EXTRACTA MATHEMATICAE Vol. 6, Núms. 2 y 3, 64 – 95 (1991)

# Time Series Analysis: Recursive Methods and their Modifications for Time Series with Outliers and Missing Observations

T. CIPRA[1], A. RUBIO[2] AND J. TRUJILLO[2]

[1] *Dept. of Statistics, Charles Univ. of Prague, Sokolovská* 83, 186 00 *Prague* 8, *Czechoslovakia*

[2] *Dpto. de Matemáticas, Fac. Veterinaria, Univ. Extremadura,* 10071 *Cáceres, Spain*

AMS *Subject Class.* (1980): 62M10, 62F15, 60G35 *Received July* 17, 1991

ABSTRACT

The recursive methods are popular in time series analysis since they are computationally efficient and flexible enough to treat various changes in character of data. This paper gives a survey of the most important type of these methods including their classification and relationships existing among them. Special attention is devoted to i) robustification of some recursive methods, capable of facing to outliers in time series, and ii) modifications of recursive methods for time series with missing observations.

## 1. INTRODUCTION AND BACKGROUND

This background is thought to be a first and introductory chapter in this paper. We will review different concepts of analysis of time series formulated by Box–Jenkins methodology, exponential smoothing models, Bayesian forecasting, state–variable representation of autoregressive and moving average process, recursive methods and robustness for time series with outliers and missing observations.

Two main approaches are traditionally used to model an univariate time series $z_1,\ldots,z_t$. In general, these models can be written as:

$$z_t = f(t;\beta) + \epsilon_t, \tag{1.1}$$

where $f(t;\beta)$ is a function of time and unknown coefficients $\beta, \epsilon_t$ is frequently assumed to be uncorrelated errors. Through the selection of appropriate fitting functions it is possible to model and represent a variety of nonseasonal and seasonal time series.

One approach mentioned above is known under the name of discounted least squares or general exponential smoothing (see Brown [24]). The classical

exponential smoothing belongs to popular smoothing and extrapolation methods. Due to its simplicity it is still used frequently in practical time series analysis although more effective, but more complicated methods have been developed.

The exponential smoothing forecast techniques are appropriate only if observations follow particular restricted time series models. In particular, if for univariate time series $z_1,\ldots,z_t$ it is possible to use the general exponential smoothing. The parameters $\beta$ in (1.1) are estimated either by ordinary least squares, which weights all observations equally, or by weighted (discounted) least squares.

Special cases of these models lead to single, double and triple exponential smoothing procedures, (see Brown and Meyer [26] and Brown [24]).

Furthermore, it should be emphasized that many time series contain seasonal components, (ciclical patterns that have the tendency to repeat themselves over a certain fixed period of time). Seasonality is modelled by either seasonal indicators or trigonometric functions of time.

The traditional approach to modelling seasonal data is to decompose the series into three components; a trend $T_t$, a seasonal component $S_t$, and an irregular (or error) component $\epsilon_t$:

$$z_t = T_t + S_t + \epsilon_t \quad \text{(the additive decomposition)} \tag{1.2}$$

$$z_t = T_t \times S_t \times \epsilon_t \quad \text{(the multiplicative decomposition)} \tag{1.3}$$

where (1.3) is readily transformed into an additive one by considering logarithmic transformation. $T_t$ is the trend component frequently modeled by a low–order polynomials of time $t$, the seasonal component $S_t$ is modeled, as we mentioned above, by either seasonal indicators or trigonometric functions of time and $\{\epsilon_t\}$ are assumed to be uncorrelated.

Least squares or the general exponential smoothing approach, employ Winter's forecasting methods to estimate the parameters of the seasonal time series (see Abraham and Ledolter [4]).

A widely used method to describe the mechanism that generates and explains univariate time series $z_1,\ldots,z_t$ is the estimation of an autoregressive integrated moving average model of orders $p$, $d$, and $q$, (ARIMA $(p,d,q)$ model), which is the basic model of Box–Jenkins methodology used for nonseasonal data (see Box–Jenkins [20]).

To model and to analyze seasonal time series by Box–Jenkins multiplicative

and nonmultiplicative models (see Box–Jenkins [20]).

The nonstationary series $z_t$ should be first transformed into a stationary one by considering relevant differences $y_t = \nabla^d z_t = (1-B)^d z_t$, where $B$ denotes the backward shift operator ($Bz_t = z_{t-1}$).

Since $y_t$ is then a stationary process, we can use the ARMA, autoregressive–moving average models to describe $y_t$.

The corresponding model can be written as:

$$\varphi(B)\, y_t = \vartheta_0 + \vartheta(B)\, \epsilon_t \tag{1.4}$$

or

$$\varphi(B)(1-B)^d z_t = \vartheta_0 + \vartheta(B)\, \epsilon_t \tag{1.5}$$

where $\varphi(B)$, $\vartheta(B)$ and $\{\epsilon_t\}$ denote the autoregressive operator, the moving average operator and a white–noise stationary process, respectively and $\vartheta_0$ is a constant usually referred to as a trend parameter.

For all subsequent considerations the stationarity of the autoregressive operator are supposed. Therefore the roots of each charasteristic equation must lie outside the complex unit circle. In the following, we will assume $d=0$ and the time series $z_t = y_t$ and concentrate on the estimation of the process ARMA $(p,q)$ (see Box–Jenkins [20]).

When the given time series is possibly contaminated by outliers, the bad performance of least squares estimators for contaminated data show the necessity of estimating time series parameter robustly, since the presence of outliers can seriously bias the estimates. There is a need of characterizing time series contaminated by outliers in appropriate probabilistic models.

Abraham and Box [3], Fox [65] and Martin [118] discuss two characterizations of outliers in the context of time series models.

In this paper we will consider (Section 8), two basic types of outliers: additive outliers (AO) and innovation outliers (IO). In order to deal with recursive methods, robust filtering and smoothing, a vector state–variable representation of ARMA processes will be described as:

$$x_t = \varphi x_{t-1} + \epsilon_t \tag{1.6}$$

where $x_1,\ldots,x_t$ is an ARMA $(p,q)$ process which has a zero–mean value and which is free, e.g., of additive outliers.

The representation of an ARMA $(p,q)$ process contaminated by additive outliers can be represented by (1.6) together with equation (1.7) (see, e.g.,

Stockinger [165]).

$$y_t = \mu + Hx_t + V_t \tag{1.7}$$

where $H=(1,0,...,0)$, $\mu$ is a location parameter and $V_t$ is the residual vector.

State space model are based on the so–called Markov–property, which implies the independence of the future of a process from its past, given the present state. In such a system the state of the process summarizes all the information from the past that is necessary to predict the future.

State space model is then described by two equations: a measurement equation, which describes the generation of the observations from a given state vector, and a system equation, which describes the evolution of the state vector.

Measurement equation: $$y_t = H_t S_t + \epsilon_t \tag{1.8}$$

System equation: $$S_t = A_t S_{t-1} + G u_{t-1} + a_t. \tag{1.9}$$

Where in equation (1.8), $S_t$ is an unobservable state vector that describes the state of the system at time $t$; $y_t$ is an observation vector; $H_t$ is a known matrix and $\epsilon_t$ is the measurement noise, which is a white–noise process with mean vector zero and covariance matrix $\mathrm{E}(\epsilon_t, \epsilon_t')$.

Equation (1.9) describes the evolution of the state vector; $u_t$ is a vector of known inputs, $A$ and $G$ are known matrices, and $a_t$ is the process noise, with mean vector zero and covariance matrix $\mathrm{E}(a_t, a_t')$.

It is assumed that the two error terms $\epsilon_t$, $a_t$ are uncorrelated at all lags; $\mathrm{E}(\epsilon_t, a_{t-k}') = 0$ for all $k$.

In order to specify the distribution of the state vector $S_{t-1}$ in (1.9), it is necessary to start with a distribution for the state vector at time zero. Kalman [101] and Kalman and Bucy [102] derived the recursive updating equations by a method of orthogonal projection.

Kalman's key idea was to fit the model recursively step by step as each new observation $Y_t$ comes in. One can proceed by assuming normality and using minimum mean square linear estimation techniques. Both give the same discussion (see Duncan and Horn [54]).

On the other hand, recursive methods play an important role in the modern time series analysis where they are successfully used for smoothing, predicting and estimation in the corresponding time series models. They work usually by updating a previous constructed value through a correction term depending on the

new observation.

The recursive methods have some advantages in comparison with nonrecursive ones. Usually they are flexible enough to treat time series that cannot be modelled by classical models with constant parameters since their character changes in time. The other important advantage consist in their numerical efficiency (e.g., in comparison with the classical regression procedures it is not necessary to invert matrices of large dimensions) and in smaller demands on the capacity of the computer memory. It is for this reason the recursive methods are recommended in situations where one must treat a large number of signals in real time (on–line).

Since their beginnings in the fifties of this century the recursive methods have been developed rapidly in many directions, in addition, various theoretical aspects of these methods have been investigated (e.g. their stability). Extensive numerical studies have compared their properties, and the corresponding software has been developed so that nowadays these methods are contained in commonly used packages of statistical programs.

This paper gives a survey information on the main directions in the development of the recursive time series methods and shows some relationships among them (see Cipra [40]). The methods presented here have been chosen in accordance with their importance for practical purposes, and their ad hoc description is preferred to theoretical derivations. The Kalman filter introduced in Section 2 seems to be a suitable methodology unifying the presentation of the majority of the recursive methods described in Sections 3–7, (exponential smoothing, general exponential smoothing, Box–Jenkins methodology, Bayesian forecasting, other adaptive techniques). Such a unifying approach may be convenient from the point of view of practical users of these methods.

Section 8 is devoted to robustification of recursive time series methods. Robust time series analysis is an important part of general robust statistics useful for data contaminated by heavy–tailed distributions (in practice, one speaks on time series with outliers), (see Huber [94] and Hampel et al. [78]). There are numerous methods suggested for treatment of such time series (e.g. M–, GM– and CMM–estimation). In this paper, attention is concentrated on the problem of combining the robust properties with recursive ones.

Finally, Section 9 deals with recursive methods for time series with missing observations. As far as the time series with missing observations are concerned,

we will deal with two types: time series containing gaps of missing observations and time series with irregularly published observations.

Missing observations (either in the form of gaps or in the form of data published at irregular time intervals) present a usual problem in practical time series analysis. The natural problem in such situations is a suitable interpolation or prediction using incomplete information only. The paper surveys some possiblities to conduct it in a recursive way.

## 2. KALMAN FILTER

Originally the Kalman filter and its applications have been developed and used as a very practical instrument for the adaptative estimation and prediction of time series not only in technical applications but also for shorter (e.g. economic) time series. On the other hand, if the Kalman filter is formulated in the Bayesian way it has a natural statistical interpretation and can be used conveniently in models of linear regression including time series models (see e.g. Anderson and Moore [10], Jazwinski [96], Priestley [150], Schneider [158] and others). A very inspiring work in this direction is one of Meinhold and Singpurwalla [132].

We shall constrain ourselves to the description of the discrete version of the Kalman filter for linear models although recently its extensions for non–linear models have appeared in literature (see e.g. Cipra [38],[43], Haggan, Heravi and Priestley [77], Jakoby and Pandit [95], Priestley [150]):

$$y_t = F_t \vartheta_t + v_t \quad \text{(the observation equation)} \tag{2.1}$$

$$\vartheta_t = G_t \vartheta_{t-1} + w_t \quad \text{(the state equation)} \tag{2.2}$$

where $y_t$ is an $(m\times 1)$ vector of observations, $F_t$ is an $(m\times n)$ observation matrix, $\vartheta_t$ is an $(n\times 1)$ vector of state variables and $G_t$ is an $(n\times n)$ state transition matrix or system matrix and $v_t$, $w_t$ are random vectors.

This model describes by means of the state equation (2.2) the time development of the state variable $\vartheta_t$ of the type $n\times 1$. The state variable characterizes the given system at time $t$ (e.g., it can be formed by the values of all parameters of a statistical model at time $t$). The observation equation (2.1) describes the relations of the state variable to the observation variable $y_t$ of the type $m\times 1$ that is at our disposal at time $t$. The matrices $F_t$ of the type $m\times n$ and $G_t$ of the type $n\times n$ are known at time $t$. The observations $y_1,\ldots,y_t$ delivered

with time $t$ will be denoted by the symbol $Y^t=\{y_1,...,y_t\}$. The residual vectors $v_t$ of the type $m\times 1$ and $w_t$ of the type $n\times 1$ fulfil in the classical case the following assumptions

$$\begin{gathered} v_t \sim \mathrm{N}(0, V_t)\ ,\quad w_t \sim \mathrm{N}(0, W_t)\ , \\ \mathrm{E}(v_s v_t') = 0\ ,\quad \mathrm{E}(w_s w_t') = 0 \quad \text{for } s \neq t\ , \\ \mathrm{E}(v_s w_t') = 0 \end{gathered} \tag{2.3}$$

with variance matrices $V_t$ and $W_t$ known at time $t$.

In our a priori knowledge on the system at time $t-1$ has the form

$$(\vartheta_{t-1} | Y^{t-1}) \sim \mathrm{N}(\hat{\vartheta}_{t-1}^{t-1}, \hat{\Sigma}_{t-1}^{t-1}) \tag{2.4}$$

then the Kalman filter provides by means of the following recursive formulas

$$\hat{\vartheta}_t^{t-1} = G_t \hat{\vartheta}_{t-1}^{t-1}\ , \tag{2.5}$$

$$\hat{\Sigma}_t^{t-1} = G_t \hat{\Sigma}_{t-1}^{t-1} G_t' + W_t \tag{2.6}$$

predictive values $\hat{\vartheta}_t^{t-1}$ and $\hat{\Sigma}_t^{t-1}$ for the distribution

$$(\vartheta_t | Y^{t-1}) \sim \mathrm{N}(\hat{\vartheta}_t^{t-1}, \hat{\Sigma}_t^{t-1})\ . \tag{2.7}$$

After obtaining the observation $y_t$ at time $t$ the Kalman filter calculates by means of the recursive formulas

$$\hat{\vartheta}_t^t = \hat{\vartheta}_t^{t-1} + \hat{\Sigma}_t^{t-1} F_t' (F_t \hat{\Sigma}_t^{t-1} F_t' + V_t)^{-1} (y_t - F_t \hat{\vartheta}_t^{t-1})\ ; \tag{2.8}$$

$$\hat{\Sigma}_t^t = \hat{\Sigma}_t^{t-1} - \hat{\Sigma}_t^{t-1} F_t' (F_t \hat{\Sigma}_t^{t-1} F_t' + V_t)^{-1} F_t \hat{\Sigma}_t^{t-1} \tag{2.9}$$

the values $\hat{\vartheta}_t^t$ and $\hat{\Sigma}_t^t$ for the a posteriori distribution

$$(\vartheta_t | Y^t) \sim \mathrm{N}(\hat{\vartheta}_t^t, \hat{\Sigma}_t^t)\ . \tag{2.10}$$

The relations (2.5) and (2.6) can be looked upon as the predictions for time $t$ from time $t-1$, while the relations (2.8) and (2.9) as the corrections of these predictions after delivering the new observations $y_t$ at time $t$. Moreover, one can also predict the variable $y$ as

$$\hat{y}_{t+1} = F_{t+1} \hat{\vartheta}_{t+1}^t\ , \tag{2.11}$$

or more generally for $k$ steps ahead as

$$\hat{y}_{t+k} = F_{t+k}\,\hat{\vartheta}^{t}_{t+k}\,, \tag{2.12}$$

where $\hat{\vartheta}^{t}_{t+k}$ can be calculated recursively as

$$\hat{\vartheta}^{t}_{t+k} = G_{t+k}\,\hat{\vartheta}^{t}_{t+k-1}\,. \tag{2.13}$$

The prediction formulas (2.11)–(2.13) can be used only under the condition that we know or are capable to estimate the values of the matrices $F$ and $G$ for time of prediction (this is specially the case if the matrices are constant with time).

The complicated formulas (2.5), (2.6), (2.8) and (2.9) usually reduce to substantially simpler and easily programmable forms in such special cases that are important from the practical point of view. It can be demonstrated by means of the linear regression model with parameters that are constant at time. The system (2.1)–(2.3) for this model can be written as

$$y_t = \sum_{i=1}^{k} b_{it}\,x_{it} + \epsilon_t = x_t'\,b_t + \epsilon_t,\quad \epsilon_t \sim \mathrm{N}(0,\sigma_t^2)\,, \tag{2.14}$$

$$b_t = b_{t-1}\,, \tag{2.15}$$

i.e. $F_t = x_t'$ ($x_t'$ is the last row of the regression design matrix known at time $t$), $G_t = I$, $V_t = \sigma_t^2$, $W_t = 0$. After substituing to (2.5), (2.6) and (2.8) we obtain

$$\hat{b}^{t}_{t} = \hat{b}^{t-1}_{t-1} + [(y_t - x_t'\hat{b}^{t-1}_{t-1})/(x_t'\hat{\Sigma}^{t-1}_{t-1}x_t + \sigma_t^2)]\hat{\Sigma}^{t-1}_{t-1}x_t\,, \tag{2.16}$$

which is the recursive version of the Least Squares method. In comparison with the nonrecursive one it is not necessary to invert any matrix.

The Kalman filter in statistics motivates numerous research works. One of the challenging problem is the estimation of the parameters in the system (2.1)–(2.3) (see e.g. Mehra [131]). The maximum likelihood method and, especially, its algorithmic form called the EM algorithm seems to give satisfactory results (see e.g. Cantarelis and Johnston [30], Harvey [87], Harvey and Peters [89], Schneider [158], Shumway and Stoffer [162] and others). Nowadays there are even some asymptotic results connected with the Kalman filtering (see e.g. Spall and Wall [164], Watanabe [174]).

## 3. EXPONENTIAL SMOOTHING

The exponential smoothing belongs to the classical recursive time series methods. In spite of this fact and its numerical simplicity it gives surprisingly good results in smoothing and predicting.

The exponential smoothing is based on the additive or multiplicative decomposition model of the time series where the series is decomposed to the trend, seasonal and random (residual) components. The trend component has usually the form of a simple mathematical curve (mostly a polynomial of lower order), the seasonal component is usually modelled by means of seasonal indices and the random component is supposed to be a white noise (i.e. a sequence of uncorrelated random variables with zero mean value and constant variance). The substantial feature of the exponential smoothing principle is its dynamic approach to time series analysis: by discounting data exponentially to the past of time series it is capable to respond in a flexible way to various changes ocurring in the character of particular decomposition components.

Let us demonstrate it by means of the simple exponential smoothing where one supposes the following model

$$y_t = b + \epsilon_t \tag{3.1}$$

i.e. the additive decomposition is formed by a constant trend and a white noise. The estimate $S_t$ of the trend at time $t$ (or in other words, the smoothed level of the time series $y_t$ at time $t$) is looked for by means of the Discounted Least Squares method minimizing the expression

$$\sum_{j=0}^{\infty}(1-\alpha)^j(y_{t-j}-b)^2 = \sum_{j=0}^{\infty}\beta^j(y_{t-j}-b)^2 ,$$

where $\beta = 1-\alpha$ is a discounting factor corresponding to a smoothing constant $\alpha$ $(0<\alpha<1)$. The approximation consisting in the exploitation of the infinite sum simplifies substantially the consequent calculations. If deriving according to $b$ we obtain

$$S_t = \alpha\sum_{j=0}^{\infty}(1-\alpha)^j y_{t-j} .$$

For the practical purposes it must be rewritten to the recursive form

$$S_t = \alpha y_t + (1-\alpha) S_{t-1} . \tag{3.2}$$

The formula (3.2) demonstrate in a distinctive way the advantages of the exponential smoothing: the numerical simplicity, small demands on memory

capacity and the single control of smoothing intensity by means of $\alpha$. The prediction at time $t$ for $k$ steps ahead in the model (3.1) obviously has the form

$$\hat{y}_{t+k} = S_t \,. \tag{3.3}$$

Let the symbol $e_t$ denote in the following text the one–step–ahead prediction error, i.e.

$$e_t = y_t - \hat{y}_t(t-1) \,. \tag{3.4}$$

The formula (3.2) can be rewritten as

$$S_t = S_{t-1} + \alpha e_t \tag{3.5}$$

showing that the mechanism of the simple exponential smoothing really consists in the recursive application of the correction term $\alpha e_t$.

Much work is devoted to various aspects of the exponential smoothing. If constraining oneself only to the monographs one can list e.g. Abraham and Ledolter [4], Armstrong [13], Bowerman and O'Connell [19], Brown [23],[24], Cipra [39], Coutie [49], Gilchrist [70], Granger and Newbold [75], Holt et al. [93], Johnson and Montgomery [97], Makridakis and Wheelwright [114], Makridakis et al. [113], Montgomery and Johnson [135] and others. One of the best survey papers devoted to the exponential smoothing is by Gardner [68] classifying the exponential smoothing methods according to the type of trend (constant, linear, exponential, damped linear), according to the presence of seasonality and according to the number of smoothing constants.

We shall present the most important models of exponential smoothing in such a way that for each model the recursive formulas of the type (3.2), the correction formulas of the type (3.5) and the prediction formulas of the type (3.3) will be given. Similarly as in Gardner [68], the symbols $S_t$, $T_t$, $I_t$ will denote the smoothed level, the smoothed slope of linear trend and the smoothed seasonal index at time $t$, respectively. The smoothing constants corresponding to $S_t$, $T_t$, $I_t$ will be denoted by $\alpha$, $\gamma$, $\delta$, respectively, and the lenght of season by $p$ (e.g. $p = 12$ for monthly observations with the annual seasonality). In addition, we shall discuss briefly the choice of initial values for the corresponding recursive formulas (see e.g. Abraham and Ledolter [4], Berry and Bliemel [16], Bowerman and O'Connell [19], Dalrymple and King [50], Flowers [64], Harrison [80], Chatfield [34], Chatfield and Yar [35], Makridakis et al. [112], Montgomery and Johnson [135]) and the choice of smoothing constants (in addition to the references

above see e.g. Gilchrist [70], Ledolter and Abraham [108], Taylor [167], Wade [173], Winters [178]). If assuming the normality of data then, in addition to the point predictions, one can also construct the prediction intervals which is closely related to the investigation of the forecast error varaince (see e.g. Bowerman and O'Connel [19], Cipra [39], McKenzie [130], Montgomery and Johnson [135], Sweet [166], Yar and Chatfield [180]).

(A) CONSTANT TREND WITHOUT SEASONALITY (SIMPLE EXPONENTIAL SMOOTHING):

See the formulas (3.2), (3.3) and (3.5). The recommended choice of the smoothing constant $\alpha$ is from the range $0.1 \leqslant \alpha \leqslant 0.3$. Its more accurate value can be determined by minimizing the mean squared error of one–step–ahead prediction over selected grid values of $\alpha$ from the interval $<0.1;0.3>$. However, some empirical and theoretical results indicate that sometimes one can use the values of $\alpha$ lying outside of this interval, e.g. the values $\alpha > 0.3$ have frequently shown to be optimal in the comparative study by Makridakis et al. [112], and the corresponding model is even stable for $0 < \alpha < 2$ according to Brenner [21]. The initial value $S_0$ is usually chosen as the arithmetic mean of several initial observations.

(B) LINEAR TREND WITHOUT SEASONALITY WITH ONE SMOOTHING CONSTANT (DOUBLE EXPONENTIAL SMOOTHING, BROWN MODEL):

$$S_t = \alpha y_t + (1-\alpha) S_{t-1} \ , \ (S_t = S_{t-1} + T_{t-1} + \alpha e_t) \tag{3.6}$$

$$T_t = \alpha (S_t - S_{t-1}) + (1-\alpha)\, T_{t-1} \ , \ (T_t = T_{t-1} + \alpha^2 e_t) \tag{3.7}$$

$$\hat{y}_{t+k}(t) = S_t + \frac{1-\alpha}{\alpha} T_t + k\, T_t \tag{3.8}$$

or equivalently

$$S_t = \alpha y_t + (1-\alpha) S_{t-1} \tag{3.9}$$

$$S_t^{[2]} = \alpha S_t + (1-\alpha) S_{t-1}^{[2]} \tag{3.10}$$

$$\hat{y}_{t+k}(t) = 2S_t - S_t^{[2]} + k \frac{\alpha}{1-\alpha} (S_t - S_t^{[2]}) \tag{3.11}$$

where $S_t^{[2]}$ is the so called smoothing statistic of the second order defined recursively by means of the formula (3.10).

(C) LINEAR TREND WITHOUT SEASONALITY WITH TWO SMOOTHING CONSTANTS (HOLT MODEL):

$$S_t = \alpha y_t + (1-\alpha)(S_{t-1} + T_{t-1}) \ , \ (S_t = S_{t-1} + T_{t-1} + \alpha e_t) \tag{3.12}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma) T_{t-1} \ , \ (T_t = T_{t-1} + \alpha\gamma e_t) \tag{3.13}$$

$$\hat{y}_{t+k}(t) = S_t + kT_t \ . \tag{3.14}$$

Practical experiences show that the introduction of two different smoothing constants for level and slope in the Holt method may improve the previous Brown method (see e.g. Gardner and Dannenbring [69], Makridakis et al. [112]).

(D) LINEAR TREND WITH ADDITIVE SEASONALITY (ADDITIVE HOLT–WINTERS MODEL):

$$S_t = \alpha(y_t - I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1}), \ (S_t = S_{t-1} + T_{t-1} + \alpha e_t) \tag{3.15}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma) T_{t-1}, \ (T_t = T_{t-1} + \alpha\gamma e_t) \tag{3.16}$$

$$I_t = \delta(y_t - S_t) + (1-\delta) I_{t-p}, \ (I_t = I_{t-p} + \delta(1-\alpha) e_t) \tag{3.17}$$

$$\hat{y}_{t+k}(t) = S_t + kT_t + I_{t-p+k} \ . \tag{3.18}$$

(E) LINEAR TREND WITH MULTIPLICATIVE SEASONALITY (MULTIPLICATIVE HOLT–WINTERS MODEL):

$$S_t = \alpha(y_t / I_{t-p}) + (1-\alpha)(S_{t-1} + T_{t-1}), \ (S_t = S_{t-1} + T_{t-1} + \alpha e_t / I_{t-p}) \tag{3.19}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma) T_{t-1}, \ (T_t = T_{t-1} + \alpha\gamma e_t / I_{t-p}) \tag{3.20}$$

$$I_t = \delta(y_t / S_t) + (1-\delta) I_{t-p}, \ (I_t = I_{t-p} + \delta(1-\alpha) e_t / S_t) \tag{3.21}$$

$$\hat{y}_{t+k}(t) = (S_t + kT_t) I_{t-p+k} \ . \tag{3.22}$$

As the choice of smoothing constants for the models with linear trend (B)–(E) is concerned similar principles as in the model (A) hold. On the other hand, more effective algorithms for the optimal choice of these constants can be used than the grid search method is (see e.g. Archibald [12], Bartolomei [14], Berry and Bliemel [16], Flowers [64]). The generally recommended values are $\alpha \sim 0.2$ or smaller ones for the Brown model and the smoothing constants not exceeding 0.3 for the Holt model and the Holt–Winters model, the Brown model being stable for $0 < \alpha < 2$ and the Holt model for $0 < \alpha < 2$, $0 < \gamma < (4-2\alpha)/\alpha$. The initial values of smoothing statistics are frequently constructed in such a way that one fits a line and seasonal indices modelled by means of zero–one regressors to a beginning segment of the time series using the classical regression techniques.

Practical experiences show that in general the models with linear trend are acceptable for short–term prediction but they incline to overestimate the real values in long–term prediction. Therefore one recommends to damp the linear trend in a suitable way if it is to be used for long–term predictions (see e.g. (G)).

(F) EXPONENTIAL TREND WITHOUT SEASONALITY (see Pegels [143], Roberts [154]):

$$S_t = \alpha y_t + (1-\alpha) S_{t-1} T_{t-1} \ , \ (S_t = S_{t-1} T_{t-1} + \alpha e_t) \tag{3.23}$$

$$T_t = \gamma (S_t / S_{t-1}) + (1-\gamma) T_{t-1} \ , \ (T_t = T_{t-1} + \alpha\gamma e_t / S_{t-1}) \tag{3.24}$$

$$\hat{y}_{t+k}(t) = S_t T_t^k \ . \tag{3.25}$$

(G) DAMPED LINEAR TREND WITHOUT SEASONALITY:

$$S_t = \alpha y_t + (1-\alpha)(S_{t-1} + \varphi T_{t-1}) \ , \ (S_t = S_{t-1} + \varphi T_{t-1} + \alpha e_t) \tag{3.26}$$

$$T_t = \gamma (S_t - S_{t-1}) + (1-\gamma)\varphi T_{t-1} \ , \ (T_t = \varphi T_{t-1} + \alpha\gamma e_t) \tag{3.27}$$

$$\hat{y}_{t+k}(t) = S_t + \textstyle\sum_{i=1}^{k} \varphi^i T_i \tag{3.28}$$

where $0 < \varphi < 1$ (for $\varphi = 1$ it converts to the Holt model (3.12)–(3.14)). The long–term predictions based on linear trends of a damped type differ substantially from the ones based on undamped trends even if $\varphi$ is near to one. The damped trends are used in the forecast system FORSYS (see Lewandowski [110]) that has been successful just for long–term predictions in the comparative study by Makridakis et al. [112].

(H) POLYNOMIAL TREND:

The Brown model (3.9)–(3.11) can be generalized for the polynomial trend of order $n$. In this case one must work with $n+1$ smoothing statistics $S_t$, $S_t^{[2]},\ldots,S_t^{[n+1]}$ defined analogously as in (3.9) and (3.10). However, practical results show that the polynomial trends of higher orders starting with $n=2$ (the so called triple exponential smoothing) provide worse results if compared with other methods (see Makridakis et al. [112]).

One can deal with a lot of further interesting topics in the framework of the exponential smoothing:

There are some multivariate generalizations suggested in literature (see e.g. Cipra [41], Enns et al. [60], Harvey [88], Jones [99], Pfefferman and Allon [146]).

In some situations the assumption of the uncorrelated random variables $\epsilon_t$

(see e.g. (3.1)) is not adequate. Therefore some authors (see e.g. Chatfield [34], Reid [152]) recommend to improve the prediction $\hat{y}_{t+k}(t)$ by adding the term $r_1^k e_t$ to it, where $r_1$ is the estimated autocorrelation between $e_t$ and $e_{t-1}$.

The important part of the exponential smoothing methodology is the monitoring of the forecast errors that may indicate some inadequacies in the procedure (see e.g. Batty [15], Bonsdorff [18], Bowerman and O'Connell [19], Brown [24],[25], Gardner [66], Golder and Settle [73], Harrison and Davies [82], Jun [100], McKenzie [128], Montgomery and Johnson [135], Trigg [168]). For such a monitoring one frequently uses the so called tracking signal treated by the cumulative sum (CUSUM) techniques. The related problem is the adaptive control that performs automatically the necessary changes of the smoothing constants. The method by Trigg and Leach [169] is very popular; it adapts the smoothing constant $\alpha$ at time $t$ to the form

$$\alpha_t = |Q_t/\Delta_t| \tag{3.29}$$

where $Q_t$ and $\Delta_t$ are calculated recursively as

$$Q_t = \alpha e_t + (1-\alpha)Q_{t-1} \ , \ \Delta_t = \alpha|e_t| + (1-\alpha)\Delta_{t-1} \ . \tag{3.30}$$

Chow [37] has suggested to treat parallely three procedures of exponential smoothing constants e.g. $\alpha - 0.05$, $\alpha$ and $\alpha + 0.05$ and to select the one with the optimal prediction property. If the selected constant at time $t$ is $\alpha + 0.05$ the one must transfer to the smoothing constants $\alpha$, $\alpha + 0.05$ and $\alpha + 0.10$. Other references are Bowerman and O'Connell [19], Eilon and Elmaleh [56], Ekern [57], Gardner [67], Montgomery [134], Raine [151], Roberts and Reed [154].

## 4. General Exponential Smoothing

The term "general exponential smoothing" (or sometimes "direct smoothing") means that the exponential smoothing principle consisting in the Discounted Least Squares method (see the discussion following the equation (3.1)) is used for a general linear model. If at time $t$ we write this model as

$$y_{t-j} = \sum_{i=1}^{k} b_{it} x_i(-j) + \epsilon_{t-j} = b_t' x(-j) + \epsilon_{t-j}, \ j=0,1,\ldots,t-1 \tag{4.1}$$

then the estimated parameter vector $\hat{b_t}$ is obtained by minimizing the discounted least squares

$$\sum_{j=0}^{t-1} \beta^j [y_{t-j} - b_t' x(-j)]^2 \ , \tag{4.2}$$

and the prediction for $k$ steps ahead in the considered time series is constructed as

$$\hat{y}_{t+k}(t) = \hat{b}_t' x(k) \,. \tag{4.3}$$

The regressors $x$ of an arbitrary form (e.g. goniometric functions) can be included in the model (4.1). However, if they fullfil the following recursive equation

$$x(t) = L x(t-1) \tag{4.4}$$

with a matrix $L$ of the type $k \times k$ constant at time $t$ then after some algebraic manipulations one obtains the following recursive formula for the extimated parameter vector

$$\hat{b}_t = L' \hat{b}_{t-1} + h e_t \,, \tag{4.5}$$

where the vector $h$ of the type $k \times 1$ depends on the form of regressors $x$ and on teh discount factor $\beta$ (see e.g. Abraham and Ledolter [4], Brown [23], Montgomery and Johnson [135] and others).

Some of the methods from Section 3 are special cases of the general exponential smoothing. E.g., the Brown model (3.6)–(3.8) can be obtained if one puts $k=2$, $x_1(t)=t$, $x_2(t)=t$ (it is $y_t = b_1 + b_2 t$) and

$$L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} .$$

On the other hand, the general exponential smoothing used for a time series with linear trend and additive seasonality gives results that differ from the ones obtained by the additive Holt–Winters method. Special modifications of the general exponential smoothing suggested for commercial purposes are known as the systems SEATREND and DOUBT (see Harrison [79], McKenzie [127]). A more general approach to the exponentially discounted estimation is presented e.g. by Ameen [8], Ameen and Harrison [9], Harrison and Akram [81].

## 5. Box–Jenkins Methodology

This methodology has become popular since it offers a systematic approach to modelling stationary and nonstationary time series with general trend and seasonality that are common in economic and technical practice. The form and construction of Box–Jenkins models is described including various details and numerical examples in many monographs (see e.g. Abraham and Ledolter [4], Anderson [11], Bowerman and O'Connell [19], Box and Jenkins [20], Cipra [39],

Granger and Newbold [75], McCleary and Hay [125], Montgomery and Johnson [135], Vandaele [171] and others). Let us only remind the general form of the model ARIMA $(p,d,q)$ that is the basic model of Box–Jenkins methodology used for nonseasonal data:

$$\varphi(B)(1-B)^d y_t = \vartheta(B)\epsilon_t\,, \tag{5.1}$$

where

$$\varphi(B)=1-\varphi_1 B-\ldots-\varphi_p B^p\,,\ \vartheta(B)=1+\vartheta_1 B+\ldots+\vartheta_q B_q \tag{5.2}$$

and $\epsilon_t$ is a white noise with $\sigma^2=\mathrm{var}(\epsilon_t)$. The symbol $B$ denotes the backward–shift operator fulfilling $B^j y_t = y_{t-j}$ $(j=0,1,\ldots)$. The operator $(1-B)^d$ forms the $d$–th difference of the original series in the model (5.1), e.g. if $d=1$ then $(1-B)y_t = y_t - y_{t-1}$. The seasonal differences of the type $(1-B^p)y_t = y_t - y_{t-p}$ are typical for the seasonal Box–Jenkins models used for seasonal data (e.g. $p=12$ for monthly data).

There are several reasons why the Box–Jenkins models are involved in this survey of recursive time series methods:

(1) Predictions based on the Box–Jenkins models can be constructed recursively. E.g., let us consider the model ARMA $(p,q)$ of the form

$$y_t = \varphi_1 y_{t-1}+\ldots+\varphi_p y_{t-p}+\epsilon_t+\vartheta_1\epsilon_{t-1}+\ldots+\varphi_q\epsilon_{t-q} \tag{5.3}$$

(see (5.1) for $d=0$). Then the prediction $\hat{y}_{t+k}(t)$ is constructed recursively as

$$\hat{y}_{t+k}(t)=\varphi_1[y_{t+k-1}]+\ldots+\varphi_p[y_{t+k-p}]+[\epsilon_{t+k}]+\ldots+\vartheta_q[\epsilon_{t+k-q}] \tag{5.4}$$

where

$$[y_{t+j}]=\hat{y}_{t+j}(t),\ [\epsilon_{t+j}]=0 \quad \text{for } j>0$$

and (5.5)

$$[y_{t+j}]=y_{t+j},\ [\epsilon_{t+j}]=y_{t+j}-\hat{y}_{t+j}(t+j-1) \quad \text{for } j\leqslant 0$$

(2) Estimation of parameters of the Box–Jenkins models can be performed recursively. E.g., let us consider the autoregresive model AR $(p,q)$ of the form

$$y_t = \varphi_1 y_{t-1}+\ldots+\varphi_p y_{t-p}+\epsilon_t=\varphi' z_t+\epsilon_t\,, \tag{5.6}$$

where $\varphi=(\varphi_1,\ldots,\varphi_p)'$ and $z_t=(y_{t-1},\ldots,y_{t-p})'$ (see (5.1) for $d=q=0$). Then one can use the following recursive estimation formulas

$$\hat{\varphi}_t=\hat{\varphi}_{t-1}+\frac{P_{t-1}z_t}{z_t'P_{t-1}z_t+1}(y_t-\hat{\varphi}_{t-1}'z_t)=\hat{\varphi}_{t-1}+P_tz_t(y_t-\hat{\varphi}_{t-1}'z_t), \tag{5.7}$$

$$P_t=P_{t-1}-\frac{P_{t-1}z_tz_t'P_{t-1}}{z_t'P_{t-1}z_t+1}, \tag{5.8}$$

$$\hat{\sigma}_t^2=\frac{1}{t-p}\left[(t-p-1)\hat{\sigma}_{t-1}^2+\frac{(y_t-\hat{\varphi}_{t-1}'z_t)^2}{z_t'P_{t-1}z_t+1}\right] \tag{5.9}$$

(see e.g. Aase [1], Cipra [39], Fahrmeier [62], Graupe [76], Young [181] and others).

(3) The majority of the exponential smoothing models provide the same prediction results as some Box–Jenkins models. We shall show the corresponding Box–Jenkins model for

– simple exponential smoothing (3.2), (3.3), (3.5), $(\beta=1-\alpha)$:

$$(1-B)y_t=(1-\beta B)\epsilon_t\,; \tag{5.10}$$

– Brown model (3.6)–(3.8), $(\beta=1-\alpha)$:

$$(1-B)^2y_t=(1-\beta B)^2\epsilon_t\,; \tag{5.11}$$

– Holt model (3.12)–(3.14) :

$$(1-B)^2y_t=[1+(\alpha+\alpha\gamma-2)B+(1-\alpha)B^2]\epsilon_t\,; \tag{5.12}$$

– additive Holt–Winters model (3.15)–(3.18):

$$(1-B)(1-B^p)y_t=\left[\sum_{i=0}^{p+1}\vartheta_iB^i\right]\epsilon_t\,, \tag{5.13}$$

where

$\vartheta_0=1,\ \vartheta_1=-1+\alpha+\alpha\gamma,$

$\vartheta_i=\alpha\gamma,\ i=2,\ldots,p-1,$

$\vartheta_p=-1+\alpha\gamma+\delta(1-\alpha),\ \vartheta_{p+1}=(1-\alpha)(1-\delta);$

– general exponential smoothing (4.1)–(4.5):

$$\varphi(B)(1-B)^dy_t=\vartheta(\beta B)\epsilon_t\,, \tag{5.14}$$

where (5.14) is a model ARIMA $(p,d,n)$ with $p+d=n$ (we shall not specify its parameters).

The related references are e.g. Abraham and Ledolter [4],[5], Chatfield [33], Cogger [48], Godolphin and Harrison [72], Ledolter and Abraham [108], Ledolter

and Box [109], McKenzie [126],[127],[129], Muth [137], Newbold [138], Pandit and Wu [140].

## 6. Bayesian Forecasting

The term Bayesian forecasting has been introduced to time series analysis due to works by Harrison and Stevens (1971,1976). It denotes the predictive methodology based on the Kalman filter (see Section 2). One of its main advantages consists in the state space representation of time series models. E.g., let us consider a functional model

$$2y_t - 3y_{t-1} + y_{t-2} = 0 \tag{6.1}$$

and its state space representation

$$y_t = \mu_t \,, \tag{6.2}$$

$$\mu_t = \frac{1}{2}\mu_{t-1} + \beta_t \,, \tag{6.3}$$

$$\beta_t = \beta_{t-1} \,. \tag{6.4}$$

Then the equations (6.2)–(6.4) imply (6.1) but the opposite implication is not valid since, in contrast to (6.1), the state space representation (6.2)–(6.4) contains the information on behaviour of level $\mu_t$ and slope $\beta_t$ of the corresponding time series. In the framework of the Bayesian forecasting the state space equations (6.2)–(6.4) must be randomized as it is the case in the following examples:

(A) Model with Constant Level (Steady Model):

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim \mathrm{N}(0,\sigma_t^2) \,, \tag{6.5}$$

$$\mu_t = \mu_{t-1} + \delta\mu_t \,, \quad \delta\mu_t \sim \mathrm{N}(0,w_t) \,. \tag{6.6}$$

In this model the level $\mu_t$ is modelled as a random walk. If denoting

$$(\mu_t | y_1,\ldots,y_t) \sim \mathrm{N}(m_t,c_t) \tag{6.7}$$

then the Kalman filter provides

$$m_t = m_{t-1} + \alpha_t(y_t - m_{t-1}) \,, \tag{6.8}$$

where

$$\alpha_t = \frac{c_{t-1} + w_t}{c_{t-1} + w_t + \sigma_t^2} \,, \quad c_t = \alpha_t \sigma_t^2 \,. \tag{6.9}$$

In other words, we have obtained the simple exponential smoothing with the smoothing constant $\alpha_t$ changing in time. If $w_t/\sigma_t^2=\tilde{w}$ does not depend on time then it implies $\alpha_t=(\alpha_{t-1}+\tilde{w})/(\alpha_{t-1}+\tilde{w}+1)$ so that in the stable state with $\alpha_t=\alpha$ it must be

$$\alpha=\tfrac{1}{2}[-\tilde{w}+(\tilde{w}+4\tilde{w})^{\frac{1}{2}}] . \tag{6.10}$$

The formula (6.10) can be used for the choice of smoothing constant in the simple exponential smoothing.

(B) MODEL WITH LINEAR TREND:

$$y_t=\mu_t+\epsilon_t, \quad \epsilon_t \sim N(0,\sigma_t^2) , \tag{6.11}$$

$$\mu_t=\mu_{t-1}+\beta_{t-1}+\delta\mu_t , \quad \delta\mu_t \sim N(0,w_t) . \tag{6.12}$$

$$\beta_t=\beta_{t-1}+\delta\beta_t , \quad \delta\beta_t \sim N(0,u_t) . \tag{6.13}$$

If $w_t/\sigma_t^2=\tilde{w}$ and $u_t/\sigma_t^2=\tilde{u}$ do not depend on time then, similarly as for the model with constant level, in the stable state the corresponding Kalman filter formulas coincide with the Holt exponential smoothing (3.12)–(3.14), and it holds

$$\tilde{u}=\gamma^2(1-\alpha\gamma)^{-1} , \quad \tilde{w}=(\alpha^2+\alpha^2\gamma-2\alpha\gamma)(1-\alpha\gamma)^{-1} . \tag{6.14}$$

(C) SEASONAL MODEL:

$$y_t=\rho_t+\epsilon_t, \quad \epsilon_t \sim N(0,\sigma_t^2) , \tag{6.15}$$

$$\textstyle\sum_{j=0}^{p-1}\rho_{t-j}=\delta\rho_t, \quad \delta\rho_t \sim N(0,r_t) . \tag{6.16}$$

Further details and models can be found e.g. in Fahrmeier (1981), Harrison and Stevens (1971,1976), Harvey [85],[86],[87], Kahl and Ledolter (1983), Morrison and Pike (1977). The Bayesian forecasting can be extended to the multi–process models that transfer with various probabilities to different states (see e.g. Bolstad [17], Harrison and Stevens (1976), West and Harrison (1986)).

## 7. OTHER ADAPTIVE TECHNIQUES

There are further recursive time series methods producing estimation and prediction in an adaptive way (see e.g. the general scheme LRE (Linear Recursive Estimators) described by Pack, Pike and Downing (1985)).

As an example of possible modifications of the basic recursive methods, let

us give the method AEP (Adaptive Estimation Procedure, see Carbonne and Longini (1977), Makridakis and Wheelwright (1978)). If in the model (2.14) we have no a priori information of the type (2.15) on the behaviour of parameters $b_t$ at time $t$ then this method uses the following recursive estimation formula

$$\hat{b}_{it}=\hat{b}_{i,t-1}+\mu\left|\frac{\hat{b}_{i,t-1}}{x_t'\,\hat{b}_{t-1}}\right|x_{it}\bar{x}_i\bar{i}^1(y_t-x_t'\hat{b}_{t-1}),\ i=1,\ldots,k, \tag{7.1}$$

where $\mu$ is a damping factor guaranteeing the stability of the method and

$$\bar{x}_{it}=\nu|x_{it}|+(1-\nu)\bar{x}_{i,t-1}\ ,\ i=1,\ldots,k \tag{7.2}$$

with a constant $\nu$ $(0<\nu<1)$. The empirical study by Bretschneider, Carbone and Longini [22] recommends very small values of $\mu$ $(0.01<\mu<0.06)$.

## 8. Robustification of Recursive Methods

The basic objective of the robust time series analysis is the identification of outliers or/and the treatment of contaminated data by means of special statistical procedures that are insensitive (robust) with respect to outliers (see e.g. Huber (1981), Stockinger and Dutter (1987) and others).

The identification of outliers can be based on special statistical models and tests or on various ad hoc procedures (see e.g. Abraham and Box [3], Chaloner and Brant (1988), Chernik, Downing and Pike (1982), Fox (1972), Hillmer (1984), Ledolter (1989), Schmid (1986), Tsay (1988)).

The choice of a robust statistical procedure for a contaminated time series depends substantially on the type of outliers. There are two basic types of outliers: additive outliers (AO) and innovation outliers (IO). If we confine ourselves to the ARMA processes (see (5.3)) then these two types of outliers can be modelled in the following way:

(1) AO–Model:

$$y_t=x_t+v_t\ , \tag{8.1}$$

$$\varphi(B)x_t=\vartheta(B)\epsilon_t,\ \ \epsilon_t\sim \mathrm{N}(0,\sigma^2)\ , \tag{8.2}$$

$$v_t\sim \text{iid}\ F=(1-\eta)\delta_0+\eta H\ ,\ \ P(\delta_0=0)=1\ , \tag{8.3}$$

where $\eta$ is a constant near to zero $(0<\eta<1)$ and $H$ is an absolute continuous distribution with heavy tails (e.g. Cauchy, Laplace, uniform, normal with a variance much larger than $\sigma^2$) so that $P(v_t=0)=1-\eta$. In other words, one

observes the ARMA process $x_t$ contaminated by outliers in a small fraction $\eta$ of observations.

(2) IO–Model:

$$\varphi(B)x_t = \vartheta(B)\epsilon_t, \tag{8.4}$$

$$\epsilon_t \sim \text{iid } G = (1-\eta)\,\mathrm{N}(0,\sigma^2) + \eta H\,,\ \ \mathrm{var}(H) >> \sigma^2\,, \tag{8.5}$$

i.e., in the IO–models the outliers arise in such a way that the variance of innovations increases considerably in a small fraction $\eta$ of observations.

There are various estimation procedures recommended for contaminated time series data (see the survey by Stockinger and Dutter (1987)): M–estimation suitable for IO–models (see e.g. Denby and Martin (1979), Martin [118],(1981), Martin and Yohai (1985), Pham Dinh Tuan (1984)), GM–estimation (Generalized M–) suitable for AO–models (see e.g. Bustos [28], Denby and Martin (1979), Martin [119], Martin and Yohai (1985)), CMM–estimation (Conditional–Mean–M–) based on ACM filtering (Approximate Conditional–Mean, see e.g. Martin [116],[117], Kleiner, Martin and Thomson (1979)) and others. The scope of the robust time series analysis is very broad nowadays, and in this overview we shall concentrate our attention only on some robustified recursive methods:

(A) ROBUST KALMAN FILTER: There are a lot of successful attempts how to robustify the Kalman filter (see Section 2) including the robustification of the Bayesian forecasting models (see Section 6). E.g., let us consider according to Cipra and Romera (1991) the Kalman filter with contaminated scalar observations (i.e. $m=1$):

$$y_t = f_t\vartheta_t + v_t\,, \tag{8.6}$$

$$\vartheta_t = G_t\vartheta_{t-1} + w_t\,, \tag{8.7}$$

$$v_t \sim (1-\eta)\,\mathrm{N}(0,r_t) + \eta H_t\,,\ \mathrm{var}(H_t) >> r_t\,,\ w_t \sim \mathrm{N}(0,W_t) \tag{8.8}$$

Then the predictive recursive formulas (2.5), (2.6) stay unchanged while the a posteriori recursive formulas (2.8), (2.9) must be robustified in the following way

$$\hat{\vartheta}_t^t = \hat{\vartheta}_t^{t-1} + \hat{\Sigma}_t^{t-1} f_t' (f_t \hat{\Sigma}_t^{t-1} f_t' + r_t/k_t)^{-1} (y_t - f_t \hat{\vartheta}_t^{t-1}), \tag{8.9}$$

$$\hat{\Sigma}_t^t = \hat{\Sigma}_t^{t-1} - \hat{\Sigma}_t^{t-1} f_t' f_t \hat{\Sigma}_t^{t-1} (f_t \hat{\Sigma}_t^{t-1} f_t' + r_t/k_t)^{-1}, \tag{8.10}$$

where

$$k_t = \psi(r_t^{-\frac{1}{2}}(y_t - f_t\hat{\vartheta}_t^{t-1}))/(r_t^{-\frac{1}{2}}(y_t - f_t\hat{\vartheta}_t^{t-1})) \tag{8.11}$$

and $\psi$ a suitable robutifying function, e.g. the Huber's function

$$\psi_H(x) = \begin{cases} x & \text{for } |x| \leqslant c \\ c\,\mathrm{sgn}(x) & \text{for } |x| > c \end{cases} \tag{8.12}$$

(the constant $c$ depends on $\eta$, e.g. one recommends $c = 1.645$ for $\eta = 0.05$). Other related references are e.g. Ershov and Liptser (1978), Girón, Martínez and Rojano (1989), Masreliez (1975), Masreliez and Martin (1977), Meinhold and Singpurwalla (1989), Peña and Guttman (1988,1989), Poljak and Tsypkin (1980), Sawitzki (1981), Servi and Ho (1981), Sorenson and Alspach (1971), Stockinger and Dutter (1987), West (1981), West, Harrison and Mignon (1985). Some investigations continue on the robustification of nonlinear Kalman filters (see e.g. Cipra (1990), Cipra and Rubio (1991)).

(B) ROBUST EXPONENTIAL SMOOTHING: It seems very useful for practical purposes torobustify exponential smoothing procedures. Exponential smoothing robustifications based on $L_1$ norm (replacing the least squares by the least absolute deviations) or on M−estimation are suggested in Cipra [42],(1991). E.g., the robustified simple exponential smoothing based on M−estimation principle replaces (3.5) by the following recursive formulas ($\beta = 1 - \alpha$)

$$S_t = S_{t-1} + M_{t-1}(M_{t-1} + \beta/k_t)^{-1} e_t\,, \tag{8.13}$$

$$M_t = \frac{1}{\beta}[M_{t-1} - M_{t-1}^2(M_{t-1} + \beta/k_t)^{-1}]\,, \tag{8.14}$$

$$k_t = \psi(e_t/\hat{\sigma}_{t-1})/(e_t/\hat{\sigma}_{t-1})\,, \tag{8.15}$$

$$\hat{\sigma}_t = \nu|e_t| + (1-\nu)\hat{\sigma}_{t-1} \tag{8.16}$$

with a robustifying function $\psi$ (see e.g. (8.12)), with a constant $\nu$ $(0 < \nu < 1)$ and with initial conditions

$$S_n = \textstyle\sum_{i=1}^n y_i/n\,,\quad M_n = 1/\mathrm{n}\,,\quad \hat{\sigma}_n = \frac{1}{n-1}\sum_{i=1}^n (y_i - S_n)^2. \tag{8.17}$$

Robust nonlinear smoothers have been suggested by Velleman (1980)

(C) ROBUST RECURSIVE BOX–JENKINS METHODOLOGY: The recursive Box–Jenkins prodedures can be derived in the framework of the robust Kalman filtering (see (A)). Some theoretical and practical results connected with such procedures are given e.g. in Campbell (1982), Cipra and Romera (1991), Englund Holst and Ruppert (1988,1989), Kitagawa (1987). A recursive construction of maximum likelihood function for the Box–Jenkins models is described e.g. by Harvey [85],[86], Harvey and Phillips (1979), Pearlman (1980), Shea (1984).

## 9. RECURSIVE METHODS WITH MISSING OBSERVATIONS

As far as the time series with missing observations are concerned, there are two cases that are important from the practical point of view:

(1) a time series $y_t$ contains gaps of missing observations such that sufficient numbers of observations are available before and after each gap;

(2) observations are published at irregular time intervals so that instead of observations $y_1,\dots,y_t$ one obtains only observations $y_{t_1},\dots,y_{t_n}$ at irregular periods $t_1,\dots,t_n$ $(t_1<\dots<t_n)$.

There is an extensive literature devoted to time series with missing observations dealing with estimation, interpolation and prediction for such series (see e.g. Abraham [2], Akaike and Ishiguro [6], Brubacher and Wilson [27], Damsleth (1980), Dunsmuir and Robinson (1981), Ferreiro (1987), Harvey and Pierse (1984), Jones (1980), Ljung (1989), Maravall and Peña (1989), Parzen (1984), Pourahmadi (1989), Rosen and Porat (1989), Shumway (1984), Shumway and Stoffer (1982), and others). Here we shall mention briefly some modifications of recursive methods suggested for time series with missing observations:

(A) KALMAN FILTER WITH MISSING OBSERVATIONS: It is a natural procedure to replace the missing observations in the Kalman filter by their predictive values both in the situation with gaps and in the one with irregularly published observations (see e.g. Cipra [42], De Jong (1989), Gordon and Smith (1990), Kohn and Ansley (1986), and others). Moreover, the Kalman filter can treat very elegantly the multivariate observations with unobservable components (see e.g. Cipra and Motyková (1987)): if only the components $y_{i_1t},\dots,y_{i_dt}$ $(1\leqslant i_1<\dots<i_d\leqslant m)$ of the $y_t$ are at our disposal at time $t$ then at this time it

suffices to replace the observation equation (2.1) by

$$y_t^* = F_t^* \vartheta_t + v_t^* , \tag{9.1}$$

where

$$y_t^* = M_t y_t , \quad F_t^* = M_t F_t , \quad V_t^* = \mathrm{var}(v_t^*) = M_t V_t M_t' \tag{9.2}$$

and $M_t$ is a matrix of the type $d \times m$ that has unities in the positions $(1,i_1),\dots,(d,i_d)$ and zeroes in the remaining positions.

(B) EXPONENTIAL SMOOTHING WITH MISSING OBSERVATIONS: The exponential smoothing for nonseasonal time series with gaps has been described by Aldrin and Damsleth [7]. E.g., in the case of the simple exponential smoothing with the gap $y_{t-k},\dots,y_{t-1}$ (i.e., the values $\dots, y_{t-k-2}, y_{t-k-1}, y_t$ are observed) they recommend at time $t$ to replace the equation (3.2) by

$$S_t = \alpha_t y_t + (1-\alpha) S_{t-k-1} , \tag{9.3}$$

where

$$\alpha_t = 1 - \alpha[1 + k(1-\alpha)^2]^{-1} , \tag{9.4}$$

and if continuing the exponential smoothing at times $t+1$, $t+2,\dots$ then to preserve the equation (3.2) (see also Cipra [41]).

The exponential smoothing for nonseasonal time series with observations published at irregular time intervals has been suggested by Wright (1986). E.g., in the case of the Holt model one must replace (3.12)–(3.14) by

$$S_n = \alpha_n y_{t_n} + (1-\alpha_n)[S_{n-1} + (t_n - t_{n-1}) T_{n-1}] , \tag{9.5}$$

$$T_n = \gamma_n (S_n - S_{n-1})(t_n - t_{n-1})^{-1} + (1-\gamma_n) T_{n-1} , \tag{9.6}$$

$$\hat{y}_{t_n+k}(t_n) = S_n + k T_n , \tag{9.7}$$

where

$$\alpha_n = \alpha_{n-1}(a_n + \alpha_{n-1})^{-1} , \quad a_n = (1-\alpha)^{t_n - t_{n-1}} , \tag{9.8}$$

$$\gamma_n = \gamma_{n-1}(c_n + \gamma_{n-1})^{-1} , \quad c_n = (1-\gamma)^{t_n - t_{n-1}} , \tag{9.9}$$

($\alpha, \gamma$ are the smoothing constants that would be used for the considered time series if no observations are missing). Moreover, one recommends to put

$$\alpha_0 = 1 - (1-\alpha)^q , \quad \gamma_0 = 1 - (1-\gamma)^q , \tag{9.10}$$

where $q$ is the average time spacing of the data.

## REFERENCES


1. K. A. AASE, Recursive estimation in non–linear time series models of autoregressive type, *Journal of the Royal Statistical Society* **B45** (1983), 228 – 237
2. B. ABRAHAM, Missing observations in time series, *Commun. Statist. – Theor. Meth.* **A10** (1976), 1643 – 1653.
3. B. ABRAHAM AND G. E. P. BOX, Bayesian analysis of some outlier problems in time series, *Biometrika* **66** (1979), 229 – 236.
4. B. ABRAHAM AND J. LEDOLTER, "Statistical Methods of Forecasting", Wiley, New York, 1983.
5. B. ABRAHAM AND J. LEDOLTER, Forecast functions implied by autoregressive–integrated moving average models and other related forecast procedures, *International Statistical Review* **54** (1986), 51 – 66.
6. H. AKAIKE AND M. ISHIGURO, Trend estimation with missing observations, *Ann. Inst. Statist. Math.* **B32** (1980), 481 – 488.
7. M. ALDRIN AND E. DAMSLETH, Forecasting non–seasonal time series with missing observations, *Journal of Forecasting* **8** (1989), 97 – 116.
8. J. R. M. AMEEN, Sequential discount estimation, *Statistician* **37** (1988), 227 – 237.
9. J. R. M. AMEEN AND P. J. HARRISON, Discounted weighted estimation, *Journal of Forecasting* **3** (1984), 285 – 296.
10. B. D. O. ANDERSON AND J. B. MOORE, "Optimal Filtering", Prentice–Hall, Englewood Cliffs, 1979.
11. O. D. ANDERSON, "Time Series Analysis: The Box–Jenkins Approach", Butterworth, London, 1976.
12. B. C. ARCHIBALD, Parameter space of the Holt–Winter's model, *International Journal of Forecasting* **6** (1990), 199 – 209.
13. J. S. ARMSTRONG, "Long–Range Forecasting", Wiley, New York, 1978.
14. S. M. BARTOLOMEI, A note on a comparison of exponential smoothing methods of forecasting seasonal series, *International Journal of Forecasting* **5** (1989), 111 –116.
15. M. BATTY, Monitoring and exponential smoothing forecasting system, *Operational Research Quarterly* **20** (1969), 319 – 325.
16. W. L. BERRY AND F. W. BLIEMEL, Selecting exponential smoothing constants: an application of pattern search, *International Journal of Production Research* **12** (1979), 483 – 499.
17. W. M. BOLSTAD, Harrison–Stevens forecasting and the multiprocess dynamic linear model, *The American Statistician* **40** (1986), 129 – 135.
18. H. BONSDORFF, A comparison of the ordinary and a varying parameter exponential smoothing, *Journal of Applied Probability* **27** (1989), 784 – 792.
19. B. L. BOWERMAN AND R. T. O'CONNELL, "Time Series and Forecasting", Duxbury Press, North Scituate, Massachusetts, 1979.
20. G. E. P. BOX AND G. M. JENKINS, "Time Series Analysis, Forecasting and Control", 1st Ed., Holden Day, San Francisco, 1970.
21. J. L. BRENNER, Difference equations in forecasting formulas, *Management Science* **14** (1968), 141 – 159.
22. S. BRETSCHNEIDER, R. CARBONE AND R. L. LONGINI, An adaptive approach to time series analysis, *Decision Sciences* **10** (1979), 233 – 244.
23. R. G. BROWN, "Statistical Forecasting for Inventory Control", McGraw–Hill, New York, 1959.
24. R. G. BROWN, "Smoothing, Forecasting and Prediction of Discrete Time Series", Prentice–Hall, Englewood Cliffs, 1963.
25. R. G. BROWN, Detection of turning points in a time series, *Decision Sciences* **2** (1971), 383 – 403.

26. R.G. BROWN AND R.F. MEYER, The fundamental theorem of exponential smoothing, *Operation Research* **9** (1961), 673–685.
27. S.R. BRUBACHER AND G.T. WILSON, Interpolating time series with application to the estimation of holiday effects on electricity demand, *Applied Statistics* **25** (1976), 107–116.
28. O.H. BUSTOS, General M–estimates for contaminated pth–order autoregressive processes: consistency and asymptotic normality, *Z. Wahrsch. verw. Gebiete* **59** (1982), 491–504.
29. K. CAMPBELL, Recursive computation of M–estimates for the paramenters of a finite autoregressive process, *Annals of Statistics* **10** (1982), 442–453.
30. N. CANTARELIS AND F.R. JOHNSTON, On–line variance estimation for the steady state Bayesian forecasting model, *Journal of Time Series Analysis* **3** (1982), 225–234.
31. R. CARBONE AND R.L. LONGINI, A feedback model for automated real estate assessment, *Management Science* **24** (1977), 241–248.
32. K. CHALONER AND R. BRANT, A Bayesian approach to outlier detection and residual analysis, *Biometrika* **75** (1988), 651–659.
33. C. CHATFIELD, Some recent developments in time series analysis, *Journal of the Royal Statistical Society* **A140** (1977), 492–510.
34. C. CHATFIELD, The Holt–Winters forecasting procedures, *Applied Statistics* **27** (1978), 264–279.
35. C. CHATFIELD AND M. YAR, Holt–Winters forecasting: some practical issues, *Statistician* **37** (1988), 129–140.
36. M..R. CHERNIK, D.J. DOWNING AND D.H. PIKE, Detecting outliers in time series data, *Journal of the American Statistical Association* **77** (1982).
37. W.M. CHOW, Adaptive control of the exponential smoothing constant, *Journal of Industrial Engineering* **16** (1965), 315–317.
38. T. CIPRA, Non linear models of time series, *Ekonomicko–matematicky obzor* **19** (1983), 164–178, in Czech.
39. T. CIPRA, "Analysis of Time Series with Applications in Economy", SNTL/ALFA, Prague, 1986, in Czech.
40. T. CIPRA, Recursive procedures in time series, *Ekonomicko–matematicky obzor* **24** (1988), 417–438, in Czech.
41. T. CIPRA, Some problems of exponential smoothing, *Aplikace matematiky* **34** (1989), 161–169.
42. T. CIPRA, Some recursive methods in time series, in "Proceedings of the Fourth Prague Symposium on Asymptotic Statistics", P. Mandl and M. Husková Eds., SNTL, Prague, 1989, 181–191.
43. T. CIPRA, Note on approximate non–Gaussian filtering with nonlinear observation relation, *Comment. Math. Univ. Carolinae* **31** (1990), 601–605.
44. T. CIPRA, Robust exponential smoothing, *Journal of Forecasting* **10** (1991), to appear.
45. T. CIPRA AND I. MOTYKOVÁ, Study on Kalman filter in time series analysis, *Comment. Math. Univ. Carolinae* **28** (1987), 549–563.
46. T. CIPRA AND R. ROMERA, Robust Kalman filter and its applications in time series analysis, *Kybernetika* **27** (1991), to appear.
47. T. CIPRA AND A. RUBIO, Kalman filter with nonlinear non–Gaussian observation relation, *Trabajos de Estadística* (1991), to appear.
48. K.O. COGGER, The optimality of general order exponential smoothing, *Operation Research* **22** (1974), 858–867.
49. G. COUTIE, "Short–Term Forecasting", Oliver and Boyd, Edinburgh, 1966.

50. D.J. DALRYMPLE AND B.E. KING, Selecting parameters for short–term forecasting techniques, *Decision Sciences* **12** (1981), 661–669.
51. E. DAMSLETH, Interpolating missing values in a time series, *Scandinavian Journal of Statistics* **7** (1980), 33–39.
52. P. DE JONG, Smoothing and interpolating with the state–space model, *Journal of the American Mathematical Association* **84** (1989), 1085–1088.
53. L. DENBY AND R.D. MARTIN, Robust estimation of the first–order autoregressive parameter, *Journal of the American Mathematical Association* **74** (1979), 140–146.
54. D.B. DUNCAN AND S.D. HORN, Linear dynamic regression extimation from the viewpoint of regression analysis, *Journal of the American Mathematical Association* **67** (1972), 815–821.
55. W. DUNSMUIR AND P.M. ROBINSON, Estimation of time series models in the presence of missing data, *Journal of the American Mathematical Association* **76** (1981), 560–568.
56. S. EILON AND J. ELMALEH, Adaptive limits for inventory control, *Management Sciences* **16** (1970), 533–548.
57. S. EKERN, Adaptive exponential smoothing revisited, *Journal of Operational Research Society* **32** (1981), 775–782.
58. J.E. ENGLUND, U. HOLST AND D. RUPPERT, Recursive M–estimators for location and scale for dependent sequences, *Scandinavian Journal of Statistics* **15** (1988), 147–159.
59. J.E. ENGLUND, U. HOLST AND D. RUPPERT, Recursive M–estimators for stationary, strong mixing processes – A representation theorem and asymptotic distribution, *Stochastic Processes and Their Applications* **31** (1989), 203–222.
60. P.G. ENNS, J.A. MACHAK, W.A. SPIVEY AND W.J. WROBLESKI, Forecasting applications of an adaptive multiple exponential smoothing model, *Management Science* **28** (1982), 1035–1044.
61. A.A. ERSHOV AND R.S. LIPTSER, Robust Kalman filter in discrete time, *Automatica and Remote Control* **39** (1978), 359–367.
62. L. FAHRMEIER, "Rekursive Algorithmen fur Zeitreihenmodelle", Vandenhoeck und Ruprecht, Gottingen, 1981.
63. O. FERREIRO, Methodologies for estimation of missing observations in time series, *Statistics and Probability Letters* **5** (1987), 65–69.
64. A.D. FLOWERS, A simulation study of smoothing constant limits for an adaptive forecasting system, *Journal of Operational Management* **2** (1980), 85–94.
65. A.J. FOX, Outliers in time series, *Journal of the Royal Statistical Society* **B34** (1972), 350–363.
66. E.S. GARDNER, Automatic monitoring of forecasts errors, *Journal of Forecasting* **2** (1983), 1–21.
67. E.S. GARDNER, Evolutionary operation of the exponential smoothing parameter: revisited, *Omega* **11** (1983), 621–623.
68. E.S. GARDNER, Exponential smoothing: the state of the art, *Journal of Forecasting* **4** (1985), 1–28.
69. E.S. GARDNER AND D.G. DANNENBRING, Forecasting with exponential smoothing: some guidelines for model selection, *Decision Sciences* **11** (1980), 370–383.
70. W.G. GILCHRIST, "Statistical Forecasting", Wiley, London, 1976.
71. F.J. GIRÓN, M.L. MARTÍNEZ AND J.C. ROJANO, Modelos lineales dinámicos y mixturas de distribuciones, *Estadística Española* **31** (1989), 165–206.

72. E.J. GODOLPHIN AND P.J. HARRISON, Equivalence theorems for polynomial–projecting predictors, *Journal of the Royal Statistical Society* **B37** (1975), 205–215.
73. E.R. GOLDER AND J.G. SETTLE, Monitoring schemes in short–term forecasting, *Operational Research Quarterly* **27** (1976), 489–501.
74. K. GORDON AND A.F.M. SMITH, Modeling and monitoring biomedical time series, *Journal of the American Statistical Association* **85** (1990), 328–337.
75. C.W.J. GRANGER AND P. NEWBOLD, "Forecasting Economic Time Series", Academic Press, New York, 1977.
76. D. GRAUPE, "Time Series Analysis, Identification and Adaptive Filtering", Krieger (2nd Ed.), Malabar, Florida, 1989.
77. V. HAGGAN, S.M. HERAVI AND M.B. PRIESTLEY, A study of the application of state–dependent models in non–linear time series analysis, *Journal pf Time Series Analysis* **5** (1984), 69–102.
78. R. HAMPEL, E.M. RONCHHETTI, P.J. ROSSEEUW AND W.A. STAHEL, "Robust Statistics, The Approach Based on Influence Functions", John Wiley & Sons, New York, 1986.
79. P.J. HARRISON, Short–term sales forecasting, *Applied Statistics* **14** (1965), 102–139.
80. P.J. HARRISON, Exponential smoothing and short–term sales forecasting, *Management Sciences* **13** (1967), 821–842.
81. P.J. HARRISON AND M. AKRAM, Generalized exponential weighted regression and parsimonious dynamic linear modelling, in "Time Series Analysis: Theory and Practice 3" (O.D. Anderson ed.), North–Holland, Amsterdam, 1983, 19–42.
82. P.J. HARRISON AND O.L. DAVIS, The use of cumulative sum (CUSUM) techniques for the control of routine forecasts of product demand, *Operations Research* **12** (1964), 325–333.
83. P.J. HARRISON AND C.F. STEVENS, A Bayesian approach to short–term forecasting, *Operational Research Quarterly* **22** (1971), 341–362.
84. P.J. HARRISON AND C.F. STEVENS, Bayesian forecasting, *Journal of the Royal Statistical Society* **B38** (1976), 205–247.
85. A.C. HARVEY, "Time Series Models", Wiley, New York, 1981.
86. A.C. HARVEY, "The Econometric Analysis of Time Series", Wiley, New York, 1981.
87. A.C. HARVEY, A unified view of statistical forecasting procedures, *Journal of Forecasting* **3** (1984), 245–275.
88. A.C. HARVEY, Analysis and generalisation of a multivariate exponential smoothing model, *Management Science* **32** (1986), 374–380.
89. A.C. HARVEY AND S. PETERS, Estimation procedures for structural time series model, *Journal of Forecasting* **9** (1990).
90. A.C. HARVEY AND G.D.A. PHILLIPS, Maximum likelihood estimation of regression models with autoregressive–moving average disturbances, *Biometrika* **66** (1979), 49–58.
91. A.C. HARVEY AND R.G. PIERSE, Estimating missing observations in economic time series, *Journal of the American Statistical Association* **79** (1984), 125–131.
92. S. HILLMER, Monitoring and adjusting forecasts in the presence of additive outliers, *Journal of Forecasting* **3** (1984), 205–215.
93. C.C. HOLT, I.F. MODIGLIANI AND H.A. SIMON, "Planing Production and Work Force", Prentice–Hall, Englewood Cliffs, 1960.
94. P.J. HUBER, "Robust Statistics", Wiley, New York, 1981.
95. W. JACOBY AND M. PANDIT, A prediction–error–method for recursive identification of nonlinear systems, *Automatica* **23** (1987), 491–496.

96. A.H. JAZWINSKI, "Stochastic Processes and Filtering Theory", Academic Press, New York, 1970.
97. L.A. JOHNSON AND D.C. MONTGOMERY, "Operation Research in Production Planning Scheduling and Inventory Control", Wiley, New York, 1974.
98. R.H. JONES, Exponential smoothing for multivariate time series, *Journal of the Royal Statistical Society* **B28** (1966), 241–251.
99. R.H. JONES, Maximum likelihood fitting of ARMA models to time series with missing observations, *Technometrics* **22** (1980), 389–395.
100. D.B. JUN, On detecting and estimating a major level or slope change in general exponential smoothing, *Journal of Forecasting* **8** (1989), 55–64.
101. R.E. KALMAN, A new approach tol linear filtering and prediction problems, *Journal of Basic Engineering* **83** (1960), 35–45.
102. R.E. KALMAN AND R.S. BURY, New results in linear filtering and prediction theory, *Journal of Basic Engineering* **83** (1961), 95–107.
103. D.R. KAHL AND J. LEDOLTER, A recursive Kalman filter forecasting approach, *Management Sciences* **29** (1983), 1325–1333.
104. G. KITAGAWA, Non–Gaussian state–space modeling of nonstationary time series, *Journal of the American Statistical Association* **82** (1987), 1032–1063.
105. R. KLEINER, R.D. MARTIN AND D.J. THOMSON, Robust estimation of power spectra, *Journal of the Royal Statistical Society* **B41** (1979), 313–338.
106. R. KOHN AND C.F. ANSLEY, Estimation, prediction and interpolation for ARIMA models with missing data, *Journal of the American Statistical Association* **81** (1986), 751–761.
107. J. LEDOLTER, The effect of additive outliers on the forecasts from ARIMA models, *International Journal of Forecasting* **5** (1989), 1–10.
108. J. LEDOLTER AND B. ABRAHAM, Somme comments on the initialization of exponential smoothing, *Journal of Forecasting* **3** (1984), 79–84.
109. J. LEDOLTER AND G.E.P. BOX, Conditions for the optimality of exponential smoothing forecast procedures, *Metrika* **25** (1978), 77–93.
110. R. LEWANDOWSKI, Sales forecasting by FORSYS, *Journal of Forecasting* **1** (1982), 205–214.
111. G.M. LJUNG, A note on the estimation of missing values in time series, *Commun. Statist.–Simula.* **18** (1989), 459–465.
112. S. MAKRIDAKIS, A. ANDERSEN, R. CARBONE, R. FILDER, M. HIBON, R. LEWANDOWSKI, I. NEWTON, E. PARZEN AND R. WINKLER, The accuracy of extrapolation (time series) methods: result of a forecasting competition, *Journal of Forecasting* **1** (1982), 111–153.
113. S. MAKRIDAKIS, S.C. WHEELWRIGHT AND V.E. MC GEE, "Forecasting Methods and Applications", 2nd ed., John Wiley & Sons, New York, 1983.
114. S. MAKRIDAKIS AND S.C. WHEELWRIGHT, "Interactive Forecasting: Univariate and Multivariate Methods", Holden Day, San Francisco, 1978.
115. A. MARAVALL AND D. PEÑA, Missing observations, additive outliers and inverse autocorrelation function, EUI Working Paper No. 89/384, European University Institute, Florence, 1989.
116. R.D. MARTIN, Robust estimation for time series autoregressions, in "Robustness in Statistics" (R.L. Launer and G.N. Wilkinson eds.), Academic Press, New York, 1979, 147–176.
117. R.D. MARTIN, Approximate conditional type of smoothers and interpolators, in "Smoothing Techniques for Curve Estimation" (T. Gasser and M. Rosenblatt eds.), Springer, New York, 1979, 117–143.
118. R.D. MARTIN, "Time Series Model Estimation, Data Analysis and Robust Procedures", Proc. Symp. Appl. Math. (R.V. Hogg ed.), vol. 23, 1980.

119. R. D. MARTIN, Robust estimation of autoregressive models, in "Directions in Time Series" (D.R. Brillinger and G.C. Tiao eds.), Inst. Math. Statist. Publications, Haywood, 1980, 228–254.
120. R. D. MARTIN, Robust methods for time series, in "Applied Time Series II" (D.F. Findley ed.), Academic Press, New York, 1981, 683–759.
121. R. D. MARTIN AND C. J. MASRELIEZ, Robust estimation via stochastic approximation, *IEEE Transactions on Information Theory* **IT–21** (1975), 263–271.
122. R. D. MARTIN AND V. J. YOHAI, Robustness in time series and estimating ARMA models, in "Handbook of Statistics 5" (E.J. Hannan, P.R. Krishnaiah and M.M. Rao eds.), North–Holland, Amsterdam, 1985, 119–155.
123. C. J. MASRELIEZ, Approximate non–Gaussian filtering with linear state and observation relations, *IEEE Transactions on Automatic Control* **AC–20** (1975), 107–110.
124. C. J. MASRELIEZ AND R. D. MARTIN, Robust Bayesian estimation for the linear model and robustifying the Kalman filter, *IEEE Transactions on Automatic Control* **AC–22** (1977), 361–371.
125. R. MCCLEARY AND R. A. HAY, "Applied Time Series Analysis for the Social Science", Beverly Hills, London, 1980.
126. E. MCKENZIE, A comparison of standar forecasting systems with the Box–Jenkins approach, *Statistician* **23** (1974), 107–116.
127. E. MCKENZIE, A comparison of some standar seasonal forecasting systems, *Statistician* **25** (1976), 3–14.
128. E. MCKENZIE, The monitoring of exponentially weighted forecasts, *Journal of the Operational Research Society* **29** (1978), 449–458.
129. E. MCKENZIE, General exponential smoothing and the equivalent ARMA process, *Journal of Forecasting* **3** (1984), 333–344.
130. E. MCKENZIE, Error analysis for Winters' additive seasonal forecasting system, *International Journal of Forecasting* **2** (1986), 373–382.
131. R. K. MEHRA, On–line identification of linear dynamic systems with applications to Kalman filtering, *IEEE Transactions on Automatic Control* **AC–16** (1971), 12–21.
132. R. J. MEINHOLD AND N. D. SINGPURWALLA, Understanding the Kalman filter, *American Statistician* **37** (1983), 123–127.
133. R. J. MEINHOLD AND N. D. SINGPURWALLA, Robustification of Kalman filter models, *Journal of the American Statistical Association* **84** (1989), 479–486.
134. D. C. MONTGOMERY, Adaptive control of exponential smoothing parameters by evolutionary operation, *AIIE Transactions* **2** (1970), 268–269.
135. D. C. MONTGOMERY AND L. A. JOHNSON, "Forecasting and Time Series Analysis", McGraw–Hill, New York, 1976.
136. G. W. MORRISON AND D. H. PIKE, Kalman filtering applied to statistical forecasting, *Management Science* **23** (1977), 768–774.
137. J. F. MUTH, Optimal properties of exponentially weighted forecasts, *Journal of the American Statistical Association* **55** (1960), 299–306.
138. P. NEWBOLD, Predictors projecting linear trend plus seasonal dummies, *Statistician* **37** (1988), 111–127.
139. D. J. PACK, D. H. PIKE AND D. J. DOWNING, The role of linear recursive estimators in time series forecasting, *Management Science* **31** (1985), 188–199.
140. S. M. PANDIT AND S. M. WU, Exponential smoothing as a special case of a linear stochastic system, *Operation Research* **22** (1974), 868–879.
141. E. PARZEN, "Time Series Analysis of Irregularly Observed Data", Lecture Notes in Statistics 25, Springer, New York, 1984.

142. J.G. PEARLMAN, An algorithm for the exact likelihood of a higher–order autoregressive–moving average process, *Biometrika* **67** (1980), 232–233.
143. C. PEGELS, Exponential forecasting: some new varaitions, *Management Science* **15** (1969), 311–315.
144. D. PEÑA AND I. GUTTMAN, A Bayesian approach to robustifying Kalman filtering, in "Bayesian Analysis of Time Series and Dynamic Models" (J.C. Spall ed.), Dekker, New York, 1988, Chap. 9.
145. D. PEÑA AND I. GUTTMAN, Optimal collapsin of mixture distributions in robust recursive estimation, *Commun. Statist. – Theo. Meth.* **18** (1989), 817–833.
146. D. PFEFFERMAN AND J. ALLON, Multivariate exponential smoothing: method and practice, *International Journal of Forecasting* **5** (1989), 83–89.
147. PHAM DINH TUAN, On robust estimation of parameters for autoregressive moving average models, in "Robust and Nonlinear Time Series Analysis" (J. Franke, W. Hardle and R.D. Martin eds.), Lecture Notes in Statistics 26, Springer, New York, 1984, 273–286.
148. B.T. POLJAK AND J.A.Z. TSYPKIN, Robust identification, *Automatica* **16** (1980), 53–63.
149. M. POURAHMADI, Estimation and interpolation of missing values of a stationary time series, *Journal of Time Series Analysis* **10** (1989), 149–169.
150. M.B. PRIESTLEY, "Spectral Analysis and Time Series", Vol. II, Academic Press, London, 1981.
151. J.E. RAINE, Self–adaptive forecasting reconsidered, *Decision Sciences* **2** (1971), 181–191.
152. D.J. REID, A review of short–term projection techniques, in ,"Practical Aspects of Forecasting" (H.A. Gordon ed.), Operational Research Society, London, 1975.
153. S.A. ROBERTS, A general class of Holt–Winters type forecasting models, *Management Sciences* **28** (1982), 808–820.
154. S.D. ROBERTS AND R. REED, The development of a self–adaptive forecasting technique, *AIIE Transactions* **1** (1969), 314–322.
155. Y. ROSEN AND B. PORAT, Optimal ARMA parameter estimation based on the sample covariances for data with missing observations, *IEEE Transactions on Information Theory* **IT–35** (1989), 342–349.
156. G. SAWITZKI, Exact filtering in exponential families: discrete time, *Math. Operationsforsch. Statistics, Ser. Statistics,* **12** (1981), 393–401
157. W. SCHMID, The multiple outlier problem in time series analysis, *Australian Journal of Statistics* **28** (1986), 400–413.
158. W. SCHNEIDER, "Der Kalmanfilter als Instrument zur Diagnose and Schatzung Variabler Parameter in Okonometrischen Modellen", Physica–Verlag, Heidelberg, 1986.
159. L.D. SERVI AND Y.C. HO, Recursive estimation in the presence of uniformly distributed measurement noise, *IEEE Transactions on Automatic Control* **AC–26** (1981), 563–565.
160. B.L. SHEA, Maximum likelihood estimation of multivariate ARMA process via the Kalman filter, in "Time Series Analysis, D.D. Anderson ed.", Elsevier, Amsterdam, 1984, 91–101.
161. R.H. SHUMWAY, Some applications of the EM algorithm to analyzing incomplete time series data, in "Time Series Analysis of Irregularly Observed Data" (E. Parzen ed.), Springer, New York. 1984, 290–324.
162. R.H. SHUMWAY AND D.S. STOFFER, An approach to time series smoothing and forecasting using the EM algorithm, *Journal of Time Series Analysis* **3** (1982), 253–264.

163. H.W. SORENSON AND D.L. ALSPACH, Recursive Bayesian estimation using Gaussian sums, *Automatica* **6** (1971), 465–479.
164. J.C. SPALL AND K.D. WALL, Asymptotic distribution theory for the Kalman filter state estimator, *Commun. Statist. – Theor. Meth.* **13** (1984), 1981–2003.
165. N. STOCKINGER AND R. DUTTER, Robust time series analysis: a survey, Supplement to the journal *Kybernetika* **23** (1987).
166. A.L. SWEET, Computing the variance of the forecast error for the Holt–Winters seasonal models, *Journal of Forecasting* **4** (1985), 235–243.
167. S.G. TAYLOR, Initializations of exponential smoothing forecasts, *AIEE Transactions* **13** (1981), 199–205.
168. D.W. TRIGG, Monitoring a forecasting system, *Operational Research Quarterly* **15** (1964), 271–274.
169. D.W. TRIGG AND A.G. LEACH, Exponential smoothing with an adaptive response rate, *Operational Research Quarterly* **18** (1967), 53–59.
170. R.S. TSAY, Outliers, level shifts and variance changes in time series, *Journal of Forecasting* **7** (1988), 1–20.
171. W. VANDAELE, "Applied Time Series and Box–Jenkins Models", Academic Press, Orlando, 1983.
172. P.F. VELLEMAN, Definition and comparison of robust nonlinear data smoothing algorithms, *Journal of the American Statistical Association* **75** (1980), 609–615.
173. R.C. WADE, A technique for initializating exponential smoothing forecasts, *Management Science* **13** (1967), 601–602.
174. N. WATANABE, Note on Kalman filter with estimated parameters, *Journal of Time Series Analysis* **6** (1985), 269–278.
175. M. WEST, Robust sequential approximate Bayesian estimation, *Journal of the Royal Statistical Society* **B43** (1981), 157–166.
176. M. WEST AND P.J. HARRISON, Monitoring and adaptation in Bayesian forecasting models, *Journal of the American Statistical Association* **81** (1986), 741–750.
177. M. WEST, P.J. HARRISON AND H.S. MIGON, Dynamic generalized linear models and Bayesian forecasting, *Journal of the American Statistical Association* **80** (1985), 73–97.
178. P.R. WINTERS, Forecasting sales by exponentially weighted moving averages, *Management Science* **6** (1960), 324–342.
179. D.J. WRIGHT, Forecasting data published at irregular time intervals using an extension of Holt's method, *Management Science* **32** (1986), 499–510.
180. M. YAR AND C. CHATFIELD, Prediction intervals for the Holt–Winters forecasting procedure, *International Journal of Forecasting* **6** (1990), 127–137.
181. P. YOUNG, Recursive Estimation and Time–Series Analysis, Springer, Berlin, 1984.