INFORMATICA, 1999, Vol. 10, No. 2, 297–312 


# Recursive Optimization of State Estimation of Dynamic Processes in the Presence of Patchy Outliers in Observations

Rimantas PUPEIKIS

*Institute of Mathematics and Informatics, Vilnius Gediminas Technical University*
*Akademijos 4, 2600 Vilnius, Lithuania*
*e-mail: pupeikis@ktl.mii.lt*

Received: March 1999

**Abstract.** In the previous paper (Pupeikis, 1998), the problem of recursive estimation of the state of linear dynamic systems, described by an autoregressive model (AR), in the presence of time-varying outliers in observations to be processed has been considered. An approach to the robust recursive state estimation has been obtained and proved by estimating the real chemical process (Box and Jenkins, 1970). The aim of the given paper is the development of the abovementioned approach for the robust recursive state estimation of an autoregressive-moving average (ARMA) process in a case of additive noises with patchy outliers. The results of numerical simulation and the state estimation of the AR model (Figs. 1–4) and the real chemical process, described by the ARMA model, which is chosen from the same book of Box and Jenkins (Figs. 5–8) are given.

**Key words:** dynamic system, Kalman filter, robustness, state estimation, optimization.

## 1. Statement of the Problem

It is known (Schick and Mitter, 1994) that some time various recursive robust techniques turn out to be efficient in the presence of rare and isolated outliers. On the other hand, it is established that always there arise special problems when the outliers occur in batches. In such a case, a generalized problem of the model of outliers which are varying in time could be solved according to Pupeikis and Huber (1997); Pupeikis (1998). Therefore, we apply here an approach to optimizing the state estimation of a dynamic process and the generalized model of time varying outliers in observations to be processed for solving the abovementioned problem.

Assume that we consider the linear discrete-time process of order $n$ with single input $\boldsymbol{\mu}_k$ and single output $\boldsymbol{x}_k$ described by the autoregressive moving average (ARMA) model of the form

$$x_k = W(q^{-1}; \alpha)\mu_k \quad \forall k = 0, 1, \ldots, \tag{1}$$

where $x_k$ and $\mu_k$ denote the unobserved values of sequences $\{x_k\}$ and $\{\mu_k\}$, respectively, $\mu_k \sim \mathcal{N}(0, \sigma_\mu^2)$,

$$W(q^{-1}; \alpha) = \frac{1 - B(q^{-1}; b)}{1 - A(q^{-1}; a)} \tag{2}$$

is a system transfer function,

$$A(q^{-1}; a) = \sum_{i=1}^{n} a_i q^{-i}, \tag{3}$$

$$B(q^{-1}; b) = \sum_{i=1}^{n} b_i q^{-i}, \tag{4}$$

are polynomials,

$$\alpha^T = (a^T, b^T), \quad a^T = (a_1, \ldots, a_n), \quad b^T = (b_1, \ldots, b_n) \tag{5}$$

are parameters of polynomials (3), (4), $q^{-1}$ is the backward shift operator defined by $x_{k-n} = x_k q^{-n}$.

Suppose that $\{x_k\}$ is observed under additive noise $\mathbf{Z}_k, i.e.$,

$$u_k = x_k + z_k, \tag{6}$$

where $u_k$ is observed value of output $\mathbf{U}_k$, $z_k$ denotes the unobserved value of the sequence $\{z_k\}$, which is the sequence of independent identically distributed variables with an "$\varepsilon$-contaminated" distribution of the form

$$p(z_k) = (1 - \varepsilon_k)\mathcal{N}(0, \sigma_\xi^2) + \varepsilon_k \mathcal{N}(0, \sigma_\nu^2), \tag{7}$$

and the variance

$$\sigma_z^2 = (1 - \varepsilon_k)\sigma_\xi^2 + \varepsilon_k \sigma_\nu^2, \tag{8}$$

$p(z_k)$ is a probability density distribution of an noise $\mathbf{Z}_k$, moreover, $\sigma_\zeta < \sigma_\nu$, $0 \leqslant \varepsilon_k \leqslant 1$ is the unknown fraction of "contamination" varying in a time.

It is supposed that the roots of $A(q^{-1}; a)$ and $B(q^{-1}; b)$ are outside the unit circle of the $q^{-1}$ plane. The true orders $n$ of polynomials (3), (4) and the true values of parameters (5) are known.

The aim of the given paper is a optimization of the robust recursive estimation of states $x_1, x_2, \ldots, x_N$ of the ARMA process (1)–(8) in the presence of patchy outliers in observations $u_1, u_2, \ldots, u_N$ of an output $\mathbf{U}_k$.

## 2. Recursive Optimization of a State Estimation

The ARMA model can be rewritten as the stochastic state-space model in the form

$$X_{k+1} = AX_k + h\mu_{k+1}, \tag{9}$$
$$u_k = c^T X_k + d^T w_k + z_k, \tag{10}$$

where

$$X_k = (x_k, x_{k-1}, \dots, x_{k-n+1})^T \tag{11}$$

is the $n \times 1$ state vector,

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \tag{12}$$

is the $n \times n$ matrix $\mu_{k+1}$ is value of the sequence $\{\mu_k\}$ at time moment $k+1$,

$$h = c = (1, 0, \dots, 0)^T, \tag{13}$$
$$d = (b_1, b_2, \dots, b_n)^T, \tag{14}$$
$$w_k = (\mu_{k-1}, \mu_{k-2}, \dots, \mu_{k-n})^T \tag{15}$$

are the $n \times 1$ vectors.

So, we have the following structure:

$$\begin{bmatrix} x_{k+1} \\ x_k \\ \vdots \\ x_{k-n+2} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-n+1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mu_{k+1}, \tag{16}$$

$$u_k = [1\,0\dots0] \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-n+1} \end{bmatrix} + [b_1\, b_2 \dots b_n] \begin{bmatrix} \mu_{k-1} \\ \mu_{k-2} \\ \vdots \\ \mu_{k-n} \end{bmatrix} + z_k. \tag{17}$$

Assume that $N$ observations $u_1, u_2, \dots, u_N$ are obtained and a bank of the state estimates $\widehat{x}_1(1), \dots, \widehat{x}_N(1), \widehat{x}_1(2), \dots, \widehat{x}_N(2), \dots, \widehat{x}_1(L), \dots, \widehat{x}_N(L)$ are calculated by processing $u_1, u_2, \dots, u_N$ by means of a bank of robust parallel $L$ Kalman filters

$$\widehat{x}^{(j)}(k+1) = A\widehat{x}^{(j)}(k) + k(k+1)\psi_j\{e_{k+1}(j)\}$$
$$\text{for} \quad j = 1, 2, \dots, L \quad \text{and} \quad k = 0, 1, \dots, N-1. \tag{18}$$

Here

$$\widehat{x}^{(j)}(k+1) = \big(\widehat{x}_{k+1}(j), \widehat{x}_k(j), \ldots, \widehat{x}_{k-n+2}(j)\big)^T \quad \text{for} \quad j = 1, 2, \ldots, L \tag{19}$$

is the state estimate at time moment $k+1$,

$$\widehat{x}^{(j)}(k) = \big(\widehat{x}_k(j), \widehat{x}_{k-1}(j), \ldots, \widehat{x}_{k-n+1}(j)\big)^T \quad \text{for} \quad j = 1, 2, \ldots, L \tag{20}$$

is the state estimate at time moment $k$, $\mathrm{k}(k)$ is a time-varying filter gain, which is obtained using the respective formulas (Schick and Mitter, 1994) and is the same one for all the $L$ filter,

$$\psi_j\big(e_{k+1}(j)\big) = \begin{cases} -\Delta_j, & \text{if } e_{k+1}(j) < -\Delta_j \\ e_{k+1}(j), & \text{if } -\Delta_j \leqslant e_{k+1}(j) \leqslant \Delta_j \\ \Delta_j, & \text{if } e_{k+1}(j) > \Delta_j \end{cases} \quad \text{for} \quad j = 1, 2, \ldots, L \tag{21}$$

is the special function, which is used to transform the residual $e_k(j)\ \forall j = 1, 2, \ldots, L$ defined by

$$e_k(j) = u_k - c^T \widehat{x}_k(j) - d^T w_k \quad \text{for} \quad j = 1, 2, \ldots, L. \tag{22}$$

The filters in the bank (18) are different because of the threshold $\Delta_j\ \forall j = 1, 2, \ldots, L$ in (21) only, which has a different value $\Delta_1 < \Delta_2 < \ldots < \Delta_L$ for each Kalman filter.

In order to select a function to be minimized in the problem of a optimization of the state estimation it is important to determine a relation between the oft used but unknown filtering error and the characteristics, which are known beforehand or could be easily calculated. Further such a relation between the filtering error (to be more precise, the average square error of prediction of the state) and the variance of reconstructed input $\mu_k$ will be used according to Pupeikis (1998).

The problem to be solved now is at each time moment $k$ from a bank of $L$ current state estimates, generated by a bank of the $L$ Kalman filter (18)–(22) to choose an estimate that guarantees the minimal current filtering error.

The current values of variance $(\sigma^2_{\mu(i)})$ for $i = 1, 2, \ldots, L,\ k = 0, 1, \ldots$ at a time moment $k+1$ are obtained by the equation

$$\big(\sigma^2_{\mu(i)}\big)_{k+1} = \Big(1 - \frac{1}{k}\Big)\big(\sigma^2_{\mu(i)}\big)_k + \frac{1}{k}\mu^2_{k+1}(i) \quad \text{for } i = 1, 2, \ldots, L,\ k = 0, 1, \ldots \tag{23}$$

using their values $(\sigma^2_{\mu(i)})_k$ for $i = 1, 2, \ldots$ at a time moment $k$ and the values of reconstructed inputs $\mu_{k+1}(i)$ for $i = 1, 2, \ldots$ at a time moment $k+1$.

Then, we can rewrite (23) in such a form

$$\big(\sigma^2_{\mu(i)}\big)_{k+1} = \Big(1 - \frac{1}{k}\Big)\big(\sigma^2_{\mu(i)}\big)_k + \frac{1}{k}\big(\mu^2_{k+1} + 2\mu_{k+1}\Delta\mu_{k+1}(i) + \Delta\mu^2_{k+1}(i)\big) \quad \text{for} \quad i = 1, 2, \ldots,\ k = 0, 1, \ldots \tag{24}$$

if $\mu_{k+1}(i)\ \forall i = 1, 2, \ldots$ in (23) is replaced by

$$\mu_{k+1}(i) = \mu_{k+1} + \Delta\mu_{k+1}(i) \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots, \tag{25}$$

where $\mu_{k+1}$ is the value of input at a time moment $k + 1$.

On the other hand,

$$\mu_{k+1}(i) = W^{-1}\big(q^{-1}; \alpha\big)\widehat{x}_{k+1}(i) \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots, \tag{26}$$

or the same

$$\mu_{k+1}(i) = \mu_{k+1} + W^{-1}\big(q^{-1}; \alpha\big)\zeta_{k+1}(i) \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots, \tag{27}$$

as

$$\widehat{x}_{k+1}(i) = x_{k+1} + \zeta_{k+1}(i) \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots \tag{28}$$

is where $\zeta_{k+1}(i) = \zeta_{k+1}(i, \Delta_i)$ for $\quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots$, moreover,

$$\big|\zeta_{k+1}(i)\big| \leqslant |x_{k+1}| \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots, \tag{29}$$

and $\zeta_k(i) \sim \mathcal{N}(0, \sigma^2_{\zeta(i)})$ for $i = 1, 2, \ldots, L$. Then, using (25) and (27) it follows that

$$\Delta\mu_{k+1}(i) = W^{-1}(q^{-1}; \alpha)\zeta_{k+1}(i) \quad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots \tag{30}$$

and correspondingly

$$\begin{aligned}\big(x_{k+1} - \widehat{x}_{k+1}(i)\big)^2 &= \big(W(q^{-1}; \alpha)(\mu_{k+1} - \mu_{k+1}(i))\big)^2 \\ &= \big(W(q^{-1}; \alpha)(-\Delta\mu_{k+1}(i))\big)^2 = \zeta^2_{k+1}(i) \\ &\qquad\qquad \text{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots.\end{aligned} \tag{31}$$

A current deviation at a time moment $k + 1$ between the value of input and that of a reconstructed input, generated by the $l$-th filter, is

$$\big|\Delta\mu_{k+1}(l)\big| = 0, \quad k = 0, 1, \ldots, \tag{32}$$

when $\zeta_{k+1}(l) = 0\ \forall k = 0, 1, \ldots$, i.e., as $\widehat{x}_{k+1}(l) = x_{k+1}\ \forall k = 0, 1, \ldots$ and hence $\mu_{k+1}(l) = \mu_{k+1}\ \forall k = 0, 1, \ldots$

**Theorem 1.** *Assume that*

$$\big|\widehat{x}_{k+1}(i)\big| = \big|x_{k+1} + \zeta_{k+1}\big| \geqslant |x_{k+1}| \quad \textit{for} \quad i = 1, 2, \ldots, L,\ k = 0, 1, \ldots, \tag{33}$$

*while*

$$\widehat{x}_k(1) \neq \widehat{x}_k(2) \neq \ldots \neq \widehat{x}_k(L-1) \neq \widehat{x}_k(L), \quad \textit{for } k = 1, 2, \ldots, \tag{34}$$

$$\left(\sigma^2_{\mu(1)}\right)_k \neq \left(\sigma^2_{\mu(2)}\right)_k \neq \ldots \neq \left(\sigma^2_{\mu(L-1)}\right)_k \neq \left(\sigma^2_{\mu(L)}\right)_k \quad \textit{for} \quad k = 1, 2, \ldots, \tag{35}$$

*and in the bank of current state estimates* (34) *there exists* $\widehat{x}_k(l) = x_k$ *for* $k = 0, 1, \ldots$ *that equality* (32) *is valid. Then*

$$\left(\sigma^2_{\mu(l)}\right)_{k+1} < \left(\sigma^2_{\mu(j)}\right)_{k+1} \quad \textit{for} \quad j = 1, 2, \ldots, L-1, \ k = 0, 1, \ldots, \tag{36}$$

*and*

$$\left(x_{k+1} - \widehat{x}_{k+1}(l)\right)^2 < \left(x_{k+1} - \widehat{x}_{k+1}(j)\right)^2 \\ \textit{for} \quad j = 1, 2, \ldots, L-1, \ k = 0, 1, \ldots, \tag{37}$$

*if*

$$\left(\sigma^2_{\mu(1)}\right)_0 = \left(\sigma^2_{\mu(2)}\right)_0 = \ldots = \left(\sigma^2_{\mu(L-1)}\right)_0 = \left(\sigma^2_{\mu(L)}\right)_0 = c. \tag{38}$$

*Here* $c \geqslant 0$.

*Proof.* Let us assume that $\zeta_{k+1}(i) = \delta_i x_{k+1}$ for $i = 1, 2, \ldots, L, \ k = 0, 1, \ldots$, where on $k+1$ iteration $\delta_i = \delta_i(k+1; \Delta_i)$ $\forall i = 1, 2, \ldots, L, \ \delta_i > 0$ for $i = 1, 2, \ldots, L-1$, while $\delta_l = 0$, and moreover $x_{k+1} \neq 0$ $\forall k = 0, 1, \ldots$.

Then, from (33) it follows that

$$\left|\widehat{x}_{k+1}(i)\right| = \left|x_{k+1} + \delta_i x_{k+1}\right| \geqslant \left|x_{k+1}\right| \\ \text{for} \quad i = 1, 2, \ldots, L, \ k = 0, 1, \ldots, \tag{39}$$

and

$$\Delta\mu_{k+1}(i) = \delta_i W^{-1}(q^{-1}; \alpha) x_{k+1} = \delta_i \mu_{k+1} \\ \text{for} \quad i = 1, 2, \ldots, L, \ k = 0, 1, \ldots. \tag{40}$$

Thus, we can rewrite equation (24) in such a form

$$\left(\sigma^2_{\mu(i)}\right)_{k+1} = \left(1 - \frac{1}{k}\right)\left(\sigma^2_{\mu(i)}\right)_k + \frac{1}{k}\mu^2_{k+1}(1+\delta_i)^2 \\ \text{for} \quad i = 1, 2, \ldots, L, \ k = 0, 1, \ldots, \tag{41}$$

while the current square filtering error (31) is

$$\left(x_{k+1} - \widehat{x}_{k+1}(i)\right)^2 = \delta_i^2 x^2_{k+1} \quad \text{for} \quad i = 1, 2, \ldots, L, \ k = 0, 1, \ldots, \tag{42}$$

and for any $\delta_j > 0$ $j = 1, 2, \ldots, L-1$ (36) and (37) are valid.

REMARK 1. The minimal variance $(\sigma^2_{\mu(l)})_{k+1}$ $\forall k = 0, 1, \dots$ guarantees a minimal current filtering error (42) which is equal to zero therefore, that in (42) $\delta_l = 0$ corresponds to the optimal current state estimate $\widehat{x}_{k+1}(l) = x_{k+1}$ $\forall k = 0, 1, \dots$. On the other hand, by increasing $\delta_j$ $j = 1, 2, \dots, L-1$ increases the current variance (41) and current filtering error (42), respectively. Hence, it is possible to choose an optimal current state estimate if $\delta_j > 0$ for $j = 1, 2, \dots, L$ in (39).

**Theorem 2.** *Assume that*

$$\left|\widehat{x}_{k+1}(i)\right| = \left|x_{k+1} + \zeta_{k+1}(i)\right| \leqslant \left|x_{k+1}\right| \quad \text{for} \quad i = 1, 2, \dots, L,\ k = 0, 1, \dots, \tag{43}$$

*while*

$$\widehat{x}_k(1) \neq \widehat{x}_k(2) \neq \dots \neq \widehat{x}_k(L-1) \neq \widehat{x}_k(L) \quad \text{for} \quad k = 1, 2, \dots, \tag{44}$$

$$\left(\sigma^2_{\mu(1)}\right)_k \neq \left(\sigma^2_{\mu(2)}\right)_k \neq \dots \neq \left(\sigma^2_{\mu(L-1)}\right)_k \neq \left(\sigma^2_{\mu(L)}\right)_k \quad \text{for} \quad k = 1, 2, \dots, \tag{45}$$

*and in the bank of current state estimates* (44) *there exists such* $\widehat{x}_k(l) = x_k$ *for* $k = 0, 1, \dots$ *that equality* (32) *is valid. Then*

$$\left(\sigma^2_{\mu(l)}\right)_{k+1} > \left(\sigma^2_{\mu(j)}\right)_{k+1} \quad \text{for} \quad j = 1, 2, \dots, L-1,\ k = 0, 1, \dots, \tag{46}$$

*and*

$$\left(x_{k+1} - \widehat{x}_{k+1}(l)\right)^2 < \left(x_{k+1} - \widehat{x}_{k+1}(j)\right)^2 \quad \text{for } j = 1, 2, \dots, L-1,\ k = 0, 1, \dots, \tag{47}$$

*if*

$$\left(\sigma^2_{\mu(1)}\right)_0 = \left(\sigma^2_{\mu(2)}\right)_0 = \dots = \left(\sigma^2_{\mu(L-1)}\right)_0 = \left(\sigma^2_{\mu(L)}\right)_0 = c. \tag{48}$$

*Here* $c \geqslant 0$.

*Proof.* Suppose that $\zeta_{k+1}(i) = \delta_i x_{k+1}$ for $i = 1, 2, \dots, L,\ k = 0, 1, \dots,$ where $\delta_i = \delta_i(k+1; \Delta_i)$ $\forall i = 1, 2, \dots, L,\ \delta_i < 0$ for $i = 1, 2, \dots, L-1$, while $\delta_l = 0$ and moreover $x_{k+1} \neq 0$ $\forall k = 0, 1, \dots$.

It follows from (43) that

$$\left|\widehat{x}_{k+1}(i)\right| = \left|x_{k+1} + \delta_i x_{k+1}\right| \leqslant \left|x_{k+1}\right| \quad \text{for} \quad i = 1, 2, \dots, L,\ k = 0, 1, \dots, \tag{49}$$

and

$$\Delta\mu_{k+1}(i) = \delta_i W^{-1}(q^{-1}; \alpha) x_{k+1} = \delta_i \mu_{k+1} \quad \text{for} \quad i = 1, 2, \dots, L,\ k = 0, 1, \dots. \tag{50}$$

Thus, one can rewrite the equation (24) in a form (41) too, while a current filtering error is of the form (42) and for any $\delta_j < 0,\ j = 1, 2, \ldots, L-1$ (46) and (47) are valid.

REMARK 2. In such a case, the current variance $(\sigma^2_{\mu(l)})_{k+1}\ \forall k = 0, 1, \ldots$, which corresponds to the current optimal state estimate $\widehat{x}_{k+1}(l)\ \forall k = 0, 1, \ldots$ will not be minimal. Hence, it is impossible to choose the optimal current state estimate using inequalityes (36) and (37) because decreased the current variance (41) and increased the current filtering error, (42), respectively. On the other hand, when inequality (43) is satisfied, e.g., when during recursive calculations all the observations are damaged by the all Kalman filters (except one filter which generates the optimal state estimate), using relatively too small thresholds $\Delta_i\ \forall i = 1, 2, \ldots, L-1$ in Huber's $\psi$-function (21), then it is possible also to choose the optimal current state estimate, which corresponds to the maximal value of current variance (24) and to the minimal value of current filtering error (42), respectively.

The respective $L$ estimates of the variance $\sigma^2_\mu$ are obtained by the formula of the form

$$\begin{pmatrix} \sigma^2_{\mu(1)} \\ \sigma^2_{\mu(2)} \\ \vdots \\ \sigma^2_{\mu(L-1)} \\ \sigma^2_{\mu(L)} \end{pmatrix}_N = \frac{1}{N-1} \begin{pmatrix} \sum\limits_{i=1}^{N} (\mu_i(1) - \overline{\mu}(1))^2 \\ \sum\limits_{i=1}^{N} (\mu_i(2) - \overline{\mu}(2))^2 \\ \vdots \\ \sum\limits_{i=1}^{N} (\mu_i(L-1) - \overline{\mu}(L-1))^2 \\ \sum\limits_{i=1}^{N} (\mu_i(L) - \overline{\mu}(L))^2 \end{pmatrix}, \tag{51}$$

using, correspondingly, the $L$ reconstructed values $\mu_k(j)\ \forall j = 1, \ldots, L$ and $k = 1, 2, \ldots, N$. Here $\overline{\mu}(1), \ldots, \overline{\mu}(L)$ are the means of reconstructed inputs, respectively.

If the variance $\sigma^2_{\mu(j)}$ is minimal, i.e.,

$$\sigma^2_{\mu(l)} < \sigma^2_{\mu(j)} \quad \forall j = 1, 2, \ldots, L-1, \tag{52}$$

then the vector $\widehat{x}_N(l) = (\widehat{x}_1(l), \widehat{x}_2(l), \ldots, \widehat{x}_N(l))^T_N$ of estimates of the states $x_1, x_2, \ldots, x_N$ is an optimal one. The abovementioned equations let us to choose the optimal state estimates after processing all observations $u_1, \ldots u_N$ only. In order to choose at each time moment $k$ the current optimal state estimate, which would guarantee the current minimal value of reconstructed input and current minimal square filtering error, respectively, at these time moments, it is necessary to rewrite formula (51) in the recursive form.

### 3. Simulation Results

The unobserved noiseless sequence $\mathbf{X}_k$ (Fig. 1, entire line) is described by AR(1) model of the form

$$x_k = 0.85x_{k-1} + \mu_k, \quad k = \overline{1, 100}. \tag{53}$$

Then the output $\mathbf{U}_k$ to be observed in the presence of patchy outliers (Fig. 1, entire-dotted line) is

$$u_k = x_k + z_k, \quad k = \overline{1, 100}, \tag{54}$$

where $u_k$, $z_k$ are the values of an output $\mathbf{U}_k$ and a noise $\mathbf{Z}_k$, respectively, at a time moment $k$, $\mathbf{Z}_k$ is a sequence of independent identically distributed variables with an "$\varepsilon$-contaminated" distribution of the form (7) and the variance (8).

The sequence (54) is used for estimation of the values $x_i$, $i = \overline{1, 100}$, of the process $\mathbf{X}_k$. In this case, for an additive noise $\mathbf{Z}_k$

$$z_k = \begin{cases} 0, & \text{if } \zeta > \varepsilon, \\ \nu_k \sigma_\nu, & \text{if } \zeta < \varepsilon \end{cases} \tag{55}$$

holds, where $\nu_k, \zeta_k$ are independent Gaussian variables with zero means and variances 1. Here $\varepsilon = 0.2$.

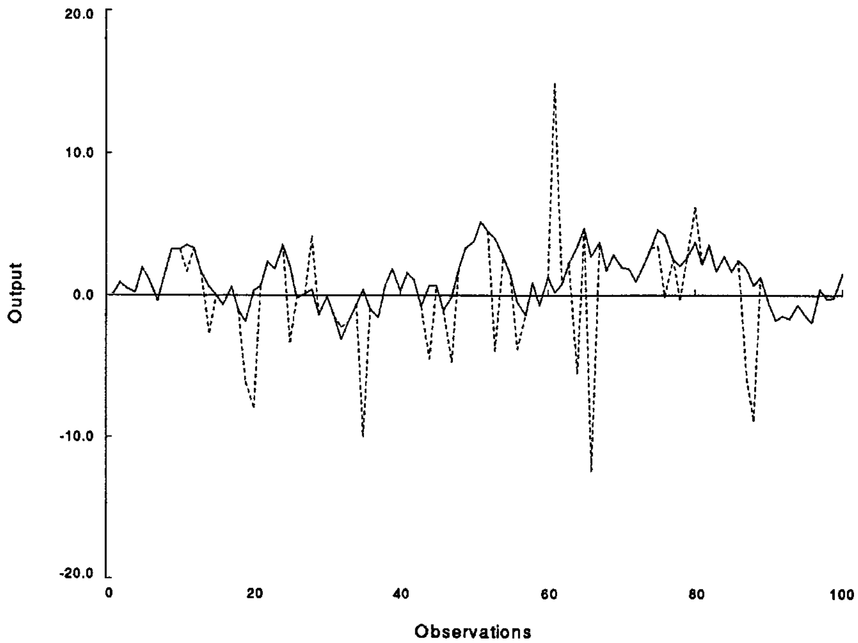


Fig. 1. Two time series in an absence (entire line) and a presence (entire-dotted line) of patchy outliers.

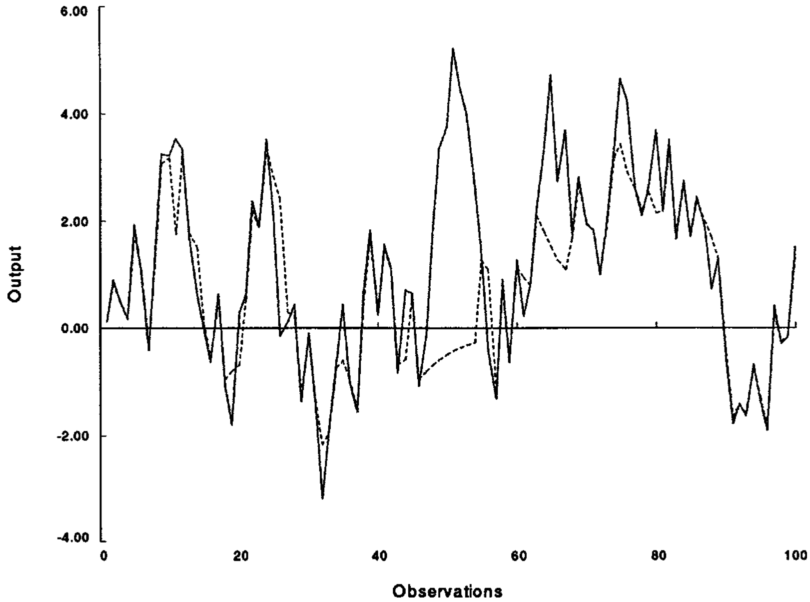


Fig. 2. Noiseless time series and its estimate, which is obtained using only one Kalman filter.

For the state estimation by processing $U_k = (u_1, \ldots, u_{100})^T$ the bank of the parallel Kalman filters (18)–(22) is used, which can be rewritten for the AR(1) process (53) with the a priori known parameter $a_1 = 0.85$ as

$$\widehat{x}^{(j)}(k+1) = a_1 \widehat{x}^{(j)}(k) + k(k+1)\psi_j \left(u_{k+1} - a_1 \widehat{x}^{(j)}(k)\right) \quad \text{for } j = 1, \ldots, L, \ \forall k = \overline{1, 100}, \tag{56}$$

where

$$\kappa_{k+1} = \phi_{k+1}(\sigma_z^2 + \phi_{k+1})^{-1}, \tag{57}$$

$$\phi_{k+1} = a_1^2 p_k + \sigma_\mu^2, \tag{58}$$

$$p_{k+1} = \phi_{k+1}(1 - \kappa_{k+1}), \tag{59}$$

$$\psi_j(e_{k+1}) = \begin{cases} \Delta_j, & \text{if } e_{k+1} < -\Delta_j, \\ e_{k+1}, & \text{if } -\Delta_j \leqslant e_{k+1} \leqslant \Delta_j, \\ \Delta_j, & \text{if } e_{k+1} > \Delta_j, \end{cases} \quad \text{for } j = \overline{1, L}, \tag{60}$$

moreover, $\Delta_1 = 1$; $\Delta_2 = 1.5$; $\Delta_3 = 2$; $\Delta_4 = 2.5$; $\Delta_5 = 2.75$; $\Delta_6 = 3$; $\Delta_7 = 3.5$; $\Delta_8 = 2000$; $L = 8$ and $p_0 = 0.1$; $\sigma_\mu^2 = 1$; $\sigma_z^2 = 1$; $\widehat{x}_0 = 0 \ \forall j = 1.8$.

In Figs. 2, 3, noiseless and really unobserved output (53) (entire line) and its two estimates (dotted lines) are presented. In Fig. 2 the first estimate is obtained using only one Kalman filter with $\Delta = 2.5$ in (60), while the second estimate in Fig. 3 is obtained using a bank of Kalman filters by optimizing the state estimation itself. In Fig. 4 the

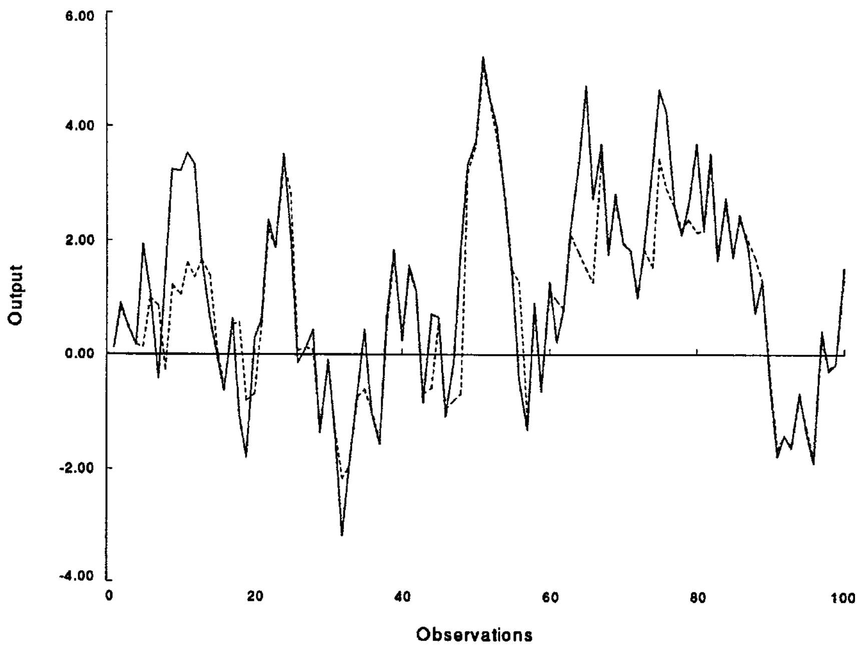


Fig. 3. Noiseless time series and its estimate, which is obtained using a bank of Kalman filters.

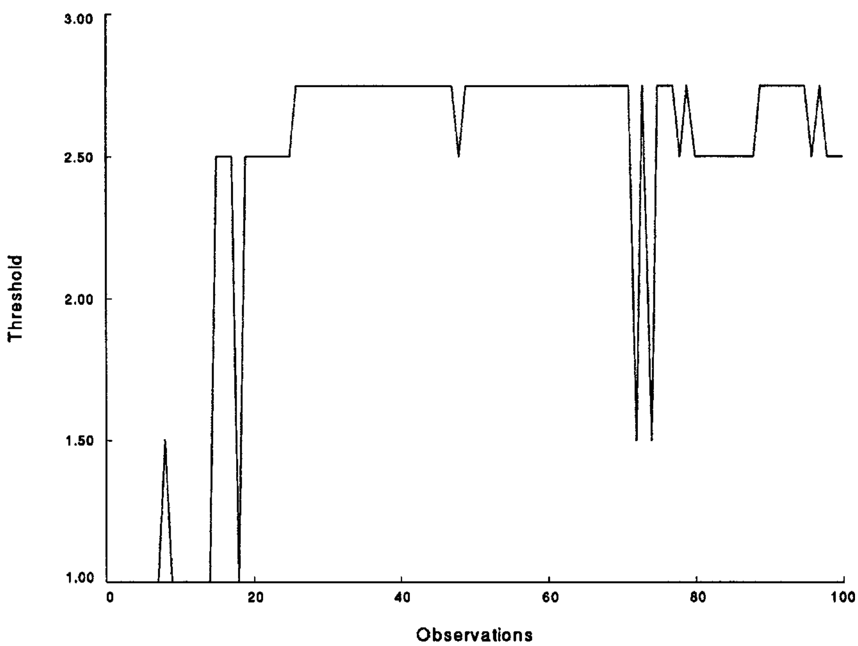


Fig. 4. Operation of the Kalman filters by processing observations with patchy outliers.

operation of the Kalman filters in time for the second estimate, respectively, is presented. From the simulation and state estimation results, presented in Figs. 2, 3, it follows, that the accuracy of state estimate, obtained using the bank of Kalman filters by optimizing the estimation itself, is higher as compared to the accuracy of such estimates based on constant $\Delta$. In the presence of patchy outliers the adaptive technique chooses 5 Kalman filters (Fig. 4) at different time moments.

## 4. Recursive State Estimation of a Chemical Process

The noiseless sequence $x_k$ is the time-series A from Box and Jenkins ( 1970), which is described by the ARMA model of the form

$$x_k^{(A)} = 1.45 + 0.92x_{k-1}^{(A)} + \mu_k - 0.58\mu_{k-1}, \quad k = \overline{1,197}, \tag{61}$$

or by the model of the form

$$\widetilde{x}_k^{(A)} = 0.92\widetilde{x}_{k-1}^{(A)} + \mu_k - 0.58\mu_{k-1}, \quad k = \overline{1,197}, \tag{62}$$

if the free term of sequence $x_k$ is eliminated.

Here $x_k^{(A)}$, $\widetilde{x}_k^{(A)}$ are the values of the abovementioned sequence at a time moment $k$.

Then the output $U_k$ to be observed in the presence of outliers (Fig. 5) is

$$\widetilde{u}_k^{(A)} = \widetilde{x}_k^{(A)} + z_k, \quad k = \overline{1,197}, \tag{63}$$

where $\widetilde{u}_k^{(A)}$, $z_k$ are the values of output $U_k$ and noise $Z_k$, $k = \overline{1,197}$, respectively, at a time moment $k$; $Z_k$ is a sequence of independent identically distributed variables with an $\varepsilon$-contaminated distribution of the form (7) with variance (8).

The sequence $\widetilde{u}_k^{(A)}$, $k = \overline{1,197}$, from (63) is used for the state estimation of the ARMA process (61) in a case of a priori known parameters $a_1 = 0.92$, $b_1 = -0.58$. In this case, for an additive noise $Z_k$, $k = \overline{1,197}$, the equality of the form is valid

$$z_k = \begin{cases} 0, & \text{if } \zeta_k > \varepsilon_k, \\ 10\nu_k, & \text{if } \zeta_k < \varepsilon_k, \end{cases} \tag{64}$$

where $\nu_k$, $\zeta_k$ are independent Gaussian variables with zero means and variances 1; $\varepsilon_k$ is a time varying "contamination" fraction of the form

$$\varepsilon_k = \begin{cases} 0.1 & \text{for } k = 1, 2, \ldots, 100, \\ 0.2 & \text{for } k = 101, 102, \ldots, 150, \\ 0.05 & \text{for } k = 151, 152, \ldots, 197. \end{cases} \tag{65}$$

By processing $\widetilde{U}_k^{(A)} = (\widetilde{u}_1^{(A)}, \ldots, \widetilde{u}_{197}^{(A)})^T$ the bank of parallel Kalman filter (18)–(22) is used for the state estimation. Moreover,

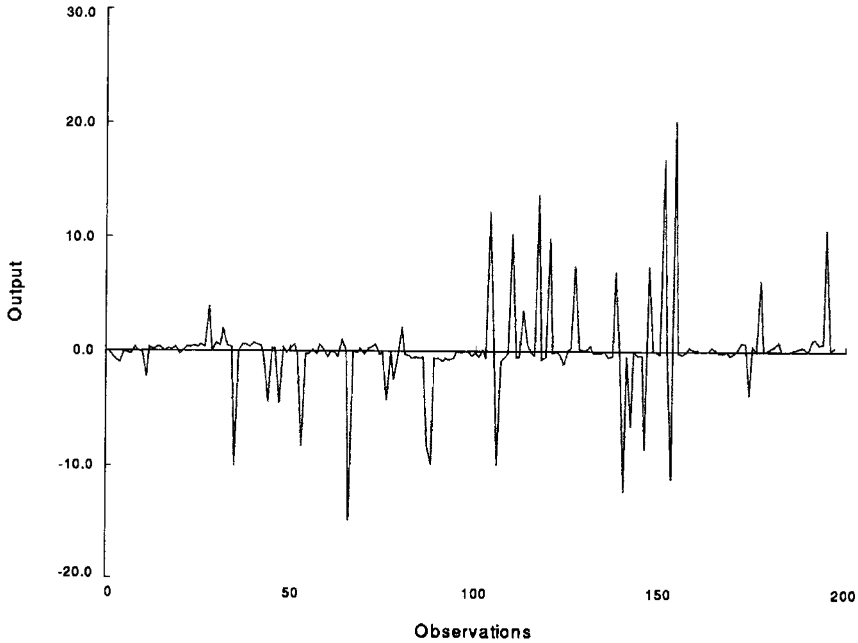


Fig. 5. The output of the chemical process (62) in the presence of patchy outliers.

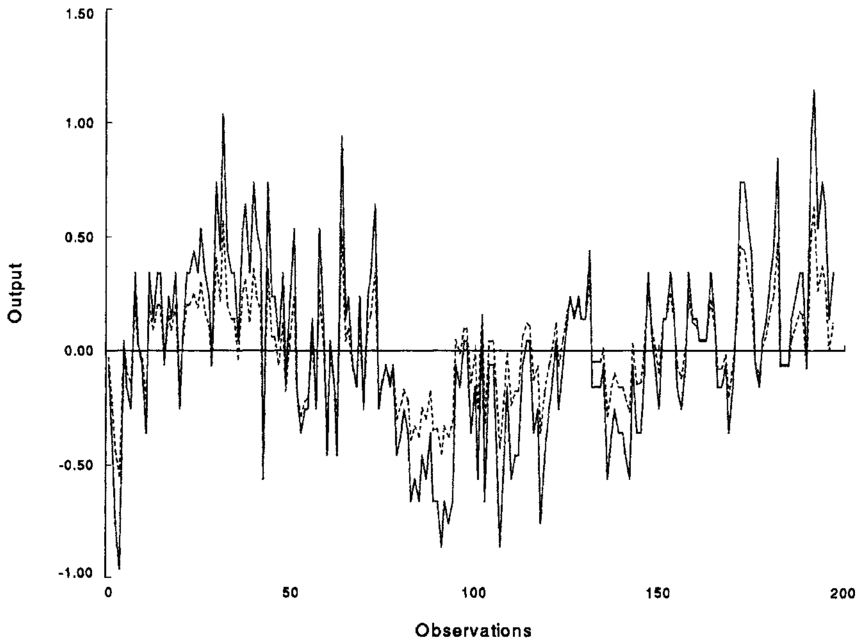


Fig. 6. Noiseless time series A and its estimate (dotted line), which is obtained using a bank of Kalman filters (18)–(22).

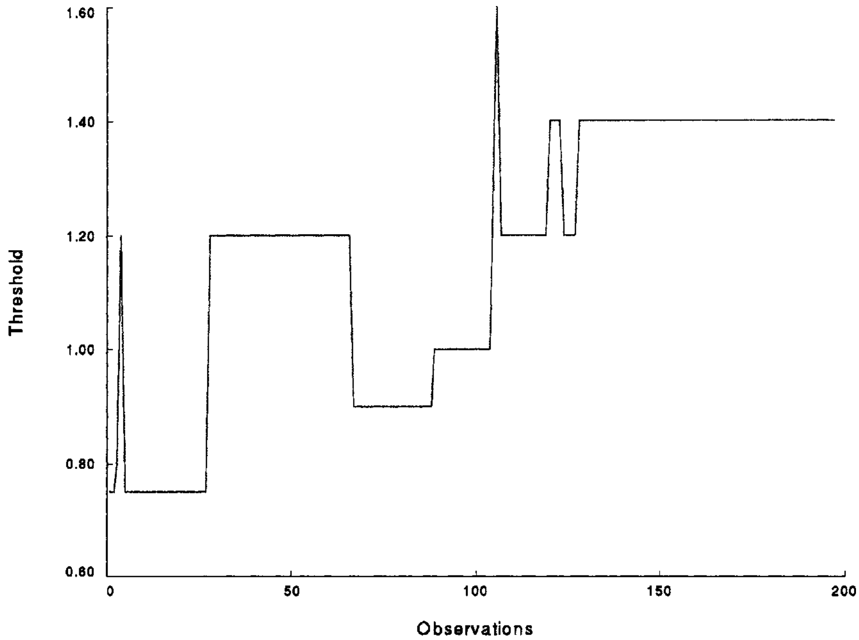


Fig. 7. Operation of the Kalman filters (18)–(22) by processing observations of the chemical process (62) with time-varying and patchy outliers.

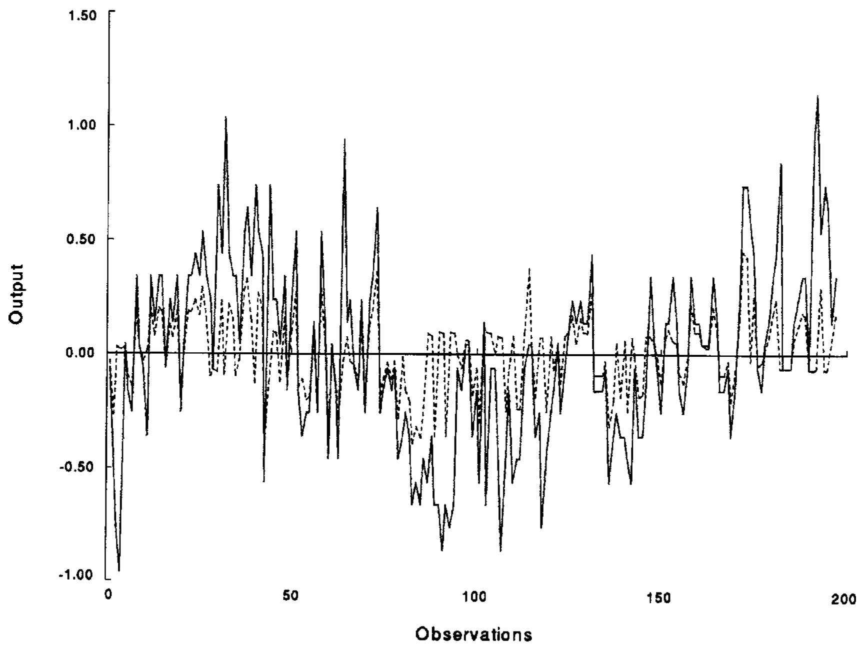


Fig. 8. Noiseless time series A and its estimate (dotted line), which is obtained using only one Kalman filter with $\Delta = 0.75$ in (21).

$\Delta_1 = 0.75$, $\Delta_2 = 0.8$, $\Delta_3 = 0.9$, $\Delta_4 = 1$, $\Delta_5 = 1.2$, $\Delta_6 = 1.4$, $\Delta_7 = 1.6$, $\Delta_8 = 1.8$, $\Delta_9 = 2$, $\Delta_{10} = 2000$ and $p_0 = 0.1$; $\sigma_\mu^2 = 1$; $\sigma_z^2 = 1$; $\widehat{x}_0(i) = 0\ \forall i = \overline{1, 10}$, $L = 10$ in (18)–(22).

In Fig. 6, a noiseless and really unobserved output (62) and its estimate are presented. The estimate (dotted line) is obtained using the bank of Kalman filters (18)–(22) by optimizing the state estimation itself. In Fig. 7, the operation of Kalman filters in time for such a current state estimate is presented. The results presented in Fig. 7 show that, in such a case, at different time moments the adaptive technique chooses seven Kalman filters with different thresholds $\Delta$ in (21). Fig. 8 shows us the unobserved output (62) and the estimate (dotted line), which is obtained using only one Kalman filter with the constant $\Delta = 0.75$ in (21). From the simulation and the state estimation results, presented in Figs. 5–8 it follows that the accuracy of the state estimates, obtained using the bank of Kalman filters by recursive optimizing of the estimation itself, is higher as compared to the accuracy of such estimates based on the constant $\Delta = 0.75$ (Fig. 8). It could be mentioned that the results averaged by 20 experiments in the presence of time varying outliers including the patchy outliers are given in (Pupeikis and Huber, 1997).

## 5. Conclusions

The classical robust Hubers theory of estimation of a location parameter uses stochastic models with time-homogeneous contamination of outliers (Huber, 1964; 1981). However, if the various robust recursive algorithms turn out to be efficient in the presence of rare and isolated outliers, then always there arise special problems in the presence of patchy outliers. In such a case, it is important to solve the generalized problem of the model of outliers, which are varying in time. Therefore, in our work, a recursive state estimation approach with an optimization of the estimation itself is applied. In theory (Theorems 1–2), it is based on a simple but important relation between the variance of a reconstructed input of the process and that of the filtering error as well as on the fact that both variables take their minimum at the same place. In practice, our approach is realized by means of a bank of parallel Kalman filters (18)–(22), consisting of simple recursive equations, which differ one from another by threshold $\Delta$ in Hubers $\psi$-function (21) only. At each recursive step a current state estimate, which guarantees the minimal filtering error, is chosen from the respective bank of current state estimates by means of optimization technique, using equations (51) and condition (52). The results of numerical simulation (Figs. 1–4) and the state estimation using a practical chemical time-series (Figs. 5–8), described by (61), (63), prove the efficiency of the model of time varying outliers and usefulness of the proposed recursive approach for the state estimation in the presence of patchy outliers in observations to be processed.

## 6. Acknowledgment

The author is grateful to Professor Peter J. Huber for his helpful comments.

## References

Box, G.E.P., and G.M. Jenkins (1970). *Time Series Analysis. Forecasting and Control*. Hoden-Day.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **35**, 73–101.

Huber, P.J. (1981). *Robust Statistics*. Wiley, New York.

Pupeikis, R., and P.J. Huber (1997). *Adaptiv-robust Zustandsschätzung mit Optimierung des Schätzprozesses. Bericht zur Projekt Echtzeit-Optimierung größer Systeme*.

Pupeikis, R. (1998). State estimation of dynamic systems in the presence of time-varying outliers in observations. *Informatica*, **9**(3), 325–342.

Schick, I.C., and S.K. Mitter (1994). Robust recursive estimation in the presence of heavy-tailed observation noise. *Ann. Math. Statist.*, **22**(2), 1045–1080.

**R. Pupeikis** received Ph.D. degree from the Kaunas Polytechnic Institute, Kaunas, Lithuania, 1979. He is a senior researcher at the Process Recognition Department of the Institute of Mathematics and Informatics. His research interest include the classical and robust approaches of dynamic system identification as well technological process control.

## Rekurentinis dinaminių procesų būvio įvertinimo optimizavimas, esant nestacionariam didelių impulsų srautui stebėjimuose

Rimantas PUPEIKIS

Straipsnyje nagrinėjamas proceso, aprašomo autoregresijos – slenkančio vidurkio (ARMA) modeliu (1)–(8), būsenų įvertinimo uždavinys, taikant lygiagrečių Kalmano filtrų banką. Įrodytos teoremos teigia, kad esant pakankamai bendroms sąlygoms mažėjant filtravimo paklaidų reikšmėms atitinkamai mažėja atkurtų įėjimų dispersijų reikšmės, kurias nesunku suskaičiuoti rekurentiškai.