Ekonomický časopis, *51*, 2003, č. 4, s. 489 – 505

Statistical, Econometric and State Space Models for Time and Savings Deposits of Households

*Dušan MARČEK**

Introduction

This article gives an applied and simple presentation of the time series models of the amount of the time and savings deposits of households useful in forecasting its data.

The amount of the time and savings deposits of households is a very important variable which can be used by commercial banks for long-terms credits, which has a great importance for the restructuring of the enterprise sphere in our conditions.

The article contains three sections. The first section presents the classical additive statistical model in which the various components in time series data are defined: trend – the downward movement of the amount of the time and savings deposits of households data over the time of four years, seasonal variations, a pattern of change in the data that completes itself within a calendar year and then is repeated on a yearly basis, and the error, the erratic movements in the data that have no definable pattern. In the second section we consider the concept of multiple regression causal (econometric) model in which time, seasonal factor and other time series variables (the independent variables) are included to explain the behaviour of the time and savings deposits of households time series. In the third section the „state-space" (or Markovian [9]) representation of the relationship between input variables is used. Especially, we consider the Kalman solution. We derive the Kalman filter as a structural model based on so called structural time series models where the trend and seasonal components are explicitly designed. A simple interpretation of the Kalman one-step recursion procedures (recursive filter, smoothing and prediction) is presented.

To develop the above mentioned modelling techniques, we will consider the changes of time and savings deposits of households time readings of the Slovak Economy from 1994 to 2001. The data was collected for the period fourth quarter, 1994 to first quarter, 2001 (see Table 1) which provides a total of 29 observations. Figure 1 displays quarterly construction time and savings deposits figures for the years 1994 through 2001.

---

* prof. Ing. Dušan MARČEK, CSc., Žilinská univerzita, Fakulta riadenia a informatiky, Katedra makro a mikroekonomiky, Moyzesova 20, 010 26 Žilina; e-mail: marcek@fria.utc.sk

Acknowledgment: This work was supported by Slovak grant foundation under the grant No. 1/9183/02.

1. Additive Decomposition Modelling Methods

Decomposition techniques are among the oldest the modelling and forecasting methods. They are among the easiest to understand and use, especially for short-term forecasting. Unfortunately, as mentioned in [3], the decomposition method is basically intuitive and there are a number of theoretical weaknesses in this approach.

Data that is reported quarterly, monthly, weekly, etc. and that demonstrates a yearly periodic pattern is said to contain a seasonal component.

The displayed changes of time and savings deposits of households on Figure 1 are clearly seasonal. The graphical presentation of the time series exhibits a downward trend suggesting that the growth is linear over the time. There is a definable drop in third quarter (summer) and a rise to a pick during the first quarter (winter). We assume that we have an additive model as follows

$$y_t = T_t + S_t + \varepsilon_t \qquad (1)$$

where T_t is trend, S_t is seasonal variation and ε_t is random error component. The Eq. (1) expresses the additive model of the time series y_t (in our case the changes of time and savings deposits of households). We will assume that the seasonal component S_t for any one season is the some each year.

Figure 1

The Data for Changes of Time and Savings Deposits of Households – DMTH

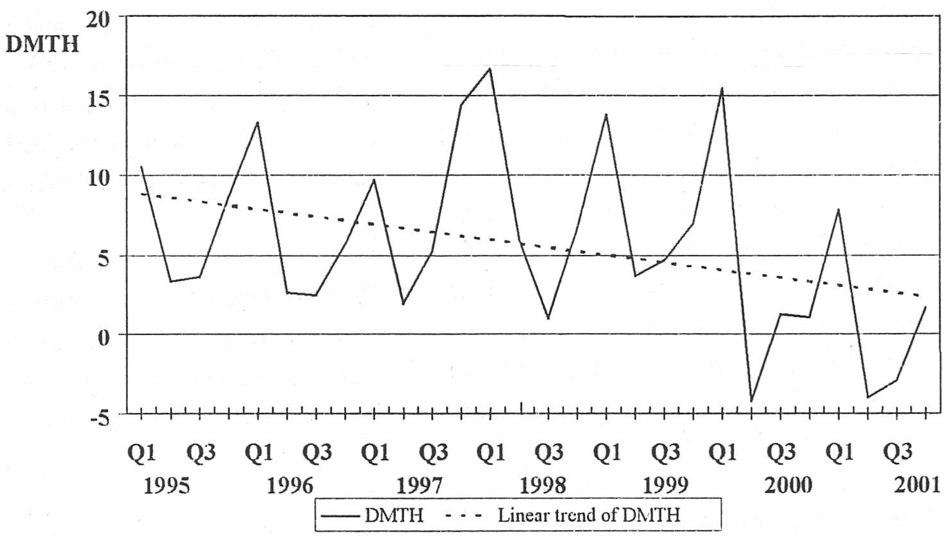


Table 1

Actual Data for the Time Series Modelling*

Year	Quarter	CP_t billion of SK	YD_t billion of SK	YDC_t billion of SK	$IRTH_t$ per cent	MTH_t billion of SK	$DMTH_t$ billion of SK
1994	Q4				14,380	90,0667	
1995	Q1	64,800	73,242	8,4423	13,990	100,6333	10,5667
	Q2	69,500	75,698	6,1977	12,650	104,0000	3,3667
	Q3	70,100	76,716	6,6160	12,540	107,6000	3,6000
	Q4	72,900	88,960	16,0600	11,080	116,3000	8,7000
1996	Q1	74,300	81,537	7,2374	10,030	129,6333	13,3333
	Q2	80,000	87,283	7,2830	8,740	132,2667	2,6333
	Q3	80,000	85,888	5,8883	8,647	134,7333	2,4667
	Q4	81,500	100,891	19,3914	8,410	140,4000	5,6667
1997	Q1	83,700	93,389	9,6895	8,740	150,1333	9,7333
	Q2	91,000	99,840	8,8397	9,423	152,1000	1,9667
	Q3	88,900	98,404	9,5038	9,800	157,3333	5,2333
	Q4	90,100	115,467	25,3671	10,420	171,8000	14,4667
1998	Q1	92,200	103,947	11,7469	11,700	188,5000	16,7000
	Q2	101,400	111,042	9,6423	12,010	194,5000	6,0000
	Q3	101,300	109,548	8,2480	11,600	195,5000	1,0000
	Q4	100,500	128,463	27,9627	12,940	202,2000	6,7000
1999	Q1	100,400	114,200	13,8000	12,570	216,0000	13,8000
	Q2	111,300	121,800	10,5000	12,680	219,7000	3,7000
	Q3	111,700	120,400	8,7000	12,270	224,3667	4,6667
	Q4	112,900	141,000	28,1000	11,920	231,3667	7,0000
2000	Q1	109,085	124,299	15,2140	10,770	246,7667	15,4000
	Q2	117,200	130,840	13,6400	8,960	242,5333	–4,2333
	Q3	118,700	129,635	10,9350	7,560	243,8667	1,3333
	Q4	124,000	152,621	28,6210	6,500	244,9667	1,1000
2001	Q1	119,900	133,952	14,0520	6,170	252,8333	7,8667
	Q2	126,600	138,436	11,8360	5,940	248,8333	–4,0000
	Q3	131,100	139,500	8,4000	5,810	245,8667	–2,9667
	Q4	137,000	165,058	28,0580	5,700	247,5667	1,7000

* CP_t – Final consumption of households in current prices
YD_t – Disposable incomes of households
YDC_t – „Unconsumed" incomes of households
$IRTH_t$ – Interest rates of time and savings deposits of households
MTH_t – Volume of time and savings deposits of households
$DMTH_t$ – Change of the volume of time and savings deposits of households

Source: Statistical abstract of economic quantities of the Slovak republic fot the period 1991 to 2001. Statistical Office of the SR, august 2002.

Let j denotes the number os seasons in a year, i denotes the year, m denotes the number of years ($m = 7$), n denotes the number of seasons ($n = 4$). It is clear that the seasonal component is in an additive model independent of the trend, i. e., the model (1) satisfy the following conditions

$$S_t = S_{ij} = a_j \text{ for } i = 1, 2, ..., m, \ j = 1, 2, 3, n \tag{2}$$

thus the magnitude of the seasonal swing is constant. Further we assume that seasonal swings should up to zero, i. e.

$$\sum_{j=1}^{n} S_{ij} = \sum_{j=1}^{n} a_j = 0 \text{ for } i = 1, 2, ..., m \tag{3}$$

The appropriate model for the trend is a linear model. The equation to model trend is

$$T_{i.} = T_{ij} = \bar{y}_{i.} = b_0 + b_1 i + \varepsilon_t \tag{4}$$

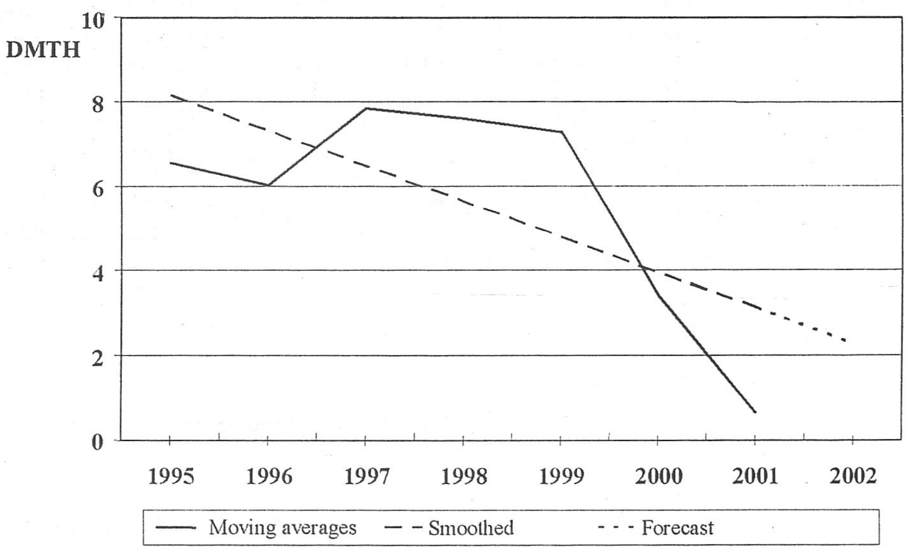
where $\bar{y}_{i.}$, $i = 1, 2, ..., m$ are moving averages calculated as $\bar{y}_{i.} = \sum_{j=1}^{n} y_{ij} / n$, $i = 1, 2, ..., m$ where the intercept b_0, and slope b_1 are unknown parameters, ε_t is random error component. The fitted simple linear regression is

$$\hat{T}_{i.} = \hat{\bar{y}}_{i.} = 8{,}9869 - 0{,}8405\, i,\ i = 1, 2, ..., m \tag{5}$$

The moving averages $\bar{y}_{i.}$ and their fitted values $\hat{\bar{y}}_{i.}$ of the model (5) are given in Table 2 and plotted in Figure 2.

Figure 2

Fitted Trend Using Deseasonalized Data



Corresponding to (2), (4) and substituting T_{ij} with $\hat{\bar{y}}_{i.}$, we can write (1) in form as

$$y_t = y_{ij} = \hat{\bar{y}}_{i.} + a_j + \varepsilon_{ij} \tag{6}$$

The least-squares of the estimates of the magnitudes of seasonal savings $\hat{a}_j$ are

$$\hat{a}_j = \hat{S}_{ij} = \bar{y}_{.j} - \bar{y} \tag{7}$$

and the fitted values $\hat{y}_t$ corresponding to (5), (7) and (1) are

$$\hat{y}_t = \hat{\bar{y}}_{i.} + \hat{a}_j \tag{8}$$

Computational results for 1995 to 2001 using the additive decomposition model (8) are shown in Table 2 and plotted in Figure 3.

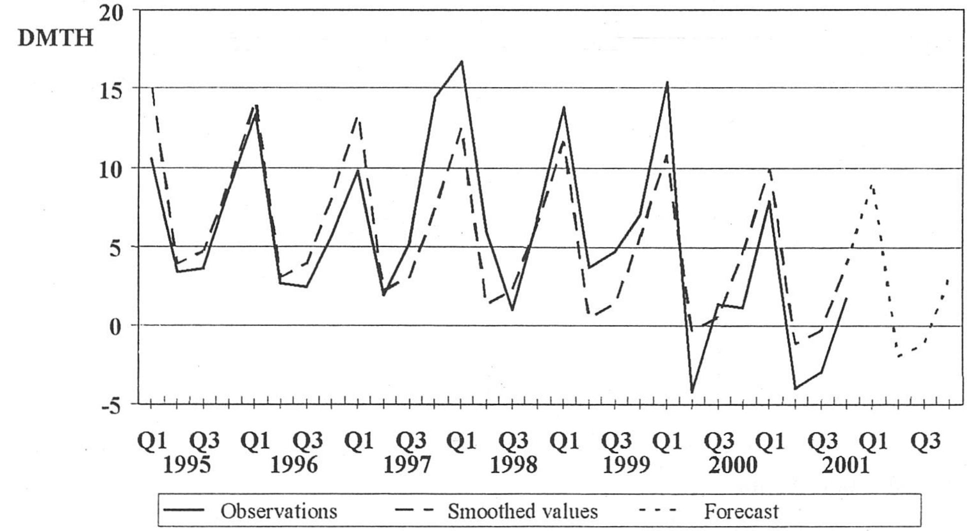
Table 2

Obtaining the Estimate for Seasonality and Trend in an Additive Decomposition Model of Construction DMTH

Year	i	y_{ij} for j				Total	$\bar{y}_{i.}$	$\hat{T}_{i.} \equiv \hat{\bar{y}}_{i.}$
		1	2	3	4			
1995	1	10,5667	3,3667	3,6	8,7	26,2334	6,55835	8,14643
1996	2	13,3333	2,6333	2,4667	5,6667	24,1	6,025	7,30595
1997	3	9,7333	1,9667	5,2333	14,4667	31,4	7,85	6,46547
1998	4	16,7	6	1	6,7	30,4	7,6	5,62499
1999	5	13,8	3,7	4,6667	7	29,1667	7,29168	4,78451
2000	6	15,4	–4,2333	1,3333	1,1	13,6	3,4	3,94403
2001	7	7,8667	–4	–2,9667	1,7	2,6	0,65	3,10355
2002	8							2,26307
Total		87,4	9,4334	15,3333	45,3334	157,5001	39,37503	39,37493
$\bar{y}_{.j}$		12,48571	1,34763	2,19047	6,4762	22,50001		
$\hat{S}_{ij} \equiv a_j$		6,86071	–4,27737	–3,43453	0,8512	0		

Figure 3

Actual, Fitted and Forecast Values: Additive Decomposition for DMTH (1995 – 2001)



2. The Econometric Model

In this section, we consider the multiple regression causal model in which causal, time and seasonal factors are included to explain the behaviour of the variable to be explain (the dependent variable change of the volume of time and savings deposits of households).

The theoretical bases for including specific independent variables in the model are economic principles. In our case a hypothesis is formulated that the unconsumed incomes of households (YDC) and the lagged interest rates of time and savings deposits of households ($IRTH_{-1}$) are important sources of time and savings deposits of households changes (DMTH) calculated as the difference between an MTH (time and savings deposits of households) and the previous MTH_{-1}. Because the data in Table 1 is reported quarterly and demonstrates a yearly periodic pattern, it contains a seasonal component. The DMTH model includes three important explanatory variables: YDC, $IRTH_{-1}$ and the seasonal component D_t. Table 1 presents data on MTH, DMTH, IRTH, $IRTH_{-1}$, YDC. The seasonal component is not directly quantifiable. We use a dummy variable to account for this factor.

An alternative specification to the decomposition model presented in previous section is an econometric model given by [8]

$$YDC_t = YD_t - CP_t, \tag{9}$$

$$DMTH_t = b_0 + b_1 YDC_t + b_2 IRTH_{t-1} + b_3 D_t + \varepsilon_t \tag{10}$$

$$MTH_t = MTH_{t-1} + DMTH_t \tag{11}$$

where D_t is seasonal variable (1Q = 1), ε_t is the disturbance term.

The econometric model contains three equations. Eq. (10) is stochastic, Eq. (9), (11) are identities.

The estimated regression Eq. (10) for the DMTH is given by

$$\hat{DMTH}_t = -8{,}729 + \underset{(0{,}09)}{0{,}241}\, YDC_t + \underset{(0{,}266)}{0{,}852}\, IRTH_t + \underset{(1{,}479)}{9{,}246}\, D_t \tag{12}$$

$$R^2 = 0{,}694 \qquad D\text{-}W = 1{,}089$$

Estimated period: 1995 – 2001

In this estimation the D-W statistic indicates the presence of positive first-order autocorrelation.

The regression Eq. (12) was reestimated using Cochrane-Orcutt. The new estimated equation is given

$$\hat{DMTH}_t = -6{,}249 + 0{,}262\ YDC_t - 0{,}110\ YDC_{t-1} + 1{,}042\ IRTH_{t-1} - 0{,}439\ IRTH_{t-2} +$$
$$\qquad (0{,}072) \qquad\qquad (0{,}388)$$
$$+ 9{,}961\ D_t - 4{,}200\ D_{t-1} + 0{,}422\ MTH_{t-1} \qquad (13)$$
$$(1{,}149)$$

$$R^2 = 0{,}769 \qquad\qquad D\text{-}W = 2{,}207$$

All explanatory variables are statistically significant and the Durbin-Watson statistic indicates no autocorrelation. So, it is possible to use this model for forecasting. Actual, fitted and forecast values for DMTH are presented in Table 3 and plotted in Figure 4.

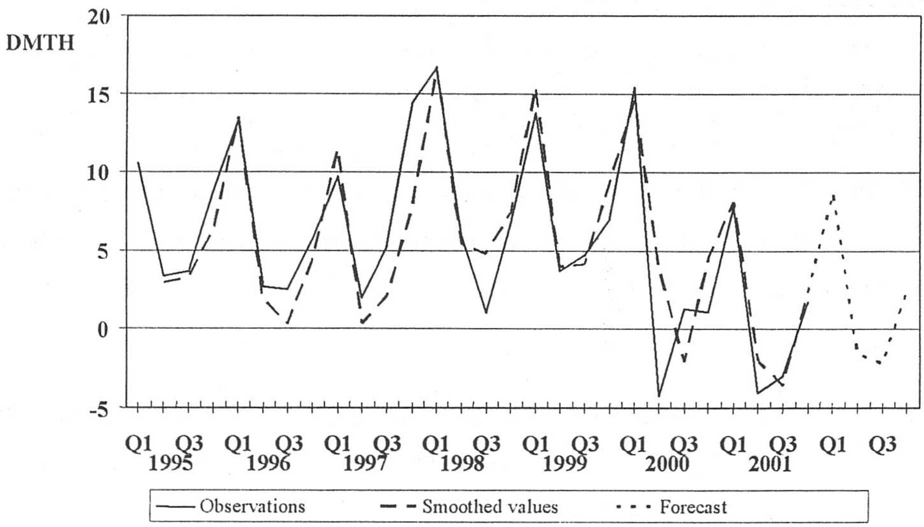
Table 3

Actual, Fitted and Forecast Values for Time and Savings Deposits of Households Changes (Cochrane-Orcutt Method)

Year	Season							
	1		2		3		4	
	$DMTH_t$	$DMTH\hat{}_t$	$DMTH_t$	$DMTH\hat{}_t$	$DMTH_t$	$DMTH\hat{}_t$	$DMTH_t$	$DMTH\hat{}_t$
1995	10,5667	–	3,3667	2,9534	3,6	3,2501	8,7	6,2465
1996	13,3333	13,5363	2,6333	1,8612	2,4667	0,2972	5,6667	4,3823
1997	9,7333	11,4599	1,9667	0,3089	5,2333	2,0682	14,4667	7,6142
1998	16,7	16,6371	6	5,4302	1	4,7467	6,7	7,3852
1999	13,8	15,4472	3,7	4,0045	4,6667	4,1162	7	9,3211
2000	15,4	14,5715	–4,2333	3,9175	1,3333	–2,0753	1,1	4,5328
2001	7,8667	8,1451	–4	–2,0138	–2,9667	–3,5663	1,7	2,3555
2002	–	8,6995	–	–1,5613	–	–2,1777	–	2,4865

Figure 4

Time Plot at the Actual, Fitted and Forecast Values for Time and Savings Deposits of Households Changes (Cochrane-Orcutt Method)



3. State-space Form of the Model and the Kalman Filter

In discussing linear models it is often more convenient to use the so-called „state-space" or „Markovian" representation of the relationship between input and output rather than the explicit form. The state-space form gives a very compact description which is valid provided the relationship between the input and output can be expressed in terms of a finite order linear difference equation [9]. The basic idea rests on the well known result that any finite order linear differential or difference equation can be expressed as a vector first order equation. In this section we will illustrate some of many time series models which can be represented in state-space form. By this we means that the series $\{Y_t, t = 1, 2, \ldots\}$ satisfy equations of the form

$$\mathbf{Y}_t = \mathbf{G}\ \mathbf{x}_t + \mathbf{W}_t \quad t = 1, 2, \ldots \tag{14}$$

where

$$\mathbf{x}_t = \mathbf{F}\ \mathbf{x}_{t-1} + \mathbf{V}_t \quad t = 1, 2, \ldots \tag{15}$$

The Eq. (15) can be interpreted as describing the evolution of the state $\mathbf{x}_t$ of a system at time t. Eq. (14) defines a sequence of observations $\mathbf{Y}_t$, which are obtained by applying a linear transformation to $\mathbf{x}_t$ and adding a random noise vector $\mathbf{W}_t$. $\mathbf{V}_t$ is sequence of random vectors. **G** is called the „observation matrix", **F** the „system matrix". It is assumed that the matrices **G** and **F** are known, jointly stationary and Gaussian zero mean white noise processes with $E(V_t W_s^{\mathrm{T}}) = 0$, all s, t.

To illustrate the versatility of state-space models, we now attend to that example trend and seasonal series with noise considered earlier in Section 1. This type of process can be represented by structural time series model, which is designed to model trend and seasonal component. For example the classical decomposition (1) expressed the time series $\{y_t\}$ as a sum of trend x_t, seasonal s_t and noise components. The seasonal component (with period $n = 4$) was sequence $\{s_t\}$ with the properties $s_{t+n} = s_t$ (seasonal component is independent of the trend) and $\sum_{j=1}^{n} s_{ij} = 0$ for $i = 1, 2, \ldots, m$. Such a process may be expressed generally in basic structural model defined by

$$Y_t = x_t + s_t + \varepsilon_t \tag{16}$$

where x_t represents the trend and s_t is the cyclic or seasonal component, while ε_t is noise part. Both x_t and s_t evolve according to their own equations. In our case, we have a simple local level model

$$Y_t = x_t + s_t + \varepsilon_t \tag{17}$$

$$x_t = x_{t-1} + \upsilon_t \tag{18}$$

$$s_t = -\sum_{j=1}^{3} s_{t-j} \tag{19}$$

where υ_t, ε_t are white noise independent variables with zero mean.

In Eq. (17), (18), (19) the true x_t and seasonality s_t are not observable. Often the task is to infer behaviour of the trend or other quantities from the time series $\{y_t\}$ we observe. This class of problems is most conveniently treated by casting the models into state-space representation and using the Kalman filtering technique. A complete discussion of the state space modelling and the details of the Kalman filter can be found in [1], [2], [9]. See also [5], [6], [10], [11].

The simple local level model (17), (18), (19) can be written in the state-space form with

$$\mathbf{Y}_t = (Y_t),\quad \mathbf{x}_t^{\mathrm{T}} = (x_t, s_t, s_{t-1}, s_{t-2}),\ \mathbf{W}_t = \varepsilon_t,\qquad \mathbf{V}_t^{\mathrm{T}} = (\upsilon_t, 0, 0, 0)$$

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} q_1 & & & \\ & 0 & & 0 \\ & 0 & & 0 \\ & & & 0 \end{pmatrix}, \mathbf{G} = (1, 1, 0, 0), \mathbf{R} = \sigma_\varepsilon^2 = r_1 =$$

$$E(\mathbf{W}_t\mathbf{W}_t^{\mathrm{T}}) \quad (20)$$

where $\mathbf{Q}$ is covariance matrix of υ_t, q_1 is variance of υ_t, $\mathbf{R}$ is covariance matrix of ε_t.

The Kalman recursions is used to obtain an estimate of the unobservable state vector $\mathbf{x}_t$ based on the information at time t and observations of $\mathbf{Y}$ up to $\mathbf{Y}_t$. A simple interpretation of the Kalman one-step recursion may be obtained in the following way [9]. Suppose that we have observed the $\mathbf{Y}$ process up to time $t-1$ only, and on the basis of these observations we have computed the optimal estimate $\hat{\mathbf{x}}_{t|t-1}$ of the state vector at time $t-1$. Using only these observations, the best estimate of $\mathbf{x}_t$ is (from Eq. (15))

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}\,\hat{\mathbf{x}}_{t-1|t-1} \tag{21}$$

and the best predictor of the next observation $\mathbf{Y}_t$ is (from Eq. (14))

$$\hat{\mathbf{Y}}_{t|t-1} = \mathbf{G}\,\hat{\mathbf{x}}_{t|t-1} \tag{22}$$

When the value of $\mathbf{Y}_t$ becomes available we can compute this value predicted from the estimate $\hat{\mathbf{x}}_{t|t-1}$ and „update“ our estimate of $\mathbf{x}_t$ by taking a linear combination of the previous $\hat{\mathbf{x}}_{t|t-1}$ and the prediction error $\{ \mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1} \} = \{ \mathbf{Y}_t - \mathbf{G}\, \hat{\mathbf{x}}_{t|t-1} \}$. Then the following equations constitute the Kalman filter

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{Y}_t - \mathbf{G}\, \hat{\mathbf{x}}_{t|t-1}) \tag{23}$$

$$\hat{\mathbf{x}}_{t+1|t} = \mathbf{F}\, \hat{\mathbf{x}}_{t|t} \tag{24}$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{G} \mathbf{P}_{t|t-1} \tag{25}$$

where

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\, \mathbf{G}^{\mathrm{T}} (\mathbf{G}\ \mathbf{P}_{t|t-1}\, \mathbf{G}^{\mathrm{T}})^{-1} \tag{26}$$

$$\mathbf{P}_{t|t-1} = E(\mathbf{e}_t \mathbf{e}_t^{\mathrm{T}}) \tag{27}$$

$$\mathbf{e}_t = \mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1} = \mathbf{G}(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) \tag{28}$$

The matrix $\mathbf{K}_t$ is called Kalman gain. One sees an obvious resemblance to first-order exponential smoothing in Eq. (23). This result was noted by Priestley [9].

The initial values of $\mathbf{x}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$ have to be given in order to start the Kalman filter. If the state vector is not stationary, we can use the first k values of $\mathbf{Y}$ (where k denotes dimension of $\mathbf{x}$) to calculate $\mathbf{x}_{k+1|k}$. To get the maximum likelihood estimate of the variance parameters, we need to minimize the log likelihood function

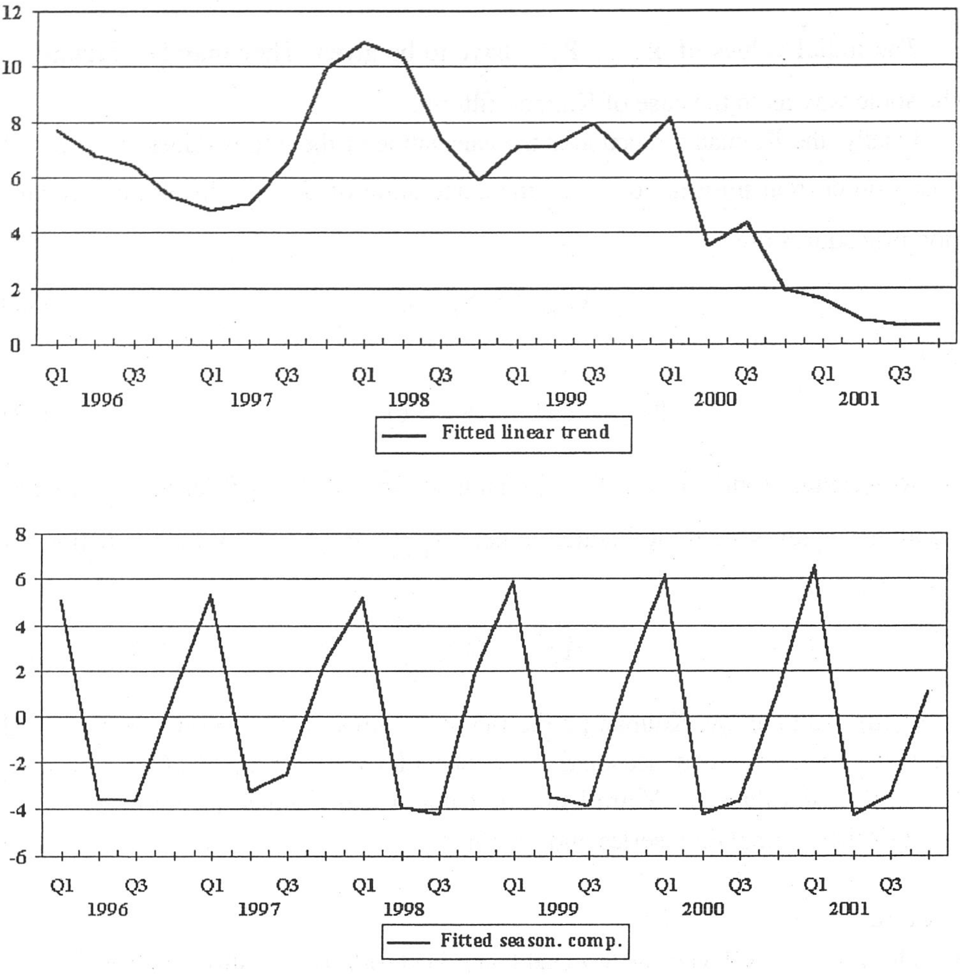
$$\ln L(q_1, r_1, \mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_N) = -\frac{1}{2}\sum_{t=1}^{N} \ln\left|\mathbf{P}_{t|t-1}\right| - \\ -\frac{1}{2}(\mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1})^{\mathrm{T}}\, \mathbf{P}_{t|t-1}^{-1} (\mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1}) + konst \tag{27}$$

In our case $k = 4$, according to the Eqs. (23), (24), the first four estimates of the state vector $\hat{\mathbf{x}}_{5|4}$ are (6,5585, 4,0085, 2,1415, –2,9585), and the initial form of the variance matrix is expressed as the function of the parameters q_1 and r_1. Then the Kalman filtering procedures (23) to (26) give the estimates of x_t (trend) and s_t (seasonal component).

These values are plotted as a function of time in Figure 5.

Figure 5

Kalman Filter: Graph of Trend and Seasonal Component Values



The Kalman filter gives the estimate of the state variable **x** at time t given the information up to t, $\hat{\mathbf{x}}_{t|t}$.

As more information is accumulated, the estimate of $\mathbf{x}_t$ can be improved using this information, which is known as Kalman smoothing. The Kalman smoothing is a way of getting the estimate of $\mathbf{x}_t$, given the information at time T, where $T > t$ [10]

$$\hat{\mathbf{x}}_{t|T} = \hat{\mathbf{x}}_{t|t} + \mathbf{J}_t(\hat{\mathbf{x}}_{t+1|T} - \hat{\mathbf{x}}_{t+1|t}) \tag{30}$$

and

$$\mathbf{P}_{t|T} = \mathbf{P}_{t|t} + \mathbf{J}_t(\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t})\mathbf{J}_t^{\mathrm{T}} \tag{31}$$

where

$$\mathbf{J}_t = \mathbf{P}_{t|t}\mathbf{F}^{\mathrm{T}}\mathbf{P}_{t+1|t}^{-1}$$

The initial values of $\mathbf{x}_{t|t-1}$, $\mathbf{P}_{t|t-1}$ have to be given. They may be obtained by the some way as in the case of Kalman filtering.

Finally, the Kalman prediction is the calculation of the future values $\mathbf{x}_{t+h}$, $h > 0$, based on current information, i. e., the calculation of $\hat{\mathbf{x}}_{t+h|t}$. The Kalman prediction procedures are

$$\hat{\mathbf{x}}_{t+h|t} = \mathbf{F}\,\hat{\mathbf{x}}_{t+h-1|t} \tag{32}$$

and

$$\mathbf{P}_{t+h|t} = \mathbf{F}\,\mathbf{P}_{t+h-1|t}\mathbf{F}^{\mathrm{T}} + \mathbf{Q}_{t+h} \tag{33}$$

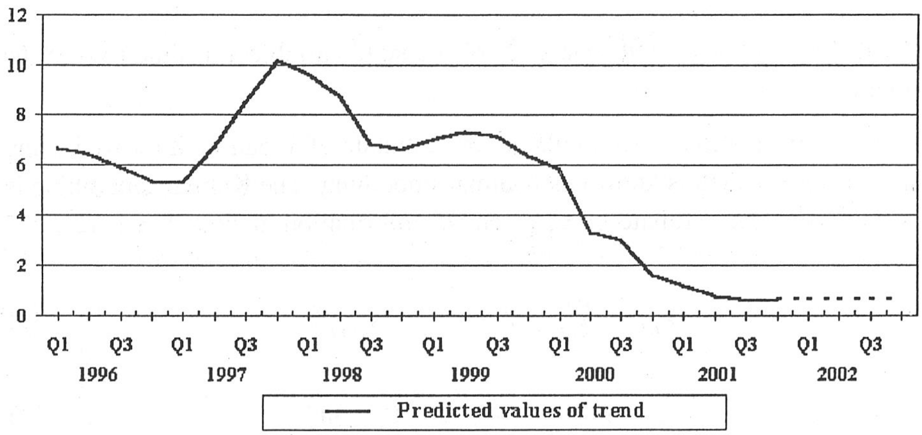
So starting from $\{\hat{\mathbf{x}}_{t+1|t},\ \mathbf{P}_{t+1|t}\}$ obtained from Kalman filtering procedures, the above equations can be iterated to get $\{\hat{\mathbf{x}}_{t+h|t},\ \mathbf{P}_{t+h|t}\}$. Then the prediction for **Y** is

$$\hat{\mathbf{Y}}_{t+h|t} = \mathbf{G}\,\hat{\mathbf{x}}_{t+h|t} \tag{34}$$

In our example, the Kalman prediction procedures (32), (33) give the predicted state-space values (trend, seasonal component) for the next 4 point (see Figure 6). The predicted values for **Y** are listed in Table 4 and graphed in Figure 7, which displays the original data series and the forecast.

Figure 6

The Plot of Predicted Trend and Seasonal Component with the 4 Predicted Values



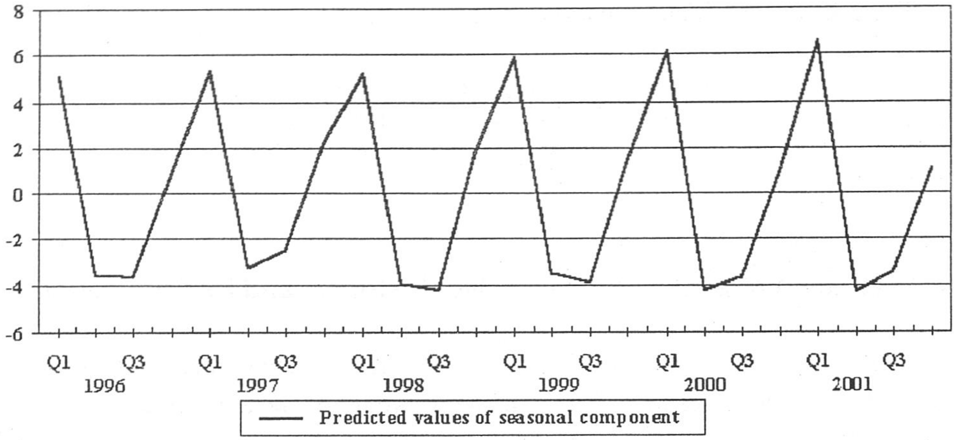


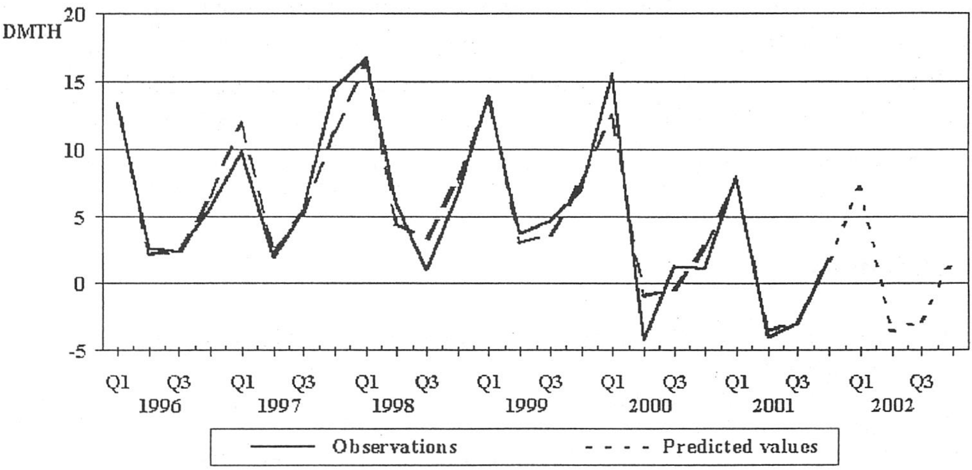
Table 4

Kalman Prediction: Original Data Series and the Kalman Prediction

Year	Season							
	1		2		3		4	
	Y_t	$\hat{Y}_t$	Y_t	$\hat{Y}_t$	Y_t	$\hat{Y}_t$	Y_t	$\hat{Y}_t$
1996	13,3333	13,2423	2,6333	2,1711	2,4667	2,4825	5,6667	6,399
1997	9,7333	11,9007	1,9667	2,4767	5,2333	5,0874	14,4667	11,1953
1998	16,7	16,2001	6	4,4882	1	3,4537	6,7	7,6616
1999	13,8	13,6088	3,7	3,0465	4,6667	3,738	7	7,4006
2000	15,4	12,4258	–4,2333	–0,9272	1,3333	–0,3679	1,1	2,6443
2001	7,8667	7,7869	–4	–3,5054	–2,9667	–2,773	1,7	1,6846
2002	–	**7,2254**	–	**–3,6149**	–	**–2,7626**	–	**1,6846**

Figure 7

The Plot of Original Data Series and the Kalman Prediction



The coefficient of determination was expressed as

$$R^2 = 1 - \frac{\sum_{t=1}^{24}(Y_t - \hat{Y}_t)^2}{\sum_{t=1}^{24}(Y_t - \bar{Y})^2} = 0{,}935948$$

4. Evaluation of Model Forecasts Over Historical and Ex Post Period

After having the option of constructing three different methodological approaches for time series modelling, which one would be the appropriate choice for the „best" forecasting? As the econometric model produces results consistent with the assumption of the classical least squares model, the answer is „the one that produces the best forecast".

Table 5 presents the forecast statistics for the historical time periods and this one for ex post periods.

Table 5

Forecast Summary Statistics: the Sample Period 1995 to 2000 for Analysis ($Y_1, \ldots, Y_{20}$) and the Ex Post Forecast ($\hat{Y}_{21}, \ldots, \hat{Y}_{24}$) Period (Year 2001)

Model	Sample period	Ex post forecast period
	RMSE	RMSE
Statistical	2,869	4,594
Econometric	2,865	2,099
State-space representation and Kalman recursion	1,636	1,366

When comparing each model by use the criteria of the RMSE, the state-space representation of the model and the Kalman filtering gives the best results. The RMSE values for predictors based on additive decomposition modelling and econometric models discussed in Section 1 and 2 are dramatically worse than the RMSE values obtained using the state space model and the Kalman recursion.

References

[1] BROCKWELL, P. J. – DAVIS, R. A.: Time Series: Theory and Methods. New York: Springer-Verlag 1987.
[2] CHUI, C. K. – CHEN, G.: Kalman Filtering with Real-time Application. Berlin: Springer-Verlag 1991.
[3] GAYNOR, P. E. – KIRPATRICK, R. C.: Introduction to Time-Series Modeling and Forecasting in Business and Economics. New York: McGraw-Hill, Inc. 1994.

[4] MARČEK, D.: Modelling Short-term Time Series with Seasonal Component by Structural Models in State Form. Slovak Statistics and Demography (Infostat Bratislava), 2001, No. 3, pp. 4 – 17 (in Slovak).

[5] MARČEK, D. – MARČEK, M.: Time Series Analysis, Modelling and Forecasting with Economic Applications. Žilina: EDIS ŽU 2001 (in Slovak).

[6] MATLAB 6: user's guide. PL, 2001.

[7] MONTGOMERY, D. C. – JOHNSTON, L. A. – GARDINER, J. S.: Forecasting and Time Series Analysis. New York: McGraw-Hill, Inc. 1990.

[8] PÁLENÍK, V. et all: Construction and Verification of Macroeconomic Model ISWE97q3. Ekonomický časopis/Journal of Ecomonics, 1998, *46*, No. 3, pp. 428 – 466.

[9] PRIESTLEY, M. B.: Spectral Analysis and Time Series. Vols. 1 and 2. New York: Academic Press 1981.

[10] Time Series Pack, Reference an User's Guide. Applications library. *Mathematica*, Wolfram Research, Inc., March 1995.

[11] VAŠÍČEK, O. – FUKAČ, M.: Macroeconomic Model of Non-Accelerating Inflation Product. Finance a úvěr, 2002, *52*, No. 6, pp. 258 – 274.

ŠTATISTICKÝ, EKONOMETRICKÝ A STAVOVÝ MODEL TERMÍNOVANÝCH VKLADOV DOMÁCNOSTÍ

Dušan MARČEK

Cieľom príspevku je prezentovať niektoré výsledky v modelovaní vývoja časového radu štvrťročných dát termínovaných vkladov domácností SR za obdobie prvého štvrťroka 1994 až štvrtého štvrťroka 2002, čo predstavuje spolu 28 pozorovaní. Ich vývoj budeme opisovať pomocou klasického štatistického modelu, ekonometrického modelu a porovnáme ich predikčnú presnosť s modelmi založenými na metodológii stavovo priestorovej reprezentácie *(state-space models)* s Kalmanovymi rekurziami.

V prvej časti je prezentovaný vývoj modelovania termínovaných vkladov domácností klasickým štatistickým modelom, druhá časť je venovaná vývoju ekonometrického modelu, tretia časť prezentuje model založený na stavovo priestorovej reprezentácii a Kalmanovej filtrácii, a napokon vo štvrtej časti sa porovnáva predikčná schopnosť jednotlivých modelov.

Jednou z najstarších modelovacích techník časových radov a ich predikcie je *dekompozičná technika*. Je vhodná na modelovanie krátkodobých časových radov a krátkodobé prognózovanie, t. j. časových radov s horizontom prognózovania kratším ako jedno ročné obdobie.

Napriek tomu, že ide o jednu z najstarších, a súčasne aj jednu z najjednoduchších techník, jej predikčné výsledky sú mnohokrát porovnateľné s inými zložitejšími technikami, napríklad s modelmi v stavovej reprezentácii, čo uvádzame napríklad v [4] pri modelovaní štvrťročných dát úspor domácností v ČSR.

Dekompozičné techniky modelovania časových radov v zásade vedú k dvom typom modelov: *aditívny* a *multiplikatívny typ.* V klasickom štatistickom aditívnom modeli sa predpokladá, že dáta sú jednoduchým súčtom jednotlivých zložiek časového radu, teda trendovej zložky, sezónnej zložky a náhodnej zložky.

Pri modelovaní *trendovej zložky* v súlade s prijatým aditívnym rozkladom časového radu na typ modelu a za predpokladu konštantnej veľkosti sezónnej zložky v každej sezóne vo všetkých rokoch sa ďalej predpokladá, že trendová zložka, t. j. medziročné zmeny úspor domácností závisia len od ročnej časovej premennej.

Pri modelovaní *sezónnej zložky* v aditívnom modeli sa vychádza z predpokladu o rovnakých účinkoch sezónnych vplyvov v j-tej sezóne v i-tom roku. Ďalej sa predpokladá, že sezónne výkyvy sa v rámci sezón v každom roku vykompenzujú.

Presnosť aproximácie časového radu aditívnym štatistickým modelom je vidieť z grafického vývoja pozorovaní *DMTH* a modelom vypočítaných hodnôt na obrázku 3. Miera aproximácie skutočných hodnôt a hodnôt vypočítaných aditívnym modelom je vyjadrená indexom determinácie $R^2 = 0{,}7199$.

V ekonometrickom modeli časové, sezónne alebo iné premenné časových radov vysvetľujú správanie časového radu *DMTH.* Ekonometrické modely sú spravidla založené na teoretickom (ekonomickom) základe vzťahov medzi premennými. Počiatočný návrh (konštrukcia) modelu sa robí na základe relevantnej ekonomickej teórie. V našom prípade ekonomická teória vecne zdôvodňuje kauzálny vzťah medzi veličinou $DMTH_t$ a veličinami úrokovej miery termínovaných vkladov z predchádzajúceho obdobia označenými $IRTH_{t\text{-}1}$ a nespotrebovaným príjmom domácností označený YDC_t. Súčasne platí identita, že YDC_t je daná rozdielom disponibilného príjmu domácností označeného symbolom YD_t a konečnej spotreby označenej symbolom CP_t. Uvedené závislosti sú formálne vyjadrené modelom s rovnicami (9) až (11). Model obsahuje tri rovnice, z ktorých druhá rovnica je *rovnicou správania*, rovnica prvá a tretia sú *identity.*

Model je špecifikovaný tromi endogénnymi premennými (*MTH, DMTH, YDC*) a tromi predeterminovanými premennými (*D, IRTH,* $MTH_{t\text{-}1}$). Premenná *D* je tzv. *dummy* premenná, ktorou sa zohľadňuje sezónnosť časového radu *DMTH.* Jej hodnoty sú v prvom štvrťroku každého roka rovné jednej, v ostatných pozorovaniach sú rovné nule. Model je rekurzívny preidentifikovaný. Parametre v rovnici správania sme odhadli metódou najmenších štvorcov. Model bol kvantifikovaný v tvare (bez uvedenia identít) Cochraneovou-Orcuttovou metódou v tvare rovnice (13)

Teoretické hodnoty *DMTH* ekonometrického modelu sú uvedené v tabuľke 3 a graf vývoja skutočných a teoretických hodnôt v čase je na obrázku 4.

Na procesy opísané uvedenými dvoma metodológiami je vhodné aplikovať namiesto explicitného určovania funkcionálnej formy ich vzťahov tzv. stavovo priestorovú (alebo Markovovu) reprezentáciu vzťahov medzi vstupmi a výstupmi. Východisková idea takejto reprezentácie sa zakladá na známom poznatku, že každá lineárna diferenčná rovnica konečného stupňa môže byť vyjadrená ako vektorová rovnica prvého stupňa. Napríklad ak budeme uvažovať uvedený štatistický aditívny model s trendovou, sezónnou a náhodnou zložkou a budeme v ňom predpokladať, že výstupná pozorovaná veličina y_t je daná ako súčet trendovej (lokálne úrovňovej) stavovej premennej x_t, sezónnej (konštantnej) stavovej zložky s_t, a náhodnej zložky ε_t, potom model možno prepísať na tvar s lokálnym konštantným trendom so sezónnou zložkou v tvare rovníc (17) až (19), v ktorých trendová zložka sa vyvíja v súlade s rovnicou (18). Rovnica (19) priamo vyjadruje podmienku kompenzácie sezónnosti v rámci 4 sezón. Model sa dá prepísať do stavovo priestorovej reprezentácie v tvare:

$$\mathbf{Y}_t = \mathbf{G}\ \mathbf{x}_t + \mathbf{W}_t$$

$$\mathbf{x}_t = \mathbf{F}\ \mathbf{x}_{t-1} + \mathbf{V}_t$$

kde $\mathbf{x}_t$ je tzv. stavový vektor. Jeho dimenzia je obvykle väčšia, ako je dimenzia pozorovaní študovaného procesu $\mathbf{Y}_t$; $\mathbf{F}$ je tzv. systémová matica, $\mathbf{G}$ je tzv. matica pozorovaní. Odhad stavového vektora $\mathbf{x}_t$ a následne vektora pozorovaní $\mathbf{Y}_t$ sa získal z tzv. Kalmanových jednostupňových rekurzií.

Na výpočet budúcich hodnôt stavových premenných $\mathbf{x}_{t+\tau}$ a na aplikáciu Kalmanovho filtra vôbec je nevyhnutná znalosť počiatočných hodnôt stavového vektora $\mathbf{x}_{1|0}$ a kovariančnej matice $\mathbf{P}_{1|0}$. Ich počiatočný odhad sme vykonali Kalmanovým filtrom z prvých ν hodnôt výstupu $\mathbf{Y}_t$. Na odhad počiatočných hodnôt matice $\mathbf{P}_{1|0}$ je potrebné ďalej numericky špecifikovať matice $\mathbf{Q}$ a $\mathbf{R}$, ktoré obsahujú na hlavnej diagonále rozptyly náhodných zložiek stavového modelu. Počiatočný odhad matice $\mathbf{P}_{1|0}$ je preto aj funkciou rozptylov vystupujúcich v maticiach $\mathbf{Q}$ a $\mathbf{R}$. Aby bol problém počiatočného odhadu matice $\mathbf{P}_{1|0}$ možný, musí sa ešte predtým vyriešiť otázka odhadu numerických dodnôt rozptylov náhodných zložiek stavového modelu.

V článku sme použili spôsob odhadu rozptylov stavového modelu z dát výstupu $\mathbf{Y}_t$ tak, aby tieto odhady zaručovali maximálnu vierohodnosť stavového modelu. Potom takto vypočítané teoretické hodnoty výstupu pre t = 5, 6, ..., 28 sú uvedené v utriedenej forme podľa jednotlivých rokov a sezón spolu so skutočnými hodnotami v tabuľke 4. V nej sú zapísané aj predikované hodnoty pre nasledujúce štyri obdobia.

Na posúdenie prognostickej vhodnosti modelov sme použili charakteristikua odmocniny z priemernej štvorcovej chyby predpovedí (RMSE). Ukázalo sa, že posledný model má kvalitatívne lepšiu modelovaciu a predikčnú schopnosť než modely založené na klasickej štatistickej, resp. ekonometrickej metodológii.