Demand Modelling in Telecommunications

Comparison of Standard Statistical Methods and Approaches Based upon Artificial Intelligence Methods Including Neural Networks

M. Chvalina

This article analyses the existing possibilities for using Standard Statistical Methods and Artificial Intelligence Methods for a short-term forecast and simulation of demand in the field of telecommunications. The most widespread methods are based on Time Series Analysis. Nowadays, approaches based on Artificial Intelligence Methods, including Neural Networks, are booming. Separate approaches will be used in the study of Demand Modelling in Telecommunications, and the results of these models will be compared with actual guaranteed values. Then we will examine the quality of Neural Network models.

Keywords: Demand, telecommunications, standard statistical methods, Box-Jenkins methodology, ARIMA, artificial intelligence methods, neural network.

1 Introduction

Demand can be defined as the relation between price and the quantity of goods that buyers are willing to purchase. This correlation is displayed in relation to the global market by the sold product quantity at one time-point. If we focus on the telecommunication services sector, we can note the development of the sale of cell phones and internet extensions (ADSL, ISDN, GPRS, Wi-Fi etc.)

Factors affecting the development of demand are for example technology, price and considerations from the fields of psychology, sociology and economies.

Generally, demand can be considered as a time series under which the demand model can be defined and development trends can be predicted.

2 Construction of the demand model

The demand model can be constructed using standard statistical methods including Decomposition Time Series, and the Box-Jenkins Metodology, or by applying Artificial Intelligence methods, including for example Neural Networks.

2.1 Decomposition time series

The series $\{y_t, t = 1, \ldots, T\}$ is gradually decomposed to several components: trend, circular component, seasonal component and residual component (unsystematic component). This method is based on work with time series systematic components. Features of time series behaviour can be better observed in separate components than in the undecomposed original time series.

In this research study from the field of standard statistical methods, exponential smoothing will be used for demand modelling.

Exponential Smoothing

The above defined time series will be written as $\{y_t, t = 1, \ldots, T\}$. Simple Exponential Smoothing is described in the recurrent form $\hat{y}_t = \alpha y_t + (1-\alpha)\hat{y}_{t-1}$, $\hat{y}_t$ is the Exponential Average in time t, $\hat{y}_{t-1}$ is the Exponential Average in time $t-1$, value α is the Smoothing Constant from the interval $\alpha \in \langle 0;1 \rangle$. The Exponential Average can be expressed on the basis of the recurrent form as:

$$\begin{aligned}\hat{y}_t &= \alpha y_t + (1-\alpha)\hat{y}_{t-1} = \alpha y_t + (1-\alpha)[\alpha y_{t-1} + (1-\alpha)\hat{y}_{t-2}] \\ &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2[\alpha y_{t-2} + (1-\alpha)\hat{y}_{t-3}] = \ldots \\ &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2 \alpha y_{t-2} + \ldots + \alpha(1-\alpha)^i y_{t-i} + \ldots \\ &\ldots + (1-\alpha)^t \hat{y}_0 = \alpha \sum_{i=0}^{t=1} (1-\alpha)^i y_{t-i} + (1-\alpha)^t \hat{y}_0.\end{aligned}$$

Brown's Simple Exponential Smoothing

The time series y_t is constructed with a stationary process in the form $y_t = \beta_0 + \varepsilon_t$, β_0 is the mean value of the process, and ε_t are random values with the features of white noise. After applying Exponential Smoothing to the Time Series $y_t = \beta_0 + \varepsilon_t$ we obtain the relation

$$\hat{y}_t = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i y_{t-i} = \alpha \sum_{i=0}^{\infty} (1-\alpha)^i (\beta_0 + \varepsilon_{t-i})$$

$$= \beta_0 + \alpha \sum_{i=0}^{\infty} (1-\alpha)^i \varepsilon_{t-i}$$

because

$$\alpha \sum_{i=0}^{\infty} (1-\alpha) = 1$$

applies to the mean value and dispersion

$$E(\hat{y}_t) = E(y_t) = \beta_0, \qquad D(\hat{y}_t) = \frac{\alpha}{2-\alpha}\sigma_a^2 = \frac{\alpha}{2-\alpha}\sigma_y^2.$$

Brown's Linear (double) Exponential Smoothing

Alternatively, we can make a double application of the Simple Exponential Smoothing method on the time series y_t expression in the form

$$\hat{y}_t^{(2)} = \alpha y_t + (1-\alpha)\hat{y}_{t-1}^{(2)}.$$

2.2 Box-Jenkins methodology

Unlike classical decompositional methods, which deal with systematic time series components (trend, circular and seasonal), the Box-Jenkins methodology deals with a residual (unsystematic) component. The method involves searching for relations of individual observations. By this method we are able to describe time series which are not manageable by standard methods.

The time series is perceived as a realization of the stochastic process which is defined as a series of accidental quantities arranged in time $\{X(s,t), s \in \mathbf{S}, t \in \mathbf{T}\}$, where $\mathbf{S}$ is a selective space and $\mathbf{T}$ is an index series. For each $s \in \mathbf{S}$ the realization of the stochastic process is defined on the index series $\mathbf{T}$.

For the Box-Jenkins methodology, the following special accesses are significant: autoregressive process AR, moving average process MA, and combined process ARMA. These processes result from the linear process by resetting all parameters till the final number. The parameters are chosen in such a way that the stationarity and invertibility of the processes will be ensured. A special non-stationary ARIMA model also exists in the Box-Jenkins methodology.

ARIMA processes

Some integrated processes may be arranged by means of differentiation to stationary and are expressed in the form of the stationary and invertible ARMA(p, q) model. The original integrated process in the form:

$$\phi_p(B)(1-B)^d y_t = \theta_q(B)\varepsilon_t$$

is called the autoregressive integrated process of sliding averages of the order p, d, q. This is called ARIMA (p, d, q). Models with $d = 1,2$ are usually used.

2.3 Artificial intelligence methods

Neural Networks

The benefit of the Artificial Neural Network lies in its ability to implement complex non-linear functions. Neural systems co-execute a large number of operations and work without an algorithm. Their activity is based upon the learning process, when the neural network gradually conforms to calculation. In the course of a learning phase we do not have to be occupied by the problem of the right selection function, because the neural network is able to make do only with practise examples.

This is the main difference in comparison with a traditional approach (e.g. comparison of traditional non-linear models and a back-propagation network). The neurone function scheme can be demonstrated graphically as follows:

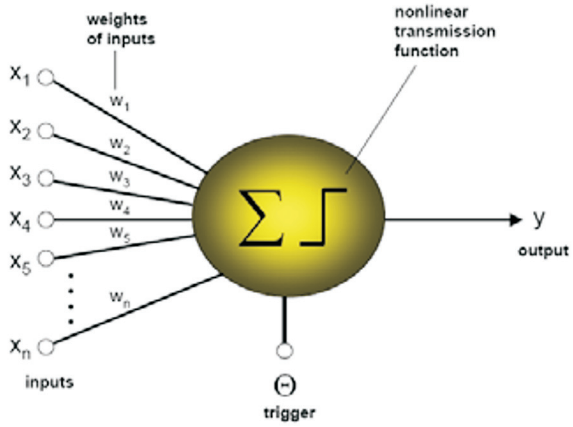

Fig. 1: Model of a neurone

$$y = S\left[\sum_{i=1}^{N} w_i x_i + \Theta\right],$$

where y is the output (neurone activity), S demonstrates a transmission function (jump, linear, non-linear), x_i is neurone input (inputs are in total N), w_i represents a level of synoptic weight, Θ describes a trigger level.

The principle of back-propagation network learning

Let us consider a neurone network with L – layers $l = 1, \ldots, L$ and as the output i-th neurone in the 1-th layer we use the indication V_i^l. V_i^0 means x_i, i.e. the i-th output. The indication w_{ij}^l expresses a connection from V_i^{l-1} to V_i^l. An algorithm can then be inscribed after separate stages as follows:

1. subranged random numbers based-Weight Initialization.
2. Insertion of $\vec{x}^P$ into the network input (layer $l = 0$), i.e. $V_k^0 = x_k^p$.
3. Network signal propagation:
   $$V_i^l = g(h_i^l) = g\left(\sum_k w_{ij}^l V_j^{l-1}\right).$$
4. Calculation of delta for the output layer:
   $$\delta_i^p = g'(h_i^P)\cdot\left[y_i^p - V_i^P\right].$$
5. Calculation of delta for previous layers by Error Back-Propagation:
   $$\delta_i^{l-1} = g'(h_i^{l-1})\sum_i w_{ji}^l\,\delta_j^l$$
6. Weight change according to formula:
   $$\Delta w_{ij}^l = \eta\,\delta_i^l\,V_j^{l-1},$$
   $$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \Delta w_{ij}.$$
7. If all samples have been submitted to the network we continue in phase no. 8, otherwise we go back to phase no. 2.
8. If the network error compared to the selected criterion value was minor or the maximum number of steps was

exhausted, then the learning process can be completed, else phase no. 2

2.4 Choice of a relevant model

An appropriate model can be determined on the basis of:

a) the graph of the time series, or from its absolute or relative characteristics,

b) interpolative criteria (a decisive deviation of residues, coefficient of determination, coefficient of autocorrelation of residues, tests of parameters),

c) extrapolative criteria (average characteristics of "ex post" forecast mistakes, graph forecasts).

Average characteristics of residues

The average square mistake – dispersion

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{a}_t^2.$$

The root mean square mistake

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\hat{a}_t^2}.$$

The average absolute mistake

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| = \frac{1}{n}\sum_{t=1}^{n}|\hat{a}_t|.$$

The lower the values of the specified characteristics, the better the chosen model is.

3 Demand modelling

The above mentioned methods will be used for demand modelling of a narrowband mobile connection (GPRS, HSCSD). The GPRS Demand Model (i.e. the number of households which are using a narrowband mobile internet connection) is based on the values from the Czech Statistical Office (between 2003 and 2009) and estimate made by an expert.

ARIMA Processes

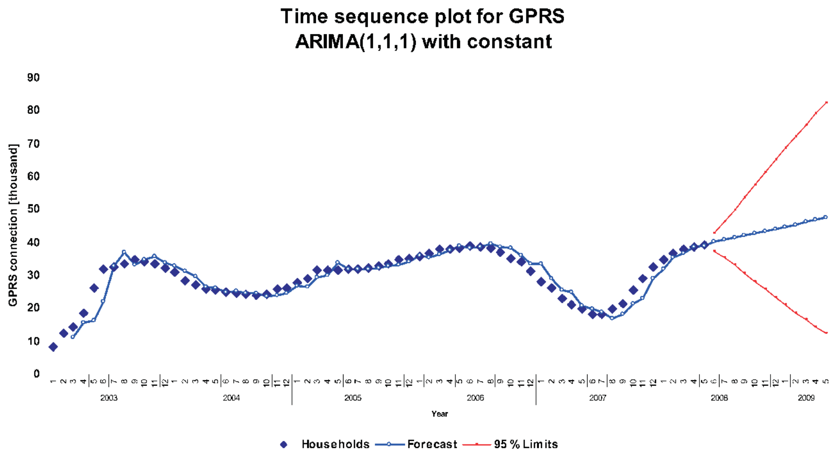

Fig. 2: ARIMA Demand modelling of GPRS

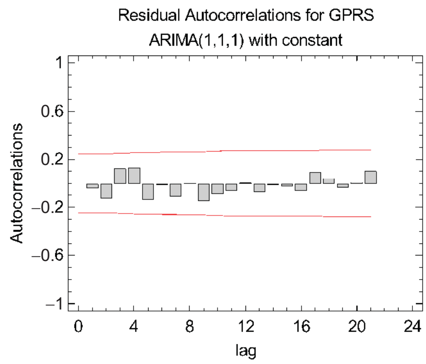

Fig. 3: Residual autocorrelation for GPRS

Brown's linear exp. smoothing with alpha = 0.9428

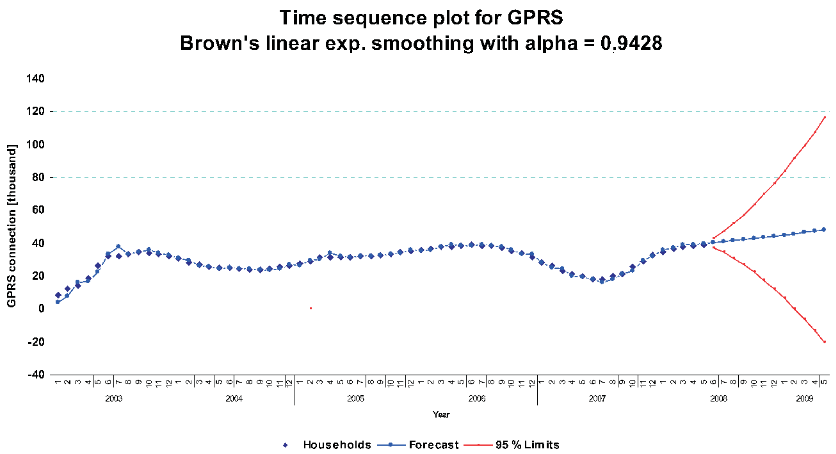

Fig. 4: Brown's linear exp. smoothing of GPRS

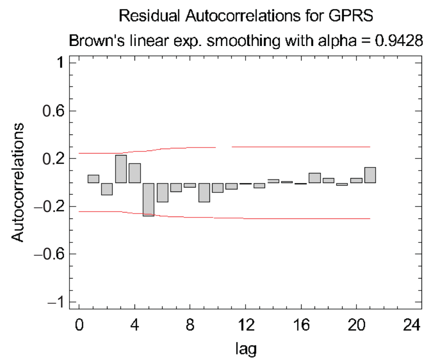

Fig. 5: Residual autocorrelation for GPRS

Back-Propagation Neural Network

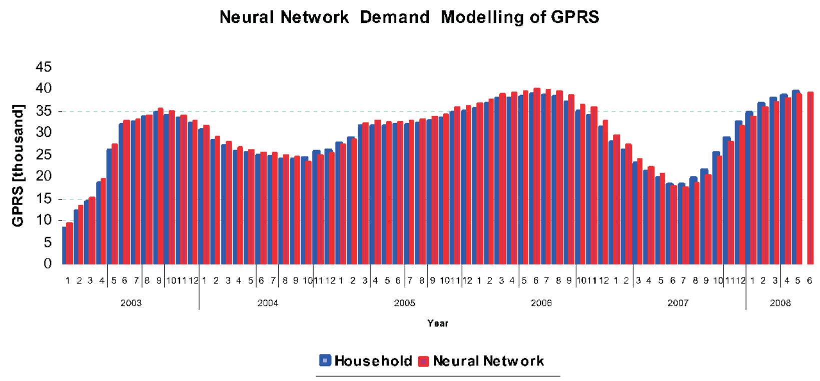

Fig. 6: Neural Network Demand Modelling

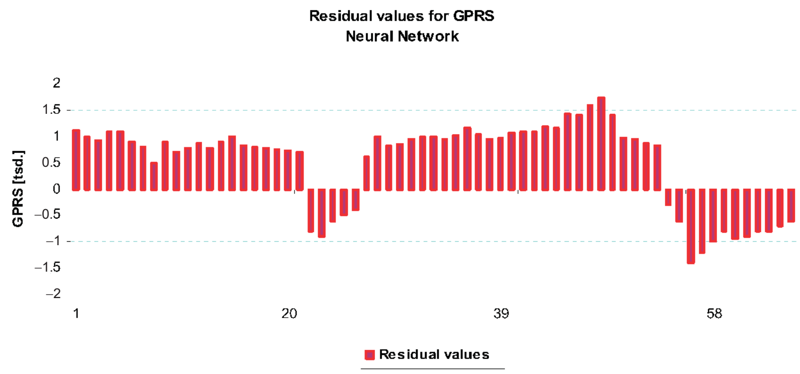

Fig. 7: Residual Values for GPRS

4 Conclusions

In this paper, we have described some methods for demand modelling in telecommunications. Within the research project on demand modelling in telecommunications we obtained the following values for the average characteristics of residues.

On the basis of the results of RMSE and MAE average characteristics of residues, the demand model based upon Neural Networks can be considered as best, but a different result arises from an evaluation of prediction quality, because the best prediction value result was obtained from a demand model based on Brown's linear exponential smoothing.

Table 1: Average characteristics of residues

	Neural Network	ARIMA (1,1,1)	Brown's exp. smoothing
RMSE	0.96	1.31	1.53
MAE	0.92	0.88	1.02

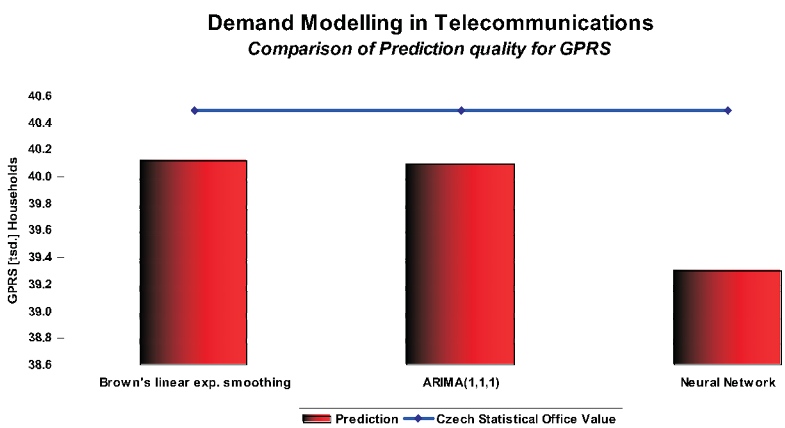

Fig. 8: Comparison of the GPRS internet connection used by households in the 2nd quarter of 2008, values published on February 10th, 2009 by the Czech Statistical Office, and results predicted by demand models for GPRS

The prediction based on the Neural Networks was also used with other demand courses in the field of telecommunications for which more accurate results were obtained in terms of prediction quality.

Future research will focus on improving the prediction results and a demand model based on the above-mentioned methods.

5 References

[1] Cipra, T.: *Analysis of Time Series with Application in Economy*, Prague: SNTL/ALFA, 1986.

[2] Kaňok, M.: *Statistical Methods in Management*, Prague: CTU in Prague, 2002.

[3] Šnorek, M.: *Neural Network and Neural Computers*, Prague: CTU in Prague, 2002.

[4] Artl, J., Artlová, M.: *Financial Time Series*, Prague: Grada Publishing a.s., 2003.

[5] Artl, J., Artlová, M., Rublíková, E.: *Analysis of Economical Time Series with Examples*, Prague: University of Economics, 2002.

[6] Özekes, S., Osman, O.: *Classsification and Prediction in Data Mining with Neural Networks*, Istanbul Commerce University, 2003.

[7] Crone, S.: *Business Forecasting with Neural Networks*, Boston: Institute of Business Forecasting, 2004.

[8] Samuelson , P., Nordhaus, D.: *Economy*, Prague: Nakladatelství Svoboda, 1991.

[9] Nau, R.: Introduction to ARIMA, Duke University Durham, 2007.

[10] Kuba, M.: *Neural Networks*, Brno: Masaryk University Brno, 1995.

Martin Chvalina
e-mail: martin.chvalina@email.cz

Department of Economics, Management and Humanities

Czech Technical University in Prague
Faculty of Electrical Engineering
Technická 2
166 27 Praha, Czech Republic