Nonlinear Dynamics and Systems Theory, **18** (4) (2018) 372–391

Bayesian Approach for Multi-Mode Kalman Filter for Abnormal Estimation

Ali Hussein Hasan [1*] Alia Muhsin Mnati [2] and Ali Obied [1]

[1] *University of Sumer, College of Computer Science and Information Technology*
[2] *University of Thiqar, College of Engineering*

Received: April 5, 2018; Revised: October 14, 2018

Abstract: The paper deals with Bayesian approach for multi-mode Kalman filter estimation for the states $x(k)$ from the set of successive observations $Y_k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations is driven. Abnormal operations may be related to fault in one of system components; sudden internal thermal noise or even missing the input signal and can be extended to the maneuver target tracking case. Whenever the abnormal operation is detected, we can start tracking the states in this mode of operation. So the main problem may be reformulated to be detection of the starting point of the abnormal operation. The numerical simulation for fault estimation of phosphor furnace in different conditions are used to show the effectiveness of the proposed approach.

Keywords: *Bayesian estimation; multi mode operation; interactive multiple model; Kalman filter fault estimation.*

Mathematics Subject Classification (2010): 93E10.

1 Introduction

Fault detection and isolation problems have many significant applications during the past three decades, such as parity space Eigen structure assignment, H_∞ filtering, H_∞ optimization, and unknown input observer [1]. It is known that multiple model systems (known as hybrid systems) are an important class of combination filtering, which are mostly used in many practical engineering and industrial fields such as maneuver target tracking systems and fault detection systems, ets. In general, the multiple model systems combine hierarchically discrete or continuous state spaces, and each state (which is called the mode) is associated with a dynamic process [2].

* Corresponding author: mailto:ali.husain@uos.edu.iq

Many algorithms have been proposed for solving the problem of hybrid systems, such as the generalized pseudo Bays [3], interacting multiple model [4], expectation propagation algorithm [5]. These algorithms estimate the states with low computational costs by approximating of the posterior state distributions in finite mixture models. The main drawback of these algorithms is that, whenever the system contains nonlinear or non-Gaussian modes, the mixture models can not approximate the distribution accurately, that leads to the estimation fail. The interactive multiple model (IMM) algorithm, see [4] and [9], may be considered one of the most important approaches for solving the switching systems with the Gaussian linear states. It is basic operation by applying the Kalman filter for each estimation mode under consideration that this mode is a correct one at this operational instant. Then, the weighted combination of state estimations by all filters is calculated to produce the final Gaussian mean and covariance [7]. This mixed estimation is used for the next estimation. The weights are calculated according to the probabilities of the models.

In this paper, a Bayesian approach for multi-mode Kalman filter estimation for the states $x(k)$ from the set of successive observations $Y_k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations is developed. The rest of the paper is organized as follows: Statement of the problem is described in Section 2. Multi-mode Kalman filter approach is described in Section 3. Proposed multi model estimation is described in Section 4. Simulation Results is described in Section 5. Conclusion is described in Section 6.

2 Statement of The Problem

Consider the classical discrete-time state space problems in case of linear model of the form:

$$x(k+1) = F(k+1,k)x(k) + BU(k) + \Gamma(k)v(k), \tag{1}$$

$$y(k) = Hx(k) + w(k), \tag{2}$$

where $x(k)$ is the discrete time state, F is a state transition matrix, B is an input matrix, the vector $U(k)$ is the input and assumed to be known at perspective times, $y(k)$ is a vector including the measurements at time k, H is the associated observation matrix, $v(k)$ is a state noise process (or an input noise), and $w(k)$ is the measurement noise; both sequences $w(k)$ and $v(k)$ are assumed to be uncorrelated with the white Gaussian noise sequence with zero means and the covariance matrices $Q(k)$ and $R(k)$, and $v(k){\sim}N(0,Q(k)), w(k){\sim}N(0,R(k))$, $[cov(v(k),v(j)) = E[v(k)v(j)^T]\delta(k,j) = Q_k\delta(k,j)]$, $\delta(k,j)$ is the Kronecker delta, $cov\,(w(k),w(j))] = E\left[w(k)w(j)^T\right]\delta(k,j) = R_k\delta(k,j)$.

It is also assumed that w and v are uncorrelated and $cov\,(w(k),v(k)) = 0$ for all k,j. The initial value of x represents a random variable with an average $\mu_x(0)$ and variance $V_x(0)$ such that $E\,[x(0)] = \mu_x(0)$ and $var\,[x(0)] = V_x(0)$. Assume that the observation noise $w(k)$ is uncorrelated with the system discrete-time state $x(k)$, such that $E\left[x(k)w(k)^T\right] = 0\ \forall\ k \geq 0$.

Now our problem is to estimate the states $x(k)$ from the set of successive observations $Y^k = \{y(1)y(2)\dots y(k)\}$ in normal and abnormal operations. Abnormal operations are related to the fault in one of the system components; sudden internal thermal noise or even missing the input signal and can be extended to the maneuver target tracking case. Whenever the abnormal operation is detected, we can start tracking the states in this mode of operation. So the main problem may be reformulated to be: detection of the starting point of the abnormal operation.

3 Multi-Mode Kalman Filter Approach

Let $\hat{x}(k/k)$ be the state estimation, then the estimation error will be:

$$\tilde{x} = x(k) - \hat{x}(k/k), \tag{3}$$

The estimation will be conditionally and unconditionally unbiased, such that

$$E\left\{\hat{x}(k/k)\,|y(k)\right\} = E\left\{x(k)|y(k)\right\}, \tag{4}$$

which leads to

$$E\left\{\hat{x}(k/k)\right\} = E\left\{x(k)\right\}, \tag{5}$$

which is a linear function of the observations $y(k)$. According to the linear unbiased estimation algorithms, we choose only the one that gives the minimal variance of error, i.e., $V_{\tilde{x}}(k/k) = var\left\{\tilde{x}(k/k)\right\}$ or simply $var\left\{\tilde{x}(k/k)|y(k)\right\}$ is as minimum as possible. For a given set of observations Y^k, the estimate based on the minimum of the mean-square error coincides with the conditional mean value of x, which is based on linear Kalman filter process.
1- Extrapolation process:

$$\widehat{x}(k/k-1) = F(k,k-1)\widehat{x}(k-1/k-1). \tag{6}$$

2- Estimation:

$$\widehat{x}(k/k) = \widehat{x}(k/k-1) + K(k)\left[y(k) - H(k)\widehat{x}(k/k-1)\right]. \tag{7}$$

3- Coefficient gain:

$$K(k) = P(k/k-1)H^T(k)\left[H(k)P(k/k-1)H^T(k) + R(k)\right]^{-1}. \tag{8}$$

4- Covariance extrapolation:

$$P(k+1/k) = F(k+1,k)P(k/k)F^T(k+1,k) + Q(k). \tag{9}$$

5- Covariance filtration:

$$P(k/k) = \left[1 - K(k)H(k)\right]P(k/k-1). \tag{10}$$

Definition 3.1 Let the innovation process noise $\vartheta(k)$ be [8]

$$\vartheta(k) = y(k) - H(k)\widehat{x}(k/k-1). \tag{11}$$

Assume the innovation process noise $\vartheta(k)$ is the white Gaussian noise with zero mean expectation in the normal operation and, also, the white Gaussian noise but with non-zero mean expectation in the abnormal operation:

$$E\left[\vartheta(k)\right] = \begin{cases} 0, \ \textit{at the normal operation}, \\ g(k-k_0,\alpha), \ \textit{at the abnormal operation}, \end{cases} \tag{12}$$

where g is a deterministic function, $k_0 + 1$ is the starting of abnormal operation and α is an intensity vector.

Definition 3.2 Let us define the covariance of the additional intensity vector α as

$$s(k) = H(k)P(k/k-1)H^T(k) + R(k). \tag{13}$$

Definition 3.3 Let the density of the probability distribution of the innovation process noise $\vartheta(k)$ be $w(\vartheta)$:

$$w(\vartheta) = \frac{1}{\sqrt{(2\pi)^n \det(s)}} e\left(\frac{-1}{2}\vartheta^T s^{-1}\vartheta\right). \tag{14}$$

The following equation is a Riccati equation

$$\begin{aligned} &P(k+1/k) = F(k+1,k) \\ &\left[P(k/k-1) - P(k/k-1)H^T(k)s(k)^{-1}H(k)P(k/k-1)\right] F(k+1/k)^T + Q(k). \end{aligned} \tag{15}$$

Note: The Riccati equation will converge to the constant matrix P as $k \to \infty$ in case of the system we are dealing with time invariant. As a result, the gain coefficient K converges to a constant small value and estimates of the parameters practically independent of the observed data. The result, is that, for small intensity α, will not be taken into account in any way.

The problem of detection the starting point of the abnormal operation k_o and estimation of its intensity vector α is a problem of detection a useful signal $g(k-k_o,\alpha)$ from the white noise and estimation of its parameters. It is known that, when solving the problem of detection, various optimality criteria (Bayesian or non-Bayesian) lead to a general decisive rule - the formation of the likelihood ratio and comparison with the threshold. The difference lies in the choice of the detection threshold. Here we assume that the abnormal operation intensity is a nonrandom process. To synthesize the meter of the intensity of the abnormal operation, it is convenient to use the maximum likelihood criterion. In this subsection we consider a general approach to the synthesis of the algorithm for simultaneous detection of the moment of the beginning and estimation of the abnormal operation. Let

$$U(k) = \begin{cases} u(k), & \text{at normal operation,} \\ u(k)+\alpha, & \text{at abnormal operation.} \end{cases} \tag{16}$$

It is assumed that the system starts the abnormal operation in the time between the moments k_0 and k_0+1, therefore the time moment, which is considered the beginning of the abnormal operation, is k_0+1. Let $\vartheta^0(k)$ and $\vartheta^1(k)$ be the updating processes corresponding to the absence and presence of an abnormal operation, then the problem of finding an abnormal operation consists of choosing one of the two alternative hypotheses.

$$H_0 : \vartheta(k) = \vartheta^0(k) \textit{ at the normal operation}, \tag{17}$$

$$H_1 : \vartheta(k) = \vartheta^1(k) = \vartheta^0(k) + g(k-k_0,\alpha) \textit{ at the abnormal operation}, \tag{18}$$

where $g(k-k_0,\alpha)$ is a useful signal that should be detected, whenever the abnormal operation is started, $g(k-k_0,\alpha)$ is introduced as an offset or bias in the innovation process at the time $k-k_0$. Since $\vartheta^0(k)$ is a random process with the white Gaussian noise with zero expectation and covariance and $\vartheta^1(k)$ is the mixing of the useful signal and the white noise $\vartheta^0(k)$, with the mathematical expectation $g(k-k_0,\alpha)$.

Theorem 3.1 *The task is detecting a vector of deterministic signal $g(k-k_0,\alpha)$ with unknown parameters (the intensity α and the moment of the onset of the abnormal operation k_0+1) from a background of the white noise $\vartheta^0(k)$.*

Proof. The optimal procedure for detection the of intensity vector α may be reduced to the formation of the likelihood ratio and compared with a predetermined threshold value λ,

$$L(k) = \frac{p\left[\vartheta(k-m+1),...,\vartheta(k)|H_1\right]}{p\left[\vartheta(k-m+1),...,\vartheta(k)|H_0\right]} \geq \lambda, \tag{19}$$

where $m = k - k_0 = 1...M$, M is the number of samples. Assume that the samples of the white Gaussian noises are statistically independent, then we have

$$p\left[\vartheta(k-m+1),...,\vartheta(k)|H_0\right] = \prod_{n=k-m+1}^{k} \frac{1}{\sqrt{(2\pi)^l \det(s(n))}} e\left(\frac{-1}{2}\vartheta^T(n)s^{-1}\vartheta(n)\right), \tag{20}$$

and

$$p\left[\vartheta(k-m+1),...,\vartheta(k)|H_1\right] = \prod_{n=k-m+1}^{k} \frac{1}{\sqrt{(2\pi)^l \det(s(n))}} e\left(\frac{-1}{2}\left[\vartheta(n) - g(n-k_0,\alpha)\right]^T s^{-1}(n)\left[\vartheta(n) - g(n-k_0,\alpha)\right]\right), \tag{21}$$

where l is the order of the vector $\vartheta(k)$ and k is the current instant of time. From the equations (19)–(21) after algebraic transformations, taking into account that s is a symmetric matrix, ϑ and g are columns vectors,

$$\vartheta^T s^{-1} g = g^T s^{-1} \vartheta, \tag{22}$$

we get

$$L(k) = e\left\{\sum_{n=k-m+1}^{k}\left[g^T(n-k_0,\alpha)s^{-1}(n)\vartheta(n) - \frac{1}{2}g^T(n-k_0,\alpha)s^{-1}(n)g(n-k_0,\alpha)\right]\right\}, \tag{23}$$

Taking the logarithm of both sides of (23), we get

$$\ln L(k) = \sum_{n=k-m+1}^{k}\left[g^T(n-k_0,\alpha)s^{-1}(n)\vartheta(n) - \frac{1}{2}g^T(n-k_0,\alpha)s^{-1}(n)g(n-k_0,\alpha)\right], \tag{24}$$

□

Let $g(m,\alpha) = \vartheta^1(k) - \vartheta^0(k) = \vartheta^1(k_0+m) - \vartheta^0(k_0+m)$. Then $\vartheta^1(k_0+m)$ is the updating process at the time $k = k_0+m$ under the condition of correspondence between the real state (presence of the abnormal operation) of the system and the model of system used by the abnormal operation (model equation with $U(k) = u(k)+\alpha$), and $\vartheta^0(k_0+m)$ is the updating process at time $k = k_0+m$ under the condition of a discrepancy between the real state (the presence of the abnormal operation) of the system and the model of system used by the abnormal operation (model equation 1 with $U(k) = u(k)$). Then

$$\begin{aligned} \widehat{g}\ (m,\alpha) &= \left[y(k_0+m) - H\,\widehat{x}^{1}\ (k_0+m/k_0+m-1)\right] \\ &\quad - \left[y(k_0+m) - H\,\widehat{x}^{0}\ (k_0+m/k_0+m-1)\right] \\ &= H\left[\widehat{x}^{0}\ (k_0+m/k_0+m-1) - \widehat{x}^{1}\ (k_0+m/k_0+m-1)\right] \\ &= HF\left[\widehat{x}^{0}\ (k_0+m/k_0+m-1) - \widehat{x}^{1}\ (k_0+m/k_0+m-1)\right]. \end{aligned} \tag{25}$$

Proof: The state estimation equation (1) at the normal and abnormal operation will be:

1. At time step 1

$$\begin{aligned} \widehat{x}^{0}\ (k_0+1/k_0+1) = {} & [F - K(k_0+1)HF]\ \widehat{x}\ (k_0/k_0) \\ & + K(k_0+1)y(k_0+1) + [B - K(k_0+1)HB]\,u(k), \end{aligned} \tag{26}$$

2. At time step 2

$$\begin{aligned} &\widehat{x}^{0}\ (k_0+2/k_0+2) = \\ &[F-K(k_0+2)HF]\,[F-K(k_0+1)HF]\ \widehat{x}\ (k_0/k_0)\,[F-K(k_0+2)HF]\,K(k_0+1)y(k_0+1) \\ &\quad + K(k_0+2)y(k_0+2)\,\{[F-K(k_0+2)HF]\,[B-K(k_0+1)HB] + [B-K(k_0+2)HB]\} \\ &u(k), \end{aligned} \tag{27}$$

3. At time step m

$$\begin{aligned} \widehat{x}^{0}\ (k_0+m/k_0+m) &= \prod_{i=0}^{m-1} [F - K(k_0+m-i)HF]\ \widehat{x}\ (k_0/k_0) \\ &+ \sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F - K(k_0+m-i)HF]\right] K(k_0+m-j-1)y(k_0+m-j-1)\right\} \\ &+ K(k_0+m)y(k_0+m) \\ &+ \sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F - K(k_0+m-i)HF]\right][B - K(k_0+m-j-1)HB]\right\} \\ &+ [B - K(k_0+m)HB]\,u(k), \end{aligned} \tag{28}$$

from the above equations (25)–(28) and after the substitution $m = k - k_0$, we obtain the expression for the "useful signal" $g(m,\alpha)$ in the following form

$$\begin{aligned} g(m,\alpha) &= HF\left\{\sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F - K(k-i)HF]\right][B - K(k-j-1)HB]\right\}\right. \\ &\qquad\qquad \left. + [B - K(k)HB]\vphantom{\sum_{j=0}^{m-2}}\right\} u(k) \\ &= G(m)u(k), \end{aligned} \tag{29}$$

where

$$G(m) = HF...$$
$$\left\{\sum_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F-K(k-i)HF]\right][B-K(k-j-1)HB]\right\}+[B-K(k)HB]\right\}, \tag{30}$$

$m \geq 1$. When $m = 1$, the term $\sum\limits_{j=0}^{m-2}\left\{\left[\prod_{i=0}^{j}[F-K(k-i)HF]\right][B-K(k-j-1)HB]\right\}$ will be equal to zero, and the matrix $G(m)$ becomes square matrix with dimensions $l \times l$, l is the order of the observation vector. Substituting equation (29) into (24), we obtain

$$\ln L(k,\alpha) = \sum_{n=k-m+1}^{k}\left[\alpha^T G^T(n-k_0)s^{-1}(n)\vartheta(n) - \frac{1}{2}\alpha^T G^T(n-k_0)s^{-1}(n)G(n-k_0)\alpha\right], \tag{31}$$

then

$$\begin{aligned}\ln L(k,\alpha) = {} & \alpha^T\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)\vartheta(n)\right]\right] \\ & - \frac{1}{2}\alpha^T\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]\alpha.\end{aligned} \tag{32}$$

From equation (32), by the criterion of maximum likelihood ratio, we find the estimate of the intensity vector,

$$\begin{aligned}\frac{\partial \ln L(k,\alpha)}{\partial \alpha} &= \sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)\vartheta(n)\right] \\ &\quad - \left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]\alpha \\ &= 0.\end{aligned} \tag{33}$$

Then

$$\widehat{\alpha}_m = \left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}\left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)\vartheta(n)\right]\right]. \tag{34}$$

Taking the second order derivative, we get

$$\frac{\partial^2 \ln L(k,\alpha)}{\partial\alpha\partial\alpha^T} = \left[\sum_{n=k-m+1}^{k}\left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right] \prec 0. \tag{35}$$

The function $\ln L(k)$ will reach its maximum value at the point $\alpha = \widehat{\alpha}_m$.

Theorem 3.2 *$var\{Xy\} = X[var\{y\}]X^T$.*

Proof. Let X be an m $\times$ n matrix and y be a n $\times$1 random vector. Then

$$var\{Xy\} = E\left[(Xy - XE(y))(Xy - XE(y))^T\right] = E\left[X(y - E(y))(y - E(y))^T X^T, \right]$$
$$= XE\left[(y - E(y))(y - E(y))^T\right] X^T = X\left[var\{y\}\right] X^T.$$

□

If all values are scaled by a constant, the variance is scaled by the square of that constant. Then, the covariance matrix of the estimation is

$$var\left\{\widehat{\alpha}_m\right\} = \left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]^{-1} \cdots$$
$$\left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)var\{\vartheta(n)\}\left[G^T(n-k_0)s^{-1}\right]^T\right]\right] \times \cdots \tag{36}$$
$$\left[\left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}\right]^T.$$

Let the covariance matrix of the innovation process be $var\{\vartheta(n)\} = s(n)$. Since $s^{-1}(n)$ is a symmetrical matrix, we have that

$$\begin{aligned}&\left[\left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1}\right]^T \\ &= \left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)G(n-k_0)\right]\right]^{-1},\end{aligned} \tag{37}$$

yields the covariance matrix of this estimate as follows

$$var\left\{\widehat{\alpha}_m\right\} = \left[\sum_{n=k-m+1}^{k} \left[G^T(n-k_0)s^{-1}(n)G^T(n-k_0)\right]\right]^{-1}, \tag{38}$$

Substituting equation (33) into (32), we get

$$\ln L(k, \widehat{\alpha}_m) = \frac{1}{2} \sum_{n=k-m+1}^{k} \vartheta^T s^{-1} G \left[\sum_{m=k-m+1}^{k} G^T s^{-1} G\right]^{-1} \sum_{m=k-m+1}^{k} G^T s^{-1} \vartheta \geq \tilde{\lambda}_m, \tag{39}$$

It is obvious that for every value of m, there exists an estimate $\widehat{\alpha}_m$. Therefore, the simultaneous detection of the start of the abnormal operation and estimating the abnormal operation is the abnormal operation in multichannel (M channels). Therefore the optimal procedure for detecting the moment of the beginning of the abnormal operation has the form

$$\max_{m=1,M} \ln L(k, \widehat{\alpha}_m) \geq \tilde{\lambda}_m, \tag{40}$$

and the moment of the beginning of the abnormal operation is the difference$(k-m+1)$, where m is the channel index, in which $\ln L(k, \widehat{\alpha}_m)$takes the maximum value, and k is the current time moment.

It follows that (equations (39)–(40)) the optimal procedure for detecting the moment of the onset of the abnormal operation is reduced to linear accumulation of the values of the quadratic forms (equation (39)) of the innovation process in m adjacent cycles from (k_0+1) to $k=(k_0+m)$th moments of time. Then, the maximum value is selected from the set of M different values and is compared with a given threshold $\tilde{\lambda}_m$.

Note: The different channels accumulate different values, so to stabilize the probability of false alarm, it is necessary to compare the maximum value of the likelihood ratio (equation (40)) with different thresholds $\tilde{\lambda}_m$, chosen from the given probability of false alarm P_{FA}. Using the decision rule (equation (40)), we can detect and evaluate, both the abnormal operation and the moment of the beginning of the abnormal operation.
To calculate the probability of a false alarm, it is necessary to know the law of distribution of the quadratic form $\ln L(k, \widehat{\alpha}_m)$ in the absence of an abnormal operation. It is known that [8] if $\vartheta(n)$ is an l -dimensional vector with independent normally distributed components, each of which has the variance $\sigma_{iq}^2(q=1,2,...,l)$ and the mathematical expectation $E[\vartheta(n)]=0$, then the probability distribution density of the quadratic form $\ln L(k, \widehat{\alpha}_m)$ is the central χ^2-distribution with $l\times m$ degrees of freedom. The corresponding density of the probability distribution is written in the form

$$\chi^2\left[\ln L(k, \widehat{\alpha}_m)\right] = \left[2^{\frac{lm}{2}}\Gamma(\frac{lm}{2})\right]^{-1}\left[\ln L(k, \widehat{\alpha}_m)\right]^{\frac{lm}{2}-1}\exp\left(\frac{1}{2}\ln L(k, \widehat{\alpha}_m)\right). \qquad (41)$$

Then the probability of false alarm is given by

$$\int_{\tilde{\lambda}_m}^{\infty}\chi^2\left[\ln L(k, \widehat{\alpha}_m)\right] d\left[\ln L(k, \widehat{\alpha}_m)\right], \qquad (42)$$

after detecting the beginning of the abnormal operation, either the Kalman filter parameters are adjusted (usually the gain factors or the covariance matrix Q (k)), or their structure is changed by using more complex models of state change taking into account the abnormal operation.

The general scheme that realizes this algorithm is shown in Figure 1 below.

In the first case, parameters adjustment of the Kalman filter will be changed according to the following formula

$$if\ \ln L(k, \widehat{\alpha}_1) \geq \tilde{\lambda}_1\ then\ Q^0(k) \rightarrow Q^1(k), \qquad (43)$$

where $Q^0(k)$ and $Q^1(k)$ are chosen on the basis of the experiment so that to better reflect the true estimate of the state, both in the absence and in the presence of an abnormal operation. According to this, the elements of the matrix $Q^0(k)$ must take small values corresponding to rectilinear and uniform tracking with weak perturbations and $Q^1(k)$ are large values corresponding to tracking with abnormal operation. In general, the value of $Q^1(k)$ is selected on the basis of the possible maximum abnormal operation intensity of the system [10], for example, one may choose $Q^0(k) = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$ and $Q^1(k) = \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix}$. The structure of the Kalman filter is changing by take the following

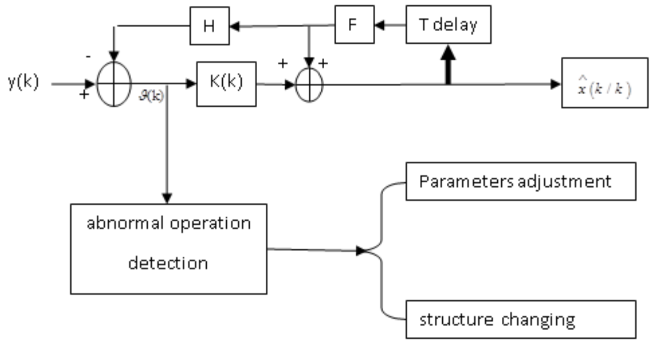

Figure 1: Kalman filter operation schema for abnormal estimation.

formula in the second case:

$$if \ \ln L(k,\widehat{\alpha}_1) \geq \tilde{\lambda}_1 \ then \ model \ U(k) = u(k) \ converted \ to \ model \ U(k) = u(k)+\alpha. \quad (44)$$

Bayes adaptive approach [9] takes into account the possibility that the adopted model is inadequate, in other words, all possible variants of the parameters or hypotheses concerning the model are taken into account.

4 Proposed Multi Model Estimation

In the problem of filtering tracking of the states, the a priori uncertainty of the statistical characteristics of the system leads to an uncertainty in the statistical characteristics of the filtered state vector. According to this, which is unknown to the observer in advance, leads to a mismatch between the real state and the model used in the estimation devices. Let ζ be a vector containing all the indeterminate parameters that represent all the undefined events associated with the hypotheses about the model.

According to this, in the general case, the equations of state of the system and the observations take the form

$$x(k+1) = F(\zeta_{k+1})x(k) + \Gamma(\zeta_{k+1})v(k), \quad (45)$$

$$y(k) = Hx(k) + w(k). \quad (46)$$

Then, in this section, we have to estimate the conditional mathematical expectation

$$\hat{x}(k/k) = E\left[\, x(k)|\, Y^k\right]. \quad (47)$$

According to the Bayesian approach, this operation is performed on the basis of the total probability theorem:

$$p\left[x(k)\left|Y^k\right.\right] = \int_{\Upsilon} p\left[x(k)\left|\zeta, Y^k\right.\right] p\left[\zeta\left|Y^k\right.\right] d\zeta, \quad (48)$$

where Υ is the set of all possible values of ζ, $p\left[x(k)\left|Y^k\right.\right]$ is the conditional probability distribution of states $x(k)$ for given observation, $p\left[x(k)\left|\zeta, Y^k\right.\right]$ is the conditional probability distribution of states $x(k)$ for given observation and accepted value of ζ for undefined parameters and events, $p\left[\zeta\left|Y^k\right.\right]$ is the posteriori distribution of vector ζ for given observations. From equations (47) and (48) the optimal mean square estimate of $\hat{x}(k/k)$ and its covariance matrix $\hat{P}(k/k)$ are

$$\hat{x}(k/k) = \int_{\Upsilon} \widehat{x}^{\zeta}(k/k) p\left[\zeta\left|Y^k\right.\right] d\zeta, \tag{49}$$

$$\hat{P}(k/k) = \int_{\Upsilon} \left\{\hat{P}^{\zeta}(k/k) + \left[\hat{x}^{\zeta}(k/k) - \hat{x}(k/k)\right]\left[\hat{x}^{\zeta}(k/k) - \hat{x}(k/k)\right]^T\right\} p\left[\zeta|Y^k\right] d\zeta, \tag{50}$$

where $\hat{x}^{\zeta}(k/k)$ and $\hat{P}^{\zeta}(k/k)$ are the partial estimate and its covariance obtained for a given value of ζ, respectively, and equal to

$$\hat{x}^{\alpha}(k/k) = E\left\{x(k)|\alpha, Y^k\right\}, \tag{51}$$

$$\hat{P}^{\zeta}(k/k) = E\left\{\left[\hat{x}^{\zeta}(k/k) - x(k)\right]\left[\hat{x}^{\zeta}(k/k) - x(k)\right]^T \left|\zeta, Y^k\right.\right\}, \tag{52}$$

Therefore, the obtained estimate can be represented as a linear combination of partial estimates $\hat{x}^{\zeta}(k/k)$, each of which is obtained under a certain hypothesis with respect to the model. The weight coefficients of this linear combination expressing the total estimate are determined by the probabilities of each hypothesis under consideration. A complete covariance matrix is calculated similarly, as a linear combination (with the same weights) of conditional matrices $\hat{P}^{\zeta}(k/k)$. Therefore, in principle, by the equations (49)–(52) we can obtain the optimal estimate $\hat{x}(k/k)$. In addition, as follows from equation (48), $p\left[x(k)|Y^k\right]$ is a linear combination of Gaussian probabilities. So, $\{\zeta_k = i\}$ or $\left\{\zeta_k^i\right\}$ the event consists the state in accordance with the state i at the k^{th} instant of time; $\Upsilon(k) = \{\Upsilon_1, \Upsilon_2, ..., \Upsilon_k\} = \left\{\{\zeta_1^i\}, \left\{\zeta_2^j\right\}, ..., \{\zeta_k^s\}\right\}, i, j, ..., s = 0, M-1$, is the set of all possible values that the discrete-random process ζ_k can take from the initial value to the k^{th} time moment. And $\Upsilon(k, l_k) = \{\Upsilon(k) \cap l_k\}$ $l_k = 1, M^k$ is a subset containing k-value. Figure 2 describes the set of possible values of Υ.

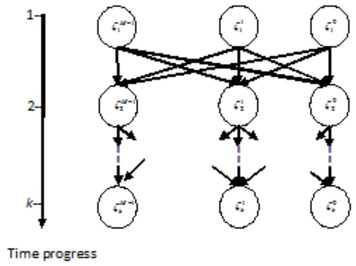

Figure 2: The set of possible values of Υ.

From Figure 2; $\Upsilon(k, l_k)$ is the ${l_k}^{th}$ branch with the duration k. Hence the number of different branches in $\Upsilon(k)$ is equal to M^k. In this case, according to the Bayes formula, we obtain the posterior probability density of the state vector as follows:

$$p\left[x(k)\left|Y^k\right.\right] = \sum_{l_k=1}^{M^k} p\left[x(k)|\Upsilon(k, l_k), Y^k\right] P(\Upsilon(k, l_k)|Y^k), \tag{53}$$

the indices $i(l_k), j(l_k), ..., s(l_k)$ are varying between the limits [0,M-1] and the index $l_k \in 1, M^k$. Let the a posteriori probability density of the l_k-th branch $\Upsilon(k, l_k)$ be

$$\begin{aligned} P(\Upsilon(k,l_k)|Y^k) &= P(\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|y(k), Y^{k-1}) \\ &= \frac{p\left[y(k)|\zeta_1^i, \zeta_2^j, ..., \zeta_k^s, Y^{k-1}\right] P\left[\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|Y^{k-1}\right]}{p\left[y(k)|Y^{k-1}\right]}, \end{aligned} \tag{54}$$

then equation (53) in base of indices will be

$$p\left[x\left(k\right)\left|Y^k\right.\right] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} ... \sum_{s=0}^{M-1} p\left[x(k)|\zeta_1^i, \zeta_2^j, ..., \zeta_k^s, Y^k\right] P(\zeta_1^i, \zeta_2^j, ..., \zeta_k^s|Y^k), \tag{55}$$

To reduce possible hypotheses, it is assumed that [4] the unknown parameter ζ_k is a simply-connected Markov chain, with a matrix of transition probabilities $P = [p_{ts}]_{t,s=0,M-1}$, where $p_{ts} = P\left(\zeta_k^s|\zeta_{k-1}^t\right)$. Since the algorithm starts working at the k^{th} time point, at the kth instant of time equation (55) will be

$$p\left[x(k)|Y^k\right] = \sum_{s=0}^{M-1} p\left[x(k)|\zeta_k^s, Y^k\right] P(\zeta_k^s|Z^k). \tag{56}$$

The Bayes description of the expression $p\left[x(k)|\zeta_k^s, Y^k\right]$ is written in the form

$$p\left[x(k)|\zeta_k^s, Y^k\right] = p\left[x(k)|\zeta_k^s, y(k), Y^{k-1}\right] = \frac{p\left[y(k)|\zeta_k^s, x(k), Y^{k-1}\right] p\left[x(k)|\zeta_k^s, Y^{k-1}\right]}{p\left[y(k)|\zeta_k^s, Y^{k-1}\right]}. \tag{57}$$

This leads to

$$p\left[x(k)|\zeta_k^s, y(k), Y^{k-1}\right] = \sum_{i=0}^{M-1} p\left[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}\right] P\left(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}\right). \tag{58}$$

Equation (58) proves that the term $p\left[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}\right]$ is the Gaussian apriori probability density of the state vector corresponding to the values $\left\{\zeta_{k-1}^i, \zeta_k^s\right\}$ between two adjacent moments. Thus

$$\begin{aligned} p[x(k)|\zeta_{k-1}^i, \zeta_k^s, Y^{k-1}] = N\Big[&x(k); F(k/k-1, \zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1), \\ &F(k/k-1, \zeta_k^s) P^i(k-1/k-1) F(k/k-1, \zeta_k^s)^T + \Gamma(\zeta_k^s) Q(k-1) \Gamma(\zeta_k^s)^T\Big]. \end{aligned} \tag{59}$$

If it is assumed that the probability density $p\left[x(k)|\zeta_k^s, Y^{k-1}\right]$of the state vector corresponding to the value ζ_k^s is Gaussian, then taking into account (58)–(59) we obtain

$$\begin{aligned} &N[x(k); \stackrel{\frown}{x}^s (k/k-1), \stackrel{\frown}{P}^s (k/k-1)] \simeq N\Bigg[x(k); \sum_{i=0}^{M-1} F(\zeta_k^s) \stackrel{\frown}{x}^i (k/k-1) P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \\ &\sum_{i=0}^{M-1} F(\zeta_k^s) \stackrel{\frown}{P}^i (k-1/k-1) F(\zeta_k^s)^T + \Gamma(\zeta_k^s) Q(k-1) \Gamma(\zeta_k^s)^T + [F(\zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1) \\ &- \hat{x}^s(k/k-1)][[F(\zeta_k^s) \stackrel{\frown}{x}^i (k-1/k-1) - \hat{x}^s(k/k-1)]^T] P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1})\Bigg], \end{aligned} \tag{60}$$

equating both sides of this expression, we obtain

$$\widehat{x}^{s}(k/k-1) = F(\zeta_k^s) \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1) P(\zeta_{k-1}^i | \zeta_k^s, Y^{k-1}), \tag{61}$$

or

$$\widehat{x}^{s}(k/k-1) = F(\zeta_k^s) \hat{x}^{0s}(k-1/k-1), \tag{62}$$

where

$$\hat{x}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1) P(\zeta_{k-1}^i | \zeta_k^s, Y^{k-1}), \tag{63}$$

and

$$\begin{aligned}\hat{P}^s(k/k-1) &= \sum_{i=0}^{M-1} F(\zeta_k^s)\, \widehat{P}^{i}(k-1/k-1) F(\zeta_k^s)^T + \Gamma(\zeta_k^s) Q(k-1) \Gamma(\zeta_k^s)^T + \\ &\left[F(\zeta_k^s)\, \widehat{x}^{i}(k-1/k-1) - \hat{x}^s(k/k-1)\right] \left[\left[F(\zeta_k^s)\, \widehat{x}^{i}(k-1/k-1) - \hat{x}^s(k/k-1)\right]^T\right] \\ &P(\zeta_{k-1}^i | \zeta_k^s, Y^{k-1}).\end{aligned} \tag{64}$$

Substituting (63) into (64), we obtain

$$[\hat{P}^s(k/k-1) = F(\zeta_k^s) \hat{P}^{0s}(k/k-1) F(\zeta_k^s)^T + \Gamma(\zeta_k^s) Q(k-1) \Gamma(\zeta_k^s)^T, \tag{65}$$

where

$$\begin{aligned}\hat{P}^{0s}(k-1/k-1) &= \sum_{i=0}^{M-1} P(\zeta_{k-1}^i | \zeta_k^s, Y^{k-1}) \ldots \\ &\left[\hat{P}^i(k-1/k-1) + \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right) \times \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right)^T\right] \frac{\partial^2 \Omega}{\partial v^2},\end{aligned} \tag{66}$$

and

$$[P(\zeta_{k-1}^i | \zeta_k^s, Y^{k-1}) = \frac{P(\zeta_k^s | \zeta_{k-1}^i) P(\zeta_{k-1}^i | Y^{k-1})}{\sum_{i=0}^{M-1} P(\zeta_k^s | \zeta_{k-1}^i) P(\zeta_{k-1}^i | Y^{k-1})}. \tag{67}$$

The output estimates is calculated from equations (59) and (60)

$$[\widehat{x}(k/k) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \cdots \sum_{s=0}^{M-1} \hat{x}^{ij\ldots s}(k/k) P(\zeta_1^i, \zeta_2^j, \cdots, \zeta_k^s | Y^k), \tag{68}$$

and its covariance matrix:

$$\begin{aligned}\hat{P}(k/k) &= \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \cdot\cdot \\ &\cdot \sum_{s=0}^{M-1} \left[\hat{P}^{ij\ldots s}(k/k) + \left[\hat{x}^{ij\ldots s}(k/k) - \hat{x}(k/k)\right] \left[\hat{x}^{ij\ldots s}(k/k) - \hat{x}(k/k)\right]^T\right] P(\zeta_1^i, \zeta_2^j, \cdots, \zeta_k^s | Y^{k)},\end{aligned} \tag{69}$$

where estimation

$$\hat{x}^{ij\ldots s}(k/k) = \hat{x}^{\zeta_1^i,\zeta_2^j,\ldots,\zeta_k^s}(k/k), \tag{70}$$

and covariance matrix

$$\hat{P}^{ij\ldots s}(k/k) = \hat{P}^{\zeta_1^i,\zeta_2^j,\ldots,\zeta_k^s}(k/k), \tag{71}$$

the corresponding ${l_k}^{th}$ branch of the algorithm M parallel Kalman filters Algorithm. The algorithm consists of M parallel Kalman filters, each of them is tuned to a single value ζ^i, as described below, Figure 3 shows the realization diagram of the algorithm.

1. Initialization:

   (a) Set the initial value of the M – estimates of the state vector $\hat{x}^i(k/k)$:

   $$\hat{x}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} \hat{x}^i(k-1/k-1)P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1}), \tag{72}$$

   (b) Set the initial value of the M – covariance $\hat{P}^i(k-1/k-1)$:

$$\hat{P}^{0s}(k-1/k-1) = \sum_{i=0}^{M-1} P(\zeta_{k-1}^i|\zeta_k^s, Y^{k-1})$$
$$\left[\hat{P}^i(k-1/k-1) + \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right) \times \left(\hat{x}^i(k-1/k-1) - \hat{x}^s(k/k-1)\right)^T\right]$$
$$\frac{\partial^2\Omega}{\partial v^2}, \tag{73}$$

2. Estimation process:

   (a) Extrapolation of the state vectors:

   $$\widehat{x}^{s}(k/k-1) = F(\zeta_k^s)\hat{x}^{0s}(k-1/k-1), \tag{74}$$

   (b) Extrapolation of correlation matrices:

   $$\hat{P}^s(k/k-1) = F(\zeta_k^s)\hat{P}^{0s}(k/k-1)F(\zeta_k^s)^T + \Gamma(\zeta_k^s)Q(k-1)\Gamma(\zeta_k^s)^T. \tag{75}$$

   (c) Coefficient gain:

   $$K^s(k) = \hat{P}^s(k/k-1)H\left[H\hat{P}^s(k/k-1)H^T + R(k)\right]^{-1}. \tag{76}$$

   (d) Estimation:

   $$\hat{x}^s(k/k) = \hat{x}^s(k/k-1) + K^s(k)\left[y(k) - H\hat{x}^s(k/k-1)\right]. \tag{77}$$

   (e) Covariance filtration

   $$P^s(k/k) = \left[I - K^s(k)H(k)\right]\hat{P}^s(k/k-1). \tag{78}$$

3. Calculate the conditional probability density of observation, corresponding to the states of ζ_k^s, we obtain

$$\begin{aligned} L^s(k) = p[y(k)|\zeta_k^s, Y^{k-1}] = p\left[y(k)|\zeta_k^s, H\hat{x}^s(k/k-1), \hat{P}^s(k/k-1)\right] \\ = \left[(2\pi)^s \det(s^s)(k)\right]^{-1/2} \exp\left[-0.5\vartheta^s(k)^T s^s(k)^{-1}\vartheta^j(k)\right] \end{aligned}, \quad (79)$$

where

$$\vartheta^s(k) = y(k) - H(k)\widehat{x}^s(k/k-1), \quad (80)$$

and

$$s^s(k) = H\hat{P}^s(k/k-1)H^T(k) + R(k). \quad (81)$$

4. Calculate the a posteriori probabilities of the state ζ_k^s, according to the Bayes formula

$$P\left[\zeta_k^s|Y^k\right] = P\left[\zeta_k^s|y(k), Y^{k-1}\right] = \frac{p\left[y(k)|\zeta_k^s, Y^{k-1}\right] P\left[\zeta_k^s|Y^{k-1}\right]}{\sum_{s=0}^{M-1} p\left[y(k)|\zeta_k^s, Y^{k-1}\right] P\left[\zeta_k^s|Y^{k-1}\right]}, \quad (82)$$

where

$$P\left[\zeta_k^s|Y^k\right] = \frac{L^s(k)\sum_{i=o}^{M-1} p_{is} P\left[\zeta_{k-1}^i|Y^{k-1}\right]}{\sum_{s=0}^{M-1} L^s(k)\sum_{i=o}^{M-1} p_{is} P\left[\zeta_{k-1}^i|Y^{k-1}\right]}. \quad (83)$$

5. The output estimate of the target state vector and its covariance were obtained on the basis of a linear combination of a posteriori probabilities

$$\widehat{x}\ (k/k) = \sum_{s=0}^{M-1} \hat{x}^s(k/k) P(\zeta_k^s|Y^k), \quad (84)$$

$$\hat{P}(k/k) = \sum_{s=0}^{M-1} \left[\hat{P}^s(k/k) + [\hat{x}^s(k/k) - \hat{x}(k/k)]\,[\hat{x}^s(k/k) - \hat{x}(k/k)]^T\right] P\left[\zeta_k^s|Y^k\right]. \quad (85)$$

5 Simulation Results

The proposed algorithm was applied to the phosphorus furnace type RKZ-80F. The system has the following specifications: a linear model of indirect control of electro-thermal processes in a three-electrode. This model is based on the well-known band structure of electric furnace baths and takes into account its geometric symmetry with respect to the three electrodes. Therefore, all the variables used in the model below will have indices i = 1,2,3; they relate to one of the three electrodes or the corresponding near-electrode region. The index 0 of the variable will indicate its relation to the whole bath of the furnace as a whole. The vector of state space $x(k) \in R^n$ $(n = 7)$ has the following states:

$$x^T(k) = \left[\ x_{len}^1(k) \quad x_{vol}^1(k) \quad x_{len}^2(k) \quad x_{vol}^2(k) \quad x_{len}^3(k) \quad x_{vol}^3(k) \quad x_{high}^0(k)\ \right],$$

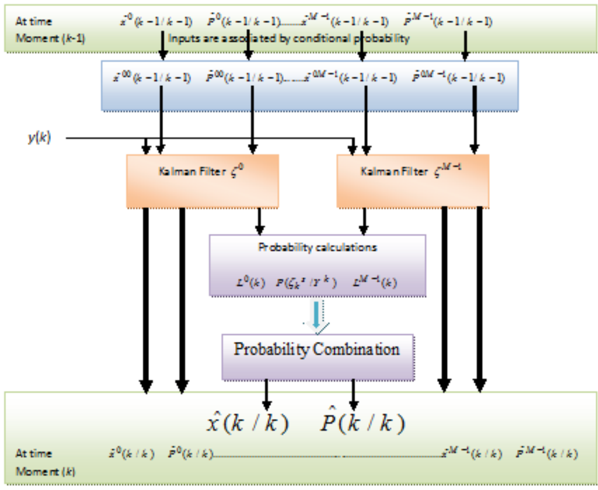

Figure 3: Proposed algorithm implementation.

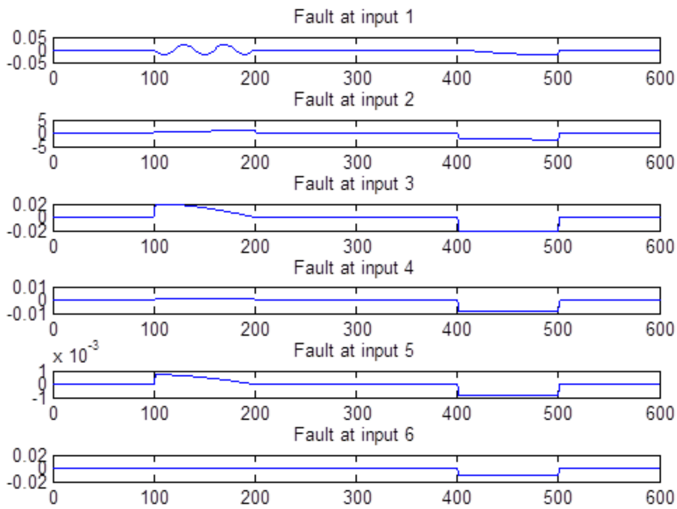

Figure 4: Fault signals forms.

where $x^i_{len}(k)$ is the length of the i-th electrode, m; $x^i_{vol}(k)$ is the volume of the crucible of the i$-$th near-electrode region, m^3; $x^0_{high}(k)$ is the height of the total working (carbon) zone in the furnace, m. The control vector $u(k) \in R^p$ $(p = 7)$ has the following structure:

$$u^T(k) = \left[\ u^1_{byp}(k) \quad u^1_{pa}(k) \quad u^2_{bup}(k) \quad u^2_{pa}(k) \quad u^3_{len}(k) \quad u^3_{byp}(k) \quad u^0_c(k) \quad \right],$$

$$u^T(k) = \left[u^1_{?5@}(k)\ u^1_P(k)\ u^2_{?5@}(k)\ u^2_P(k)\ u^3_{?5@}(k)\ u^3_P(k)\ u^0_C(k)\right].$$

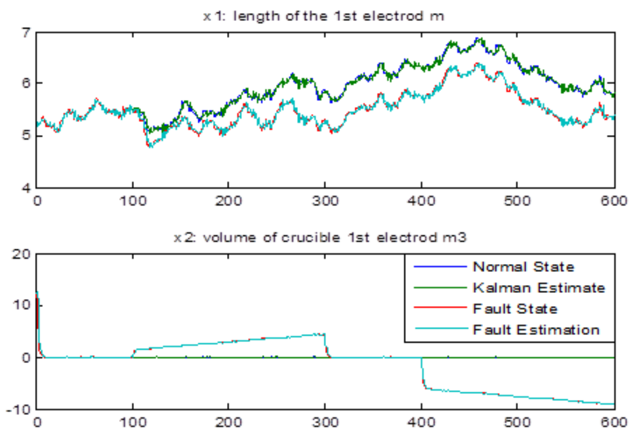

Figure 5: Estimations of the length and volume of the first electrode.

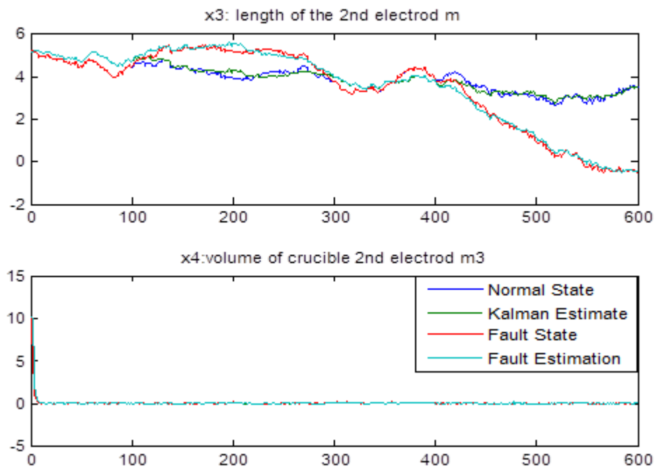

Figure 6: Estimations of the length and volume of the second electrode.

$u^i_{byp}(k$ is the bypass of the ith electrode, m;
$u^i_{pa}(k)$ is the average useful active power, in the i-th near-electrode region, MW;
$u^0_c(k)$ is the the amount of carbon entering the furnace with the charge, T.
The vector of observation $y(k) \in R^m$ $(m = 15)$ has the following structure:

$$y^T(k) = \left[y^1_1(k)y^1_2(k)y^1_3(k)y^1_4(k)y^2_1(k)y^2_2(k)y^2_3(k)y^2_4(k)y^3_1(k)y^3_2(k)y^3_3(k)y^3_4(k)y^0_1(k)y^0_2(k)y^0_3(k)\right],$$

where
$y^i_1(k)$ is the position of the electrode holder of the i-th electrode relative to the slag tap, m;
$y^i_2(k)$ is the active resistance of the i-th phase, mΩ;
$y^i_3(k)$ is the temperature under the arch at the i-th electrode,$^\circ C$;

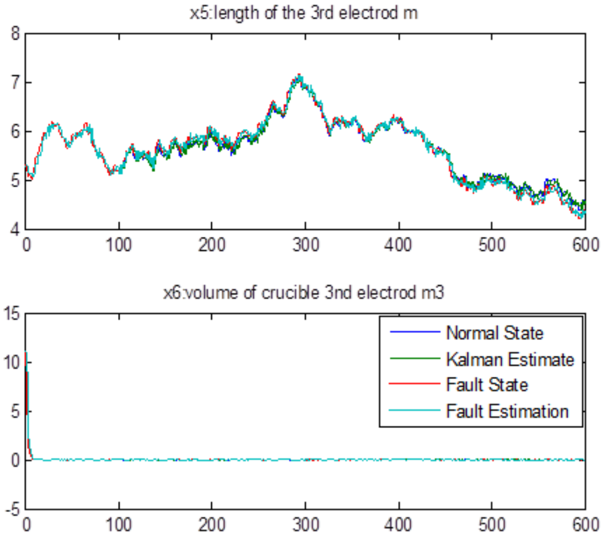

Figure 7: Estimations of the length and volume of the third electrode.

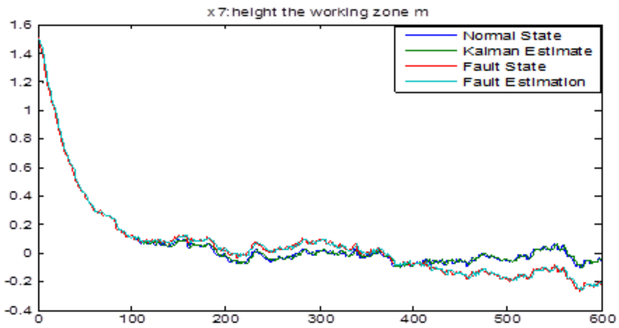

Figure 8: Estimations of the high of the working area.

$y_4^i(k)$ is the gathering of charge for the i-th electrode, T;
$y_1^0(k)$ is the the average temperature under the roof of the entire furnace, $°C$;
$y_2^0(k)$ is the current furnace capacity, T;
$y_3^0(k)$ is the relative recoverability of the product, (%P_2O_5 in the batch /% P_2O_5 in the slag).

The matrices parameters Φ, Γ and H are as follows

$$\Phi = \begin{bmatrix} \Phi_1 & 0 & 0 & f_{01} \\ 0 & \Phi_1 & 0 & f_{01} \\ 0 & 0 & \Phi_1 & f_{01} \\ f_{10} & f_{10} & f_{10} & f_{00} \end{bmatrix}, \Gamma = \begin{bmatrix} \Gamma_1 & 0 & 0 & 0 \\ 0 & \Gamma_1 & 0 & 0 \\ 0 & 0 & \Gamma_1 & 0 \\ 0 & 0 & 0 & g_{00} \end{bmatrix}, H = \begin{bmatrix} H_1 & 0 & 0 & h_{01} \\ 0 & H_1 & 0 & h_{01} \\ 0 & 0 & H_1 & h_{01} \\ H_{10} & H_{10} & H_{10} & h_{00} \end{bmatrix},$$

where $\Phi_1 = \begin{bmatrix} a_1 & -a_2 \\ 0 & a_4 \end{bmatrix}$, $f_{01} = \begin{bmatrix} \frac{a_3}{a_5} \end{bmatrix}$, $f_{10} = \begin{bmatrix} 0 & a_6 \end{bmatrix}$, $f_{00} = a_7, g_{00} = \beta_2$, $\Gamma_1 = \begin{bmatrix} 1 & 0 \\ 0 & \beta_1 \end{bmatrix}$, $H_1 = \begin{bmatrix} 1 & 0 \\ \gamma_2 & 0 \\ 0 & \gamma_4 \\ 0 & \gamma_6 \end{bmatrix}$, $h_{01} = \begin{bmatrix} \gamma_1 \\ \gamma_3 \\ \gamma_5 \\ 0 \end{bmatrix}$, $H_{10} = \begin{bmatrix} 0 & \gamma_7 \\ 0 & \gamma_9 \\ 0 & 0 \end{bmatrix}$, $h_{00} = \begin{bmatrix} \gamma_8 \\ 0 \\ \gamma_{10} \end{bmatrix}$,

$\theta = [a_1, ..., a_7, \beta_1, \beta_2, \gamma_1, ..., \gamma_{10}]$. The parameters values are described in Table below:

No.	Parameter	Physical dimension	Admissible values	Recommended initial approximation
1	α_1	$-$	$0.95 - 1.0$	0.999
2	α_2	$1/M^2$	$0.0002 - 0.002$	0.001
3	α_3	$-$	$0.0001 - 0.002$	0.001
4	α_4	$-$	$0.0 - 1.0$	0.5
5	α_5	M^2	$0.0 - 1.0$	0.5
6	α_6	$1/M^2$	$0.0001 - 0.005$	0.001
7	α_7	$-$	$0.8 - 1.0$	0.9
8	β_1	M^3/MW	$0.2 - 0.4$	0.28
9	β_2	M/T	$0.004 - 0.0045$	0.0042
10	γ_1	-	$0.5 - 1.0$	0.75
11	γ_2	$m\Omega/M$	$0.05 - 0.5$	0.1
12	γ_3	$m\Omega/M$	$0.2 - 0.8$	0.5
13	γ_4	$^\circ C/M^3$	$1.0 - 10.0$	5.0
14	γ_5	$^\circ C/M$	$300 - 500$	400
15	γ_6	t/M^3	$1.0 - 2.0$	1.5
16	γ_7	$^\circ C/M^3$	$1.0 - 10.0$	5
17	γ_8	$^\circ C/M$	$250 - 400$	300
18	γ_9	T/M^3	$0.05 - 0.2$	0.1
19	γ_{10}	$1/M$	$20 - 40$	30

The initial values of states are: x = $[5.25\ 25\ 5.18\ 20\ 5.36\ 22\ 1.5]^T$; the number of samples is 600 samples; the input vector is generated as random signal in the interval [-0.05, 0.05]; different faults scenarios are added to the input vectors. The simulation results are shown in figures below, in which Figure 4 represents the fault signals forms. Figure 5 shows that the algorithm starts tracking in the abnormal state in sample time number 100 for both states estimations and could not return to the normal operation even the input signal return without fault. Figures 6 to 8 show the switching between modes to track the changes in states directly after the abnormal operation is detected to start the fault tracking.

6 Conclusion

In this paper we proposed a multi-operational mode based on Bayes approach to select the best Kalman filter estimator. The estimation was focused on internal state estimation during the fault operation which may be caused by many different resources. The simulation fault testing inputs signals failure, which were lost by composed signals. The fault scenarios started at time sample 100 with different shapes and values for each input and ended at time sample 200. Another faults signals were inserted at time sample 400 and ended at time sample 500. From the simulation results of states estimations, we noted that the estimation of the lengths of the electrodes were very satisfied in which most of estimations values were very near from the estimations in normal operation. The volume of the first electrode during the fault was not estimated correctly where a negative values was estimated. While the volumes of the second and third electrodes during the fault was estimated correctly. Finally, estimations of the high of the working area may be accepted with some error percentage. As a result the proposed algorithm provide a good results with some acceptable range of error during the faults operations.

References

[1] Maiying Zhong, Hao Ye, Steven X. Ding, Guizeng Wang, and Zhou Donghua. Fault detection filter for linear time-delay systems. *Nonlinear Dynamics and Systems Theory* **5** (3) (2005) 273–284.

[2] Wu, H. and Noe, F. A flat Dirichlet process switching model for Bayesian estimation of hybrid systems. *IEEE Transaction on Automatic Control* **15** (1) (2002) 10–17.

[3] Ackerson, G. and Fu, K. On state estimation in switching environments. *Procedia Computer Science* (4) (2011) 1393–1402.

[4] Blom, H. and Bar-Shalom, Y. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transaction on Automatic Control* (17) (2001) 362–369.

[5] Minka, T. Expectation propagation for approximate Bayesian inference. *Proceedings of Uncertainty in Artificial Intelligence* **33** (8) (2002) 780–783.

[6] Bar-Shalom, Y., Kirubasrajan, T. and Li, X.R. *Estimation with Applications to Tracking and Navigation*. Wiley, New York, 2002.

[7] Evgenia Suzdaleva and Ivan Nagy. Recursive state estimation for hybrid systems. *Applied Mathematical Modelling* (36) (2012) 1347–1358.

[8] F. van der Heijden, Duin, R.P.W., D. de Ridder and Tax, D.M.J. *Classification, Parameter Estimation and State Estimation*. John Wiley & Sons, 2004.

[9] Bar-Shalom, Y. and Fortmann, T.E. *Tracking Data Association*. Academic Press, 1988.

[10] Ali Hussein Hasan and Grachev, A.N. Adaptive $\alpha - \beta-$filter for target tracking using real time genetic algorithm. *Journal of Electrical and Control Engineering* **3** (4) (203) 32–38.