

## Research Article

# Crop Yield Forecasting Using Artificial Neural Networks: A Comparison between Spatial and Temporal Models

William W. Guo<sup>1</sup> and Heru Xue<sup>2</sup>

<sup>1</sup> School of Engineering and Technology, Central Queensland University, Rockhampton, QLD 4702, Australia

<sup>2</sup> College of Computer & Information Engineering, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia 010018, China

Correspondence should be addressed to Heru Xue; [xuehr@imau.edu.cn](mailto:xuehr@imau.edu.cn)

Received 31 October 2013; Accepted 27 November 2013; Published 23 January 2014

Academic Editor: Chih-Cheng Hung

Copyright © 2014 W. W. Guo and H. Xue. 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.

Our recent study using historic data of wheat yield and associated plantation area, rainfall, and temperature has shown that incorporating statistics and artificial neural networks can produce highly satisfactory forecasting of wheat yield. However, no comparison has been made between the outcomes from the spatial neural network model and commonly used temporal neural network models in crop forecasting. This paper presents the latest research outcomes from using both the spatial and temporal neural network models in crop forecasting. Our simulation shows that the spatial NN model is able to predict the wheat yield with respect to a given plantation area with a high accuracy compared with the temporal NARNN and NARXNN models. However, the high accuracy of the spatial NN model in crop yield forecasting is limited to the forecasting of crop yield only within normal ranges. Users must be cautious when using either NARNN or NARXNN for crop yield forecasting due to their inconsistency between the results of training and forecasting.

## 1. Introduction

Crop yield forecasting plays an important role in farming planning and management, domestic food supply, international food trade, ecosystem sustainability, and so on [1–3]. For instance, China has the largest population in the world but with limited agricultural land so accurate crop forecasting helps the government provide sufficient food supply to the people. Australia has a small population with vast agricultural land so its concern on crop production is how to optimize revenue from international crop export to countries like China.

There are many factors that have an influence on crop yield, such as plantation area, efficiency of irrigation systems, variations in rainfall and temperature, quality of crop seeds, topographic attributes, soil quality and fertilisation, and disease occurrences [4–8]. Crop growing follows seasonal cycles but many of the factors above are largely irrelevant to the temporal factor. For example, plantation area, rainfall, fertilising, and disease occurrence vary yearly; efficiency of irrigation systems, quality of crop seeds, and soil quality may be improved or degraded from year to year; and topographic

attributes may largely remain the same for a long period of time.

Effort has been made in using either statistics to identify relationships or neural networks to establish mappings between crop yield and some of these factors [4–10]. Our recent study using historic data of wheat yield and associated plantation area, rainfall, and temperature in Queensland, Australia, has shown that incorporating statistics and artificial neural networks can produce a high level of satisfactory forecasting of wheat yield [10]. The neural network employed in this study was a spatial model that treats the wheat plantation areas and yields as mutual mappings, rather than yearly time series. Doubts have been raised about the lack of comparison between the outcomes from this spatial neural network model and commonly used temporal neural network models in crop forecasting. To address this issue, using the wheat yield in Queensland as a reference, this paper presents our research outcomes from using both the spatial and temporal neural network models in crop forecasting. Comparison and discussion are made in terms of their usefulness in crop yield forecasting.

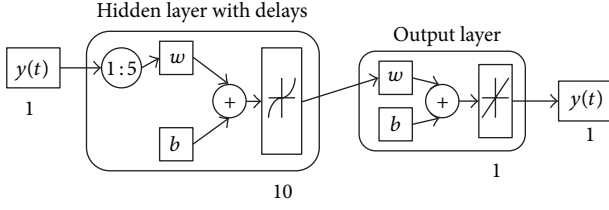


FIGURE 1: Structure of NARNN with 10 hidden neurons and 5 delays.

This is a comparative study between two types of NN-based forecasting models so our strategy is to focus on examining the performances of existing models that have been applied to forecasting, rather than introducing new forecasting models. Readers who are interested in fundamentals of NN forecasting can refer to [9–14] for details.

## 2. Neural Network-Based Forecasting Models

**2.1. Nonlinear Autoregressive Neural Network (NARNN) Model.** Nonlinear autoregressive (NAR) is a widely used statistical forecasting model for time series [15, 16]. The forecasting model takes the form as follows:

$$y(t) \approx f(y(t-1), y(t-2), \dots, y(t-d)), \quad (1)$$

where  $y(t)$  is the forecasted output and  $f$  is an unknown function of  $d$  previous known outputs. Traditionally, function  $f$  is determined by statistical optimization processes, such as the minimum mean square method.

The feedforward neural network has been used to establish NAR models, in which the traditional function  $f$  is replaced by a number of neurons that work together to implicitly approximate the same functionality [11, 17, 18] as

$$y(t) \approx f(y(t-1), y(t-2), \dots, y(t-d)) = \sum_i b_i \psi \left( \sum_{j=1}^d a_{ji} y(t-j) \right), \quad (2)$$

where  $\psi$  is the transfer functions;  $a_{ji}$  denotes the input-to-hidden layer weights at the hidden neuron  $j$ ; and  $b_i$  is the hidden-to-output layer weight.

This is a time-delay and recurrent neural network model. The input is the known time series which is fed to the hidden layer as input according to the number of time delay. This model is visually illustrated in Figure 1.

**2.2. Nonlinear Autoregressive with External Input Neural Network (NARXNN) Model.** Nonlinear autoregressive with external input (NARX) is a modified NAR model by including another relevant time series as extra input to the forecasting model [19–21], which can be expressed as

$$y(t) \approx f(x(t-1), x(t-2), \dots, x(t-d), y(t-1), y(t-2), \dots, y(t-d)), \quad (3)$$

where  $x[t]$  is the external input to the forecasting model with the same number of time delay as  $y[t]$ .

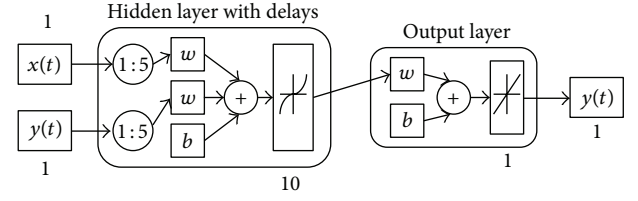


FIGURE 2: Structure of NARXNN with 10 hidden neurons and 5 delays.

Similarly the feedforward neural network is able to establish NARX models, which can be expressed as

$$y(t) \approx \sum_i c_i \psi \left( \sum_{j=1}^d (a_{ji} x(t-j) + b_{ji} y(t-j)) \right), \quad (4)$$

where  $\psi$  is the transfer functions;  $a_{ji}$  and  $b_{ji}$  denote the input-to-hidden layer weights at the hidden neuron  $j$ ; and  $c_i$  is the hidden-to-output layer weight.

This time-delay recurrent neural network model uses two known time series as independent inputs to the hidden layer according to the same number of time delay. This model is visually illustrated in Figure 2.

**2.3. Spatial Feedforward Neural Network Forecasting Model.** Multilayer perceptron (MLP) model belongs to feedforward neural networks. In terms of functionality, MLP has no difference from the neural networks used in both NARNN and NARXNN models if the input is time series. Additionally MLPs have been proven to be able to approximate any continuous function by adjusting the number of nodes in the hidden layer [12], with numerous cases of successful applications [13, 14, 22–24]. Figure 3 illustrates the general structure of a three-layer MLP with one hidden layer of  $L$  nodes, a  $p$ -dimensional input vector  $\mathbf{X}$ , and a  $q$ -dimensional output vector  $\mathbf{Y}$ . The relationship between the input and output components for this MLP can be generally expressed as

$$y_k = \varphi \left( \sum_{j=1}^L b_{kj} \psi \left( \sum_i a_{ji} x_i \right) \right), \quad (5)$$

where  $\varphi$  and  $\psi$  are the transfer functions;  $a_{ji}$  denotes the input-to-hidden layer weights at the hidden neuron  $j$ ; and  $b_{kj}$  is the hidden-to-output layer weights at the output unit  $k$ .

There are at least two relevant time series used in the NARXNN model, the internal series  $y[t]$  and external series  $x[t]$ . Time series analysis emphasises on the appearance of consecutive events. However, for example, in crop yield forecasting, the current plantation area should have a much higher impact on the forthcoming crop yield than the historic yields of any past years.

Treating crop yield and plantation area as a correlated pair, MLPs have been used to approach the nonlinear relation that may exist between the two sequences in a correlated “spatial” manner [9, 10], rather than a correlated temporal mode. This has resulted in some encouraging outcomes.

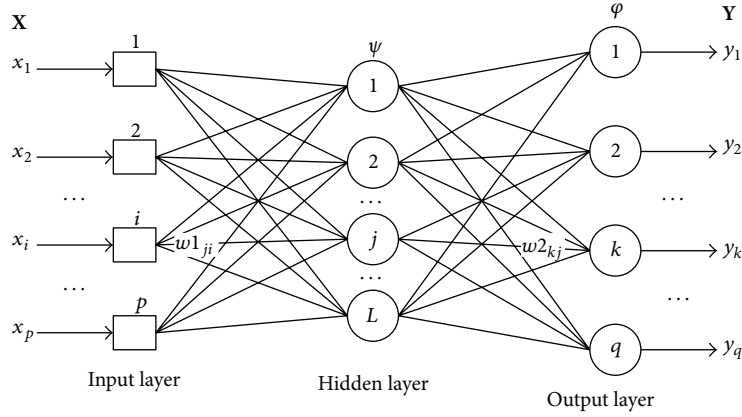


FIGURE 3: Three layer multilayer perceptron (MLP) neural network.

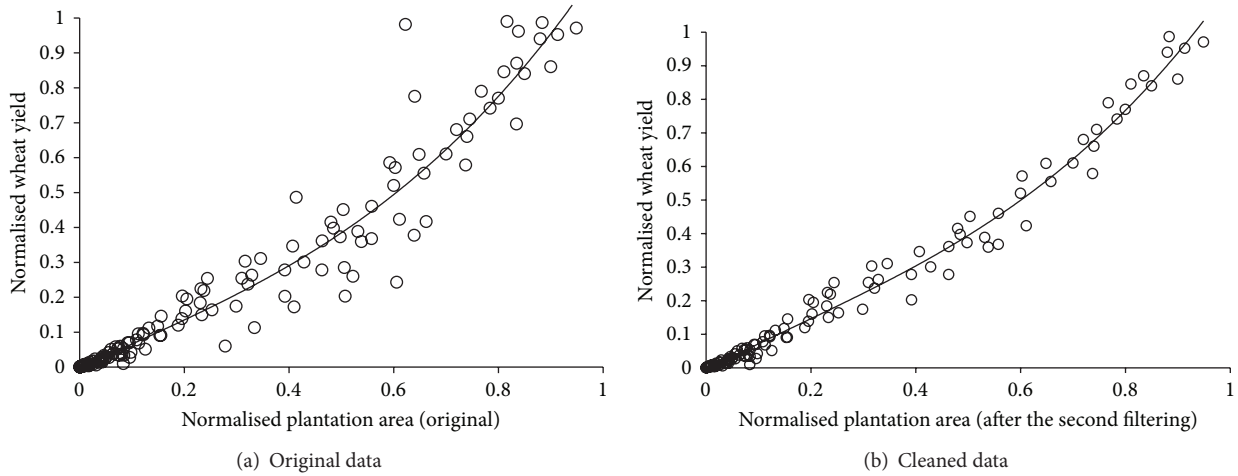


FIGURE 4: Correlations between normalised wheat yield and plantation area.

### 3. Crop Data for Neural Network Simulation

**3.1. Historic Crop Data.** The Queensland historic wheat plantation area in hectare and wheat yield in ton from 1861 to 2007 are extracted from the report of Australian Bureau of Statistics [25], which gives a total of 135 entries over the past 147 years. Both plantation area in hectare and wheat yield in ton are listed in their approximated absolute values each year in the original data. We normalise both factors with their ceilings in the order of millions. The ceiling for plantation area is 13 million hectares and that for wheat yield is 20 million tons. Plot of these two normalised factors is shown in Figure 4(a). After two rounds of outlier detection and exclusion, a third-order polynomial correlation has been defined as

$$w = 0.8197a^3 - 0.5102a^2 + 0.8511a - 0.0073, \quad (6)$$

where  $w$  represents the normalised annual wheat yield and  $a$  is the normalised plantation area. This correlation fits the filtered data well (Figure 4(b)) with a coefficient of 0.9904. This nonlinear correlation indicates that, through properly training, a neural network system can be used to approach

such nonlinear relation between the crop production and plantation area.

**3.2. Data for Training and Testing Neural Network Models.** Neural network training requires a sufficient amount of data for achieving a high reliability. For MLPs, since temporal factor does not play any role in correlation (6), this correlation can be used to generate more data without changing the general trend. By keeping the same pattern, a moving window operator with different sizes is repeatedly applied to these cleaned data so as to generate more entries to fill the gaps where the original entries are scarce. The final data to train and test the selected neural networks are compiled by mixing the cleaned and regenerated entries together.

For both NARNN and NARXNN models, such data expansion cannot be applied because the training data must be a sequence ordered by time. Among the 135 datasets, the first three are with an interval of 5 years and thus excluded from both NARNN and NARXNN training and testing. Time series data for NARNN model is the normalised annual wheat yield by the corresponding plantation area because NARNN

TABLE 1: Statistical results of neural network training and testing.

|               | Training |        |        | Testing |        |        |
|---------------|----------|--------|--------|---------|--------|--------|
|               | $N$      | MSE    | $R$    | $N$     | MSE    | $R$    |
| MLP (100)     | 329      | 0.0001 | 0.9996 | 40      | 0.0001 | 0.9943 |
| MLP (200)     | 329      | 0.0001 | 0.9998 | 40      | 0.0001 | 0.9981 |
| NARNN (5 d)   | 105      | 0.1035 | 0.9682 | 20      | 0.0231 | 0.8800 |
| NARNN (10 d)  | 105      | 0.0069 | 0.9753 | 20      | 0.0188 | 0.9298 |
| NARXNN (5 d)  | 105      | 0.0174 | 0.908  | 20      | 0.0528 | 0.6945 |
| NARXNN (10 d) | 105      | 0.0271 | 0.8574 | 20      | 0.0628 | 0.6801 |

takes only one time series as the input. To some extent, this normalised series actually absorbs the effect of plantation area on crop yield into the forecasting. For NARXNNs, the normalised annual wheat yield by wheat ceiling is the internal input  $y[t]$  and the normalised plantation area by plantation ceiling is the external input  $x[t]$ . If the normalised wheat yield by plantation area is used as the internal input  $y[t]$ , the effect of plantation area on wheat yield will be doubly accounted in the forecasting.

#### 4. Results of Neural Network Simulation

Two MLP models are used for training and simulation. By running the process ten times with both the 100-node hidden-layer MLP and 200-node hidden-layer MLP, the simulations produce a highly satisfactory average outcome for both training and testing (Table 1). The difference is that the latter achieves a slightly lower MSE and a higher correlation than the former but both show a high consistency between the results of training and simulation or testing.

Two NARNN models are also used in training and simulation. Both models have ten hidden neurons but with 5 delays and 10 delays, respectively. Running ten repetitions for each model has resulted in a fairly satisfactory outcome on average shown in Table 1. Since the model changes data partition for training and testing dynamically between separate runs, a highly satisfactory outcome from training does not always produce a highly satisfactory outcome from testing. In general, the result of simulation is always inferior to that of training, with the 10-delay model performing slightly better than the 5-delay model. Similar trends are found from the results of training and testing the two NARXNN models, whose performance is even worse than that of NARNNs (Table 1).

#### 5. Discussion

In terms of consistency between the performances of training and testing using the same model, MLPs are able to achieve the highest consistency and hence produce the best simulation results among the three forecasting models. This is mainly because the data used to train MLPs have been subject to outlier cleaning, which means the abnormal wheat yields outside the statistical trend have no impact on the training and testing. In addition, without the temporal constraint, the expanded dataset ensures that the MLPs are adequately

trained and tested with multiple crossing validations. Since the original data have been cleaned, in theory the MLPs should only be effective for crop forecasting of any “normal” year.

NARNNs exhibit a highly satisfactory performance in training but the simulation is highly dependent on the selection of testing dataset; hence, the range of forecasting error is large. This indicates that a well trained NARNN model is not able to produce consistently accurate forecasting. This inconsistency between the training and testing is clearly illustrated in Figure 5. Our experiments also show that changes in number of hidden neuron and length of delay ( $>3$ ) for NARNN do not make significant improvement to the performance of forecasting. Although the NARNNs are not consistent in forecasting, they use the whole data without excluding “abnormal” datasets in both training and testing. This is a complement to MLPs to some extent.

NARXNNs exhibit similar inconsistent patterns between training and testing but even worse than NARNNs (Figure 6). This may be caused by the double impacts on forecasting exerted by the “anomalies” in both the wheat yield series and plantation area series, which were not excluded through data cleaning.

Data cleaning for both NARNN and NARXNN is very challenging since both models use time series as the input. Excluding some temporal events will leave irregular gaps in the time series, which in turn influences the training and testing. The other possible reason that contributed to the inconsistency between the training and testing of both NARNN and NARXNN may be the inadequacy of the historic data in this case. Since we cannot artificially create extra yearly crop yields, like using interpolation to generate extra spatial datasets [9, 26], using time series based NN models to forecast crop yield may be immature at this stage.

#### 6. Conclusion

The spatial NN model is able to predict the wheat yield with respect to a given plantation area with a high accuracy, compared with the temporal NN models such as NARNN and NARXNN. However, the high accuracy of the spatial NN model in crop yield forecasting is only applicable to the forecasting of crop yield within normal ranges because the model is trained using the cleaned and expanded data following a third-order polynomial trend between the crop yield and

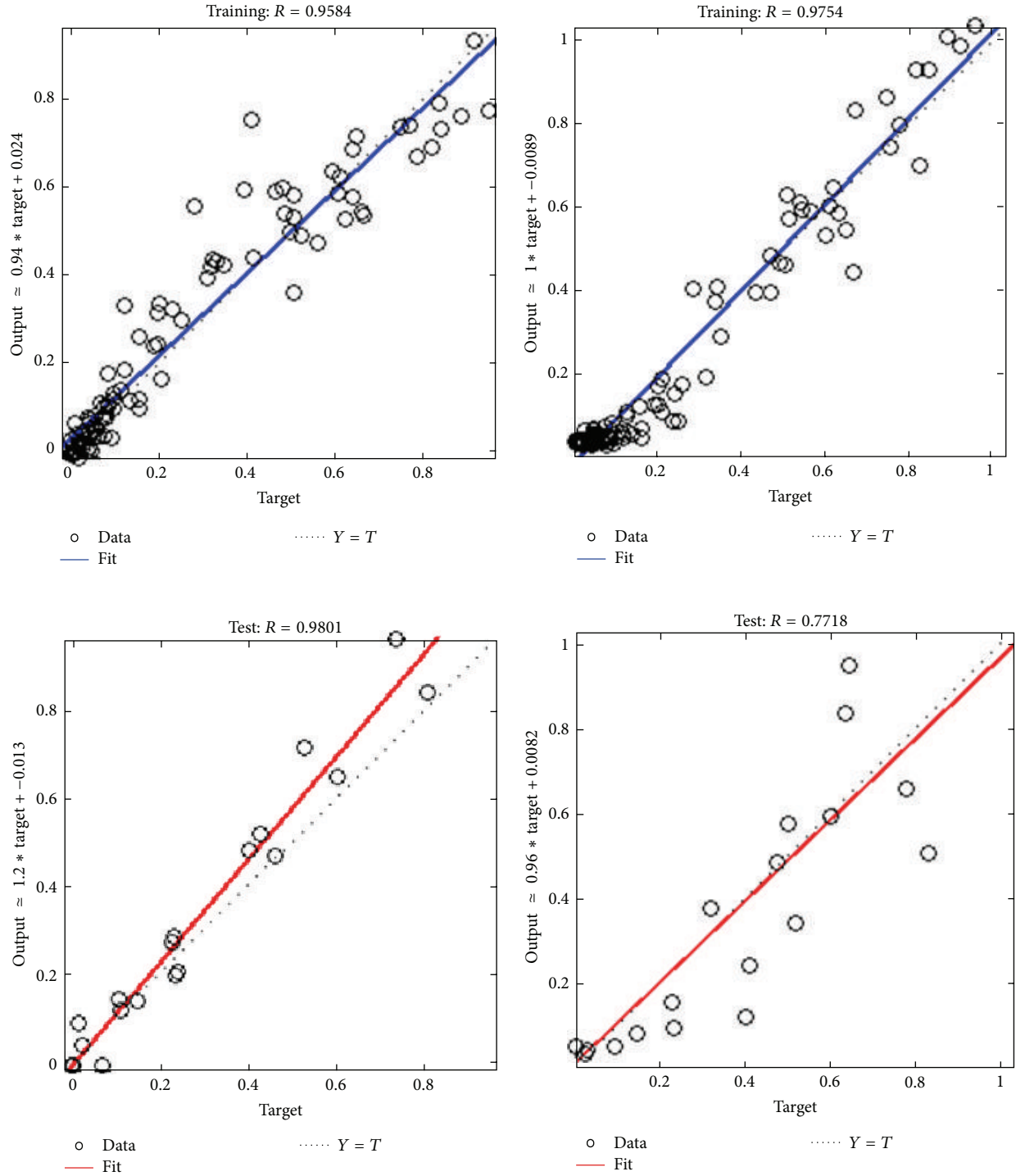


FIGURE 5: Linear regression for two separate training and testing runs of NARNN.

plantation area. NARNNs may be used as a complementary means to the MLPs due to its usage of the whole data. Users must be cautious when using either NARNN or NARXNN for crop yield forecasting due to the inconsistency between training and forecasting.

In the future, other factors, such as efficiency of irrigation systems, variations in rainfall and temperature, quality of crop seeds, topographic attributes, soil quality and fertilisation, and disease occurrences, should be incorporated in forecasting model building and simulation.



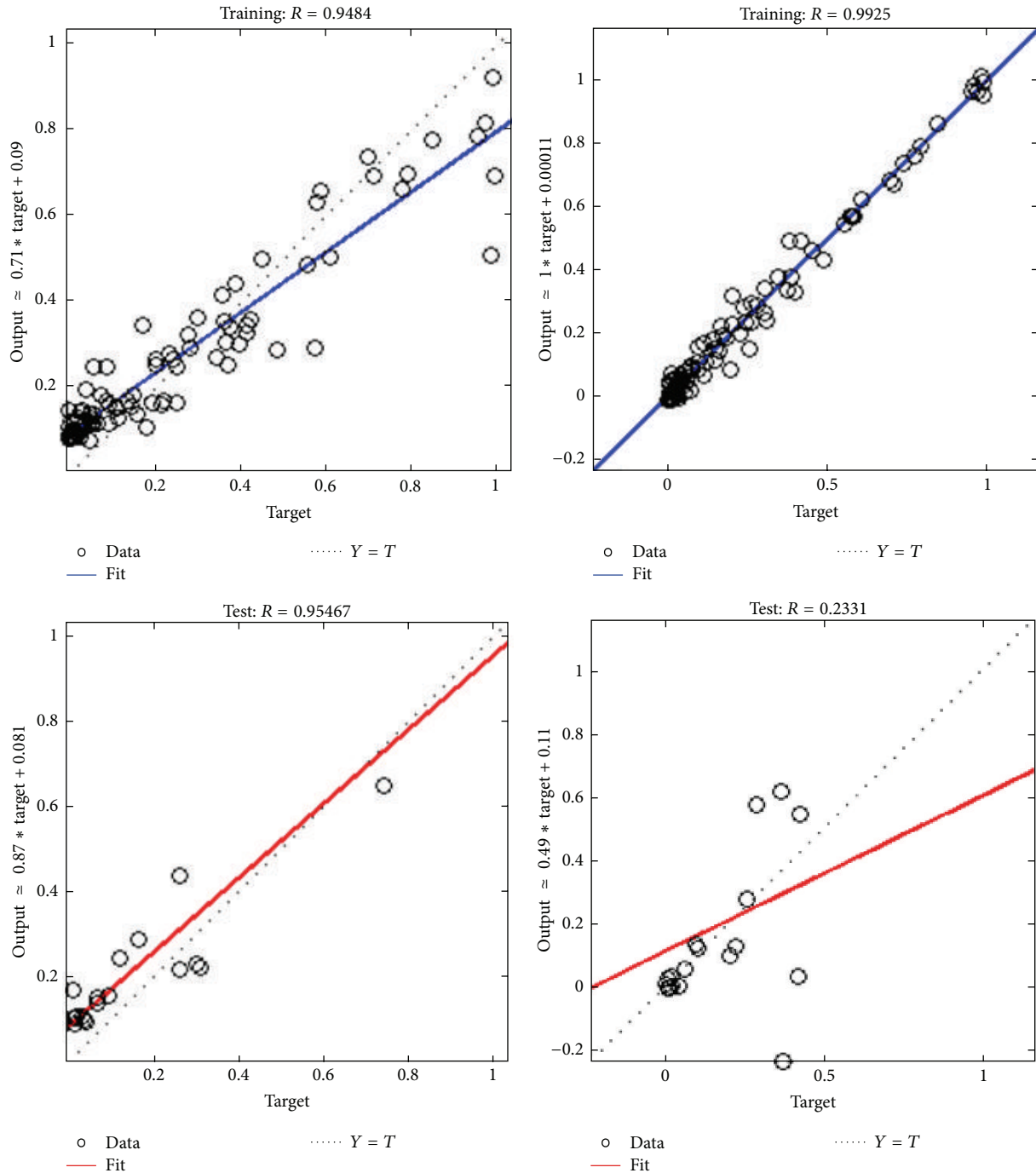


FIGURE 6: Linear regression for two separate training and testing runs of NARXNN.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] Y. W. Jame and H. W. Cutforth, "Crop growth models for decision support systems," *Canadian Journal of Plant Science*, vol. 76, no. 1, pp. 9–19, 1996.
- [2] J. I. Yun, "Predicting regional rice production in South Korea using spatial data and crop-growth modeling," *Agricultural Systems*, vol. 77, no. 1, pp. 23–38, 2003.
- [3] F. Tao, M. Yokozawa, Z. Zhang, Y. Xu, and Y. Hayashi, "Remote sensing of crop production in China by production efficiency models: models comparisons, estimates and uncertainties," *Ecological Modelling*, vol. 183, no. 4, pp. 385–396, 2005.
- [4] G. Hoogenboom, "Contribution of agrometeorology to the simulation of crop production and its applications," *Agricultural and Forest Meteorology*, vol. 103, no. 1-2, pp. 137–157, 2000.

- [5] Y. He, Y. Zhang, S. Zhang, and H. Fang, "Application of artificial neural network on relationship analysis between wheat yield and soil nutrients," in *Proceedings of the IEEE 27th Annual Conference on Engineering in Medicine and Biology*, pp. 4530–4533, Shanghai, China, September 2005.
- [6] T. R. Green, J. D. Salas, A. Martinez, and R. H. Erskine, "Relating crop yield to topographic attributes using spatial analysis neural networks and regression," *Geoderma*, vol. 139, no. 1-2, pp. 23–37, 2007.
- [7] E. de Jong and D. A. Rennie, "Effect of soil profile type and fertilizer on moisture use by wheat grown on fallow or stubble land," *Canadian Journal of Soil Science*, vol. 49, no. 2, pp. 189–197, 1969.
- [8] C. A. Campbell, R. P. Zentner, and P. J. Johnson, "Effect of crop rotation and fertilization on the quantitative relationship between spring wheat yield and moisture use in southwestern Saskatchewan," *Canadian Journal of Soil Science*, vol. 68, no. 1, pp. 1–16, 1988.
- [9] W. W. Guo, L. D. Li, and G. Whymark, "Simulating wheat yield in New South Wales of Australia using interpolation and neural networks," in *Neural Information Processing. Models and Applications*, vol. 6444 of *Lecture Notes in Computer Science*, pp. 708–715, Springer, Berlin, Germany, 2010.
- [10] W. W. Guo and H. Xue, "An incorporative statistic and neural approach for crop yield modelling and forecasting," *Neural Computing and Applications*, vol. 21, no. 1, pp. 109–117, 2012.
- [11] F. Liang, "Bayesian neural networks for nonlinear time series forecasting," *Statistics and Computing*, vol. 15, no. 1, pp. 13–29, 2005.
- [12] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [13] W. W. Guo, M. M. Li, G. Whymark, and Z.-X. Li, "Mutual complement between statistical and neural network approaches for rock magnetism data analysis," *Expert Systems with Applications*, vol. 36, no. 6, pp. 9678–9682, 2009.
- [14] M. M. Li, W. Guo, B. Verma, K. Tickle, and J. O'Connor, "Intelligent methods for solving inverse problems of backscattering spectra with noise: a comparison between neural networks and simulated annealing," *Neural Computing and Applications*, vol. 18, no. 5, pp. 423–430, 2009.
- [15] S. Y. Hwang and I. V. Basawa, "Large sample inference based on multiple observations from nonlinear autoregressive processes," *Stochastic Processes and their Applications*, vol. 49, no. 1, pp. 127–140, 1994.
- [16] G. Kapetanios, "Nonlinear autoregressive models and long memory," *Economics Letters*, vol. 91, no. 3, pp. 360–368, 2006.
- [17] T. Taskaya-Temizel and M. Casey, "A comparative study of autoregressive neural network hybrids," *Neural Networks*, vol. 18, no. 5-6, pp. 781–789, 2005.
- [18] W. Pawlus, H. R. Karimi, and K. G. Robbersmyr, "Data-based modeling of vehicle collisions by nonlinear autoregressive model and feedforward neural network," *Information Sciences*, vol. 35, pp. 65–79, 2012.
- [19] S. Chatterjee, S. Nigam, J. B. Singh, and L. N. Upadhyaya, "Software fault prediction using Nonlinear Autoregressive with exogenous Inputs (NARX) network," *Applied Intelligence*, vol. 37, no. 1, pp. 121–129, 2012.
- [20] S. H. Arbain and A. Wibowo, "Neural networks based nonlinear time series regression for water level forecasting of Dungun River," *Journal of Computer Science*, vol. 8, no. 9, pp. 1506–1513, 2012.
- [21] F. Huo and A.-N. Poo, "Nonlinear autoregressive network with exogenous inputs based contour error reduction in CNC machines," *International Journal of Machine Tools and Manufacture*, vol. 67, pp. 45–52, 2013.
- [22] W. W. Guo, "Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3358–3365, 2010.
- [23] E. Bojórquez, J. Bojórquez, S. E. Ruiz, and A. Reyes-Salazar, "Prediction of inelastic response spectra using artificial neural networks," *Mathematical Problems in Engineering*, vol. 2012, Article ID 937480, 5 pages, 2012.
- [24] Y. E. Shao, "Prediction of currency volume issued in Taiwan using a hybrid artificial neural network and multiple regression approach," *Mathematical Problems in Engineering*, vol. 2013, Article ID 676742, 9 pages, 2013.
- [25] Australian Bureau of Statistics, *Agricultural Commodities—Historical Data*, Commonwealth of Australia, Canberra, Australia, 2008.
- [26] W. W. Guo, "A novel application of neural networks for instant iron-ore grade estimation," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8729–8735, 2010.