

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Data Mining-Based Identification of Nonlinear Systems

*Natalia Bakhtadze, Vladimir Lototsky, Valery Pyatetsky
and Alexey Lototsky*

Abstract

This chapter presents identification methods using associative search of analogs and wavelet analysis. It investigates the properties of data mining-based identification algorithms which allow to predict: (i) the approach of process variables to critical values and (ii) process transition to chaotic dynamics. The methods proposed are based on the modeling of human operator decision-making. The effectiveness of the methods is illustrated with an example of product quality prediction in oil refining. The development of fuzzy analogs of associative identification models is further discussed. Fuzzy approach expands the application area of associative techniques. Finally, state prediction techniques for manufacturing resources are developed on the basis of binary models and a machine learning procedure, which is named associative rules search.

Keywords: process identification, knowledge base, associative search models, wavelet analysis

1. Introduction

The reduction of uncertainty in object description in terms of adjustable model has been a key conceptual direction in the identification theory and applications for a long time. In the statistical description of uncertainty, consistent estimates of plant's characteristics can be obtained by analyzing the convergence of the empirical distribution functional with the corresponding "theoretical" values, but this entails appropriate increase of the sample size. The difficulties in implementing this approach, especially for nonlinear and nonstationary objects, along with the increased possibilities of plant history analysis resulted in the advent of identification methods based on data mining [1].

The use of additional a priori information on the system for its training is considered by some authors today to be one of the key trends in the theory and practice of identification [2, 3].

One method that implements this approach to identification is the associative search method based on the design of predictive models [2]. They are based on inductive learning, that is, on associative search of analogs by means of intelligent analysis of process history and knowledge base development. The development of a predictive model for a dynamic object by associative search technique (i.e., by building a new model at every time step) is based on the generated and updated

knowledge about the system. This approach allows to use any available a priori information about the plant [3].

The stability of a model built using the associative search techniques is investigated in terms of the spectrum analysis of a *multi-scale wavelet expansion* [4]. Methods based on the wavelet analysis open up a unique possibility to select “frequency-domain windows” as against the well-known windowed Fourier transform.

The development of intelligent identification algorithms for nonlinear and nonstationary objects is important for various applications, in particular, in chemical, oil refining, and power (smart grids) industries; transportation and logistics system; and trading processes (Bakhtadze et al. [1, 2, 4–7]).

2. Control system identification

Consider a traditional problem of dynamic object identification. For input vectors meeting Gauss-Markov assumptions, the least squares parameter estimates are consistent, unbiased, and efficient. However, the development of a closed-loop control system (for identification-based control system synthesis) faces considerable challenges. In a closed loop, the system state depends on control values at earlier time instants, which results in a degeneration problem.

To develop an informational model of control system’s dynamics in a degenerate case, the Moore-Penrose method [8, 9] can be used for getting pseudo-solutions to a linear system by means of least squares techniques.

For a wide class of objects and, in particular, processes, control based on a linear model identification is not satisfactory. At the same time, models constructed by the method of associative search frequently are highly accurate even for nonlinear objects. However, some processes can be characterized by certain “irregularities” in certain time intervals, which affect the accuracy and adequacy of associative models.

Examples of such irregularities (which are often oscillatory in engineering systems) can be:

- seasonal and daily load oscillations in power networks that affect directly the optimization of power transmission control modes;
- ups and downs of stock market caused by various economic reasons;
- feed source changes in industrial process, and so on.

3. Associative search as intelligent modeling method

The difference between the associative search method based on data mining and traditional identification techniques is as follows. The method does not approximate process dynamics in time; it rather builds a new predictive model of the dynamic object (a “virtual model”) at each time step using historical data sets (“associations”) generated at the training phase.

As a result, at any time step, process control decision-making by a human individual (process operator, supervisor, plant or enterprise manager, trading operator, etc.) is modeled on the base of his/her knowledge and emerging associations.

Clustering (self-organizing learning) is an effective way to form associations.

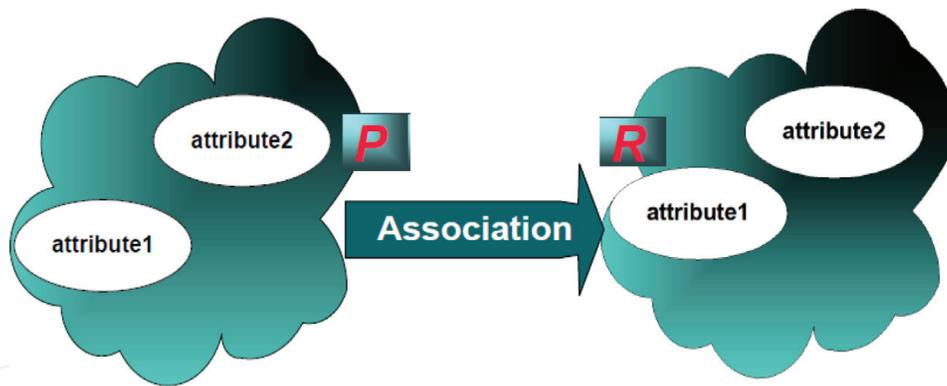


Figure 1.
Model of a human associative memory.

Knowledge in intelligent systems is of two types [10]. The first type of knowledge, that is, declarative knowledge, by means of appropriate ontologies describes different facts, events, and observation. A formal description of skills is called procedural knowledge. Depending on the level of this knowledge, users can be referred to as beginners or experts [11]. These two groups have different structures and ways of thinking. Beginners use so-called *inverse reasoning* in the procedure for decision-making. They make decisions based on the analysis of the information obtained in the previous step. In contrast to the beginners, experts at an intuitive, subconscious level form the so-called direct reasoning. Thus, cognitive psychology defines knowledge as a collection of symbols stored in the memory of a particular person [12]. The symbols, in turn, can be determined by their structure and the nature of neuron links [13].

Knowledge processing in an intelligent system consists in the recovery (associative search) of knowledge by its fragment [14]. The knowledge can be defined as an associative link between images (**Figure 1**). As an image, we will use “feature sets,” that is, components of input vectors or input variables. The set of all associations over the set of images forms the memory of the intelligent system’s knowledge base.

The associative search process can be either an image reconstruction procedure by a feature set (this set may not be complete; this approach is often used in models of a human associative memory), or the search procedure of other images in the archive, similar to the image under study by a certain criterion.

In Ref. [14], a model of decision-making search by the human operator is proposed, representing the process of associative thinking as a sequence of sets of associations. Association is a pair of images (the image-source and the image-output), wherein each image is described by a set of features. This approach is intermediate between neural networks and logical models in the classical theory of artificial intelligence.

The criterion for the similarity of two images in the general case can be represented as a logical function—a predicate. In the particular case, the features have a numerical expression. The feature sets that form the image are vectors in n -dimensional space. In this case, as a criterion of image similarity can be a metric in the space.

4. Associative search technique

Associative search method consists in constructing *virtual* predictive models. The term “virtual” should be understood as “ad hoc” [2]. The method presumes the construction of a predictive model for a dynamic object as follows. A traditional

identification algorithm approximates real process in time. As against such algorithms, our method builds a new model at each time step t based on the analysis of the history data set (“associations”) formed at the stage of learning and further adaptively corrected in accordance to certain criteria.

Within the present context, linear dynamic model is of the form:

$$y_N = \sum_{i=1}^m a_i y_{N-i} + \sum_{j=1}^{r_s} \sum_{s=1}^S b_{j,s} x_{N-j,s}, \forall j = 1, \dots, N, \quad (1)$$

where y_N is the prediction of the object’s output at the time instant N , x_N is the input vector, m is the memory depth in the output, r_s is the memory depth in the input, S is the dimension of the input vectors, and a_i and $b_{j,s}$ are tuning coefficients of the model. Model (1) is a regression whose structure is determined by a criterion of similarity of images forming the association.

In general, a new structure is formed for each time instant. The associative model is virtual in the sense that for each time step, it formed a new structure. For each current input vector, the corresponding input vectors and their corresponding outputs are selected from the archive. Further, a system of linear equations with respect to the adjustable coefficients is formed. Its decision in accordance with the least squares method determines the **point linear model of a nonlinear object**, as well as the output forecast.

Thus, each point of the global nonlinear regression surface is formed as a result of using linear “local” models at each new time step.

The set of values of inputs at each fixed point and the corresponding output replenish the procedural knowledge base.

Unlike classical regression models, for each fixed time instant from the process history, input vectors are selected close to the current input vector in the sense of a certain criterion (rather than the chronological sequence as in regression models). Thus, in Eq. (1), r_s is the number of vectors from the archive (from the time instant 1 to the time instant N), selected in accordance to the associative search criterion. A certain set of vectors r_s , $1 \leq r_s \leq N$, is selected at each time segment $[N - 1, N]$. The criterion for selecting the input vectors from the archive is described below (Figure 2).

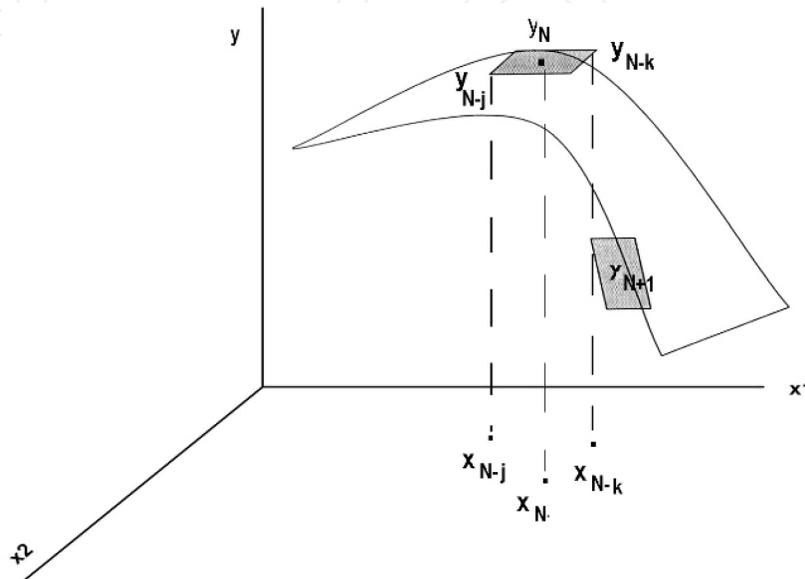


Figure 2.
Approximating hypersurface design.

As a distance (a norm in R^S) between points of the S -dimensional space of inputs, we introduce the value:

$$d_{N,N-j} = \sum_{s=1}^S |x_{N,s} - x_{N-j,s}|, \forall j = 1, \dots, N, \quad (2)$$

where $x_{N,s}$ are the components of the input vector at the current time instant N . By virtue of a property of the norm (“the triangle inequality”), we have:

$$d_{N,N-j} \leq \sum_{s=1}^S |x_{N,s}| + \sum_{s=1}^S |x_{N-j,s}|, \forall j = 1, \dots, N, \quad (3)$$

Let for the current input vector x_N :

$$\sum_{s=1}^S |x_{N,s}| = d_N. \quad (4)$$

To derive an approximating hypersurface for the vector x_N , we select from the archive of the input data such vectors $x_{N-j}, j = 1, \dots, N$ that for a set D_N the condition:

$$d_{N,N-j} \leq d_N + \sum_{s=1}^S |x_{N-j,s}| \leq D_N, \forall j = 1, \dots, N, \quad (5)$$

holds, where D_N may be selected, for instance, from the condition (**Figure 3**):

$$D_N \geq 2d_N^{\max} = 2 \max_j \sum_{s=1}^S |x_{N-j,s}|. \quad (6)$$

Under the assumptions that the inputs meet the Gauss-Markov conditions, the estimates obtained via the LS method are unbiased and statistically effective.

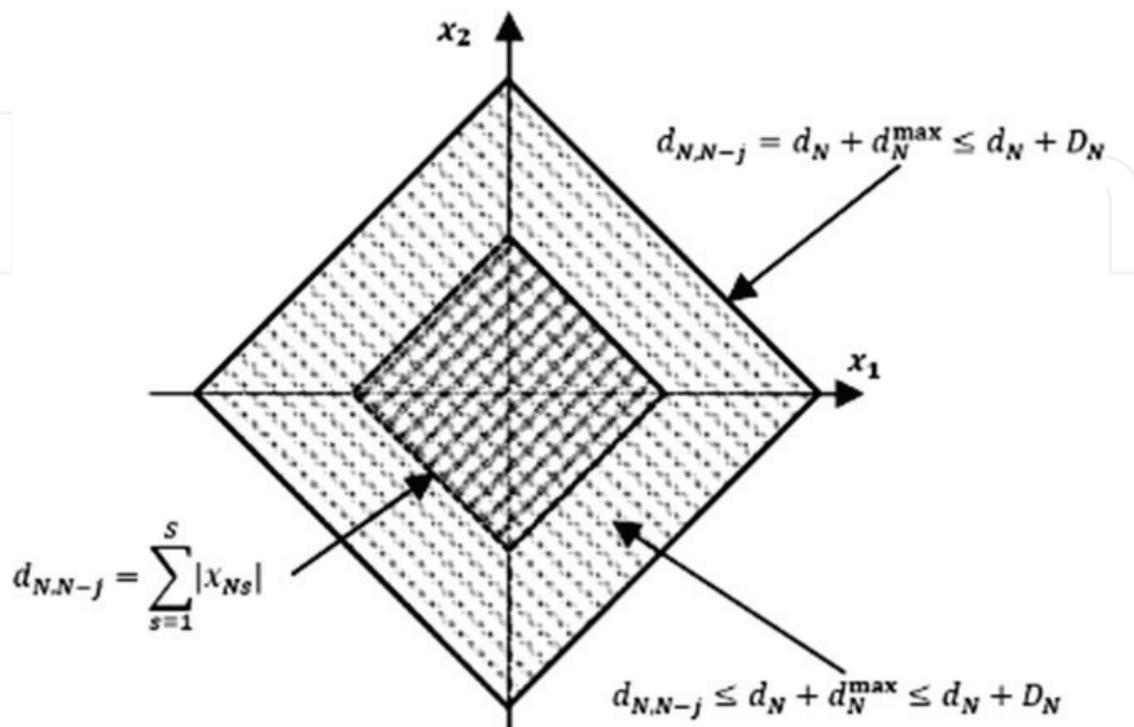


Figure 3.
 Approximating hypersurface building.

5. Fuzzy virtual models

Fuzzy models under uncertainty are advisable to apply in decision-making systems in the following cases [3]:

- dynamics of the investigated quality index is described by a complex nonlinear dependence; and
- one or more factors of this dynamics are weakly or not formalized.

In fuzzy systems, the most commonly used technique is the production rule one. The production rule consists of antecedent (or several premises) and consequent. In the general case, the premises are connected by logical operators AND and OR.

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

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

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

The fuzzy system based on the production rules has the form:

A fuzzy model with $n + m$ input variables $\mathbf{X} = \{X_1, X_2, \dots, X_{n+m}\}$ defined in space $D\mathbf{X} = DX_1 \times DX_2 \times \dots \times DX_{n+m}$ and with one-dimensional output Y is defined in the domain of reasoning DY .

Clear values of fuzzy variables X_i and Y are denoted by x_i and y , respectively.

$LX_i = \{LX_{i,1}, \dots, LX_{i,l_i}\}$ is the fuzzy domain of definition of the i -th input variable and X_i is the number of linguistic terms on which this fuzzy variable is defined.

$LY = \{LY_1, \dots, LY_l\}$ is the domain of the fuzzy output variable.

l is the number of fuzzy values.

LY_j is the name of the output linguistic term.

The rule base in the fuzzy *Mamdani system* is a set of fuzzy rules such as:

$$R_j : LX_{1,j_1} \text{ AND } \dots \text{ AND } LX_{n,j_n} \rightarrow LY_j. \quad (7)$$

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

$$R_j : LX_{1,j_1} \text{ AND } \dots \text{ AND } LX_{n,j_n} \rightarrow r_j \quad (8)$$

where r_j is a real number to estimate the output y .

The j -th rule in the Takagi-Sugeno model [15] looks as follows:

$$R_j : LX_{1,j_1} \text{ AND } \dots \text{ AND } LX_{n+m,j_{n+m}} \rightarrow r_{0j} + r_{1j}x_1 + r_{2j}x_2 + \dots + r_{(n+m)j}x_{n+m} \quad (9)$$

where the output y is estimated by a linear function.

Thus, the fuzzy system performs the mapping $L : \mathfrak{R}^{n+m} \rightarrow \mathfrak{R}$.

The grade of crisp variable x_i membership in the fuzzy notion LX_{ij} is determined by membership functions $\mu_{LX_{ij}}(x_i)$. The rule base is formed by the criterion of minimum output error which can be defined by the following expressions:

$$\frac{\sum_{i=1}^K |f(x_i) - L(x_i)|}{K}, \frac{\sqrt{\sum_{i=1}^K (f(x_i) - L(x_i))^2}}{K}, \max_{i \in K} |f(x_i) - L(x_i)| \quad (10)$$

where K is the number of samples.

Depending on the features of the object and the purpose of identification, various fuzzy models can be formed. Thus, the Takagi-Sugeno model is most suitable for objects with complex nonlinear dynamics, such as moving objects, in the control of which the accuracy requirements prevail.

A fuzzy model of the Mamdani type is suitable for problems in the solution of which it is important to form knowledge based on data analysis.

The singleton-type system may be used in both identification and knowledge-formation tasks.

Singleton-type fuzzy model performs the mapping $L : \mathfrak{R}^{n+m} \rightarrow \mathfrak{R}$ where the fuzzy conjunction operator is replaced by a product, and the operator of fuzzy rules aggregation, that is, by summation. The mapping L is defined by the following expression:

$$L(x) = \frac{\sum_{i=1}^q \mu_{LX_{1i}}(x_1) \cdot \mu_{LX_{2i}}(x_2) \cdot \dots \cdot \mu_{LX_{(n+m)i}}(x_{n+m}) \cdot r_i}{\sum_{i=1}^q \mu_{LX_{1i}}(x_1) \cdot \mu_{LX_{2i}}(x_2) \cdot \dots \cdot \mu_{LX_{(n+m)i}}(x_{n+m})} \quad (11)$$

where $\mathbf{x} = [x_1, \dots, x_{n+m}]^T \in \mathfrak{R}^{n+m}$; q is the number of rules in a fuzzy model; $n + m$ is the number of input variables in the model; and $\mu_{LX_{ij}}(x_{ij})$ is the membership function.

The expression for L mapping in the Takagi-Sugeno model looks as follows:

$$L(x) = \frac{\sum_{i=1}^q \mu_{LX_{1i}}(x_1) \cdot \mu_{LX_{2i}}(x_2) \cdot \dots \cdot \mu_{LX_{(n+m)i}}(x_{n+m}) \cdot (r_{0i} + r_{1i}x_1 + r_{2i}x_2 + \dots + r_{(n+m)i}x_{n+m})}{\sum_{i=1}^q \mu_{LX_{1i}}(x_1) \cdot \mu_{LX_{2i}}(x_2) \cdot \dots \cdot \mu_{LX_{(n+m)i}}(x_{n+m})} \quad (12)$$

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

Let the variables in (1) be fuzzy. In this case, (1) can be represented as a fuzzy model of Takagi-Sugeno (TS) [15].

To form the model, product rules with linear finite-difference equations on the right-hand side are defined (for simplicity, we consider one-input case, i.e., $P = 1$):

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

$$y^\theta(t) = a_0^\theta + \sum_{k=1}^r a_k^\theta y(t-k) + \sum_{l=0}^s b_l^\theta x(t-l), \theta = 1, \dots, n, \quad (13)$$

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

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

$$(U_1^\theta(u_1(t)), \dots, U_m^\theta(u_m(t))) = (Y_1^\theta(y(t-1)), \dots, Y_r^\theta(y(t-r)), X_0^\theta(x(t)), \dots, X_s^\theta(x(t-s))),$$

where $m = r + s + 1$, one obtains the analytic form of the fuzzy model, intended for calculating the output $\hat{y}(t)$:

$$\hat{y}(t) = c^T \tilde{u}(t), \quad (14)$$

where $c = (c_0^1, \dots, c_0^n, \dots, c_m^1, \dots, c_m^n)^T$ is the vector of the adjustable parameters; $\tilde{u}^T(t) = (u_0(t)\beta^1(t), \dots, u_0(t)\beta^\theta(t), \dots, u_m(t)\beta^1(t), \dots, u_m(t)\beta^n(t))$ is the extended input vector;

$$\beta^\theta(t) = \frac{U_1^\theta(u_1(t)) \otimes \dots \otimes U_m^\theta(u_m(t))}{\sum_{\theta=1}^N (U_1^\theta(u_1(t)) \otimes \dots \otimes U_m^\theta(u_m(t)))} \quad (15)$$

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

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

$$\begin{aligned} c(t) &= c(t-1) + Q(t)\tilde{u}(t)[y(t) - c^T(t-1)\tilde{u}(t)] \\ Q(t) &= Q(t-1) - \frac{Q(t-1)\tilde{u}(t)\tilde{u}^T(t)Q(t-1)}{1 + \tilde{u}^T(t)Q(t-1)\tilde{u}(t)} \end{aligned} \quad (16)$$

$Q(0) = \gamma I$, $\gamma > 1$, 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 which 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 $\Xi = \{\Xi_i(R_0^a, R^a, T^a)\}$, so that the number of production rules in the TS model is significantly reduced according to some criterion.

For example, the following matrix:

$$\begin{matrix} \beta_1^{\theta_t} & \dots & \beta_P^{\theta_t} \\ \dots & \dots & \dots \\ \beta_1^{\theta_{t-s}} & \dots & \beta_P^{\theta_{t-s}} \end{matrix} \quad (17)$$

can be defined for P -dimensional input vectors at time steps $t-j$, $j = 1, \dots, s$. If the rows of this matrix are ranged, say, w.r.t. $\sum_{p=1}^P |\beta_p^{\theta_i}|$ decrease and a certain number of rows are selected, then such selection combined with condition (4) will determine the predicate Ξ 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 $\sum_{p=1}^P |\beta_p^{\theta_i}|$, and select a certain number of rows. Such selection combined with condition (4) defines the predicate $\Xi = \{\Xi_i(R_0^a, R^a, T^a)\}$, and, respectively, the image selection criterion (sets of input vectors) from the archive.

5.1 Fuzzy associative search

Notwithstanding all benefits delivered by fuzzy techniques, their application significantly reduces the calculation 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:

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

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

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.

6. Solving the associative search problem by means of clusterization techniques

The associative search problem is solved by clustering technique (both crisp and fuzzy) in the following way.

The current vector under investigation is attributed to a certain cluster per the criterion of minimum distance to the center:

$$\min_k \sum_{k=1}^K \|g_k - \acute{x}_N\|^2,$$

where $\acute{x}_N \in X$ is the current input vector of the control plant under investigation.

Within this cluster, the vectors are sought that satisfy the assigned associative criterion. It may turn out that one cannot find within this cluster the number of vectors necessary to solve the problem of forecasting using the method of least squares. In this case, one of the known methods of combining two clusters with the minimum distance between any two of their members can be applied. This approach provides significant savings in computing resources compared to searching through a full search. However, such a combination of clusters does not yet guarantee the solution of the problem. The approach described below looks the most reasonable.

6.1 Virtual clustering (“impostor” method)

The current input vector at any particular time can be assigned to a specific cluster. This can, for example, be done by the criterion of the minimum distance to the center.

Let

$$\min_k \sum_{k=1}^K \|g_k - \acute{x}_N\|^2$$

be satisfied for $k = r$.

Let \acute{x}_N denote the center of the cluster A_r . If additional selection of input vectors from the archive is required (to form a system of a sufficient number of equations to identify the system using the associative search method), clusters with the minimum distance between their centers and \acute{x}_N are selected for the join. This approach allows not only to discard a significant number of vectors removed from \acute{x}_N , but also to select from the archive the maximum possible number of vectors satisfying the criterion of associative search.

After the completion of this procedure, assigning x_N as the cluster center A_r is canceled, and the procedure of the formation of virtual (relevant to the certain time instant) models continues using conventional clustering algorithms.

7. Case study: oil refining product quality modeling

Key process equipment of an atmospheric distillation unit comprises of cold and hot crude oil preheat trains, desalter, a flash drum or, instead, a pre-flash column with an overhead reflux drum, atmospheric heaters, and an atmospheric distillation column with a reflux drum and three side stripping columns for middle distillates (typically, kerosene, light diesel and heavy diesel aka atmospheric gas oil). The naphtha streams from both reflux drums are re-combined and further sent to downstream stabilization and rerun facilities. The atmospheric residuum from the bottom of the main atmospheric column is typically streamed to a vacuum distillation section.

To obtain a soft sensor model for the 10% distillation point of a kerosene stream, the lab data for this quality were collected along with process data from the atmospheric column. The predictive model is formed by means of the associative search method. The process data were analyzed, and process variables measured by plant instruments were selected for modeling along with the distillation point sampled at the plant and measured in the refinery's laboratory. Based on the preliminary data analysis, the following linear predictive model was developed:

$$T(t) = \sum_{i=1}^4 b_i F_i(t-1) + b_5 F_5(t-3) + b_6 F_6(t-5) + \sum_{i=7}^{12} b_i F_i(t-7), \quad (18)$$

where $T(t)$ is the desired estimate; $F_i(t-j)$ are various process parameters, such as flows, temperatures, and pressures, measured directly at the plant; and b_1, \dots, b_{12} are model's coefficients.

The forecast was calculated per linear and associative models for 10,525 time steps (1 step = 10 min). **Figure 4** shows simulation results for the steps $t = 102, \dots, 301$.

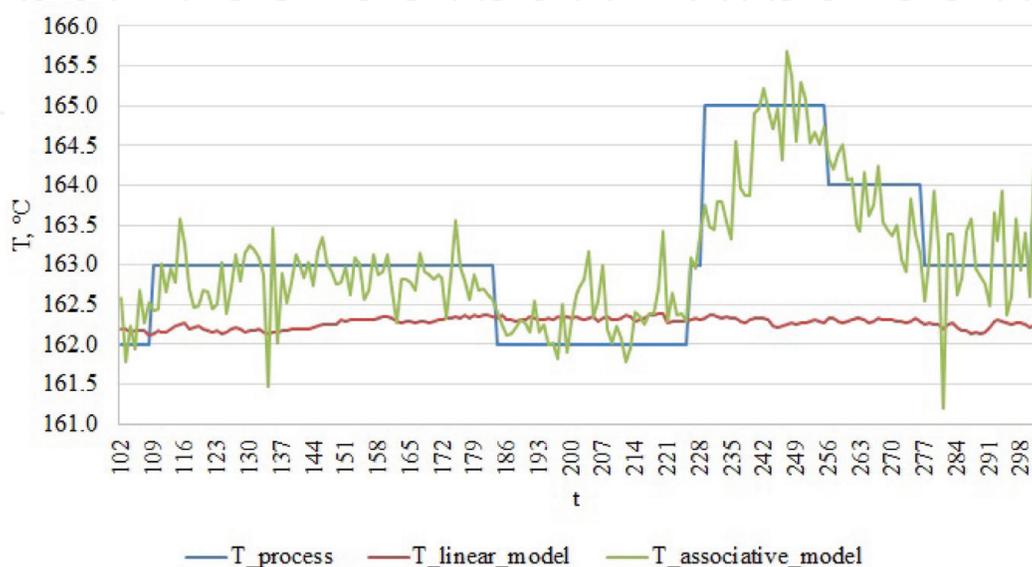


Figure 4.
Kerosene 10% distillation point forecast.

8. Application of wavelet approach to the analysis of nonstationary processes

Within the last two decades, applying wavelet transform (WT) to the analysis of nonstationary processes has been widely used. The wavelet transform of signals is a generalization of the spectral analysis, for instance, with regard to the Fourier transform.

First papers on the wavelet analysis of time (spatial) series with a pronounced heterogeneity appeared in the end of 1980s [16, 17]. The method was positioned as an alternative to the Fourier transform, localizing the frequencies but not providing the time extension of a process under study. In sequel, the theory of wavelets has appeared and is developed, as well as its numerous applications.

The scope of wavelet analysis today is very wide: it includes the synthesis and processing of nonstationary signals, compression and coding of information, image recognition and image analysis, the study of functions and time-dependent signals and inhomogeneity in space. The approach is effective for tasks where the results of the analysis should contain not only the characteristics of the frequency signal (signal power distribution by frequency components) but also information about local coordinates in which certain groups of frequency components manifest themselves or in which rapid changes in the frequency components of the signal occur. A significant number of practical applications have been created, including in health care, the study of geophysical fields, temporary meteorological series, and prediction of earthquakes [18].

The wavelet analysis method consists in applying a special linear conversion of signals. In particular, it becomes possible to study the physical properties or dynamics of real objects and processes in depth. For example, it can be processes in manufacturing. The wavelet transform (WT) of a one-dimensional signal is its representation in the form of a generalized Fourier series (or Fourier integral) over a system of basis functions called the “wavelet.” A wavelet is characterized by the fact that the function that forms it (a wavelet-formation function or a wavelet matrix) is distinguished by a certain scale (frequency) and localization in time based on the time shift and the change in the time scale.

The time scale is analogous to the oscillation period, that is, it is inverse one with regard to the frequency, and the shift interprets the displacement of the signal over the time axis.

The wavelet transform performs the projection of a one-dimensional process into a two-dimensional surface in three-dimensional space. The frequency and time are treated as independent variables.

At the same time, it becomes realistic to simultaneously study the properties of the process being studied both in the time domain and in the frequency domain. It becomes possible to investigate the dynamics of the frequency process and its local features. This allows us to identify the coordinates at which certain frequencies manifest themselves most significantly.

The graphical representation of the wavelet analysis can be displayed in the form of isolines, illustrating the change in the intensities of wavelet transform coefficients at different time scales, and also for revealing local extrema of surfaces.

If a function is used in the Fourier transform that generates an orthonormal basis of space by means of a scale transformation, then the wavelet transform is formed using a basis function localized in a bounded domain, although defined on the whole numerical axis.

The wavelet transform, as a mathematical tool, serves mainly to analyze data in the time and frequency domains.

Wavelet transformation, as a mathematical tool, provides the ability to analyze data in the time and frequency domains simultaneously. The wavelet transform can

provide time-frequency information about a function that in many practical situations is more relevant than information obtained through standard Fourier analysis.

There are examples of the use of wavelet analysis in identification problems [5]. In the literature, it is noted that wavelets are used mainly to identify nonlinear systems with a certain structure, where unknown time-varying coefficients can be represented as a linear combination of basis wavelet functions [6, 7]. It was stated that along with the usual (“direct”) wavelet analysis, biorthogonal bursts [18], wavelet frames [19], or wavelet networks [20] can be used to identify the system.

There exist many different ways of applying wavelets for linear system identification. In Ref. [21], the identification of systems with a specific input/output structure was studied, in which the parameters are identified via spline-wavelets and their derivatives. In paper [22], an extended use of an orthonormal transformation least squares method is presented in order to reveal useful information from data.

9. Conditions of the associative model stability in the aspect of the analysis of the spectrum of multi-scale wavelet expansion

Let (1) be an associative search model. We represent the multi-scale wavelet decomposition for the current input vector $x(t)$ for a fixed level of detail L [7]:

$$\begin{aligned} x(t) &= \sum_{k=1}^N c_{L,k}^x(t) \varphi_{L,k}(t) + \sum_{l=1}^L \sum_{k=1}^N d_{l,k}^x(t) \psi_{l,k}(t), \\ y(t) &= \sum_{k=1}^N c_{L,k}^y(t) \varphi_{L,k}(t) + \sum_{l=1}^L \sum_{k=1}^N d_{l,k}^y(t) \psi_{l,k}(t), \end{aligned} \quad (19)$$

where L is the depth of the multi-scale expansion; $\varphi_{L,k}(t)$ are scaling functions; $\psi_{l,k}(t)$ are the wavelet functions that are obtained from the mother wavelets by tension/compression and shift

$$\psi_{l,k}(t) = 2^{l/2} \psi_{\text{mother}}(2^l t - k)$$

(as the mother wavelets, in the present case, we consider the Haar wavelets); l is the level of data detailing; $c_{L,k}$ are the scaling coefficients; and $d_{l,k}$ are the detailing coefficients. The coefficients are calculated by use of the Mallat algorithm [17].

Let us expand Eq. (1) over wavelets:

$$\begin{aligned} \sum_{k=1}^N c_{Lk}^y(t) \varphi_{Lk}(t) + \sum_{l=1}^L \sum_{k=1}^N d_{lk}^y(t) \psi_{lk}(t) &= \sum_{k=1}^N \left(\sum_{i=1}^m a_i c_{Lk}^y(t-i) \varphi_{Lk}(t-i) \right) \\ &+ \sum_{l=1}^L \sum_{k=1}^N \left(\sum_{i=1}^m a_i d_{lk}^y(t-i) \psi_{lk}(t-i) \right) + \sum_{k=1}^N \left(\sum_{s=1}^S \sum_{j=1}^{r_s} b_{sj} c_{Lk}^s(t-j) \varphi_{Lk}(t-j) \right) \\ &+ \sum_{l=1}^L \sum_{k=1}^N \left(\sum_{s=1}^S \sum_{j=1}^{r_s} b_{sj} d_{lk}^s(t-j) \psi_{lk}(t-j) \right) \end{aligned}$$

Let us consider individually the detailing and approximating parts correspondingly:

$$(t) \psi_{lk}(t) = \sum_{i=1}^m a_i d_{lk}^y(t-i) \psi_{lk}(t-i) + \sum_{s=1}^S \sum_{j=1}^{r_s} b_{sj} d_{lk}^s(t-j) \psi_{lk}(t-j), \quad (20)$$

$$c_{Lk}^y(t) \varphi_{Lk}(t) = \sum_{i=1}^m \hat{a}_i c_{Lk}^y(t-i) \varphi_{Lk}(t-i) + \sum_{s=1}^S \sum_{j=1}^{r_s} \hat{b}_{sj} c_{Lk}^s(t-j) \varphi_{Lk}(t-j). \quad (21)$$

In [7], it was shown that a sufficient condition for the stability of plant (1) is as follows: for $\forall k = 1, \dots, N$ meeting the inequalities is to be provided:

1. if $m > R$, $R = \max_{s=1, S} r_s$, then the condition for the detailing coefficients:

$$\begin{aligned} & \left| \frac{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{l,k}^y(t)} \right| < 1, \\ & \left| \frac{-a_2 d_{l,k}^y(t-2) + \sum_{s=1}^S b_{s,2} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)} \right| < 1, \dots, \\ & \left| \frac{-a_{R+1} d_{l,k}^y(t-R-1)}{a_R d_{l,k}^y(t-R) + \sum_{s=1}^S b_{s,R} d_{l,k}^{x_s}(t-R)} \right| < 1, \quad (22) \\ & \left| \frac{-a_{R+2} d_{l,k}^y(t-R-2)}{a_{R+1} d_{l,k}^y(t-R-1)} \right| < 1, \dots, \\ & \left| \frac{-2a_m d_{l,k}^y(t-m)}{a_{m-1} d_{l,k}^y(t-m+1)} \right| < 1 \end{aligned}$$

for the approximating coefficients:

$$\begin{aligned} & \left| \frac{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^y(t)} \right| < 1, \\ & \left| \frac{-a_2 c_{L,k}^y(t-2) + \sum_{s=1}^S b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)} \right| < 1, \dots, \\ & \left| \frac{-a_{R+1} c_{L,k}^y(t-R-1)}{a_R c_{L,k}^y(t-R) + \sum_{s=1}^S b_{s,R} c_{L,k}^{x_s}(t-R)} \right| < 1, \quad (23) \\ & \left| \frac{-a_{R+2} c_{L,k}^y(t-R-2)}{a_{R+1} c_{L,k}^y(t-R-1)} \right| < 1, \dots, \\ & \left| \frac{-2a_m c_{L,k}^y(t-m)}{a_{m-1} c_{L,k}^y(t-m+1)} \right| < 1; \end{aligned}$$

2. if $m < R$, $R = \max_{s=1, S} r_s$, then the condition for the detailing coefficients:

$$\begin{aligned} & \left| \frac{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{l,k}^y(t)} \right| < 1, \\ & \left| \frac{-a_2 d_{l,k}^y(t-2) + \sum_{s=1}^S b_{s,2} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)} \right| < 1, \dots, \\ & \left| \frac{-\sum_{s=1}^S b_{s,m+1} d_{l,k}^{x_s}(t-m-1)}{a_m d_{l,k}^y(t-m) + \sum_{s=1}^S b_{s,m} d_{l,k}^{x_s}(t-m)} \right| < 1, \\ & \left| \frac{-\sum_{s=1}^S b_{s,m+2} d_{l,k}^{x_s}(t-m-2)}{\sum_{s=1}^S b_{s,m+1} d_{l,k}^{x_s}(t-m-1)} \right| < 1, \dots, \\ & \left| \frac{-2\sum_{s=1}^S b_{s,R} d_{l,k}^{x_s}(t-R)}{\sum_{s=1}^S b_{s,R-1} d_{l,k}^{x_s}(t-R+1)} \right| < 1 \end{aligned}$$

for the approximating coefficients:

$$\begin{aligned}
 & \left| \frac{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^y(t)} \right| < 1, \\
 & \left| \frac{-a_2 c_{L,k}^y(t-2) + \sum_{s=1}^S b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)} \right| < 1, \dots, \\
 & \left| \frac{-\sum_{s=1}^S b_{s,m+1} c_{L,k}^{x_s}(t-m-1)}{a_m c_{L,k}^y(t-m) + \sum_{s=1}^S b_{s,m} c_{L,k}^{x_s}(t-m)} \right| < 1, \dots, \quad (24) \\
 & \left| \frac{-\sum_{s=1}^S b_{s,m+2} c_{L,k}^{x_s}(t-m-2)}{\sum_{s=1}^S b_{s,m+1} c_{L,k}^{x_s}(t-m-1)} \right| < 1, \dots, \\
 & \left| \frac{-2 \sum_{s=1}^S b_{s,R} c_{L,k}^{x_s}(t-R)}{\sum_{s=1}^S b_{s,R-1} c_{L,k}^{x_s}(t-R+1)} \right| < 1;
 \end{aligned}$$

3. if $m = R \neq 1, R = \max_{s=1, \dots, S}$, then the condition of the stability for the detailing coefficients:

$$\begin{aligned}
 & \left| \frac{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)}{2d_{l,k}^y(t)} \right| < 1, \\
 & \left| \frac{-a_2 d_{l,k}^y(t-2) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-2)}{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)} \right| < 1, \dots, \quad (25) \\
 & \left| \frac{-2 \left[a_m d_{l,k}^y(t-m) + \sum_{s=1}^S b_{s,m} d_{l,k}^{x_s}(t-m) \right]}{a_{m-1} d_{l,k}^y(t-m+1) + \sum_{s=1}^S b_{s,m-1} d_{l,k}^{x_s}(t-m+1)} \right| < 1
 \end{aligned}$$

for the approximating coefficients:

$$\begin{aligned}
 & \left| \frac{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)}{2c_{L,k}^y(t)} \right| < 1, \\
 & \left| \frac{-a_2 c_{L,k}^y(t-2) + \sum_{s=1}^S b_{s,2} c_{L,k}^{x_s}(t-2)}{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)} \right| < 1, \dots, \quad (26) \\
 & \left| \frac{-2 \left[a_m c_{L,k}^y(t-m) + \sum_{s=1}^S \hat{b}_{s,m} c_{L,k}^{x_s}(t-m) \right]}{a_{m-1} c_{L,k}^y(t-m+1) + \sum_{s=1}^S \hat{b}_{s,m-1} c_{L,k}^{x_s}(t-m+1)} \right| < 1;
 \end{aligned}$$

4. if $m = R = 1, R = \max_{s=1, \dots, S}$, then the condition of the stability for the detailing coefficients:

$$\left| \frac{a_1 d_{l,k}^y(t-1) + \sum_{s=1}^S b_{s,1} d_{l,k}^{x_s}(t-1)}{d_{l,k}^y(t)} \right| < 1$$

for the approximating coefficients:

$$\left| \frac{a_1 c_{L,k}^y(t-1) + \sum_{s=1}^S b_{s,1} c_{L,k}^{x_s}(t-1)}{c_{L,k}^y(t)} \right| < 1.$$

10. Prediction of the transfer to chaos

The chaotic system dynamics is characterized by considerable dependence on initial conditions, when as close as needed at the initial time instant trajectories during certain time are diverge by a finite distance. The main characteristics of the chaotic behavior are the speed of divergence of the trajectories defined by the senior Lyapunov exponent. This speed is determined by the Lyapunov exponent whose value represents the degree of instability or degree of sensitivity to the original data.

For a linear system with a constant matrix, the senior Lyapunov exponent is $\chi_1 = \max \Re \lambda_i$, where λ_i are the eigenvalues of the system matrix. In other words, $|\chi_1|$ coincides with the conventional degree of the system stability [23].

Thus, (23) and (24) are sufficient conditions for chaotic dynamics prediction, what is a key condition under implementing phase transfers of technological processes under study.

11. Prediction of manufacturing situations

Optimal routine enterprise resource planning and scheduling are currently based on detailed mathematical models of production processes [24]. Rescheduling requires model update subject to the current production information.

Present-day industrial sites feature interrelated multi-variable production processes and sophisticated material flow networks; scheduling at such sites poses nonlinear NP-hard optimization problems.

The state of manufacturing resources should be nevertheless assessed and predicted both to improve control agility and to foresee the situations where schedule execution becomes problematic or impossible. Such situations will be further referred to as incidents.

It may make sense to develop intelligent predictive models describing the overall current state of resources employed to execute all production operations of a specific production process.

The term “production resources” will hereafter mean the following:

- input flows characterized by formal properties dependent on production specificity; and
- production equipment.

$$d_{ij}, i = 1, \dots, N; j = 1, \dots, M$$

and other facilities used for performing the j -th operation;

- human resources

$$h_{ij}, i = 1, \dots, H; j = 1, \dots, M$$

involved in the j -th operation;

- other factors

$$f_{ij}, k = 1, \dots, N; j = 1, \dots, M$$

affecting the j -th operation such as energy resources and a variety of formal indices and factors related with the production process.

Production resources may be described differently.

1. Some have qualitative characteristics which take on specific values that may be checked against norms at any moment.
2. The state of others such as certain equipment pieces may be exclusively either “working” or “not working.” The remaining life time may be known or not for such resources. The process historian may however keep failure statistics for a specific equipment piece; maintenance downtime statistics may be also available for a specific piece or similar kind of equipment.
3. One more resource type (including human resources) is not subject to maintenance. In case of outage, such resources should be immediately replaced from the backlog. The replacement process is typically fast; therefore, no values other than 1 (OK) and 0 (not OK) should be assigned to such resource.

Assume a model of a specific manufacturing situation as a dynamic schedule fragment comprising the following components:

$$r_{ij}(t) = \{ \langle C_1 \rangle \langle C_2 \rangle \langle C_3 \rangle \langle C_4 \rangle \langle C_5 \rangle \}_{ijt} \quad (27)$$

where

$\langle C_1 \rangle \stackrel{\text{def}}{=} \langle ijt \rangle$ is a *resource identifier* including the resource number, the operation number, and the time stamp (the number of characteristics may be increased).

Other components of the resource state vector at the time moment t may be represented by a binary code.

$\langle C_2 \rangle$ is the code of the numerical value of a state variable; this code is different for each of the above-listed resource types.

$\langle C_3 \rangle$, $\langle C_4 \rangle$, and $\langle C_5 \rangle$ will be discussed further.

Consider the resources whose state may be described by some quantitative characteristic, such as inlet flow rate or temperature for chemical processes or an average equipment failure number.

For a specific resource, we assume that the characteristic of its state possesses the values on the half-interval $[0; 1)$ (this half-interval was chosen as an example for simplicity, the results can be easily spread to any other).

This half-interval can be represented as the union

$[0; 0.5) \cup [0.5, 1)$. We will further correspond the symbols $\{0; 1\}$ to the left and right half-intervals respectively, namely, 0 to the left half-interval, and 1 to the right one.

Each of the two subintervals can be further split in the same way, and, again, the values 0 and 1 can be assigned to the left and the right parts, respectively.

In that way, a finite chain of symbols from $\{0; 1\}$ has a one-to-one correspondence with a half-interval embedded in $[0; 1)$. For a binary partition, a chain of n symbols corresponds to a half-interval with the length $\frac{1}{2^n}$.

This way, for each value of a numerical characteristic at the current time moment, we obtain a code of 0s and 1s. The number of positions, as we show further, will determine the accuracy of prediction.

For the resources from the categories 2 and 3, the respective codes will have the same value in all positions (either 1 or 0). $\langle C_3 \rangle$ is the code of the time before the maintenance end. If a resource is available and operated, the respective code consists of 1s. $\langle C_4 \rangle$ is the code of the time before the equipment piece fails with the probability close to 1 (remaining life).

In the scheduling practice, this time is not less than the operating time. However, resource replacement just during the operation may be sometimes more cost-effective. Moreover, the equipment piece may fail unexpectedly. For resource types from categories 1 and 3, $\langle C_4 \rangle$ has 1s in all positions.

$\langle C_5 \rangle$ is the time before the scheduled end of the operation. In real-life manufacturing situations, time may be wasted (with the need in schedule update) for the reasons neither stipulated in the production model nor caused by equipment failures.

Generally, it is hardly possible to formalize all such causes of schedule disruption. Therefore, their consolidation as the “remaining plan execution time” is a way to allow for these hidden factors in the production state model.

For the developed binary chain, a forecast may be obtained using data mining techniques. It makes sense to apply the methods named **association rules search** [25].

A forecast of a state described by a binary chain with an identifier can be obtained by revealing the most probable combination of two binary sets of values at a fixed time instant and at the next instant (a one-step forecast). A more distant prediction horizon is also possible.

12. Conclusion

Modern information technologies offer new possibilities for solving identification problems for control and decision-making systems. Data mining methods allow to solve problems that in the general case could not be solved by classical methods, or required heuristic approaches.

In this chapter, associative search techniques are presented. The techniques allow the identification of nonlinear systems, without the need to build a bunch of Wiener-Hammerstein models, etc. An alternative is to analyze the current state of the system using the knowledge base and training system. This approach allows the best use of a priori information on the object.

The algorithms may be successfully applied in the identification of nonlinear nonstationary processes. For these purposes, the multi-scale wavelet expansion is used. By investigating the dynamics of the coefficients of this expansion, one can predict the approach of process parameters to stability limits. Finally, sufficient conditions of stability are derived.

The high accuracy of forecasting by associative search technique makes it relevant for studying the dynamics of processes and predicting the transition to chaos. Also, it becomes possible to predict the contingencies of production processes. For this, the method of searching for associative rules is applied.

IntechOpen

Author details

Natalia Bakhtadze^{1,2*}, Vladimir Lototsky¹, Valery Pyatetsky³ and Alexey Lototsky¹

1 V.A. Trapeznikov Institute of Control Sciences, Moscow, Russia

2 Bauman Moscow State Technical University, Moscow, Russia

3 National University of Science and Technology “MISIS”, Moscow, Russia

*Address all correspondence to: sung7@yandex.ru

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Peretzki D, Isaksson A, Carvalho A, Bittencourt C, Forsman K. Data Mining of Historic Data for Process Identification. Sweden: Linköping University Electronic Press; 2014. <http://manualzz.com/doc/8482583/modeling-and-diagnosis-of-friction-and-wear-in-industrial>
- [2] Bakhtadze N, Kulba V, Lototsky V, Maximov E. Identification-based approach to soft sensors design. IFAC-PapersOnLine. 2007;**10**:302-307. DOI: 10.3182/20100701-2-PT-4011.00052
- [3] Bakhtadze N, Maximov E, Valiakhmetov R. Fuzzy soft sensors for chemical and oil refining processes. IFAC Proceedings Volumes. 2008;**41**: 4246-4250. DOI: 10.3182/20080706-5-KR-1001.00017
- [4] Bakhtadze N, Lototsky V, Vlasov S, Sakrutina E. Associative search and wavelet analysis techniques in system identification. IFAC Proceedings Volumes. 2012;**45**:1227-1232. DOI: 10.3182/20120711-3-BE-2027.00242
- [5] Bakhtadze N, Sakrutina A. The intelligent identification technique with associative search. International Journal of Mathematical Models and Methods in Applied Sciences. 2015;**9**:418-431. ISSN: 1998-0140
- [6] Bakhtadze N, Lototsky V. Knowledge-based models of nonlinear systems based on inductive learning. In: New Frontiers in Information and Production Systems Modelling and Analysis Incentive Mechanisms, Competence Management, Knowledge-based Production. Heidelberg: Springer; 2016. pp. 85-104. ISBN 978-3-319-23338-3. DOI: 10.1007/978-3-319-23338-3
- [7] Bakhtadze N, Sakrutina E, Pyatetsky V. Predicting oil product properties with intelligent soft sensors. IFAC-PapersOnLine. 2017;**50**:14632-14637. DOI: 10.1016/j.ifacol.2017.08.1742
- [8] Moore E. On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society. 1920; **26**:394-395
- [9] Penrose R. A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society. 1955;**51**(3): 406-413
- [10] Larichev OI, Asanov A, Naryzhny Y, Strahov S. Expert system for the diagnostics of acute drug poisonings, applications and innovations in intelligent systems IX. In: Macintosh A, Moulton M, Preece A, editors. Proceedings of the 21 SGES International Conference on Knowledge Based Systems and Applied Artificial Intelligence. Cambridge, UK: Springer-Verlag; 2001. pp. 159-168
- [11] Patel V, Ramoni M. Cognitive models of directional inference in expert medical reasoning. In: Feltovich P, Ford K, Hofman R, editors. Expertise in Context: Human and Machine. Menlo Parc, CA: AAAI Press; 1997
- [12] Hunt E. Cognitive science: Definition, status and questions. Annual Review of Psychology. 1989;**40**:603-629
- [13] Newell A, Simon HA. Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall Inc.; 1972
- [14] Gavrilov A. The model of associative memory of intelligent system. In: Proceedings of 6-th Russian-Korean International Symposium on Science and Technology. Novosibirsk. Vol. 1. 2002. pp. 174-177
- [15] Takagi T, Sugeno M. Fuzzy identification of systems and its

applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics. 1985;**26**:116-132

[16] Daubechies I, Lagarias J. Two-scale difference equations I: Existence and global regularity of solutions. SIAM Journal on Mathematical Analysis. 1991;**22**:1388-1410

[17] Mallat S. In: Barlaud M, editor. Wavelet Tour of Signal Processing. San Diego; CA: Academic Press; 1999. 635p

[18] Vána Z, Preisig H. System identification in frequency domain using wavelets: Conceptual remarks. Systems & Control Letters. 2012;**61**(10): 1041-1051

[19] Ho K, Blunt S. Adaptive sparse system identification using wavelets. IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing. 2003;**49**(10):656-667

[20] Sureshbabu N, Farrell JA. Wavelet-based system identification for nonlinear control. IEEE Transactions on Automatic Control. 1999;**44**(2):412-417

[21] Preisig HA. Parameter estimation using multi-wavelets. Computer Aided Chemical Engineering. 2010;**28**:367-372

[22] Carrier J, Stephanopoulos G. Wavelet-based modulation in control-relevant process identification. AIChE Journal. 1998;**44**(2):341-360

[23] Fradkov A, Evans R. Control of chaos: Survey—1997–2000. IFAC Proceedings Volumes. 2002;**35**:131-142

[24] Al-Otabi GA, Stewart MD. Simulation model determines optimal tank farm design. Oil & Gas Journal. 2004;**102**(7):50-55

[25] Qin JS, Badgwell TA. A survey of industrial model predictive control technology. Control Engineering Practice. 2003;**11**(7):733-764