A Novel Nonlinear Neural Network Ensemble Model Using K-PLSR for Rainfall Forecasting

Chun Meng and Jiansheng Wu

Department of Mathematics and Computer, Liuzhou Teacher College, Liuzhou, Guangxi, 545004, China
biaji2003@yahoo.com.cn,
wjsh2002168@163.com

Abstract. In this paper, a novel hybrid Radial Basis Function Neural Network (RBF–NN) ensemble model is proposed for rainfall forecasting based on Kernel Partial Least Squares Regression (K–PLSR). In the process of ensemble modeling, the first stage the initial data set is divided into different training sets by used Bagging and Boosting technology. In the second stage, these training sets are input to the RBF–NN models of different kernel function, and then various single RBF–NN predictors are produced. Finally, K–PLSR is used for ensemble of the prediction purpose. Our findings reveal that the K–PLSR ensemble model can be used as an alternative forecasting tool for a Meteorological application in achieving greater forecasting accuracy.

Keywords: Neural network ensemble, kernel partial least square regression, rainfall forecasting.

1 Introduction

Accurate and timely rainfall forecasting is a major challenge for the scientific community because it can help prevent casualties and damages caused by natural disasters [1,2]. However, neural network are a kind of unstable learning methods, i.e., small changes in the training set and/or parameter selection can produce large changes in the predicted output. This diversity of neural networks is a naturally by-product of the randomness of the inherent data and training process, and also of the intrinsic non–identifiability of the model. For example, the results of many experiments have shown that the generalization of single neural network is not unique [3,4].

In order to overcome the main limitations of neural network, recently a novel ensemble forecasting model, i.e. artificial neural network ensemble (NNE), has been developed. Because of combining multiple neural networks learned from the same training samples, NNE can remarkably enhance the forecasting ability and outperform any individual neural network. It is an effective approach to the development of a high performance forecasting system [5].

In general, a neural network ensemble is constructed in two steps, i.e. training a number of component neural networks and then combining the component

D.-S. Huang et al. (Eds.): ICIC 2011, LNBI 6840, pp. 41–48, 2012.

predictions. Different from the previous work, this paper proposes a novel hybrid Radial Basis Function Neural Network (RBF-NN) ensemble forecasting method in terms of Kernel Partial Least Squares Regression (K–PLSR) principle, namely RBF–K–PLSR. The rainfall data of Guangxi is predicted as a case study for our proposed method. An actual case of forecasting monthly rainfall is illustrated to show the improvement in predictive accuracy and capability of generalization achieved by our proposed RBF–K–PLSR model.

The rest of this study is organized as follows. Section 2 describes the proposed RBF–LS–SVR, ideas and procedures. For further illustration, this work employs the method set up a prediction model for rainfall forecasting in Section 3. Discussions are presented in Section 4 and conclusions are drawn in the final Section.

2 The Building Process of the Neural Network Ensemble Model

In this section, a triple–phase nonlinear RBF–NN ensemble model is proposed for rainfall forecasting. First of all, many individual RBF–NN predictors are generated in terms of different kernel function. Then RBF–NN predictors are ensemble into an aggregated predictor by K–PLSR.

2.1 Radial Basis Function Neural Network

Radial basis function was introduced into the neural network literature by Broomhead and Lowe [6]. RBF–NN are widely used for approximating functions from given input–output patterns. The performance of a trained RBF network depends on the number and locations of the radial basis functions, their shape and the method used for learning the input–output mapping. The network is generally composed of three layers, the architecture of RBF–NN is presented in Figure 1.

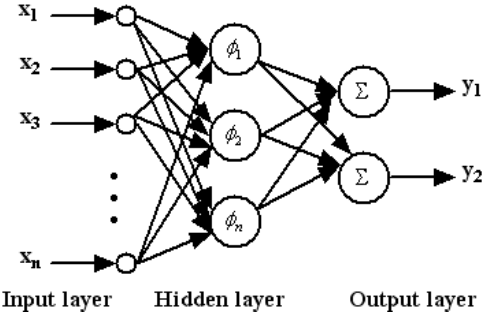

Fig. 1. The RBF–NN architecture

The output of the RBF–NN is calculated according to

$$y_i = f_i(x) = \sum_{k=1}^{N} w_{ik}\phi_k(\|x - c_k\|), i = 1, 2, \cdots, m \tag{1}$$

where $x \in \Re^{n\times 1}$ is an input vector, $\phi_k(\cdot)$ is a function from $\Re^+$ to $\Re$, $\|\cdot\|_2$ denotes the Euclidean norm, w_{ik} are the weights in the output layer, n is the number of neurons in the hidden layer, and $c_k \in \Re^{n\times 1}$ are the centers in the input vector space. The functional form of $\phi_k(\cdot)$ is assumed to have been given, and some typical choices are shown in Table 1.

Table 1. Types of function name and Function formula

Denote	Functional name	Function formula
RBF1	Cubic approximation	$\phi(x) = x^3$
RBF2	Reflected sigmoid	$\phi(x) = (1 + exp(x^2/\sigma^2))^{-1}$
RBF3	Thin-plate-spline function	$\phi(x) = x^2 lnx$
RBF4	Guassian function	$\phi(x) = exp(-x^2/\sigma^2)$
RBF5	Multi-quadratic function	$\phi(x) = \sqrt{x^2 + \sigma^2}$
RBF6	Inverse multi-quadratic function	$\phi(x) = \frac{1}{\sqrt{x^2+\sigma^2}}$

The training procedure of the RBF networks is a complex process, this procedure requires the training of all parameters including the centers of the hidden layer units $(c_i, i = 1, 2, \cdots, m)$, the widths (σ_i) of the corresponding Gaussian functions, and the weights $(\omega_i, i = 0, 1, \cdots, m)$ between the hidden layer and output layer. In this paper, the the orthogonal least squares algorithm (OLS) is used to training RBF based on the minimizing of SSE. The more detailed about algorithm is described by the related literature [7,8].

2.2 Kernel PLS Regression

Partial least squares (PLS) regression analysis was developed in the late seventies by Herman O. A. Wold [9]. It has been a popular modeling, regression, discrimination and classification technique in its domain of origin–chemometrics. It is a statistical tool that has been specifically designed to deal with multiple regression problems where the number of observations is limited, missing data are numerous and the correlations between the predictor variables are high.

PLS regression is is a technique for modeling a linear relationship between a set of output variables (responses) $\{Y_i, i = 1, 2, \cdots, N, Y \in R^L\}$ and a set of input variables (regressors) $\{X_i, i = 1, 2, \cdots, N, Y \in R^N\}$. The K–PLSR methodology was proposed by Roman Rosipal [10]. In kernel PLS regression a linear PLS regression model in a feature space F is considered. The data set y variables into a feature F_1 space. This simply means that $K_1 = YY^T$ and F_1 is the original Euclidian R^L space. In agreement with the standard linear PLS

model it is assumed that the score variables $\{t_i\}_{i=1}^{p}$ are good predictors of Y. Further, a linear inner relation between the scores of t and u is also assumed; that is,

$$\hat{g}^m(x, d^m) = \sum_{i=1}^{N} d_i^m K(x, x_i) \tag{2}$$

which agrees with the solution of the regularized form of regression in RKHS given by the Representer theorem [11,12]. Using equation (10) the kernel PLS model can also be interpreted as a linear regression model of the form

$$\hat{g}^m(x, c^m) = c_1^m t_1(x) + c_2^m t_2(x) + \cdots + c_N^m t_N(x) = \sum_{i=1}^{N} c_i^m t_i(x) \tag{3}$$

where $\{t_i(x)\}_{i=1}^{p}$ are the projections of the data point x onto the extracted p score vectors and $c^m = T^T Y^m$ is the vector of weights for the m-th regression model. The more detailed about K–PLSR algorithm is described by the related literature [10,12].

2.3 Our Proposed Novel Hybrid RBF–K–PLS

To summarize, the proposed hybrid RBF–NN ensemble model consists of three main steps. Generally speaking, firstly, the initial data set is divided into different training sets by used Bagging and Boosting technology. Secondly, these training sets are input to the different individual RBF–NN models, and then various single RBF–NN predictors are produced based on different kernel function. Thirdly, K–PLSR is used to aggregate the ensemble results. The basic flow diagram can be shown in Figure 2.

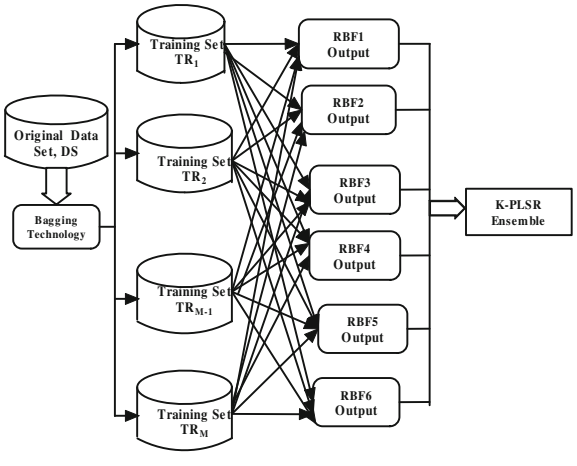

Fig. 2. A Flow Diagram of The Proposed Ensemble Forecasting Model

3 Experiments Analysis

3.1 Empirical Data

This study has investigated Modeling RBF–K–PLSR to predict average monthly precipitation from 1965 to 2009 on July in Guangxi (including Ziyuan, Guilin, Liuzhou, Bose, Wuzhou, Nanning, Yulin and Beihai). Thus the data set contained 529 data points in time series, 489 samples were used to train and the other 40 samples were used to test for generalization ability.

Firstly, the candidate forecasting factors are selected from the numerical forecast products based on 48h forecast field, which includes: the 23 conventional meteorological elements, physical elements from the T213 numerical products of China Meteorological Administration, the 500–hPa monthly mean geopotential height field of the Northern Hemisphere data and the sea surface temperature anomalies in the Pacific Ocean data. We get 10 variables as the predictors, which the original rainfall data is used as the predicted variables.

Method of modeling is one-step ahead prediction, that is, the forecast is only one sample each time and the training samples is an additional one each time on the base of the previous training.

3.2 Performance Evaluation of Model

In order to measure the effectiveness of the proposed method, three types of errors are used in this paper, such as, Normalized Mean Squared Error (NMSE), the Mean Absolute Percentage Error (MAPE) and Pearson Relative Coefficient (PRC), which be found in many paper [4]. The minimum values of NMSE indicate that the deviations between original values and forecast values are very small. The minimum values of MAPE indicate the smallest variability from sample to sample, which it is expressed in generic percentage terms that can be understandable to a wide range of users. The accurate efficiency of the proposed model is measured as PRC, The higher values of PRC (maximum value is 1) indicate that the forecasting performance of the proposed model is effective, which can capture the average change tendency of the cumulative rainfall data.

According to the previous literature, there are a variety of methods for rainfall forecasting model in the past studies. For the purpose of comparison, we have also built other three rainfall forecasting models: ARIMA and stacked regression (SR) ensemble [13] method based on RBF–NN.

The authors used Eviews statistical packages to formulate the ARIMA model. Akaike information criterion (AIC) was used to determine the best model. The model generated from the data set is AR(4). The equation used is presented in Equation 5.

$$x_t = 1 + 0.7326x_{t-1} - 0.6118x_{t-2} + 0.0231x_{t-3} + 0.0173x_{t-4} \tag{4}$$

The standard RBF–NN with Gaussian-type activation functions in hidden layer were trained for each training set, then tested as an ensemble for each method

for the testing set. Each network was trained using the neural network toolbox provided by Matlab software package. In addition, the best single RBF neural network using cross–validation method [14] (i.e., select the individual RBF network by minimizing the MSE on cross–validation) is chosen as a benchmark model for comparison.

3.3 Analysis of the Results

Table 2 illustrates the fitting accuracy and efficiency of the model in terms of various evaluation indices for 489 training samples. From the table, we can generally see that learning ability of RBF–K–PLSR outperforms the other two models under the same network input. The more important factor to measure performance of a method is to check its forecasting ability of testing samples in order for actual rainfall application.

Table 2. Fitting result of three different models about training samples

Moel	NMSE	MAPE	PRC
AR(4)	0.0146	0.2451	0.8765
SR Ensemble	0.0120	0.1890	0.9431
RBF-K-PLSR	0.0113	0.1209	0.9756

Figure 3 shows the forecasting results of three different models for 30 testing samples, we can see that the forecasting results of RBF–K–PLSR model are best in all models. Table 3 shows that the forecasting performance of three different models from different perspectives in terms of various evaluation indices. From the graphs and table, we can generally see that the forecasting results are very promising in the rainfall forecasting under the research where either the measurement of fitting performance is goodness or where the forecasting performance is effectiveness.

From more details, the NMSE of the AR(4) model is 0.2164. Similarly, the NMSE of the SR ensemble model is 0.1762; however the NMSE of the RBF–K–PLSR model reaches 0.0235. The NMSE result of the model has obvious advantages over two other models. Subsequently, for MAPE efficiency index, the proposed RBF–K–PLSR model is also the smallest.

Similarly, for PRC efficiency index, the proposed RBF–K–PLSR model is also deserved to be confident. As shown in Table 3, we can see that the forecasting rainfall values from RBF–K–PLSR model have higher correlative relationship with actual rainfall values; As for the testing samples, the PRC for the AR(4) model is only 0.8971, for the SR ensemble model PRC is 0.9032; while for the RBF–K–PLSR forecasting models, the PRC reaches 0.9765. It show that the PRC of RBF–K–PLSR model is close to their real values in different models and the model is capable to capture the average change tendency of the daily rainfall data.

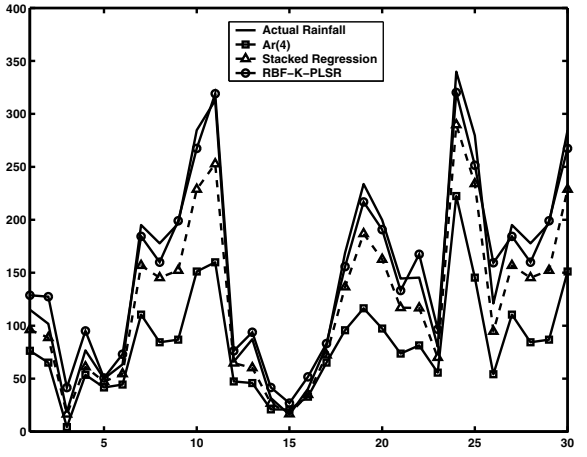

Fig. 3. The Forecasting Results of Testing Samples

Table 3. Forecasting result of three different models about testing samples

Moel	NMSE	MAPE	PRC
AR(4)	0.2164	0.3876	0.8971
SR Ensemble	0.1762	0.2912	0.9032
RBF-K-PLSR	0.0235	0.1400	0.9765

4 Conclusion

Accurate rainfall forecasting is crucial for a frequent unanticipated flash flood region to avoid life losing and economic loses. This paper proposes a novel hybrid Radial Basis Function Neural Network ensemble forecasting method in terms of Kernel Partial Least Squares Regression principle. This model was applied to the forecasting fields of monthly rainfall in Guangxi. The results show that using different the kernel function of RBF can establish the effective nonlinear mapping for rainfall forecasting. It demonstrated that K–PLSR is used to combine the selected individual forecasting results into a nonlinear ensemble model, which keeps the flexibility of the nonlinear model. So the proposed nonlinear ensemble model can be used as a feasible approach to rainfall forecasting.

Acknowledgments. The authors would like to express their sincere thanks to the editor and anonymous reviewer's comments and suggestions for the improvement of this paper. This work was supported in part by the Natural Science Foundation of China under Grant No.41065002, and in part by the Guangxi Natural Science Foundation under Grant No.0832019Z.

References

1. Nasseri, M., Asghari, K., Abedini, M.J.: Optimized Scenario for Rainfall Forecasting Using Genetic Algorithm Coupled with Artificial Neural Network. Expert Systems with Application 35, 1414–1421 (2008)
2. Yingni, J.: Prediction of Monthly Mean Daily Diffuse Solar Radiation Using Artificial Neural Networks and Comparison with Other Empirical Models. Energy Policy 36, 3833–3837 (2008)
3. Kannan, M., Prabhakaran, S., Ramachandran, P.: Rainfall Forecasting Using Data Mining Technique. International Journal of Engineering and Technology 2(6), 397–401 (2010)
4. Wu, J.S., Chen, E.h.: A Novel Nonparametric Regression Ensemble for Rainfall Forecasting Using Particle Swarm Optimization Technique Coupled with Artificial Neural Network. In: Yu, W., He, H., Zhang, N. (eds.) ISNN 2009. LNCS, vol. 5553, pp. 49–58. Springer, Heidelberg (2009)
5. French, M.N., Krajewski, W.F., Cuykendal, R.R.: Rainfall Forecasting in Space and Time Using a Neural Network. Journal of Hydrology 137, 1–37 (1992)
6. Broomhead, D.S., Lowe, D.: Multivariable Functional Interpolation and Adaptive Networks. Complex Systems 26, 321–355 (1988)
7. Moravej, Z., Vishwakarma, D.N., Singh, S.P.: Application of Radial Basis Function Neural Network for Differential Relaying of a Power Transformer. Computers and Electrical Engineering 29, 421–434 (2003)
8. Ham, F.M., Ivica, K.: Principles of Neurocomputing for Science & Engineering. The McGraw-Hill Companies, New York (2001)
9. Wold, S., Ruhe, A., Wold, H., Dunn, W.J.: The Collinearity Problem in Linear Regression: Rhe Partial Least Squares Aapproach To Generalized Inverses. Journal on Scientific and Statistical Computing 5(3), 735–743 (1984)
10. Rosipal, R., Trejo, L.J.: Kernel Partial Least Squares Regression in Reproducing Kernel Hilbert Space. Journal of Machine Learning Research 2, 97–123 (2001)
11. Wahba, G.: Splines Models of Observational Data. Series in Applied Mathematics. SIAM, Philadelphia (1990)
12. Rosipal, R., Trejo, L.J., Matthews, B.: Kernel PLS-SVC for Linear and Nonlinear Classication. In: Proceedings of the Twentieth International Conference on Machine Learning (ICML 2003), Washington DC (2003)
13. Yu, L., Wang, S.Y., Lai, K.K.: A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and ANN for Foreign Exchange Rates. Computers & Operations Research 32, 2523–2541 (2005)
14. Krogh, A., Vedelsby, J.: Neural Network Ensembles, Cross Validation, and Active Learning. In: Tesauro, G., Touretzky, D., Leen, T. (eds.) Advances in Neural Information Processing Systems, vol. 7, pp. 231–238. The MIT Press, Cambridge (1995)