

Intelligent Ensemble Systems for Modeling NASDAQ Microstructure: A Comparative Study

Salim Lahmiri

ESCA School of Management
7 El Kindy Street, BD Moulay Youssef, Casablanca, Morocco
slahmiri@esca.ma

Abstract. In this study, four neural networks (NN) ensemble systems are presented and compared for NASDAQ returns prediction. They are the conventional feed-forward back-propagation neural network (FFNN) ensemble which widely used in the literature, time-delay neural network (TDNN) ensemble, nonlinear auto-regressive with exogenous inputs (NARX) ensemble and the radial basis neural network (RBFNN) ensemble. Each component of the NN ensemble is used to learn specific patterns related to a given NASDAQ submarket. Based on the mean of absolute errors (MAE), the experiments show that ensemble models based on advanced NN architectures such as TDNN, NARX, and RBFNN ensemble all achieve lower forecasting errors than traditional FFNN ensemble system. In addition, the RBFNN ensemble outperformed all other NN ensembles under study.

Keywords: Neural networks, ensemble, stock market microstructure, forecasting.

1 Introduction

Successful prediction models for stock market trading are of great interest for investors to make profits. Therefore, there is a need for effective stock price forecasting systems capable to limit personal bias and mistakes. However, the financial market is complex, evolutionary, and non-linear dynamical system [1,2]. As a result, financial market data are noisy and nonstationary [1,2].

Financial market prediction had received a large attention in the literature where different methods and algorithms for automated stock market prediction were proposed; including artificial neural networks, fuzzy logic, and expert systems to name a few. Indeed, such studies were surveyed in Atsalakis and Valavanis [3] and Bahrammirzaee [4]. In general, the literature [3,4] used two broad classes to predict stock market prices [5]: fundamental analysis and technical analysis. The fundamental analysis depends on knowledge of microeconomics and macroeconomics factors, whilst technical analysis is based on historical patterns to predict market prices [5]. However, fundamental analysis based knowledge is usually not readily available, and historical patterns are not always evident because of the noise [5].

The purpose of this study is to design a neural network (NN) committee (ensemble) to model stock market microstructure and achieve better prediction accuracy of the market returns. Indeed, neural network ensemble systems were found to be effective in forecasting stock market [5-8], and also in other time series forecasting applications such as customer purchase behavior [9], drug dissolution [10], weather [11], climate [12], and software reliability [13].

A neural network (NN) ensemble or committee is a learning paradigm where a collection of several neural networks is trained for the same task. It is expected to provide the following advantages over the traditional neural network. First, a NN is an adaptive nonlinear soft computing system that can learn from patterns and capture hidden functional relationships in a given data even if the functional relationships are not known or difficult to identify [14]. Indeed, it is capable of parallel processing of the information with no prior assumption about the model form and the process that generates the data. In addition, it is robust to noisy data; hence the network is capable to model non-stationary and dynamic data [15]. Furthermore, a NN can theoretically approximate a continuous function to an arbitrary accuracy on any compact set [16-18]. Second, a NN ensemble or committee can produce even more accurate results than any of the individual neural networks by making up the ensemble and thus intensifying discriminant capability of neural networks [19]. Third, a NN is prone to overfitting when it is too closely adjusted to the training set [19]. Therefore, its generalization error tends to increase when it is applied to previously unseen samples [19]. As a solution, a committee or an ensemble approach can make base neural networks robust to overfitting and thus reduce generalization error [19]. Fourth, a NN committee is used to deal with sampling and modeling uncertainties that may otherwise impair individual NN forecasting accuracy and robustness [20]. Fifth, previous works demonstrated the ability of NN based committee system in modeling and predicting various types of time series [5-13].

All of the aforementioned studies [5-8] only consider single neural network type for financial time series prediction. However, the performance of a neural network system significantly depends on the type of NN used to design the committee system. For instance, previous studies relied on the well known feed-forward back-propagation neural networks as basic systems. Indeed, although NN committees are essential for providing accurate forecasts the improvements in the construction of such committees is important by considering other type of neural networks.

In this paper, the performance of the NN committee is investigated depending on the type NN used to form the ensemble. In particular, four different NN architectures are considered to form the NN committee and are compared. They are the conventional feed-forward back-propagation neural network which widely used in the literature, time-delay neural network, nonlinear auto-regressive with exogenous inputs (NARX) network, and the radial basis neural network.

Despite the potential benefits of using fundamental or technical analysis knowledge for ensemble training and prediction task, we rely on information from

microstructure of the stock market to train NN committees. In particular, we aim to model stock market returns based on the dynamics of each sub-market (component) that composes the whole stock market. In other words, our emphasis is on the importance of information related to price movements in each sub-market in determination of the overall stock market behavior. Indeed, we rely on microstructure information for two main reasons. First, price formation in each sub-market is key information that determines the stock market trend. Therefore, modeling the stock market microstructure from an informational point of view would be helpful to improve forecasting accuracy when predicting stock market future return. Second, unlike fundamental and technical analysis based information, microstructure information is always available and not affected with noisy information.

The rest of this paper is organized as follows. Section 2 presents the committee system, and gives a brief introduction to each type of artificial neural network adopted to form the neural network committee. Section 3 presents the empirical results from the NASDAQ market data. The paper is concluded in Section 4.

2 Methods

In this section, we present the proposed system for stock market return prediction based on neural network ensembles based on market microstructure information. Neural network ensemble was proposed by originating Hansen and Salamon [21] as a learning paradigm where several neural networks are trained for the same task. The purpose is to improve the generalization performance of NN system in comparison with using a single neural network.

The prediction system based on neural network ensemble (committee) is shown in Figure 1. The input of the system is the financial return of a given sub-market and the output of the system is the predicted return of the aggregate stock market. For instance, each NN is used as the basic prediction system making up ensemble and each is trained with a sub-market specific price returns. Finally, the output NN is used to generate a single system to produce the output of the ensemble by combining the predictions of multiple neural networks.

In this study, the neural network committee consists of K component neural networks where each component neural network is a three-layer single input single-output NN with two nodes in the hidden layer. Each component neural network is trained with different initial weights connecting three-layers. The outputs of the component neural networks are combined using the output NN as shown in Figure 1. In this study, a neural network (NN) could be the conventional feed-forward back-propagation neural network which widely used in the literature, time-delay neural network, NARX network, or the radial basis neural network. They are described in next sub-sections. The accuracy of each single NN and the NN committee is evaluated based on the mean absolute error (MAE) statistic.

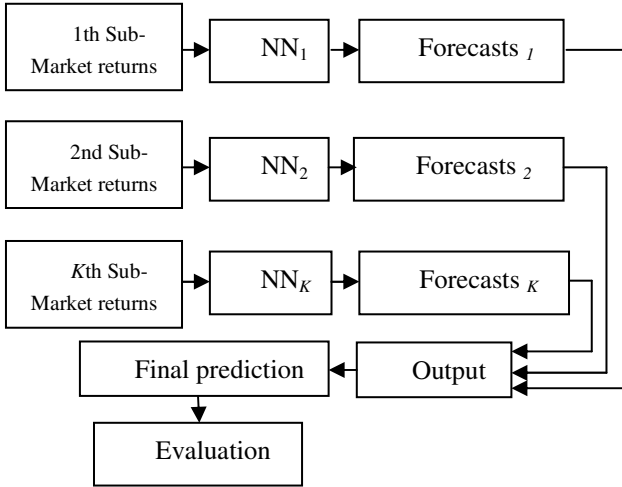


Fig. 1. A generic committee system for modeling and forecasting stock market returns

2.1 Feed-Forward Back-Propagation Neural Network

An artificial neural network (NN) [15] is a generic model for data process data that uses a brain metaphor. The feed-forward back-propagation neural network is a popular architecture that has one input layer with x predictive variables, one or more hidden layers that fulfill the input-output mapping, and an output layer with the predicted variable y . The output y is computed as:

$$y_i = f \left(\sum_{j=1}^m x_j w_{ij} + \theta_i \right) \tag{1}$$

where w_{ij} is a connecting weight from neural j to neural i , θ denotes the bias, and $f(\bullet)$ is an activation function employed to control the amplitude of the output. In this study, the commonly used sigmoid function is used for activation. It is given by:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

The neural network is trained with the Levenberg-Marquardt (L-M) algorithm [9][22] where weights are adjusted based on the gradient method as follows:

$$\Delta w_k = - \left(J_k^T J_k + \mu I \right)^{-1} J_k e_k \tag{3}$$

where J is the Jacobian matrix (first derivatives) of weights, I is the identity matrix, μ is an adaptive learning parameter arbitrarily set to 0.001, and e is a vector of network errors.

The neural network with this type of error training is often called back-propagation neural network [15] or multi-layer feed-forward neural network (FFNN).

2.2 Time-Delay Neural Networks

Unlike the conventional FFNN neural network, the time-delay neural network [22] is a recurrent neural network which is dynamic and possesses a memory to perform temporal processing of the input space. Like the FFNN, TDNN has feed forward connections, but has multiple connections between the individual neurons: each neuron consists of the outputs of earlier neurons during both the current time epoch and fixed number of previous ones ($t-1, t-2, \dots, t-n$). Then, each neuron possesses a memory to remember previous layer outputs for n periods of time. Similar to the FFNN, the TDNN uses the back-propagation algorithm optimized based on the gradients method. The time delay n is set to three in this study.

2.3 NARX Neural Network

The Nonlinear Auto-Regressive with Exogenous inputs (NARX) neural network which was introduced by Leontaritis and Billings [23] is a class of discrete-time nonlinear systems that can represent a variety of nonlinear dynamic systems. In particular, the NARX network is a dynamic neural network that contains recurrent feedbacks from several layers of the network to the input layer. It can be mathematically represented as follows:

$$y(t+1) = f(y(t), \dots, y(t-d_y), u(t-1), \dots, u(t-d_u); W) \tag{4}$$

where $u(t)$ and $y(t)$ are respectively the input and the output of the system at time t , $d_u \geq 1, d_y \geq 1, d_y \geq d_u$, W is a weights matrix, and f is an unknown nonlinear function to be approximated by a FFNN neural network. As a result, the general NARX network equation can be written as:

$$y(t+1) = f_0 \left(\begin{matrix} b_0 + \sum_{h=1}^{N_h} w_{h0} f_h \\ \left(b_h + \sum_{i=0}^{d_u} w_{ih} u(t-i) + \sum_{j=0}^{d_y} w_{jh} (t-j) \right) \end{matrix} \right) \tag{5}$$

where $i = 1, \dots, d_u; j = 1, \dots, d_y; h = 1, \dots, N_h$, f_h and f_o are the hidden and output functions, w_{ih}, w_{jh} and w_{ho} are the weights, and b_h and b_o are biases. In this study, the time delay d is set to three in this study.

2.4 Radial Basis Function Neural Network

The radial basis function neural network (RBFNN) [16] is suitable to model flexible in dynamic environment because of its ability to quickly learn data local complex

patterns and adapt to changes. The RBFNN system consists of three layers; namely the input, hidden and output layer. The input layer distributes the input data among the hidden nodes (units) which are fully connected to the previous layer. In other words, the input variables are each assigned to a node in the input layer and pass directly to the hidden layer without weights. The hidden nodes contain the radial basis functions (RBF) represented by Gaussian kernels and used as transfer functions to process information contained in input layer nodes. For instance, each neuron in hidden layer computes local response to its input. Finally, the neuron in the output layer only sums up the outputs of the hidden neurons. Mathematically, the output of the j th unit, $H_j(x)$, in the hidden layer for an input x_i is computed as follows:

$$H_j(x) = H_j(\|x_i - c_j\|) = \exp\left(\frac{-(x_i - c_j)^2}{2\delta_j^2}\right) \quad (6)$$

where x_i is the first difference of natural logarithm of market price (for example the market return defined as first difference of logarithmic price), c_j represents the position of the center of the j th Gaussian function, and δ is the width parameter controlling the smoothness of the Gaussian function. Finally, the output y of the system is calculated by a linear combination of the K radial basis functions plus the bias w_0 as follows:

$$y(x_j) = \sum_{j=1}^K w_j H_j(\|x_i - c_j\|) + w_0 \quad (7)$$

In this paper, the width of the Gaussian kernel is set to 0.55.

In our study, the number of neurons in the input layer and hidden layer of a single NN (FFNN, TDNN, NARX, RBFNN) is set to one and two respectively. They are set to seven in both input and hidden layer in the case of the NN used to combine all forecasts. In all cases, the output layer has one neuron corresponding to the predicted return.

2.5 Performance Measure

Each single NN and NN committee accuracy is evaluated by computing the mean absolute error (MAE) which is defined as follows:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - p_i| \quad (8)$$

where y is the observed value, p is the predicted value, and m is the total number of observations in the testing data. The lower is the MAE, the better is the accuracy.

3 Data and Results

The empirical study involved the prediction of the NASDAQ price return. Submarket set includes banking (first submarket), biotechnology (second submarket), insurance (third submarket), other finance (fourth submarket), Telecom (fifth submarket), transport (sixth submarket), and computers (seventh submarket). The data were daily price values from 3 January 2007 to 15 November 2013. The purpose is to predict the aggregate market (NASDAQ) return series. They are computed as first difference of log-price. Figures 2 to 8 provide banking, biotechnology, insurance, other finance, Telecom, transport, and computers submarket return series respectively. The aggregate market NASDAQ return series are depicted in Figure 9. The learning phase consisted of 80% of the observations, while the testing phase consisted of the remaining 20%. All the obtained results were compared and evaluated by the mean absolute error (MAE) statistic. The empirical results are depicted in Table 1. It indicates that in all cases the NN ensemble (committee) generated the lowest errors than single NN. This result is in accordance with previous works found in the literature [5-13].

Comparing accuracies between FFNN ensemble, TDNN ensemble, NARX ensemble, and RBFNN ensemble, there are differences in favor of the latter when looking at MAE statistic. For instance, they respectively achieved 0.0036, 0.033, 0.0028, 0.0016 MAE. This result indicates that TDNN ensemble, NARX ensemble, and RBFNN ensemble all outperformed the conventional FFNN ensemble used as the main reference NN ensemble. Thus, ensemble system composed of more advanced NN architecture yield to lower prediction error. Among the NN ensembles considered in this study, the RBFNN achieved the lowest forecasting error. This could be explained by the fact that

RBF neural networks have advantages of easy design, good generalization, strong tolerance to input noise, and online learning ability in comparison with traditional neural networks including sophisticated fuzzy inference systems [24]. In addition, contrary to the traditional neural networks trained with back-propagation algorithm each hidden unit of the RBFNN acts locally by computing a score for the match between the input vector and its centers. As a result, the basis units are highly specialized to detect patterns in the underlying data.

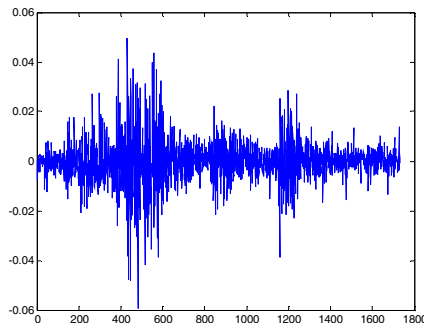


Fig. 2. Banking sub-market return series

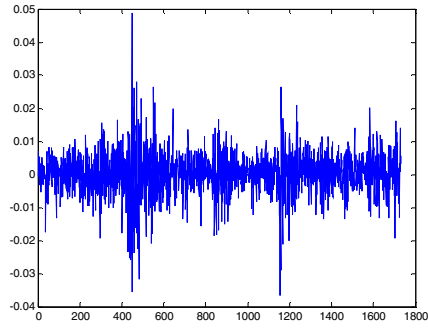


Fig. 3. Biotechnology sub-market return series

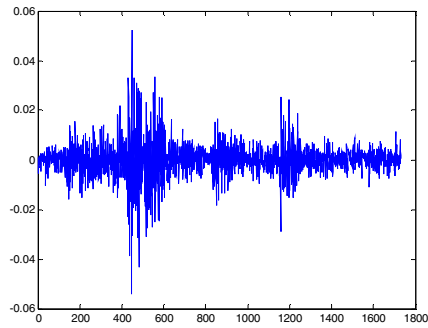


Fig. 4. Insurance sub-market return series

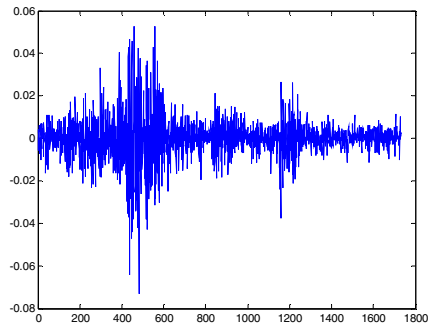


Fig. 5. Other finance sub-market return series

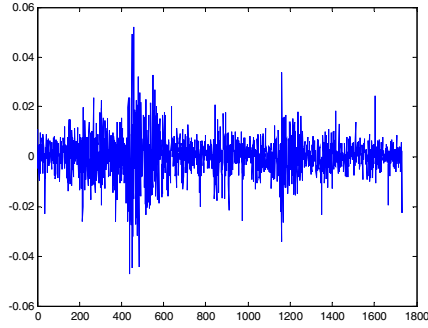


Fig. 6. Telecom sub-market return series

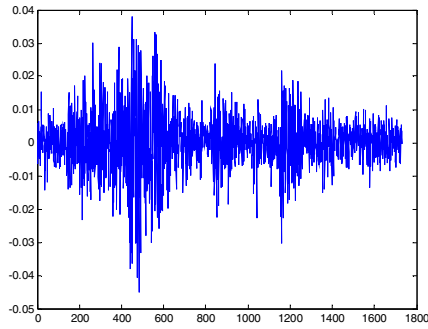


Fig. 7. Transport sub-market return series

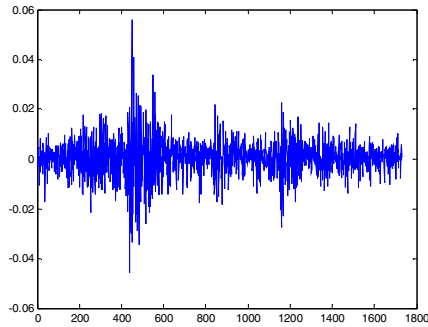


Fig. 8. Computers sub-market return series

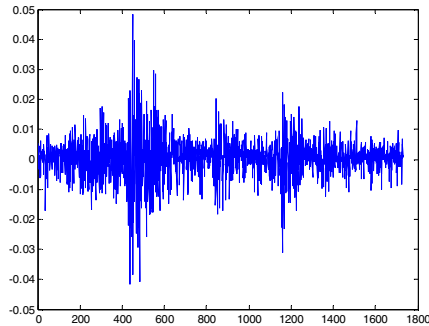


Fig. 9. NASDAQ return series

Table 1. MAE results

	FFNN	TDNN	NARX	RBFNN
Component 1	0.0152	0.0416	0.0080	0.0025
Component 2	0.0189	0.0176	0.0325	0.0025
Component 3	0.0379	0.0431	0.0095	0.0019
Component 4	0.0248	0.0349	0.0084	0.0023
Component 5	0.0220	0.0327	0.0410	0.0021
Component 6	0.0279	0.0288	0.0063	0.0023
Component 7	0.0192	0.0264	0.0252	0.0018
Ensemble	0.0036	0.0033	0.0028	0.0016

4 Conclusion

This paper evaluates four neural network ensembles each with different neural network architecture with application in forecasting NASDAQ returns. They are the conventional feed-forward back-propagation neural network which widely used in the literature, time-delay neural network, NARX network, and the radial basis neural network. Each ensemble component is used to model the relationship between NASDAQ returns and a given NASDAQ submarket returns and to provide a forecast. The outputs of the component neural networks are combined together by the combination module to produce the final output of the ensemble. The combination module is represented by a neural network. This result indicated that TDNN ensemble, NARX ensemble, and RBFNN ensemble all outperformed the conventional FFNN ensemble which was used as the main reference NN ensemble. In addition, the RBFNN achieved the lowest forecasting error.

In summary, our findings are in accordance with those of the literature: ensemble neural networks perform better than single neural networks. In addition, this work demonstrated that ensemble systems composed of more advanced NN architectures yield to lower prediction error than neural network ensemble composed of conventional feed-forward neural networks widely used in the literature. Indeed, such sophisticated

architectures provided significant improvement in the prediction accuracy of the NASDAQ return series.

Future research directions include considering other types of time series forecasting problems for better generalization of the results, and considering classification problems.

References

1. Francis, E.H., Chao, L.J.: Modified Support Vector Machine in Financial Time Series Forecasting. *Neurocomputing* 48, 847–861 (2002)
2. Zhang, G.P., Berardi, V.L.: Time Series Forecasting with Neural Network Ensembles: An Application for Exchange Rate Prediction. *Journal of the Operational Research Society* 52, 652–664 (2001)
3. Atsalakis, G.S., Valavanis, K.P.: Surveying Stock Market Forecasting Techniques – Part II: Soft Computing Methods. *Expert Systems with Applications* 36, 5932–5941 (2009)
4. Bahrammirzaee, A.: A Comparative Survey of Artificial Intelligence Applications in Finance: Artificial Neural Networks, Expert System and Hybrid Intelligent Systems. *Neural Computing & Applications* 19, 1165–1195 (2010)
5. Xiao, Y., Xiao, J., Lu, F., Wang, S.: Ensemble ANNs-PSO-GA Approach for Day-Ahead Stock Exchange Prices Forecasting. *International Journal of Computational Intelligence Systems* 6, 96–114 (2013)
6. Kim, M.J., Min, S.H., Han, I.: An Evolutionary Approach to The Combination of Multiple Classifiers to Predict a Stock Price Index. *Expert Systems with Applications* 31, 241–247 (2006)
7. Tsai, C.-F., Lin, Y.-C., Yen, D.C., Chen, Y.-M.: Predicting Stock Returns by Classifier Ensembles. *Applied Soft Computing* 11, 2452–2459 (2011)
8. Wang, L., Wu, J.: Neural Network Ensemble Model using PPR and LS-SVR for Stock Market Forecasting. In: Huang, D.-S., Gan, Y., Bevilacqua, V., Figueroa, J.C. (eds.) *ICIC 2011*. LNCS, vol. 6838, pp. 1–8. Springer, Heidelberg (2011)
9. Kim, E., Kim, W., Lee, Y.: Combination of Multiple Classifiers for The Customer's Purchase Behavior Prediction. *Decision Support Systems* 34, 167–175 (2003)
10. Goh, W.Y., Lim, C.P., Peh, K.K.: Predicting Drug Dissolution Profiles with An Ensemble of Boosted Neural Networks: A Time Series Approach. *IEEE Transactions on Neural Networks* 14, 459–463 (2003)
11. Langella, G., Basile, A., Bonfante, A., Terribile, F.: High-Resolution Space–Time Rainfall Analysis using Integrated ANN Inference Systems. *Journal of Hydrology* 387, 328–342 (2010)
12. Fildes, R., Kourentzes, N.: Validation and Forecasting Accuracy in Models of Climate Change. *International Journal of Forecasting* 27, 968–995 (2011)
13. Zheng, J.: Predicting Software Reliability with Neural Network Ensembles. *Expert Systems with Applications* 36, 2116–2122 (2009)
14. Zhang, G., Patuwo, B.E., Hu, M.Y.: Forecasting with Artificial Neural Networks: The State of The Art. *International Journal of Forecasting* 14, 35–62 (1998)
15. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning Representations by Back-Propagating Errors. *Nature* 323, 533–536 (1986)
16. Cybenko, G.: Approximation by Superpositions of Sigmoidal Function. *Math. Contr. Signals Syst.* 2, 303–314 (1989)

17. Funahashi, K.-I.: On The Approximate Realization of Continuous Mappings by Neural Networks. *Neural Networks* 2, 183–192 (1989)
18. Hornik, K.: Approximation Capabilities of Multilayer Feedforward Networks. *Neural Networks* 4, 251–257 (1991)
19. Kim, M.-J., Kang, D.-K.: Ensemble with Neural Networks for Bankruptcy Prediction. *Expert Systems with Applications* 37, 3373–3379 (2010)
20. Kourentzes, N., Barrow, D.K., Crone, S.F.: Neural Network Ensemble Operators for Time Series Forecasting. *Expert Systems with Applications* 41, 4235–4244 (2014)
21. Hansen, L.K., Salamon, P.: Neural Network Ensembles. *IEEE Trans. Pattern Analysis and Machine Intelligence* 12, 993–1001 (1990)
22. Kim, S.S.: Time-Delay Recurrent Neural Network for Temporal Correlations and Prediction. *Neurocomputing* 20, 253–263 (1998)
23. Leontaritis, I.J., Billings, S.A.: Input–Output Parametric Models for Non-Linear Systems. Part I: Deterministic Non-Linear Systems. *International Journal of Control* 41, 303–328 (1985)
24. Hao, Y., TianTian, X., Paszczynski, S., Wilamowski, B.M.: Advantages of Radial Basis Function Networks for Dynamic System Design. *IEEE Transactions on Industrial Electronics* 58, 5438–5450 (2011)