P. X. Wang, S. Y. Zhang, and D. H. Zhu, "Drought forecasting based on the remote sensing data using ARIMA models," *Mathematical and Computer Modelling*, vol. 51, no. 11-12, pp. 1398–1403, 2010.
- [14] E. E. Moreira, C. A. Coelho, A. A. Paulo, L. S. Pereira, and J. T. Mexia, "SPI-based drought category prediction using loglinear models," *Journal of Hydrology*, vol. 354, no. 1–4, pp. 116–130, 2008.
- [15] A. K. Mishra and V. R. Desai, "Drought forecasting using feed-forward recursive neural network," *Ecological Modelling*, vol. 198, no. 1-2, pp. 127–138, 2006.
- [16] H. E. Hurst, "Long term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 770–799, 1951.
- [17] B. B. Mandelbrot and J. R. Wallis, "Noah, Joseph and operational hydrology," *Water Resource Research*, vol. 4, pp. 908–918, 1968.
- [18] B. B. Mandelbrot and J. R. Wallis, "Computer experiments with fractional Gaussian noises," *Water Resource Research*, vol. 5, pp. 228–267, 1969.
- [19] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, pp. 422–437, 1968.
- [20] J. Beran, "Statistical method for data with long-range dependence," *Statistical Science*, vol. 7, no. 4, pp. 404–416, 1992.
- [21] M. Li and J. Y. Li, "On the predictability of long-range dependent series," *Mathematical Problems in Engineering*, vol. 2010, Article ID 397454, 2010.
- [22] M. Li, "Fractal time series—a tutorial review," *Mathematical Problems in Engineering*, Article ID 157264, 26 pages, 2010.
- [23] M. Li and S. C. Lim, "Modeling network traffic using generalized Cauchy process," *Physica A*, vol. 387, no. 11, pp. 2584–2594, 2008.
- [24] M. Li and W. Zhao, "Visiting power laws in cyber-physical networking systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 302786, 13 pages, 2012.
- [25] J. D. Pelletier and D. L. Turcotte, "Long-range persistence in climatological and hydrological time series: analysis, modeling and application to drought hazard assessment," *Journal of Hydrology*, vol. 203, no. 1–4, pp. 198–208, 1997.
- [26] M. Radziejewski and Z. W. Kundzewicz, "Fractal analysis of flow of the river Warta," *Journal of Hydrology*, vol. 200, no. 1–4, pp. 280–294, 1997.
- [27] M. Li, C. Cattani, and S. Y. Chen, "Viewing sea level by a one-dimensional random function with long memory," *Mathematical Problems in Engineering*, vol. 2011, Article ID 654284, 2011.
- [28] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

- [29] Y. Kang and S. Choi, "Restricted deep belief networks for multi-view learning," in *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD '11)*, vol. 6912 of *Lecture Notes in Computer Science*, pp. 130–145, 2011.
- [30] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [31] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26th International Conference On Machine Learning (ICML '09)*, pp. 609–616, June 2009.
- [32] T. Liu, "A novel text classification approach based on deep belief network," in *Proceedings of the 17th International Conference on Neural Information Processing (ICONIP '10)*, vol. 6443 of *Lecture Notes in Computer Science*, pp. 314–321, 2010.
- [33] S. Zhou, Q. Chen, and X. Wang, "Discriminative Deep Belief networks for image classification," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 1561–1564, Hong Kong, 2010.
- [34] J. Chao, F. Shen, and J. Zhao, "Forecasting exchange rate with deep belief networks," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '11)*, pp. 1259–1266, San Jose, Calif, USA, 2011.
- [35] L. Zhai and Q. Feng, "Spatial and temporal pattern of precipitation and drought in Gansu Province, Northwest China," *Natural Hazards*, vol. 49, no. 1, pp. 1–24, 2009.
- [36] H. Wu, M. D. Svoboda, M. J. Hayes, D. A. Wilhite, and F. Wen, "Appropriate application of the Standardized Precipitation Index in arid locations and dry seasons," *International Journal of Climatology*, vol. 27, no. 1, pp. 65–79, 2007.
- [37] Y. P. Xu, S. J. Lin, Y. Huang, Q. Q. Zhang, and Q. H. Ran, "Drought analysis using multi-scale standardized precipitation index in the Han River Basin, China," *Journal of Zhejiang University*, vol. 12, no. 6, pp. 483–494, 2011.
- [38] N. B. Guttman, "Accepting the standardized precipitation index: a calculation algorithm," *Journal of the American Water Resources Association*, vol. 35, no. 2, pp. 311–322, 1999.
- [39] S. M. Vicente-Serrano, "Differences in spatial patterns of drought on different time scales: an analysis of the Iberian Peninsula," *Water Resources Management*, vol. 20, no. 1, pp. 37–60, 2006.
- [40] D. C. Edwards and T. B. McKee, "Characteristics of 20th Century drought in the United States at multiple time scales," *Climatology Report 97-2*, Colorado State University, Fort Collins, Colo, USA, 1997.
- [41] B. Lloyd-Hughes and M. A. Saunders, "A drought climatology for Europe," *International Journal of Climatology*, vol. 22, no. 13, pp. 1571–1592, 2002.
- [42] L. Zhang and G. Subbarayan, "An evaluation of back-propagation neural networks for the optimal design of structural systems: part I. Training procedures," *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 25-26, pp. 2873–2886, 2002.
- [43] L. H. Jiang, A. G. Wang, N. Y. Tian, W. C. Zhang, and Q. L. Fan, "BP neural network of continuous casting technological parameters and secondary dendrite arm spacing of spring steel," *Journal of Iron and Steel Research International*, vol. 18, no. 8, pp. 25–29, 2011.
- [44] Q. Y. Tang and M. G. Feng, *DPS Data Processing System for Trial Design, Statistical Analysis and Data Mining*, Chinese Science Press, 2007.
- [45] M. C. Wang, Y. H. Yan, Y. L. Han et al., "Application of neural network in the surface quality testing of cool rolled strip," *Machinery*, vol. 12, p. 71, 2006 (Chinese).
- [46] Z. X. Ge and Z. Q. Sun, *Neural Network Theory and MATLAB R2007 Realization*, Electronic Industry Press, Beijing, China, 2007.
- [47] V. Nair and G. E. Hinton, "Rectified linear units improve Restricted Boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML '10)*, pp. 807–814, June 2010.
- [48] H. Chen and A. F. Murray, "A continuous restricted Boltzmann machine with hardware-amenable learning algorithm," in *Proceedings of the 12th International Conference on Artificial Neural Networks (ICANN '02)*, pp. 358–363, Madrid, Spain, 2002.
- [49] B. J. Frey, "Continuous sigmoidal belief networks trained using slice sampling," *Advance in Neural Information processing Systems*, vol. 9, pp. 452–458, 1997.
- [50] G. E. Hinton, "Learning multiple layers of representation," *Trends in Cognitive Sciences*, vol. 11, no. 10, pp. 428–434, 2007.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

