Exploiting Neural Networks to Enhance Trend Forecasting for Hotels Reservations

Athanasius Zakhary[1], Neamat El Gayar[1,2], and Sanaa El-Ola. H. Ahmed[1]

[1] Faculty of Computers and Information, Cairo University, 12613 Giza, Egypt
{athanasius.it,sana.ola}@gmail.com
[2] Center for Informatics Science, School of Communication and
Information Technology, Nile University, Giza, Egypt
nelgayar@nileuniversity.edu.eg

Abstract. Hotel revenue management is perceived as a managerial tool for room revenue maximization. A typical revenue management system contains two main components: Forecasting and Optimization. A forecasting component that gives accurate forecasts is a cornerstone in any revenue management system. It simply draws a good picture for the future demand. The output of the forecast component is then used for optimization and allocation in such a way that maximizes revenue. This shows how it is important to have a reliable and precise forecasting system. Neural Networks have been successful in forecasting in many fields. In this paper, we propose the use of NN to enhance the accuracy of a Simulation based Forecasting system, that was developed in an earlier work. In particular a neural network is used for modeling the *trend* component in the simulation based forecasting model. In the original model, *Holt's* technique was used to forecast the trend. In our experiments using real hotel data we demonstrate that the proposed neural network approach outperforms the Holt's technique. The proposed enhancement also resulted in better arrivals and occupancy forecasting when incorporated in the simulation based forecasting system.

1 Introduction

The importance of hotel arrivals and hotel occupancy forecasting stems from its need when designing a hotel revenue management system. Revenue management is the science of managing a limited amount of supply to maximize revenue by dynamically controlling the price/quantity offered [1] [2]. While cutting edge revenue management systems are currently in use by the airline industry, only recently has it been considered for the hotel industry. Because of the large number of existing hotels, the amount of possible total savings, if one implements optimized revenue management systems, is potentially large.

Revenue management forecasting methods fall into one of three types: Historical booking models, advanced booking models and combined models. Historical booking models only consider the final number of rooms or arrivals on a particular stay night. Advanced booking models only include the build up of reservations

over time for a particular stay night. Combined models use either regression or a weighted average of historical data and advanced booking models to develop forecasts. A review of forecasting methods for all three types is found in [3] and [6]. In this study, a particular interest is devoted to reservations data as they are very rich and contain very useful information indicating the actual demand to come.

A simple but efficient forecasting technique that is based on reservations data is the *Pickup* method. The *Pickup* forecasting model is a popular advanced booking model which exploits the unique characteristics of reservations data instead of relying only on complete arrival histories to make better forecasts. The main idea of using the pickup method is to estimate the *increments* of bookings (to come) and then aggregate these increments to obtain a forecast of total demand in the future [1]. *Pickup* is defined as the number of reservations picked up from a given point of time to a different point of time over the booking process [3]. A good comparison among the different variations of the pickup forecasting method and applying it on simulated hotel data can be found in [4].

In the theory of forecasting there have been two competing philosophies. The first one is based on developing an empirical formula that relates the value to be forecasted with the recent history (for example ARIMA-type or exponential smoothing models). The other approach focuses on developing a model from first principles that relates the value in question with the available variables/parameters etc, and simulate that model forward to obtain the forecast. Because the majority of real-world systems are either intractable or very complex to model, most forecasting applications follow the first approach. In contrast, Zakhary et al [5] present a forecasting application following the second approach, namely *Forecasting Hotel Arrivals and Occupancy Using Monte Carlo Simulation*.

In this paper, we propose an enhancement to the above mentioned model. In particular, a neural networks model is proposed to provide a multi-step ahead forecast for the *trend* component of the simulation model. In our experiments we demonstrate that the proposed neural network based model provides a more accurate estimate for the trend and results in better arrivals and occupancy forecasting when incorporated in the overall forecasting system. Our case study was based on real reservations data from Plaza hotel [1], Alexandria, Egypt.

The paper is organized as follows: Next section briefly describes the simulation based forecasting system, specifically the estimation of the system components. Section 3 reviews Neural Networks and its usage in forecasting. Section 4 presents the details of the proposed method for forecasting the trend. Section 5 shows the experiments and finally the paper is concluded in Section 6.

2   Forecasting Arrivals and Occupancy Using Simulation

This section briefly describes the simulation based forecasting model. For more details refer to [5]. The forecasting simulator consists of two main modules: *the*

---

[1] Plaza Hotel, Alexandria, Egypt is a 4 stars, Mid-size, business as well as sea-side hotel. `http://www.plazaegypt.com/home.htm`

analysis module and *the simulation module.* The analysis module takes as an input the historical reservation records. It then analyzes these data and uses it to extract many parameters and components like: trend, seasonality, booking curves, cancellations dynamics, length of stay, etc. Distributions of these components are deduced from the data directly. For every distribution or parameter, a suitable approach is devised to estimate it. For example, the seasonal index is estimated using the multiplicative seasonal decomposition [6]. The length of stay and room type distributions are estimated using a simple frequency based distribution estimator. The effects of the average daily reservations (i days prior to arrival) and the trend and seasonality are estimated by assuming that the net reservation rate $b(i)$ equals a normalized reservation rate $b_n(i)$ times a trend component $t_r(t)$ times a seasonal component $s(t)$, where t is the arrival date, as follows:

$$b(i) = b_n(i) * t_r(t) * s(t) \tag{1}$$

The reservation rates $b_n(i)$ are estimated by grouping the reservation curves to high, medium and low according to their corresponding seasonality indexes, and estimating a template for each seasonal index. The simulation module takes the parameters and components estimated by the analysis module as an input to generate forward reservation records that would take place in the future. Analyzing these generated reservation cases; one can obtain realistic predictions for occupancy, arrivals, and revenue in the future. One advantage of this simulation approach is that one obtains distributions of future key parameters (reservations, arrivals, etc). This is performed by running the simulator many times from the current (deterministic) starting point (determined by the current snapshot of existing reservations), thereby producing many paths, each obtained by the different random components of the future reservations process.

3   NN in Forecasting

A neural network (NN) is a semi-parametric model, inspired by how the brain processes information. It consists of a network of neurons (or nodes) that perform a weighted sum operation, followed by applying a nonlinear squashing function. Some of these nodes are hidden nodes that perform intermediate computations, and feed into the output node (which produces the final output). There are typically many free parameters in the network (called the weights). If the number of parameters or the number of hidden nodes is left unchecked, the network can overfit the data [7].

NNs can be useful for nonlinear processes that have an unknown functional relationship and as a result are difficult to fit [8]. The main idea with NNs is that inputs, or dependent variables, get filtered through one or more hidden layers each of which consist of hidden units, or nodes, before they reach the output variable. Next the intermediate output is related to the final output [9]. One major application area of NNs is forecasting. Refer to Kline et al [10]. for a good survey of the literature. A general problem with nonlinear models is the "curse of model complexity and model over parametrization".

NNs offers a natural alternative to traditional forecasting techniques. They have three great advantages over traditional forecasting methods:

1. They have universal approximation capabilities,
2. They can recognize "on their own" implicit dependencies and relationships in data,
3. They can learn to adapt their behavior (their prediction) to changed conditions quickly and without complication.

4   Proposed Trend Component Forecasting Using NN

As shown in Equation 1, the trend $t_r(t)$ is a key component in forecasting the reservations and then arrivals and occupancy. Enhancing the accuracy of estimating the trend will definitely lead to a more precise forecasting model. For this reason a forecasting model is applied for predicting the trend in the considered forecast horizon. Due to the importance of this component to the whole system, Artificial Neural Networks is used to forecast it.

Investigating the nature of the trend curve shown in Figure 1, it is obvious that it changes very frequently and abruptly. In addition, the training process in neural networks usually envolves intensive computations. For these reasons, it is better to simplify the forecasting problem. This could be done by converting
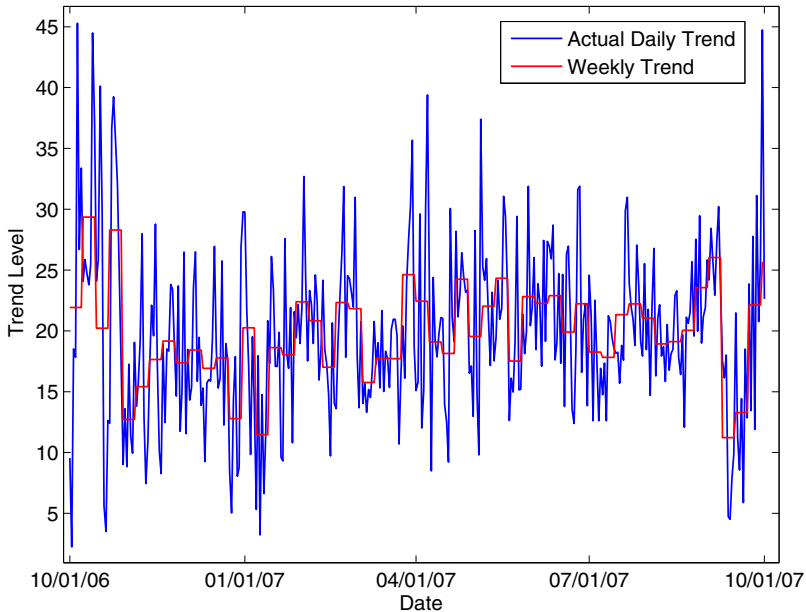

Fig. 1. The daily trend extracted from Plaza Hotel data versus a weekly version

the daily trend into a weekly trend by averaging the trend values of each week. In this case, the points to be forecasted will shrink by a factor of seven. For example, to forecast three months, we need to forecast 14 weekly points instead of 92 daily points. Each forecasted point is then considered a forecast for the corresponding week.

The proposed approach is to train a set of different neural networks for forecasting the future trend levels. These neural networks will vary in forecast horizon, number of inputs, formation of inputs, and the number of hidden nodes. For each forecast horizon, the best three neural networks are used in forecasting. The forecast value is the average of their forecasts. The proposed method can be described in steps as follows:

1. Prepare the inputs for Neural Networks
2. Adjust the NN parameters (Number of hidden units, Learning rate, ...)
3. Train the different NNs
4. Forecast using the best trained NNs

In the following subsections, we describe the sequence in more details.

4.1 Prepare Inputs

In multi step ahead forecasting using NN there are three well-known ways: the *iterative* approach, the *direct* approach, and the *parallel* approach [10]. In the *iterative* approach the model is trained on a one-step ahead basis. The trained model is then used to forecast one step ahead (i.e. one week ahead). The forecasted value is then fed as input to the model to forecast the subsequent point, and continue in this manner until the end of the forecast horizon. In the direct approach, a different network is used for every point in the future to be forecasted. For example, if the forecast horizon is 14 weeks, so 14 networks are built, each one forecasting a specific week in the future. When forecasting, these 14 networks are applied to obtain the forecast of the required forecast horizon. Of course, every network is trained separately. In the parallel approach, only one network is used with a number of outputs equal to the length of the horizon to be forecasted. The network is trained in a way so that output number k produces the kth step ahead forecast. In this work, we mainly focus on the *direct* approach. Neural networks are trained with a number of inputs that varies from two inputs to six inputs. Input/output pairs used in training can be built using different ways. As follows, we describe the different ways we suggest to build the input/output pairs:

- M-I: input lags are consecutive irrespective of the forecast horizon
- M-II: input lags are interleaved in proportion to step-ahead being forecasted
- M-III: consists of sampling in the same way as M-II. Instead of skipping time series values, averages are formed. The target output is also averaged
- M-IV: is same like M-III without averaging output
- M-V: is same like M-I , but one of the inputs is the average of eight consequent inputs. This is to account for the effect of longer term average

Given the time series $x(t-m), ..., x(t-3), x(t-2), x(t-1), x(t), x(t+1), x(t+2), x(t+3), ..., x(t+n)$. The generated input/output pairs for horizon=2 and number of lags =2 for this series using the different aforementioned methods are:

- M-I: $x(t-1), x(t) \Rightarrow x(t+2)$
- M-II: $x(t-2), x(t) \Rightarrow x(t+2)$
- M-III: $average(x(t-3), x(t-2)), average(x(t-1), x(t)) \Rightarrow average(x(t+1), x(t+2))$
- M-IV: $average(x(t-3), x(t-2)), average(x(t-1), x(t)) \Rightarrow x(t+2)$
- M-V: $average(x(t-8), ..., x(t-1)), x(t) \Rightarrow x(t+2)$

4.2   Adjust Parameters and Train Neural Networks

A five-fold validation procedure is used to select the best number of hidden nodes [11]. The following candidate values are considered for the number of hidden nodes: NH = [0, 1, 3, 5]. Note there is a possibility of having zero hidden nodes (NH = 0), meaning simply a linear model. Balkin and Ord [12] have shown that the possibility of switching to a linear model for some time series improved performance. Concerning the other less key parameters and model details, The sigmoid is used as activation functions for the hidden layer, and a linear for output layer. Training is performed using the Levenberg-Marquardt algorithm for 500 epochs, using a momentum term 0.2, and an adaptive learning rate with initial value 0.01, an increase step of 1.05 and a decrease step of 0.7.

Trend variable $s_{des}(t)$ is converted to weekly averages. Neural Networks are built with variations in forming training data (five variations) and number of lags (five variations). For each forecast horizon, the best three NNs are chosen.

4.3   Forecast Trend Using Trained NNs

For each forecast horizon, formulate the input according to the selected NN structure. Forecast the value of the required horizon in weeks. The output of each forecast horizon is the mean value of the three selected NN of that horizon. The next step is to convert the weekly forecasted points to daily forecasts. This is done by simply repeating each forecast seven times. The last step is to additively tune the level of the forecast. This is done by fitting a straight line on the past data and then getting the forecast of this linear regression for the interval to be forecasted. The mean value of the NN forecast is additively set to be that of the regressed forecast. This is done by subtracting the NN forecast mean and adding the linear regression mean.

Next section describes the results of applying the proposed NN model versus the Holt's technique used primarily in the simulator based forecasting system to forecast the trend obtained from Plaza Hotel for three intervals.

5   Experiments and Results

The proposed model is trained using the trend data of a complete year. The model is tested by forecasting the subsequent 3 months. Table 1 lists the 3

Table 1. The in-sample and the three months ahead forecast periods for the three forecasted snapshots

Snapshot No.	In-Sample Period	Forecast Period
1	1-Oct-2006 - 30-Sep-2007	1-Oct-2007 - 31-Dec-2007
2	1-Oct-2006 - 31-Oct-2007	1-Nov-2007 - 31-Jan-2008
3	1-Oct-2006 - 30-Nov-2007	1-Dec-2007 - 29-Feb-2008

different training and testing periods that were used in our experiment. We compare the performance of the proposed model to Holt's forecasting method. A detailed discussion to this technique can be found in [13]. We have chosen to compare the proposed NN Forecasting technique to Holt's forecasting method for two reasons:

1. The Holt's technique is used for forecasting the trend in our original simulation based forecasting system [5],
2. By definition Holt's is a suitable technique for forecasting data that have trend and no seasonality[13].

We use as error measure the symmetric mean absolute percentage error, defined as follows for both the the arrivals and occupancy time series:

$$SMAPE = \frac{1}{M} \sum_{m=1}^{M} \frac{|\hat{y}_m - y_m|}{(|\hat{y}_m| + |y_m|)/2} \tag{2}$$

where y_m and $\hat{y}_m$ are the actual time series value and the forecasted value respectively. Also, M is the total number of points that are forecasted.

Due to the limited length of the actual data (17 complete months), only three different forecast snapshots could be used. Thus we could not apply test of significance on the results. Figure 2 depicts the trend versus the NN forecast and Holt's forecast for the first interval. Table 2 compares the SMAPE error measures of the proposed model to the Holt's method. Analyzing the results, it is obvious that Neural Network forecasting method has outperformed Holt's methods.

In other cases, trend can go upward or downward. To test the performance of the NN model in such cases, an upward trend is generated by multiplying the original trend of plaza in a straight line with an upward trend. The model is trained using the modified trend. Forecast of the first interval against the modified trend is shown in Figure 3. Table 3 shows the corresponding SMAPE error measures. Similarly, trend has been modified with a downward trend and the results are shown in Figure 4 and Table 4. Investigating the results one can see that NN model has stable performance whether the trend is stationary, upward or downward. Finally in Table 5 we compare the overall forecast SMAPE for the Out-of-Sample periods for the original Monte Carlo Model [5] to the enhanced model that uses the proposed NN trend forecasting.
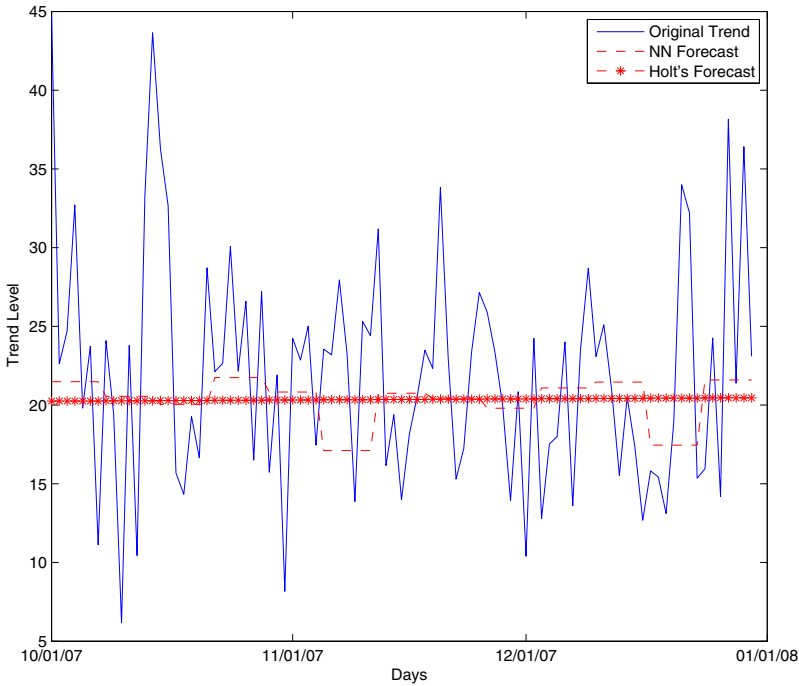
Fig. 2. The original trend component of Plaza Hotel versus its forecasts using the proposed NN model and Holt's in the 1st interval

Table 2. The SMAPE error measures of the proposed NN model versus Holt's for forecasting the trend of Plaza Hotel

	NN model	Holts
1st Interval	0.2574	0.2697
2nd Interval	0.2537	0.2589
3rd Interval	0.2625	0.2808
Average	0.2541	0.270867

Table 3. The SMAPE error measures of the proposed NN model versus and Holt's for forecasting an upward modified version of trend of Plaza Hotel

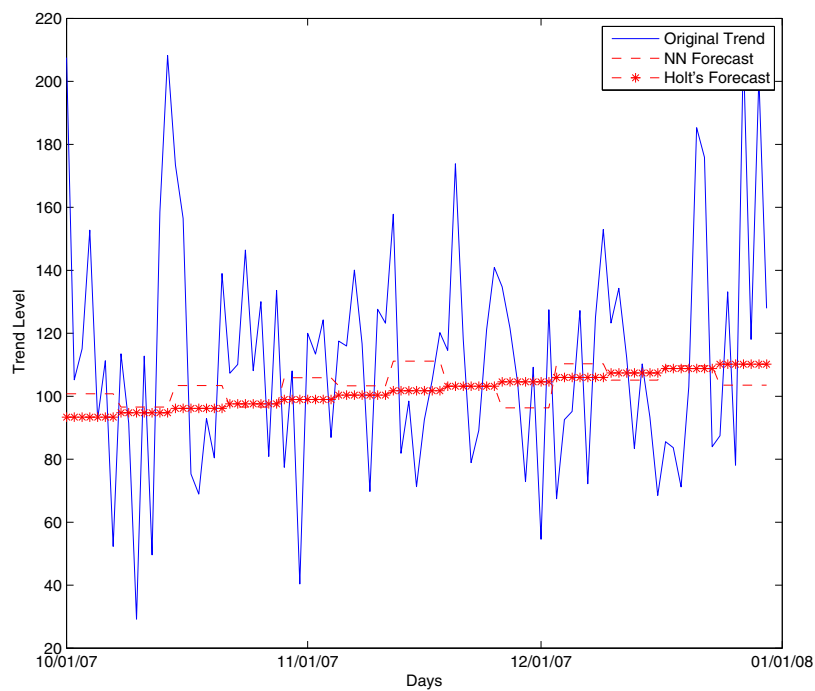|	NN model	Holts
1st Interval	0.2674	0.2729
2nd Interval	0.2424	0.2589
3rd Interval	0.2625	0.2808
Average	0.25743	0.2709

Fig. 3. The modified upward trend component of Plaza Hotel versus its forecasts using the proposed NN model and Holt's in the 1st interval

Table 4. The SMAPE error measures of the proposed NN model versus and Holt's for forecasting an downward modified version of trend of Plaza Hotel

	NN model	Holts
1st Interval	0.2737	0.2846
2nd Interval	0.2364	0.2482
3rd Interval	0.2786	0.2921
Average	0.2629	0.2750

Table 5. The Overall Forecast SMAPE for the Out-of-Sample Periods for the Original Monte Carlo Model versus Modified Model

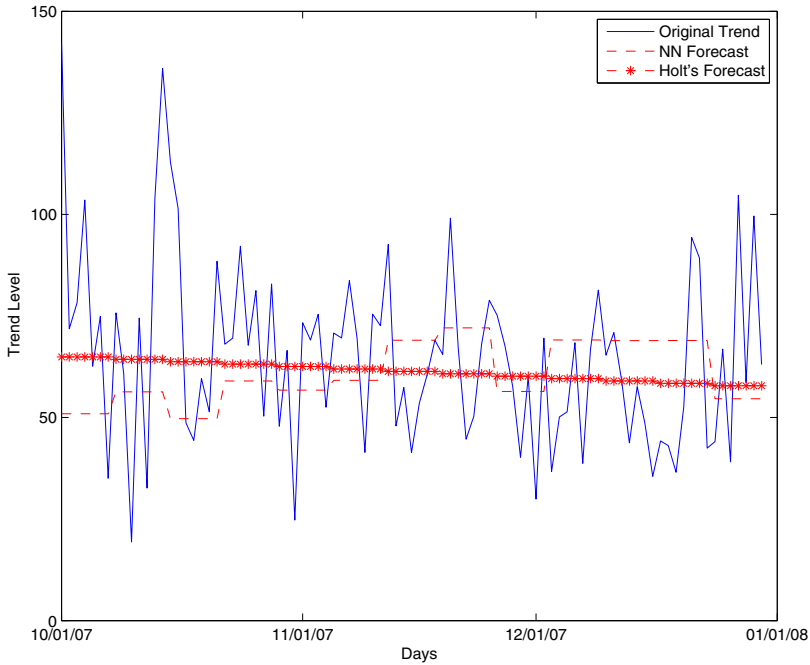|	Arrivals	Occupancy
Original Monte Carlo Model	43.9	37.7
Modified Monte Carlo Model	42.21	37.1

Fig. 4. The modified downward trend component of Plaza Hotel versus its forecasts using the proposed NN model and Holt's in the 1st interval

6   Conclusions

In this paper we propose an enhancement to a simulation based forecasting model previously developed. In particular we suggest using NNs to estimate the trend component, which is one of the key components that the forecasting model is based upon. In the original model, Holt's method is used for trend forecasting. Experimental results on real hotel data show that the use of NN increases the accuracy of trend forecasting by 5%. This improvement is also reflected in the overall arrivals and occupancy results of the overall simulation based forecasting system.

We believe that the gain from using NN for estimating the trend component in the forecasting system can be more significant in hotel environments in which the trend component is the main contributor to the reservation process. This is of particular importance to us since we plan to test and deploy the developed forecasting system in various hotel types (resort, city hotels, airport hotels, ...etc) that exhibit different reservation patterns.

Acknowledgment

References

1. Talluri, K.T., Van Ryzin, G.J.: The Theory and Practice of Revenue Management. Springer Science+Buisness Media, Inc. (2005)
2. Ingold, A., McMahon-Beattie, U., Yeoman, I. (eds.): Yield Management. Continuum, 2nd edn. (2003)
3. Weatherford, L.R., Kimes, S.E.: A comparison of forecasting methods for hotel revenue management. International Journal of Forecasting 99, 401–415 (2003)
4. Zakhary, A., El Gayar, N., Atiya, A.F.: A comparative study of the pickup method and its variations using a simulated hotel reservation data. ICGST International Journal on Artificial Intelligence and Machine Learning 8, 15–21 (2008)
5. Zakhary, A., Atyia, A., El-Shishiny, H., El Gayar, N.: Forecasting hotel arrivals and occupancy using monte carlo simulation. Journal of Revenue and Pricing Management (to appear)
6. Frechtling, D.: Forecasting Tourism Demand: Methods and Strategies. Butterworth Heinemann, Oxford (2001)
7. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice-Hall, Englewood Cliffs (1999)
8. Darbellay, G.A., Slama, M.: Forecasting the short-term demand for electricity: Do neural networks stand a better chance? International Journal of Forecasting 16, 71–83 (2000)
9. Gooijer, J.G.D., Hyndman, R.J.: 25 years of IIF time series forecasting: A selective review. Monash Econometrics and Business Statistics Working Papers 12/05, Monash University, Department of Econometrics and Business Statistics (2005)
10. Kline, D.M., Zhang, G.P.: Methods for multi-step time series forecasting with neural networks. Neural Networks for Business Forecasting, 226–250 (2004)
11. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings International Joint Conference on Artificial Intelligence, IJCAI (1995)
12. Sandy, J.K.O., Balkin, D.: Automatic neural network modeling for univariate time series. International Journal of Forecasting 16(4), 509–515 (2000)
13. Gardner, E.S.: Exponential smoothing: The state of the art Part II. International Journal of Forecasting 22, 637–666 (2006)