The Intelligent Identification Technique with Associative Search

Nataliya N. Bakhtadze and Ekaterina A. Sakrutina

Abstract— In modern control systems, identification is an integral part of adaptive control where process models are adjusted using real-time operation data and control actions optimal with respect to some performance criterion are developed. The methods for developing predictive models in control systems and decision-making support for nonlinear non-stationary objects are proposed. The methods are based on the application of associative search procedure to virtual model identification as well as on wavelet analysis techniques.

The paper presents novel associative search techniques enabling the development of a new dynamic object's model on each time step rather than plant approximation pertaining to time. The model is build using the data samples from process history (associations) developed at the learning phase. The new techniques employs the models of human individual's (process operator's, stock analyst's or trader's) behavior based on professional knowledge formalization. Application examples from oil refining and chemical industries, power engineering, and banking are adduced..

Keywords— associative search models, identification, knowledgebase, process model predictive control, wavelet analysis.

I. Introduction

THE design of control systems with identifier is the problem of significant practical value. The control plant identification is fulfilled by identifier which realizes both the function of parametric disturbances generation for control plant and controller parameters adjusting. In control systems with identifier the adjusted model is used not only for controller design but for solution of auxiliary problems (e.g., prediction of slow parametric failures), so demonstrating the universality of systems with identifier comparing to direct control approach.

The background of identification-based control approach was formulated in 50th – 60th years on the basis of Bayesian estimation [1]. The first industrial ASI (adaptive system with identifier), where identifier in feedback loop was used for compensator adjusting was implemented for tube rolling mill control [2]. The strict theoretical results in dynamic plants adaptive control were obtained in the beginning of 80th. The identifier typically used least mean squares technique and its modifications [3]. Later the progress was achieved in optimizing the control with respect to (w.r.t.) control cost function for the plants both with parametric and non-parametric uncertainty [4].

The theoretic results are accompanied by related soft products MATLAB system (System Identification Toolbox), based on widely used algorithms [5].

The up-to-date reality is design and implementation of advanced control systems which use all advantages of microprocessors and have great economic efficiency. The concept of advanced process control (APC) systems include large class beginning with generalized controllers (compensators, balance control systems, Smith's predictors, etc.) up to multivariable control of large-scale processes.

The last include sets of virtual analyzers, so providing the straight automatic control of marketable products properties. Among various approaches to APC the most frequently used in wide industrial areas is the one based on predicting models use, being well suited for robust control of large-scale plants with multiple complex interconnections. The main trend of APC systems development is using of sophisticated techniques of design and improvement of virtual analyzers, based on modern achievements of applied statistics, robust control, etc.

In modern decision-making support systems, the identification is done aiming to investigate the process properties what result in optimal control decision making by human operator.

The prediction model technique is one of so called substitution methods: the unknown parameter is implicitly determined as the unique point minimizing cost functional which is substituted by empirical one through the estimates calculation.

The use of stochastic approximation algorithms (both gradient and pseudo-gradient ones) for minimizing the empirical functional gives the opportunity to adjust the model in real time. For example, with quadratic cost function one gets recursive generalized LMS technique, while more slow increasing cost functions result in different estimation algorithms, robust w.r.t. innovation distribution [6].

The predicting model technique can be effectively implemented for Gaussian ARMA disturbances. For non-Gaussian disturbances, the optimal available control strategy is generally formed by nonlinear backward loop [7], being evidently much more hard task.

In this paper, the identification technique based on virtual models design is presented. The term "virtual" is to be understood as "ad hoc". This technique named "associative

Natalia Bakhtadze is with the V.A. Trapeznikov Institute of Control Sciences, 65, Profsoyuznaya, Moscow, 117997, RUSSIA, tel. +7 9165442259, e-mail: sung7@yndex.ru

Ekaterina A. Sakrutina is with the V.A. Trapeznikov Institute of Control Sciences, 65, Profsoyuznaya, Moscow, 117997, RUSSIA, e-mail: kathrina@inbox.ru

search" suggests predicting model design for dynamic plant in each time époque using the sets of archive data (associations) obtained on learning stage rather than real process approximation in time. Such approach is close to Ljung's idea [8] to use additional a priori information about the control plant in learning process. Fuzzy logic technique is used to develop the algorithms. So suggested methods imply for real-life plant identification to use simulation of human operator behavior based on technology knowledge formalization.

II. Nonlinear Prediction Algorithm Based on Virtual Models Development

A method using prediction models based on the imitation of analyst's associative thinking can be considered.

Identification algorithms employed in modern control systems often use expert knowledge both from human expert and from a knowledgebase. In the second case, an operator can choose between recommended control action and a forecast based on process state monitoring.

Two knowledge types are distinguished: declarative and procedural [9]. The first type includes the description of various facts, events, and observations, while skills and experience refer to the second type. Experts differ from novices by their structure and way of thinking and, in particular, the searching strategy [10]. If a person is not experienced, he/she would use the so-called 'backward reasoning'. He/she reviews different possible answers and makes a decision in favor of a specific answer based on the information received from the process at the current time step. On the contrary, an expert does not need to analyze current information in the process of decision-making, rather he/she uses the so-called 'forward reasoning' method which implies that the decision-making strategy is created subconsciously and this strategy is nonverbal. Therefore, in terms of the method of computational view of thought [11] the effectiveness of system will to a great extent be determined by expert's qualification and by the available a priori information.

Within the framework of this method, the cognitive psychology determines knowledge as a certain set of actually existing elements-symbols stored in human memory, processed during thinking and determining the behavior. The symbols, in turn, could be determined by their structure and the nature of neuron links [12].

Knowledge processing in an intelligent system consists in the recovery (associative search) of knowledge by its fragment [13]. The knowledge can be defined as an associative link between images. The associative search process can take place either as a process of image recovery using partially specified symptom (or knowledge fragment recovery by incomplete information; this process is usually emulated in various associative memory models) or as searching others images (linked associatively with the input image) related with other time steps. Those images have sense of a cause or an effect of an input image.

Gavrilov [13] offered a model which describes the associative thinking process as a sequential process of remembering based on associations – pairs of images defined by a set of symptoms. Such model can be considered as an intermediate level between neuron network models and logical models used in classical artificial intelligence systems. In this paper, we discuss an approach to developing on-line support of trader's decision-making based on the dynamic simulation of associative search and the identification technique based of virtual models. An identification algorithm for complex nonlinear dynamic objects such as continuous and batch processes was presented in [14]. The identification algorithm with continuous real-time self-tuning is based on virtual models design.

At every time step, a new virtual model is created. To build a model for a specific time step, a temporary "ad hoc" database of historic and current process data is generated. After calculating the output forecast based on object's current state, the database is deleted without saving.

The linear dynamical prediction model looks as follows:

$$y_t = a_0 + \sum_{i=1}^{r} a_i y_{t-i} + \sum_{j=1}^{s} \sum_{p=1}^{P} b_{jk} x_{t-j,p} \,,$$

where y_t is the object's output forecast at the t-th step, x_t is the input vector, r is the output memory depth, s is the input memory depth, P is the input vector length.

The original dynamic algorithm consists in the design of an approximating hyper surface of input vector space and the related one-dimensional outputs at every time step (see Fig. 1). To build a virtual model for a specific time step, the points close in a manner to the current input vector are selected. The output value at the next step is further calculated using least mean squares (LMS).

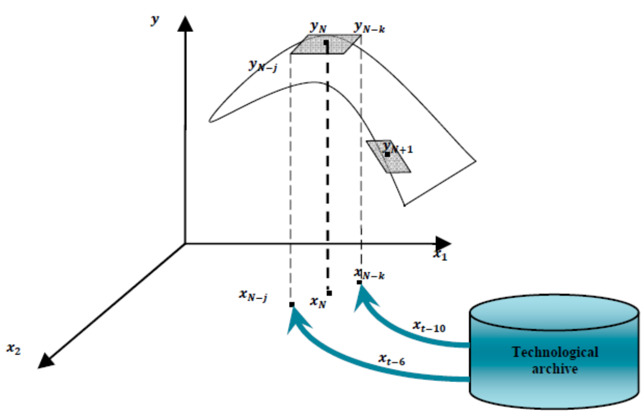

Fig. 1. Approximating hyper surface design

We use associative search technique for virtual models design - a method based on the associative thinking model.

High-speed approximating hyper surface design algorithms enabling the usage of fuzzy models for various process applications were offered in [15].

The following quantity

$$d_{t,t-j} = \sum_{p=1}^{P} \left| x_{tp} - x_{t-j,p} \right|, \; j = 1, \ldots, s$$

was introduced as distance (metric in $\Re^P$) between points of P-dimensional input space, where, generally, $s < t$, and x_{tp} are the components of the input vector at the current time step t.
Assume that for the current input vector $\mathbf{x}_t$:

$$\sum_{p=1}^{P} \left| x_{tp} \right| = d_t \qquad (1)$$

To build an approximating hyper surface for xt, we select such vectors xt-j, j = 1,…,s from the input data archive that for a given D_t the following condition will hold:

$$d_{t,t-j} \le d_t + \sum_{p=1}^{P} \left| x_{t-j,p} \right| \le d_t + D_t,\ j = 1,\dots,s \qquad (2)$$

The 2-D case is illustrated below (Fig. 2).

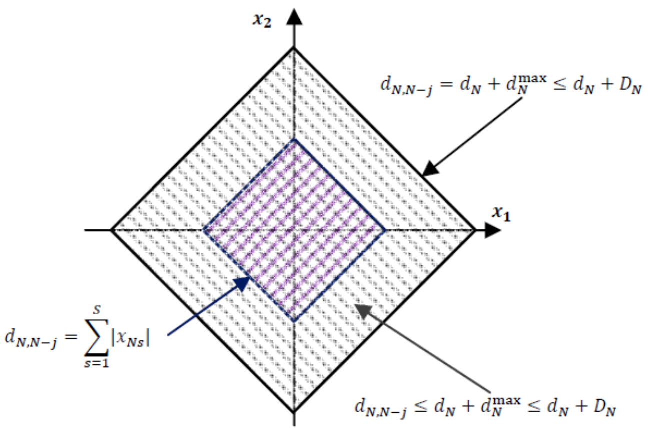

Fig.2. Approximating hyper surface building

The preliminary value of D_j is determined on the basis of process knowledge. If the selected domain does not contain enough inputs for applying LMS, i.e., the corresponding SLAE has no solution, then the chosen points selection criterion can be slackened by increasing the threshold D_t.

To increase the speed of the virtual models-based algorithm, an approach is applied based on employing a model of analyst's or trader's associative thinking for predicting.
For modeling the associative search procedure imitating the intuitive prediction of process status by a trader, we assume that the sets of process variable values, which are the components of an input vector, as well as the system outputs at previous time steps altogether create a set of symptoms, making an image of the object output at the next step.

The associative search process consists in the recovery of all symptoms describing the specific object based on its images. Denote the image initiating the associative search by R_0 and the corresponding resulting image of the associative search by R. A pair of images (R_0,R) will be further called association A or $A(R_0,R)$. The set of all associations on the set of images forms the memory of the knowledgebase of the intelligent system.

At the system learning phase, an archive of images is created. In our case, a set of n input vectors selected form the process history according to the algorithm described above will be considered as an image. At the prediction stage, the input vector x_t will be considered as an initial image R_0^a of the associative search, while approximating hypersurface formed by the input vectors from the process history built with the help of the associative search algorithm will be the final image R^a of the associative search. This means that to build a virtual model, one should select the existing hypersurfaces stored in the archive at the learning phase rather than individual vectors close to x_t. The selected hypersurface is an image of the current input vector which is used for output prediction. The algorithm implements the process of image R^a recovery based on R_0^a, i.e., the associative search process, and can be described by a predicate $\Xi = \{\Xi_i(R_{0i}^a, R_i^a, T^a)\}$ where $R_{0i}^a \subset R_0$, $R_i^a \subset R$, and T^a is the duration of the associative search.

For the algorithm described in this section, this predicate is a function asserting the truth or the falsity of input vector's membership of the specific domain in the inputs space. Therefore, the associative search process is reduced to the selection of a certain set of input vectors satisfying the condition (4) from the process archive. If the process history contains no image satisfying (4), then either the threshold D_t should be increased, or for a certain image of our input vector, some symptom should be replaced with a more relevant one. Formally, this means that the "worst" (i.e., located farther away from the current input then the rest ones w.r.t. the chosen criterion) vector from the process history will be deleted and replaced with a more relevant one, and so on.

Therefore, the analyst's decision-making process about to buy or to sell at any time step t could be constructed as associative search (process of remembering) of images (similar situations). The coordinates of approximating hypersurfaces used at previous steps are kept in archive.

III. ASSOCIATIVE SEARCH PROCEDURES IN SHORT TERM FORECASTING

In short-term prediction, not only the current situation but also object dynamics is very important. The conventional regression models are not precise enough to handle this problem.

We apply the associative search procedure with more complicated requirements to approximating hypersurface selection. We select from the archive such hypersurface (corresponding to some x_{t-j}, $j = 1,\dots,s$), that (i) it contains input vector at the current time step t, and (ii) the hypersurface corresponding to x_{N-j-1}, contains the input vector at the previous time step t-1. Formally, this means that the predicate becomes more complex:

$$\Xi(R_0^a, R^a, T^a) = \left\{ \sum_{p=1}^{P} \left| x_{t-j,p} \right| \le D_t - \sum_{p=1}^{P} \left| x_{tp} \right|;\ \sum_{p=1}^{P} \left| x_{t-j-1,p} \right| \le D_{t-1} - \sum_{p=1}^{P} \left| x_{tp} \right| \right\} \qquad (3)$$

There is principal opportunity to find more precise rules in the process of output changing by increasing the memory, say, to l

steps (l<t).

$$\Xi(R_0^a, R^a, T^a) = \left\{ \begin{array}{l} \sum_{p=1}^{P} |x_{t-j,p}| \le D_t - \sum_{p=1}^{P} |x_{tp}|; \\ \sum_{p=1}^{P} |x_{t-j-1,p}| \le D_{t-1} - \sum_{p=1}^{P} |x_{tp}|; \ldots; \sum_{p=1}^{P} |x_{t-l,p}| \le D_{t-l} - \sum_{p=1}^{P} |x_{tp}| \end{array} \right\} \quad (4)$$

IV. Model development by means of associative search technique using Wavelet Analysis

There are a variety of processes, which cannot be controlled using linear predictive models. Associative search technique offers a constructive solution for nonlinear processes. However, such processes may feature irregularities at certain instants. In engineering systems, such irregularities often demonstrate oscillating nature. The variability of feed properties due to feed supplier changes in process industries is a typical example. Another example is seasonal and daily load oscillations in power networks that directly effect the optimization of power transmission control modes. The ups and downs of stock market caused by the variety of economic reasons are also well known. Therefore, the design of predictive models by means of associative search technique for such type of time-varying processes looks relevant.

Over the past 20 years, the wavelet transform technique has been widely applied for time-varying process analysis [16]. Further we consider an approach to wavelet analysis application in identification tasks, in particular, in model development by means of associative search [17]. Such approach looks promising under transient conditions both for time-varying input signal and unmodeled dynamics of the control object.

V. Associative search in case of time-varying input vector

In most of industrial applications, in particular, in control systems with an identifier, the input signal is a vector. Let each component of the input vector meet Gauss-Markov conditions, in particular, the independence of sequence members. Also, suppose that at any time instant, vector's components are mutually independent.

At the same time, we suppose [18] that each component of the input vector is, generally, a time-varying sequence, but its singularities apparent at various instants are similar in a certain sense or identical. For example, stock market analysis techniques enables the detection of such "regularities" in market dynamics. In engineering systems, some repeatability (not necessarily periodical) of input signal's properties may be detected by applying statistical analysis to historical process data. Some processes, such as load dynamics in power networks, feature evident cyclicity [19].

In order to apply associative search algorithm for predicting the dynamics of such processes, we will, as usual, need to select from process history the vectors close to the current one in the sense of the selected criterion (associative impulse).

Case 1: a SISO system. As soon as the input sequence is time-varying, it makes sense to examine discrete wavelet expansions of input signals. In the general case, such expansion can be represented by the following expression:

$$f(t) = \sum_{k=-\infty}^{\infty} c_{jk}\varphi_{jk}(t) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d_{jk}\psi_{jk}(t) \quad (5)$$

where j is the depth of multiresolution expansion, which specifies the detailing depth, c_{jk} are scale factors, d_{jk} are detail coefficients.

Suppose $\psi(t)$ are Haar wavelets:

$$\psi(t) = \begin{cases} 1, & 0 \le t < 0.5, \\ -1 & 0.5 \le t \le 1, \\ 0 & t < 0,\ t \ge 1, \end{cases}$$

$$\varphi_{jk} = 2^{j/2}\varphi(2^j t - k),\ \psi_{jk} = 2^{j/2}\psi(2^j t - k).$$

The coefficients are calculated by means of Mallat algorithm (1999).

In the associative search procedure, we will choose from the archive such inputs $x(t^*)$ whose wavelet expansion with the same depth L

$$x(t^*) = \sum_{k=1}^{N/2^L} c^*{}_k \varphi_{L,k}(t) + \sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d^*{}_{jk}\psi_{j,k}(t^*)$$

meets the following 2 conditions:

1) $$\left|\sum_{k=1}^{N/2^L} c_k^* \varphi_{Lk}(t^*)\right| + \left|\sum_{k=1}^{N/2^L} c_k \varphi_{Lk}(t)\right| \le c_L, \quad (6)$$

where c_L is a positive number, $c_L \underset{L\to\infty}{\rightarrow} 0$.

2) for all coefficients $d_{jk}, k = 1, \ldots N/2^j$ there exists such constant $\tilde{c}_{jk}$ that:

$$\left|d_{kj}^*\right| \le \frac{\tilde{c}_{jk} - |\psi_{kj}(t)||d_{kj}|}{|\psi_{kj}(t^*)|} \quad (7)$$

Now, by the triangle inequality, we get

$$\left|\sum_{k=1}^{N/2^j} d^*_{jk}\psi_{jk}(t^*) - \sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(t)\right| \le$$

$$\left|\sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(t^*)\right| + \left|\sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(t)\right| \le \tilde{c}_j,$$

$$\tilde{c}_j = N/2^{j-1} \cdot \max_{k=1,\ldots,N/2^j} \tilde{c}_{jk}.$$

Combining (6) and (7) we obtain

$$\left|x(t^*)-x(t)\right| \le \left|x(t^*)\right| + \left|x(t)\right| =$$
$$\left|\sum_{k=1}^{N/2^L} c_k^*\phi_k(t^*) + \sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{jk}^*\psi_{jk}(t^*)\right| +$$
$$+\left|\sum_{k=1}^{N/2L} c_k\phi_k(t) + \sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(t)\right| \le$$

$$\le \left|\sum_{k=1}^{N/2^L} c_k^*\phi_k(t^*)\right| + \left|\sum_{k=1}^{N/2^L} c_k\phi_k(t)\right| + \sum_{j=1}^{L}\left|\sum_{k=1}^{N/2^j} d_{jk}\psi_{jk}(t)\right| +$$
$$+\sum_{j=1}^{L}\left|\sum_{k=1}^{N/2^j} d_{jk}^*\psi_{jk}(t^*)\right| \le c_L + L\tilde{c}_j = C = const > 0.$$

Case 2. Multi-dimensional input case.

Consider the case where the conditions of Gauss-Markov theorem are met. The predicted output value in the associative search procedure is obtained by the LSM. In view of the independence of input vector's components, the associative search criterion should be built as the existence condition of such constants $\tilde{c}_{ij}^s$, where $s=\overline{1,S}$ is the input vector's length, that the following inequalities

2a) $\left|d_{kj}^{s*}\right| \le \dfrac{\tilde{c}_{jk}^s - \left|\psi_{kj}^s(t)\right|\left|d_{kj}^s\right|}{\left|\psi_{kj}^s(t^*)\right|}$ hold true simultaneously for all $s=\overline{1,S}$. P known outputs meeting virtual linear model equations $y(t^*)=\sum_{s=1}^{S}\hat{h}^s x^s(t^*)$ correspond to the P vectors $x(t^*)$ selected from the history. The predicted value $y(t)$ is calculated by the LSM.

VI. The case of object's unmodeled internal dynamics

At each step in the associative search procedure, we build a new linear model $y(t)=\sum_{s=1}^{S}\hat{h}^s(t)x^s(t)$ of the nonlinear plant. In the unsteady-state case, the coefficients of the transfer function, generally, depend on *t*.

Suppose, at the first stage of the associative search procedure, P input vectors $x(t^*)$ meeting the conditions (1) and (2a) are selected. The resulting virtual models look as follows:

$$y(t^*)=\sum_{s=1}^{S}\hat{h}^s(t^*)x^s(t^*).$$

In view of the wavelet expansion of the input signals, we have

$$y(t^*)=\sum_{s=1}^{S}\hat{h}^s(t^*)\left[\sum_{k=1}^{N/2^L}c_k^{s*}\varphi_{Lk}^s(t^*)+\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*)\right]=$$
$$\sum_{s=1}^{S}\hat{h}^s(t^*)\sum_{k=1}^{N/2^L}c_{kj}^{s*}\varphi_{jk}^s(t^*)+\sum_{s=1}^{S}\hat{h}^s(t^*)\sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty}d_{jk}^{s*}\psi_{jk}^s(t^*).$$

It should be noted that $\hat{h}^s(t)$ and $\hat{h}^s(t^*)$ are, generally, different. Consider 2 special cases where this difference can be formalized.

Additive shift

$\hat{h}^s(t^*)=\hat{h}^s(t)+a^s(t^*)$, where

$\left|a^s(t^*)\right| \le C_a^{add} = const > 0 \quad \forall s=\overline{1,S}$.

For simplicity, we confine ourselves by the case where the degree of the detailing of Multiresolution Analysis is rather high. That results in the nulling of the scale factors c_{jk}^s, $s=\overline{1,S}$. In this case, we have

$$\sum_{s=1}^{S}(\hat{h}^s(t)+a^s(t^*))\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*)=$$
$$\sum_{s=1}^{S}\hat{h}^s(t)\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*)+\sum_{s=1}^{S}a^s(t^*)\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*).$$

In particular, at

$\left|a^s(t^*)\right| = C_a^{add} = const$ we have

$$y(t^*)-SC_a^{add}\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*)=$$
$$=\sum_{s=1}^{S}\hat{h}^s(t)\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*).$$

Denote

$$\tilde{y}(t^*)= y(t^*)-SC_a^{add}\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*)= =\sum_{s=1}^{S}\hat{h}^s(t)\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}d_{jk}^{s*}\psi_{jk}^s(t^*). \tag{8}$$

P known outputs meeting virtual linear model equations (8) correspond to the P vectors $x(t^*)$ selected from the history. The predicted value $y(t)$ is calculated by the LSM. It should be noted that all equations now include the coefficients $\hat{h}^s(t)$ as unknown quantities.

Multiplicative shift.

Now let $\hat{h}^s(t^*)=\hat{h}^s(t)\cdot a^s(t^*)$, where

$\left|a^s(t^*)\right| \le C_a^{mult} = const > 0 \quad \forall s=\overline{1,S}$.

Then the following equalities hold true:

$$y(t^*) = \sum_{s=1}^{S} \hat{h}^s(t^*) \sum_{j=1}^{L} \sum_{k=1}^{N/2^j} d_{jk}^{s*} \psi_{jk}^s(t^*) =$$

$$\sum_{s=1}^{S} \hat{h}^s(t) a^s(t^*) \sum_{j=1}^{L} \sum_{k=1}^{N/2^j} d_{jk}^{s*} \psi_{jk}^s(t^*) =$$

$$\sum_{s=1}^{S} \hat{h}^s(t) \sum_{j=1}^{L} \sum_{k=1}^{N/2^j} a^s(t^*) d_{jk}^{s*} \psi_{jk}^s(t^*) .$$

In particular, at $\left|a^s(t^*)\right| = C_a^{mult}$

$$y(t^*) = \sum_{s=1}^{S} \hat{h}^s(t) \sum_{j=1}^{L} \sum_{k=1}^{N/2^j} C_a^{mult} d_{jk}^{s*} \psi_{jk}^s(t^*) \qquad (9)$$

Equations (8) and (9) are used in the prediction algorithm directly together with the equation for $y(t)$. All equations now include the coefficients $\hat{h}^s(t)$ as unknown quantities.

VII. Fuzzy Virtual Models

The application of fuzzy models for decision-making under fuzziness and uncertainty conditions is justified in the following cases:

- if one or more factors of the quality index's dynamics are weakly or not formalized
- if the dynamics under investigation is described by sophisticated nonlinear relationships.

Production rules technique allowing to a certain extent to model human individual's thinking style is a key method of knowledge representation in modern systems employing expert knowledge. Any production rule consists of premises and a conclusion. Several premises in a rule are allowed; in such case they are combined by logical operators AND, OR.

Fuzzy systems are based on production-type rules with linguistic variables used as premise and conclusion in the rule [20].

By renaming the variables, the linear dynamic plant's model (1) can be represented as follows:

$$Y_N = \sum_{i=1}^{n+m} a_i X_i$$

We define a fuzzy model as a system with $n+m$ input variables $\mathbf{X} = \{X_1, X_2,..., X_{n+m}\}$ defined over the input reasoning domain $DX = DX_1 \times DX_2 \times ... \times DX_n$, and a single output variable Y defined over the output reasoning domain DY. Crisp values of X_i and Y will be denoted by x_i and y respectively.

Fuzzy definitional domain of the i-the input variable X_i is denoted by $LX_i = \{LX_{i,1},..., LX_{i,l_i}\}$ where l_i is the number of linguistic terms on which the input variable is defined; LX_{ik} specifies the name of the k-th linguistic term. Similarly, $LY = \{LY_1,..., LY_{ly}\}$ is the fuzzy definition domain of the output variable, l is the number of fuzzy values; LY_j is the name of the output linguistic term.

The rulebase in the fuzzy Mamdani system is a set of fuzzy rules such as:

$$R_j : LX_{1,j_1} \text{ AND } ... \text{ AND } LX_{n,j_n} \to LY_j \qquad (10)$$

The j-th fuzzy rule in the singleton-type system looks as follows:

$$R_j : LX_{1,j_1} \text{ AND } ... \text{ AND } LX_{n,j_n} \to r_j \qquad (11)$$

where r_j is real number to estimate the output y.

The j-th rule in Takagi–Sugeno model looks as follows:

$$R_j : LX_{1,j_1} \text{ AND } ... \text{ AND } LX_{n+m,j_{n+m}} \to r_{0j} + r_{1j}x_1 + r_{2j}x_2 + ... + r_{(n+m)j}x_{n+m} \qquad (12)$$

where the output y is estimated by a linear function.

Thus, the fuzzy system performs the mapping $L : \Re^{n+m} \to \Re$.

The grade of crisp variable x_i membership in the fuzzy notion LX_{ij} is determined by membership functions $\mu_{LXij}(x_i)$. The rulebase is determined by the criterion of minimum output error defined by one of the following expressions:

$$\frac{\sum_{i=1}^{K}\left|f(\mathbf{x}_i) - L(\mathbf{x}_i)\right|}{K}, \ \frac{\sqrt{\sum_{i}^{K}(f(\mathbf{x}_i) - L(\mathbf{x}_i))^2}}{K}, \ \max_{i \in K}\left|f(\mathbf{x}_i) - L(\mathbf{x}_i)\right| \qquad (13)$$

where K is the number of samples.

The choice of a fuzzy model depends on the plant's type and identification objective. For complex nonlinear dynamic plants, such as moving objects where the accuracy requirements are predominant, the choice of Takagi-Sugeno model looks reasonable. In the problems of knowledge formation from data (as linguistic rules) or the search of associative relations in a dataset, the Mamdani fuzzy system must be used. The singleton-type system may be used in both identification and knowledge formation tasks.

Singleton-type fuzzy model specified by the rules (14) performs the mapping $L : \Re^{n+m} \to \Re$ where the fuzzy conjunction operator is replaced by a product, and the operator of fuzzy rules aggregation – by summation. The mapping L is defined by the following expression:

$$L(\mathbf{x}) = \frac{\sum_{i=1}^{q} \mu_{LX\,1i}(x_1) \cdot \mu_{LX\,2i}(x_2) \cdot ... \cdot \mu_{LX\,(n+m)i}(x_{n+m}) \cdot r_i}{\sum_{i=1}^{q} \mu_{LX\,1i}(x_1) \cdot \mu_{LX\,2i}(x_2) \cdot ... \cdot \mu_{LX\,(n+m)i}(x_{n+m})}, \qquad (14)$$

where: $\mathbf{x} = [x_1,..., x_{n+m}]^T \in \Re^{n+m}$;

q is the number of rules in a fuzzy model;

$n+m$ is the number of input variables in the model;

$\mu_{LX\,ij}$ is the membership function.

The expression for L mapping in a Takagi-Sugeno model looks as folows:

$$L(\mathbf{x}) = \frac{\sum_{i=1}^{q} \mu_{LX\,1i}(x_1)\cdot \mu_{LX\,2i}(x_2)\cdot\ldots\cdot \mu_{LX\,(n+m)i}(x_{n+m})\cdot(r_{0i}+r_{1i}x_1+r_{2i}x_2+\ldots+r_{(n+m)i}x_{n+m})}{\sum_{i=1}^{q} \mu_{LX\,1i}(x_1)\cdot \mu_{LX\,2i}(x_2)\cdot\ldots\cdot \mu_{LX\,(n+m)i}(x_{n+m})} \qquad (15)$$

In Mamdani fuzzy systems, fuzzy logic techniques are used for describing the input vector's $\boldsymbol{x}$ mapping into the output value y, for example, Mamdani approximation or a method based on a formal logical proof.

Assume that one or more variables in (1) are fuzzy. In the real life, this may mean the fuzzification of weekly recommendation provided by major investment banks that in this case are considered as experts.

Generally, (1) can be represented as a fuzzy Takagi–Sugeno (TS) model [21].The fuzzy TS model consists of a set of production rules with linear finite difference equations in the right-hand member (for simplicity, a single input case, i.e., P=1, is considered):

If $y(t-1)$ is $Y_1^\theta,\ldots,$ $y(t-r)$ is Y_r^θ, $x(t)$ is $X_0^\theta,\ldots,$ $x(t-s)$ is X_r^θ, then

$$y^\theta(t) = a_0^\theta + \sum_{k=1}^{r} a_k^\theta y(t-k) + \sum_{l=0}^{s} b_l^\theta x(t-l), \qquad (16)$$

$$\theta = 1,\ldots,n$$

where $\boldsymbol{a}^\theta = (a_0^\theta, a_1^\theta, \ldots, a_r^\theta)$, $\boldsymbol{b}^\theta = (b_0^\theta, b_1^\theta, \ldots, b_s^\theta)$ are adjustable parameter vectors; $\mathbf{y}(t-r) = (1, y(t-1),\ldots, y(t-r))$ is state vector; $\boldsymbol{x}(t-s) = (x(t), x(t-1),\ldots, x(t-s))$ is an input sequence; $Y_1^\theta,\ldots,Y_r^\theta$, $X_0^\theta,\ldots,X_r^\theta$ are fuzzy sets.

By re-denoting input variables: $(u_0(t), u_1(t), \ldots, u_m(t)) = (1, y(t-1), \ldots, y(t-r), x(t),\ldots, x(t-s))$, finite difference equation's coefficients: $(c_0^\theta, c_1^\theta,\ldots,c_m^\theta) = (a_0^\theta, a_1^\theta,\ldots,a_r^\theta, b_1^\theta,\ldots,b_s^\theta)$, and membership functions:

$$\left(U_1^\theta(u_1(t)),\ldots,U_m^\theta(u_m(t))\right) =$$

$$\left(Y_1^\theta(y(t-1)),\ldots,Y_r^\theta(y(t-r)),X_0^\theta(x(t)),\ldots,X_s^\theta(x(t-s))\right)$$

where $m=r+s+1$,

one obtains the analytic form of the fuzzy model (4), intended for calculating the output $\hat{y}(t)$:

$$\hat{y}(t) = \boldsymbol{c}^T \tilde{\boldsymbol{u}}(t), \qquad (17)$$

where: $\boldsymbol{c} = (c_0^1,\ldots,c_0^n,\ldots,c_m^1,\ldots,c_m^n)^T$ is the vector of the adjustable parameters;

$$\tilde{\boldsymbol{u}}^T(t) = (u_0(t)\beta^1(t),\ldots,\ u_0(t)\beta^\theta(t),$$
$$\ldots,u_m(t)\beta^1(t),\ldots,u_m(t)\beta^n(t))$$

is the extended input vector;

$$\beta^\theta(t) = \frac{U_1^\theta(u_1(t))\otimes\ldots\otimes U_m^\theta(u_m(t))}{\sum_{\theta=1}^{N}\left(U_1^\theta(u_1(t))\otimes\ldots\otimes U_m^\theta(u_m(t))\right)} \qquad (18)$$

is a fuzzy function where $\otimes$ denotes the minimization operation of fuzzy product.

If for t = 0, the vector $\mathbf{c}(0) = 0$, the correcting $nm\times nm$ matrix $Q(0)$ (m is the number input vectors, n is the number of production rules), and the values of $u(t)$, $t = 1,\ldots, N$ are specified, the parameter vector $\mathbf{c}(t)$ is calculated using the known multi-step LSM :

$$\boldsymbol{c}(t) = \boldsymbol{c}(t-1) + Q(t)\tilde{\boldsymbol{u}}(t)[y(t) - \boldsymbol{c}^T(t-1)\tilde{\boldsymbol{u}}(t)]$$

$$Q(t) = Q(t-1) - \frac{Q(t-1)\tilde{\boldsymbol{u}}(t)\tilde{\boldsymbol{u}}^T(t)Q(t-1)}{1+\tilde{\boldsymbol{u}}^T(t)Q(t-1)\tilde{\boldsymbol{u}}(t)}$$

$Q(0) = \gamma I$, $\gamma >> 1$ where I is the unit matrix.

The above equations show that even in case of one-dimensional input and few production rules, a lot of observations are needed to apply LSM that makes the fuzzy model too unwieldy. Therefore, only a part of the whole set of rules ($r < n$) should be chosen according to a certain criterion.

The application of the associative search techniques where one or more model parameters are fuzzy, is reduced to such determination of the predicate $\Xi = \left\{\Xi_i(R_0^a, R^a, T^a)\right\}$, that the number of production rules in the TS model is significantly reduced according to some criterion.

For example, the following matrix:

$$\begin{matrix} \beta_1^{\Theta_t} & \ldots & \beta_P^{\Theta_t} \\ \ldots & \ldots & \ldots \\ \beta_1^{\Theta_{t-s}} & \ldots & \beta_P^{\Theta_{t-s}} \end{matrix}$$

can be defined for P-dimensional input vectors at time steps $t-j$, $j = 1,\ldots,s$. If the rows of this matrix are ranged, say, w.r.t. $\sum_{p=1}^{P}\left|\beta_p^{\Theta_i}\right|$ decrease and a certain number of rows are selected then such selection combined with the condition of the additive shift will determine the predicate $\boldsymbol{\Xi} = \left\{\Xi_i(R_0^a, R^a, T^a)\right\}$ and, respectively, the criterion for selecting the images (sets of input vector) from the history.

Let us range the rows of this matrix, for example, subject to the criterion of descending the values $\sum_{p=1}^{P}\left|\beta_p^{\Theta_i}\right|$, and select a certain number of rows. Such selection combined with the condition of the additive shift defines the predicate $\Xi = \left\{\Xi_i(R_0^a, R^a, T^a)\right\}$, and, respectively, the image selection criterion (sets of input vectors) from the archive.

VIII. Fuzzy associative search

Notwithstanding all benefits delivered by fuzzy techniques, their application reduces significantly the calculations speed that is critical for predicting the dynamics of some plants. This consideration coupled with the principal impossibility of formalizing some factors necessitated the development of algorithms that could combine all advantages of fuzzy approach and associative search algorithms.

Assume the associative search procedure is determined by the predicate $\Xi(P^a, R^a)$ which interprets input variables' limits (specified, say, by process specifications) as a fuzzy conjunction of input variables:

$$\Xi(P^a,R^a)= \left\{ (X_1 : x_1 \subset A_1) \wedge (X_2 : x_2 \subset A_2) ... (X_n : x_n \subset A_n) \right\}$$

for all $X_1, X_2, ... X_n$ from $DX = DX_1 \times DX_2 \times ... \times DX_n$.

Then the production rules where fuzzy variables possess such values that $\Xi(P^a, R^a)$ possesses the value FALSE, will be discarded automatically. This reduces drastically the number of production rules employed in the fuzzy model and thus increases significantly the algorithms' speed.

IX. Associative model's stability conditions derived by means of multiresolution spectrum analysis

Let a predictive associative model of a time-varying nonlinear object is described by eq. (1).

For the specified detailing level L, we obtain the following multiresolution decomposition of the current input vector $x(t)$ [16]:

$$x(t) = \sum_{k=1}^{N/2^L} c_t^x(t)\phi_{L,k}(t) + \sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^x(t)\psi_{j,k}(t), \quad N \geq 2^L \qquad (19)$$

$$y(t) = \sum_{k=1}^{N/2^L} c_k^y(t)\phi_{L,k}(t) + \sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^y(t)\psi_{j,k}(t), \quad N \geq 2^L, \qquad (20)$$

where: L is the depth of the multiresolution decomposition ($k<t$); $\phi_{L,k}(t)$ are scaling functions; $\psi_{j,k}(t)$ are wavelet functions generated form mother wavelets by stretching/compressing and shifting,

$$\psi_{jk}(t) = 2^{j/2}\psi_{mother}(2^j t - k) \qquad (21)$$

(Haar wavelets are further considered as mother ones); j is the analysis detailing level; c_{jk} and $d_{j,k}$ are scaling and detailing coefficients respectively. These coefficients are calculated by Mallat [22].

In view of the definition of discrete system's transfer function, eq. (1) is equivalent to:

$$P(z)\cdot y(t) = W(z)\cdot x(t),$$

where $P(z)$ and $W(z)$ are transfer matrices whose elements are polynomials of degrees m and r respectively ($L<t-m$), z is a one step backward shift operator.

Let $W(z)$ be a diagonal matrix whose elements are polynomials with coefficients $b_{0s},...,b_{rs}$, $s=1,...,S$.

Then

$$\begin{aligned}&[a_0 - a_1 z - a_2 z^2 - ... - a_m z^m]y(t) = \\ &\sum_{s=1}^{S}[b_0 + b_{1s}z + b_{2s}z^2 + ... + b_{rs}z^r]x_s(t)\end{aligned} \qquad (22)$$

Instead of the signals $x(t)$ and $y(t)$ we substitute multiresolution decompositions of input and output signals in this equation. As far as the first sums in (5) and (6) vanish and can be omitted at sufficiently large detailing level, (5) and (6) imply

$$P(z)[\sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^y(t)\psi_{j,k}(t)] = W(z)[\sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^x(t)\psi_{j,k}(t)]$$

or, in view of (22):

$$\begin{aligned}&a_0\sum_{j=1}^{L}\sum_{k=1}^{N/2^L} d_{j,k}^y(t)\psi_{jk}(t) = \\ &\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}\{a_1\cdot d_{j,k}^y(t-1) + \sum_{s=1}^{S} b_{1s}d_{j,k}^{x_s}(t-1)\}\psi_{j,k}(t-1) + ... \\ &+\sum_{j=1}^{L}\sum_{k=1}^{N/2^j}\{a_m d_{j,k}^y(t-m) + \sum_{s=1}^{S} b_{ms}d_{j,k}^{x_s}(t-m)\}\psi_{j,k}(t-m) + \\ &\sum_{s=1}^{S} b_{m+1,s}\sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^x(t-m-1)\psi_{j,k}(t-m-1) - \\ &+ ... + \sum_{s=1}^{S} b_{rs}\sum_{j=1}^{L}\sum_{k=1}^{N/2^j} d_{j,k}^x(t-r)\psi_{j,k}(t-r)\end{aligned} \qquad (23)$$

where x_s denotes the belonging of the corresponding coefficients to the s-th component of the input vector.

The dynamic object described by eq. (23) will be stable if the following objects (corresponding to the following expressions for each summand with respect to k $(k = 1,...,N/2^j)$ and j, $(j = 1,...,L)$ in the left and right sides of eq. (23):

$$\begin{aligned} a_0 d_{jk}^y(t)\psi_{jk}(t) &= \{a_1 \cdot d_{j,k}^y(t-1) + \sum_{s=1}^S b_{1s} d_{j,k}^{x_s}(t-1)\}\psi_{j,k}(t-1) + ... \\ &+\{a_m d_{j,k}^y(t-m) + \sum_{s=1}^S b_{ms} d_{j,k}^{x_s}(t-m)\}\psi_{j,k}(t-m) + ... \\ &+\sum_{s=1}^S b_{m+1,s} d_{j,k}^{x_s}(t-m-1)\psi_{j,k}(t-m-1) - \\ &+ ... + \sum_{s=1}^S b_{rs} d_{j,k}^{x_s}(t-r)\psi_{j,k}(t-r) \end{aligned} \quad (24)$$

are stable altogether.

In other terms, as far as $a_0 d_k^y(t) \neq 0$, we have:

$$\begin{aligned} \psi_{jk}(t) &= \{[a_1 \cdot d_{j,k}^y(t-1) + \sum_{s=1}^S b_{1s} d_{j,k}^{x_s}(t-1)] / a_0 d_k^y(t)\}\psi_{j,k}(t-1) + ... \\ &+\{[a_m d_{j,k}^y(t-m) + \sum_{s=1}^S b_{ms} d_{j,k}^{x_s}(t-m)] / a_0 d_k^y(t)\}\psi_{j,k}(t-m) + ... \\ &+[\sum_{s=1}^S b_{m+1,s} d_{j,k}^{x_s}(t-m-1) / a_0 d_k^y(t)]\psi_{j,k}(t-m-1) + \\ &+ ... + [\sum_{s=1}^S b_{rs} d_{j,k}^{x_s}(t-r) / a_0 d_k^y(t)]\psi_{j,k}(t-r) \ . \end{aligned} \quad (25)$$

To make it simple, we temporarily omit the indices k and j in (25) and denote:

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} [\tilde{x}_1(t)] \\ ... \\ [\tilde{x}_r(t-r+1)] \end{bmatrix}, \quad \tilde{\mathbf{x}}(t) \in \mathbf{R}^r;$$

Where

$$\begin{aligned} \tilde{x}_1(t) &= \psi(t) \\ \tilde{x}_2(t) &= \psi(t-1); \end{aligned} \quad (26)$$

$\tilde{x}_r(t) = \psi(t-r+1)$, then:

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} [\psi(t)] \\ ... \\ [\psi(t-r+1)] \end{bmatrix}; \quad \tilde{\mathbf{x}}(t-1) = \begin{bmatrix} [\psi(t-1)] \\ ... \\ [\psi(t-r)] \end{bmatrix}.$$

Further, we denote:

$$\begin{aligned} c_1 &= \{a_1 d^y(t-1) + \sum_{s=1}^S b_{1s} d^{x_s}(t-1)\} / a_0 d^y(t); ... \\ c_m &= \{a_m d^y(t-m) + \sum_{s=1}^S b_{ms} d^{x_s}(t-m)\} / a_0 d^y(t); \\ c_{m+1} &= \sum_{s=1}^S b_{m+1,s} d^{x_s}(t-m-1) / a_0 d^y(t); ... c_m = \\ &= \sum_{s=1}^S b_{rs} d^{x_s}(t-r) / a_0 d^y(t) \end{aligned} \quad (27)$$

and rewrite (25) as:

$$\begin{aligned} &\psi(t) - \frac{c_1}{2}\psi(t-1) - ... - \frac{c_{r-1}}{2}\psi(t-r+1) = \frac{c_1}{2}\psi(t-1) + ... + \frac{c_r}{2} \\ &+c_m \cdot \psi(t-m) + c_{m+1}\psi(t-m-1) + c_r \cdot \psi(t-r) \ . \end{aligned} \quad (28)$$

Simultaneous fulfillment of the equalities

$$\begin{aligned} &\psi(t) = \frac{c_1}{2}\psi(t-1); -\frac{c_1}{2}\psi(t-1) = \frac{c_2}{2}\psi(t-2); \\ &-\frac{c_{m-1}}{2}\psi(t-m+1) = \frac{c_m}{2}\psi(t-m); ... \\ &... - \frac{c_{r-1}}{2}\psi(t-m+1) = \frac{c_r}{2} \cdot \psi(t-r) \end{aligned}$$

is a sufficient condition for the fulfillment of eq. (28).
With the above notification we have

$$\begin{bmatrix} 1 & & 0 & 0 \\ 0 & -\frac{c_1}{2} & & ... \\ ... & & ... & 0 \\ 0 & & & -\frac{c_{r-1}}{2} \end{bmatrix} \tilde{\mathbf{x}}(t) = \begin{bmatrix} \frac{c_1}{2} & & 0 & 0 \\ 0 & ... & & ... \\ ... & & ... & 0 \\ 0 & & & -\frac{c_r}{2} \end{bmatrix} \cdot \tilde{\mathbf{x}}(t-1) \quad (29)$$

Let the matrix in the left side of (29) be invertible. Then:

$$\tilde{\mathbf{x}}(t)=\begin{bmatrix}1 & 0 & 0 \\ 0 & -\frac{c_1}{2} & \dots \\ \dots & \dots & 0 \\ 0 & & -\frac{c_{r-1}}{2}\end{bmatrix}^{-1}\cdot\begin{bmatrix}\frac{c_1}{2} & 0 & 0 \\ 0 & \frac{c_2}{2} & \dots \\ \dots & \dots & 0 \\ 0 & & \frac{c_r}{2}\end{bmatrix}\cdot\tilde{\mathbf{x}}(t-1)=$$

$$=\begin{bmatrix}1 & 0 & 0 \\ 0 & -\frac{2}{c_1} & \dots \\ \dots & \dots & 0 \\ 0 & & -\frac{2}{c_{r-1}}\end{bmatrix}^{-1}\cdot\begin{bmatrix}\frac{c_1}{2} & 0 & 0 \\ 0 & \frac{c_2}{2} & \dots \\ \dots & \dots & 0 \\ 0 & & \frac{c_r}{2}\end{bmatrix}\cdot\tilde{\mathbf{x}}(t-1)=$$

$$=\begin{bmatrix}\frac{c_1}{2} & 0 & 0 \\ 0 & -\frac{c_2}{c_1} & \dots \\ \dots & \dots & 0 \\ 0 & & -\frac{c_r}{c_{r-1}}\end{bmatrix}\cdot\tilde{\mathbf{x}}(t-1) \tag{30}$$

We denote $\hat{\mathbf{x}}(t)=\tilde{\mathbf{x}}(t)\cdot d^y(t)$. By multiplying both sides of eq. (30) by $d^y(t)$ and then multiplying and dividing its right side by $d^y(t-1)$ we obtain:

$$\hat{\mathbf{x}}(t)=\begin{bmatrix}\frac{c_1}{2}\cdot\frac{d^y(t)}{d^y(t-1)} & 0 & 0 \\ 0 & -\frac{c_2}{c_1}\cdot\frac{d^y(t)}{d^y(t-1)} & \dots \\ \dots & .. & 0 \\ 0 & & -\frac{c_r}{c_{r-1}}\cdot\frac{d^y(t)}{d^y(t-1)}\end{bmatrix}\cdot\hat{\mathbf{x}}(t-1) \tag{31}$$

Eq. (31) can be interpreted as a system's state-space representation. The system's stability is determined by the properties of the characteristic polynomial of the respective diagonal matrix.

Thus, a sufficient stability condition for the object (25) (and, hence, (23) for any $k=1,...,N/2^j$) is ensured by the fulfillment of all inequalities

$$\left|\frac{c_1}{2}\frac{d^y(t)}{d^y(t-1)}\right|<1 \ ;\text{ or } |c_1|\cdot\left|d^y(t)\right|<2\cdot\left|d^y(t-1)\right|;$$

$$\left|\frac{-c_2}{c_1}(t)\frac{d^y(t)}{d^y(t-1)}\right|<1,...,\text{или } |c_2|\left|d^y(t)\right|<|c_1|\left|d^y(t-1)\right|, \tag{32}$$

$$\left|\frac{-c_r}{c_{r-1}}\frac{d^y(t)}{d^y(t-1)}\right|<1,...,\text{или } |c_r|\left|d^y(t)\right|<|c_{r-1}|\left|d^y(t-1)\right|,$$

or

$$\left|a_1d^y(t-1)+\sum_{s=1}^{S}b_{1s}d^{x_s}(t-1)\right|/\left|a_0d^y(t)\right|\cdot\left|d^y(t)\right|<$$
$$<2\left|d^y(t-1)\right|;$$
$$\left|a_md^y(t-m)+\sum_{s=1}^{S}b_{ms}d^{x_s}(t-m)\right|/\left|a_0d^y(t)\right|\cdot\left|d^y(t)\right|<$$
$$\left|a_{m-1}d^y(t-m+1)+\sum_{s=1}^{S}b_{m-1s}d^{x_s}(t-m+1)\right|/$$
$$/\left|a_0d^y(t)\right|\cdot\left|d^y(t-1)\right|;$$
$$\left|\sum_{s=1}^{S}b_{m+1,s}d^{x_s}(t-m-1)/a_0d^y(t)\right|\left|d^y(t)\right|<$$
$$<\left|\sum_{s=1}^{S}b_{m,s}d^{x_s}(t-m)/a_0d^y(t)\right|\left|d^y(t-1)\right|;$$
$$\dots\left|\sum_{s=1}^{S}b_{rs}d^{x_s}(t-r)/a_0d^y(t)\right|\left|d^y(t)\right|<$$
$$<\left|\sum_{s=1}^{S}b_{r-1,s}d^{x_s}(t-r+1)/a_0d^y(t)\right|\left|d^y(t-1)\right|.$$

Therefore, the sufficient stability conditions for the object (22) for all coefficients of multiresolution wavelet decomposition of its input and output at a specified time step t look as follows:

$$\left|a_1d^y(t-1)+\sum_{s=1}^{S}b_{1s}d^{x_s}(t-1)\right|<2\left|a_0d^y(t-1)\right|;$$
$$\left|a_2d^y(t-2)+\sum_{s=1}^{S}b_{2s}d^{x_s}(t-2)\right|\left|d^y(t)\right|<$$
$$\left|a_1d^y(t-1)+\sum_{s=1}^{S}b_{1s}d^{x_s}(t-1)\right|\left|d^y(t-1)\right|;$$
$$\dots$$
$$\left|a_md^y(t-m)+\sum_{s=1}^{S}b_{ms}d^{x_s}(t-m)\right|\left|d^y(t)\right|<$$
$$\left|a_{m-1}d^y(t-m+1)+\sum_{s=1}^{S}b_{m-1s}d^{x_s}(t-m+1)\right|\left|d^y(t-1)\right|;$$
$$\left|\sum_{s=1}^{S}b_{m+1,s}d^{x_s}(t-m-1)\right|\left|d^y(t)\right|<$$
$$<\left|\sum_{s=1}^{S}b_{m,s}d^{x_s}(t-m)\right|\left|d^y(t-1)\right|;$$
$$\dots$$
$$\left|\sum_{s=1}^{S}b_{rs}d^{x_s}(t-r)\right|\left|d^y(t)\right|<$$
$$<\left|\sum_{s=1}^{S}b_{r-1,s}d^{x_s}(t-r+1)\right|\left|d^y(t-1)\right|.$$

X. STABILITY CONDITIONS FOR AN ASSOCIATIVE MODEL OF A MULTIMODAL OBJECT

A system approach to the study of the operation of large manufacturing and power plants can be effectively interpreted in terms of multimodal objects.

A feature of multimodal objects is the possibility of predicting their future operation based on the knowledge generated at the learning stage. These knowledge formalize the dynamics of process variables specific to a certain operating mode. The "belonging" to the operating mode is determined when a predictive model is built using clusterization.

Examples of multimodal objects are: product lifecycle (manufacturing and logistical aspects), the operation of a power generation system or its subsystems in normal, pre-emergency, emergency and post-emergency modes, oil pumping through a piping system with changing topology, etc.

The key attributes of multimodal systems are independent production stages and/or multiple process modes.

In both cases, the plant can be described by an extended input vector $\widehat{\mathbf{x}} = x_{t1}^{1} \ldots x_{tS}^{1} \ldots x_{t1}^{m} \ldots x_{tS}^{m} \ldots x_{t1}^{M} \ldots x_{tS}^{M}$, where the index t denotes the instant of discrete system operation, m is the mode (manufacturing stage, operating mode, etc.) number, s is the number of the input vector's component. The mode is characterized by input vector's belonging to one of disjoint domains in R^S space.

Here:

- plant operation in a specific mode m characterized by certain input parameters such as flowrates, pressures or temperatures presumes that all input vector's components except for $x_{t1}^{m} \ldots x_{tS}^{m}$ are equal to zero;
- for a multistage process, the input vector's components may have different nature and possess different values dependent on the process stage. In particular, specific groups of input variables may change while the others remain constant or exceed the bounds of a certain subspace.

In order to predict the object's violation of safe operating boundaries the following scheme is applied:

1) Associative search technique is used to build a linear predictive model of a nonlinear object with extended input vector $\widehat{\mathbf{x}}$.
2) The object with predicted properties is investigated.

In case of time-varying plant, it makes sense to examine the spectrum of a multiresolution wavelet decomposition of the forecast of the object's output signal.

The stability of linear dynamic objects can be investigated by means of the method described above or using Gramian technqiue. According to the latter, the approach to the operating boundary is associated with unrestricted growth of Frobenius controllability gramian for object's certain state-space realization.

Associative search allows to apply this technique to the nonlinear plant study. It was proved in [23] that the investigation of the Frobenius norm of dynamic plant's transfer function would be sufficient.

Sometimes it makes sense to investigate the operation of nonlinear dynamic plants by analyzing their wavelet representations and multiresolution wavelet decompositions.

In the present work, the sufficient conditions of the stability of time-varying object's predicted output were formulated in terms of transfer function's wavelet spectrum.

Wavelet transformation coefficients reveal the signal's fluctuating structure in various scales and instants of time. This is especially important for time-varying plants. The sufficient conditions of stability formulated in this paper were based on the investigation of associative predictive model's wavelet spectrum.

XI. APPLICATION

The presented methods were successfully applied in soft sensor design for chemical and oil refining processes. The approach proposed in [14] is based on virtual models and associative search techniques. A fuzzy model is applied in combination with production knowledgebase to compensate for the lack of lab data.

A. Soft sensor describing the dynamic behavior of power grid's generation facilities participation in the overall primary frequency regulation in contingencies

The paper [15] presents a soft sensor describing the dynamic behavior of power grid's generation facilities participation in the overall primary frequency regulation in contingencies. The soft sensor is based on generating capacity and frequency time series and was developed by the authors for the Control Center of Russia's Unified Energy System (RAO UES) of Russia.

The establishment of a power plant's participation and the estimation of the degree of its contribution to the overall primary frequency regulation (OPFR) is performed at the frequencies exceeding 0.2 Hz. When the grid is operated in design mode (with frequency deviations less than 0.2 Hz), the control is purely qualitative and informative.

At the same time, the qualitative assessment of generating facilities' participation in the OPFR at abrupt frequency excursions in the grid in the range 0.05…0.2 Hz, the systematic (more than 50% cases over a year) nonparticipation in the primary regulation of the generating facilities from a number of heat power plants was detected due to the lack of the requisite power adjustments to compensate for the frequency deviations.

Along with the establishment the fact of specific generating facility's participation (or nonparticipation) in the OPFR, the technology of RAO UES Control Center's historical data processing enables the evaluation of the degree of plant's participation in the regulation.

To rank the generating facilities' participation in the OPFR, identification models and algorithms describing power grid's dynamics were developed.

When building the identification model, essential non-linearity of the object under investigation was allowed for. Therefore, it was found rational to use associative search models in the soft sensor design.

The identification algorithms show the generation facilities ranking w.r.t. the probabilities of their violation of the OPFR participation requirements to the generating facilities. The indicators of specific generating facilities influence on the OPFR were evaluated that contributed to the quality improvement of the secondary frequency regulation.

An intelligent system intended for dynamic state estimation of a complex power grid was created per the EU project "INTELLIGENT COORDINATION OF OPERATION AND EMERGENCY CONTROL OF EU AND RUSSIAN POWER GRIDS"(ICOEUR, FP7-ENERGY) will be running from 2007 to 2013. The system is underpinned by intelligent algorithms of grid dynamics identification with automatic on-line self-tuning based on the data from monitoring systems.

State estimation models for power facilities with on-line model tuning are based on data monitoring and application of a new predictive method for state estimation – the associative search method.

The acquisition, storage, processing, displaying, analysis and documenting of the information are executed in real time based on the data from automated power generation, distribution and consumption systems and supervisory control, monitoring and accounting.

The development of intelligent dynamic state estimation algorithms based on the use of process knowledge for important power plant and power network control tasks such as the optimization of generating equipment and power grid optimization is in sight.

Those will be underlain by fuzzy models; virtual object models using the associative search method will be also employed.

State estimation models, customizable during real-time operation, can be used both in the automatic mode of a control system, and to support management decisions.

Based on the dynamically configurable state models, power facility operation modes can be optimized over the whole of power grid including all power market participants subject to reliability and profitability criteria.

B.Predicting the qualities of delayed coking unit distillates

Quality control of delayed coker distillates for their subsequent utilization as hydrotreater's feedstock is a challenge. The reason is that coker's fractionator was originally designed for some average feed rate and quality, while in real life both change several times per day sufficiently for making a serious disturbance for downstream process equipment.

The traditional control strategy for a fractionator under disturbances is temperature profile stabilization closer to the steady-state values established at design phase for average feed rate and quality, and their further slight adjustment subject to lab data. The product samples are analyzed by refinery lab 3–6 times per day. This makes the control strategy ineffective because the object's state cannot be identified unambiguously from such scanty samples.

Soft sensors (SS)-based virtual models were built for this plant using both process history and lab data. Those models enabled on-line calculation of desirable product qualities with sufficient accuracy. This resulted in process unit's throughput increase combined with more consistent product quality.

The SS-based quality analyzers were built for coker naphtha IBP, 50%, 90% distillation points, and EP and coker gas oil IBP, 50%, and 90% distillation points. Based on these, a predictive model structure for distillation points of key product streams was obtained as well as the forecast accuracy estimate.

The forecast was calculated using a mathematical model whose inputs were process variable measured on-line. The forecast accuracy depends on right selection of informative variables, memory depth, and the amount of available plant data.

Typically, the precise forecast is impossible for such complex objects as fractionator because the existing measurements do not observe all factors affecting the product qualities. For example, there were no tools at the process unit to measure feedstock makeup changes. In such case, the informative variables had to be selected from the vast amount of data. This was done using process history.

At design operation of the model-based predictor, its adaptation to plant dynamics and input properties changes is executed automatically with the changes of the nonlinear model's structure, while its dynamic depth remains the same.

Figure 3 shows an example of a predictive model for coker naphtha 50% distillation point (ASTM D86).

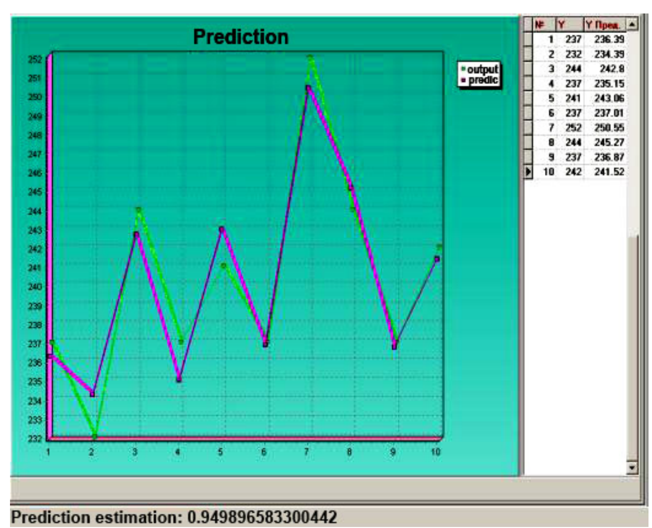

Fig. 3. Naphtha 50% distillation point model

The SS-based control system can calculate control actions with adaptive models adjustment.

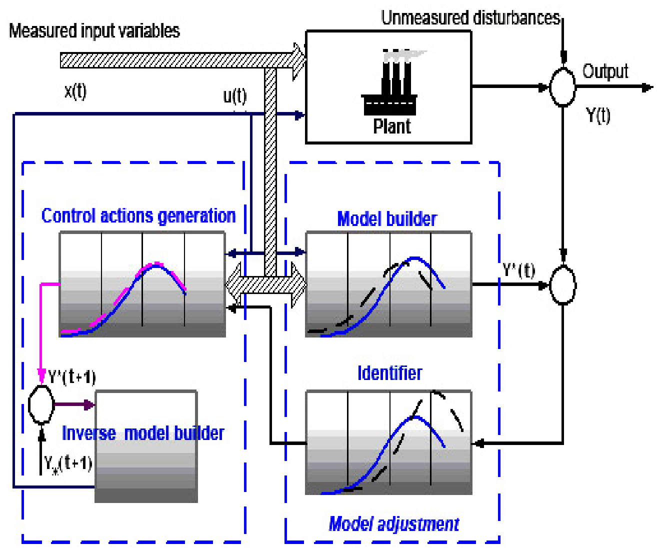

Fig. 4. Adaptive control system with identifier

After site acceptance tests in advising mode are complete, the recommended control action can be further used in the closed loop, i.e., in the automatic control system with an identifier (Figure 4).

CONCLUSION

Model predictive control is an effective method both in terms of efficiency and safety of technical systems.

In [21] the intellectualization of identification methods, in particular, the employment of additional a priori information about the object for system teaching, is outlined as a key trend. Associative search algorithms based on the design of virtual predictive models, which use intelligent analysis of historical process data for dynamic tuning of identification models, can be attributed to this type of methods.

The methods proposed for building predictive models of complex multimodal production processes prevent dangerous approaches to stability threshold even for nonlinear non-stationary objects.

REFERENCES

[1] Bellman R. and Kalaba R. (1960). Dynamic programming and adaptive control processes: mathematical foundations. IRE Trans. on Automatic Control, vol. AC-5, pp. 5–10.

[2] Bunich, A.L. and Lototsky V.A. (2002). Fast convergent identification algorithm for limited noise. In: Proc. of 15th IFAC World Congress, Barcelona, Spain.

[3] Goodwin, G., Ramadge, P., and Caines, P. (1981). Discrete time stochastic adaptive control, SIAM Jornal on Control and Optimization, vol. 19, pp. 829 853.

[4] Juditsky A., Nazin A. (2001). On minimax approach to nonparametric adaptive control. Int.J. Adapt. Control & Signal Process, vol.15, pp.153 – 168.

[5] Ljung, L. (1999). System identification. Theory for user. Prentice Hall, Upper Saddle River, N.J., 2nd edition, 609 p.

[6] Poznyak, A. S. and Kaddour N. (1997). Learning Automata and Stochastic Optimization. Springer Verlag, New York, N.Y.

[7] Chen, H.F. (1984). Recursive system identification and adaptive control by use of the modified least squares algorithm. SIAM J. Control and Optimization, vol.22, no.5, pp.758 – 776.

[8] Ljung, L. (2008). Perspectives on System Identification / Plenary on 17th IFAC World Congress, Seoul, Korea.

[9] Larichev, O. I., Asanov, A., Naryzhny, Y., and Strahov, S. (2001). Expert system for the diagnostics of acute drug poisonings. In: Preprints of the 21th SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence, pp. 159—168.

[10] Patel, V.L. and Ramoni M.F. (1997). Cognitive models of directional inference in expert medical reasoning. In: Feltovich P., Ford K., Hofman R. (ed.), Expertise in Context: Human and Machine. AAAI Press, Menlo Parc, CA.

[11] Hunt E. (1989). Cognitive Science: Definition, Status and Questions. Annual Review of Psychology, vol. 40, pp. 603-629.

[12] Simon, H.A. (1997). The Sciences of the Artificial (3rd ed.). The MIT Press, Cambridge, MA.

[13] Gavrilov A.V. (2002). The model of associative memory of intelligent system. In: Preprints of the 6-th Russian-Korean International Symposium on Science and Technology, vol. 1, pp. 174-177.

[14] Bakhtadze N., Lototsky V., Maximov E., and Pavlov, B. (2007). Associative search models in industrial systems. In: Proc. of IFAC workshop on intelligent manufacturing systems, pp. 156-161.

[15] Bakhtadze, N., Yadikin, I., Kulba, V., Lototsky, V., and Maximov, E. (2008). Intelligent Power Generation Systems Control. Proc. of 9th IFAC Workshop on Intelligent Manufacturing Systems, Szczecin, Poland .P. 24-32.

[16] Daubechies, I. and Lagarias, J. (1991). Two-scale diference equations I. Existence and global regularity of solutions. SIAM J. Mathematical Analysis, volume no 22 Issue 5, pp. 1388-1410.

[17] Bakhtadze, N., Kulba, V., Lototsky, V., and Maximov, E. (2011). Identification Methods Based on Associative Search Procedure. Control and Cybernetics, vol. 2, no. 3, pp.6-18.

[18] Bakhtadze N.N., Pavlov B.V., Sakrutina E.A.(2013). Development of Intelligent Identification Models and their Applications to Predict the Submarine Dynamics by Use of Computer Simulation Complexes // Proceedings of the 7-th IFAC Conference on Manufacturing Modelling, Management, and Control (MIM`2013), Saint Petersburg: IFAC Publication. P. 1244-1249.

[19] Bakhtadze N.N., Lototsky V.A., Vlasov S.A., Sakrutina E.A. Associative Search and Wavelet Analysis Techniques in System Identification // Proceedings of the 16-th IFAC Symposium on System Identification, 2012. Brussels, Belgium: IFAC Publication. P. 1227-1232.

[20] M. J. Er and Y. Gao (2003). Online Adaptive Fuzzy Neural Identification and Control of Nonlinear Dynamic Systems. In Fusion of Soft Computing and Hard Computing for Autonomous Robotic Systems edited by C. Zhou, D. Maravall and D. Ruan, New York, Heidelberg: Physica-Verlag, pp. 373-402.

[21]Bakhtadze, N., Kulba, V., Lototsky, V., and Maximov, E. (2011). Identification Methods Based on Associative Search Procedure. Control and Cybernetics, vol. 2, no. 3, pp.6-18.

[22] Mallat, S. (1999). Wavelet Tour of Signal Processing. Barlaud, M. (Ed.), 635 p. Academic Press, San Diego, CA.

[23] Sukhanov, O.A., Novitsky, D.A. and Yadikin, 2012. I. Method of Steady-state Stability Analysis in Large Electrical Power Systems / Proceedings of 17th Power Systems Computation Conference PSCC2011, Stockholm. Stockholm: Curran Associates, Inc. T.2. P. 934-941.

Professor Natalia N Bakhtadze is Head of Identification Laboratory of V V.A. Trapeznikov Institute of Control Sciences, Moscow, RUSSIA. PhD'1993; defended her doctoral thesis in 2004.
She is author of about two hundred scientific publications, including
1) Virtual Analyzers: Identification Approach
Automation and Remote Control, 2004, 65:11, 1691–1709 ..

2) Identification methods based on associative search procedure
N. N. Bakhtadze, V. V. Kulba, I. B. Yadikin, V. A. Lototsky, more
Management and Production Engineering Review > 2011 > Vol. 2, No. 3 > 6-18

3) An approach to coordination of power generating units. Natalia N. Bakhtadze, Igor B. Yadikin, Vladimir V. Kulba, Vladimir A. Lototsky and Evgeny M. Maximov Control and Cybernetics, vol. 39 (2010) No. 1

Prof. N.Bakhtadze is Vice-chair of IFAC