Appl. Math. and Comp. Sci., 1993, vol.3, No.3, 487–507

PROBLEMS OF COMPUTER–AIDED DATA PROCESSING IN MACHINE DIAGNOSTICS AND MONITORING

JAN KAŹMIERCZAK*

The first part of the paper presents a structure of tasks which are performed in systems of monitoring and diagnosing of states of machines. Basing on the list of such tasks the areas of utilizing particular methods and algorithms for the needs of data processing in monitoring and diagnostic systems are then defined. The paper shows some common features of data processing from the view–point of computer aiding. This part of the paper is illustrated by on example of forecasting as a particular problem in diagnostic research. By means of this example some consequences due to the use of general methods in specific field are discussed. In conclusions some proposals are introduced which concern the managing of input and output data in diagnostic and monitoring systems, supporting the choice of computational algorithms and problems of visualization of input data as well as results of its processing in various phases of functioning of the system as the tasks which can be supported by means of computer tools. This part of the paper is illustrated by some results of research works which have been carried out in the Chair of Fundamentals of Machine Design, Silesian University of Technology in Gliwice.

1. Introduction

The problem of machine and industrial installation functioning in order to identify their current state appears in many branches of technology. If we assume that a machine is the object of investigations which is to be observed by means of external effects called *signals* we are then in the field of the *machine diagnostics.* It is possible to propose a general approach in order to analyze a structure of the research which may be carried out in this field (Kaźmierczak, 1989).

First of all we have to state that the processes which occur in the machines and which are the subjects of investigations have usually continuous nature. The commonly accepted assumption in the machine investigations is that the continuous processes which occur in the machine and are projected in signals may be treated as sequences of discrete elements (*events*). A continuous process could be represented by a chain of two kinds of events: *states and changes of the states* in the process.

Let's distinguish the following *basic* states which are considered in the diagnostic investigations of machine functioning:

* Chair of Fundamentals of Machine Design, Silesian University of Technology, Gliwice, Poland

- instantaneous state of the u–th machine in the class of U machines at the time point t which can be marked as $s_{u,t}$,
- instantaneous state of the signal related to the state $s_{u,t}$, marked as $z_{u,t}$.

Applying these two parallel types of the states one may describe all particular research problems which have to be solved in diagnostics. Of course, the states of one type form a unique sequence (representing respectively the processes of machine functioning and the related signal). Each pair of different instantaneous states is connected by an interaction. The problem of identifying the connections between elements of such a *net of states* can be considered in terms of relations.

2. Relations Between States of Machines and States of Signals

Let's assume in the initial step that two separate sets $\boldsymbol{A} = \{a_i : i = 1, 2....\}$ and $\boldsymbol{B} = \{b_i : i = 1, 2....\}$ are given. The interaction which connects the elements of these two sets, called relation, can be treated as a set R of selected pairs:

$$R = \{(a, b) : a \in A,\ b \in B)\} \tag{1}$$

In particular the *relation* can be described as a subset in the Cartesian product of sets $\boldsymbol{A}$ and $\boldsymbol{B}$:

$$R(A, B) \subset A \times B \tag{2}$$

Thus we are able to identify such relations between the states of the investigated machine $\{s_{u,t} : s_{u,t} \in S_U\}$ and values of signal features $\{z_{u,t} : z_{u,t} \in Z_U\}$. According to (2) we can note such relations as:

$$R1(S, Z) \subset S_U \times Z_U \tag{3}$$

If we are to consider the relationships which occur along the time axis in the processes then the set of time points has to be taken into account. For the purpose of defining such a relation between the elements of a single process $\{A(t)\}$ and elements of the set of time we may select the set $\boldsymbol{T}$ of time points $\boldsymbol{t}$ with two distinct subsets $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ which fulfil the following conditions:

$$T_1 \subset T \qquad T_2 \subset T \qquad T_1 \cap T_2 = 0 \tag{4}$$

In this case the relation should be also shown as the Cartesian product:

$$R(A) \subset A_1 \times A_2 \tag{5}$$

where:

$$\begin{aligned} A_1 &= \{a_t : t \in T_1 \subset T\} \subset A \\ A_2 &= \{a_t : t \in T_2 \subset T\} \subset A \end{aligned} \tag{6}$$

According to (5) one may note respectively the relations:

$$R2(S) \subset \{S_U, t_1\} \times \{S_U, t_2\} \tag{7}$$

for the set of states of the machine and:

$$R2(Z) \subset \{Z_U, t_1\} \times \{Z_U, t_2\} \tag{8}$$

for the set of states of the signal.

In order to distinguish different types of the introduced relations they are denoted as **R**1 (see (3)) or **R**2 (see (7) and (8)). Following the above introduced considerations, the basic *net of relations* in diagnostic research of machines could be shown as in Figure 1.

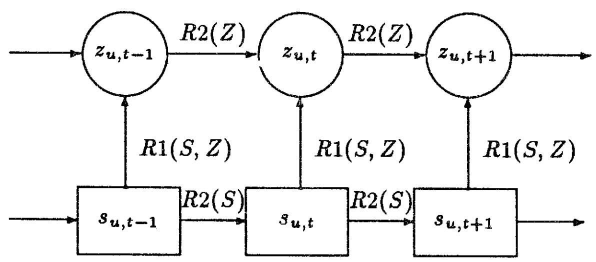


Fig. 1. Net of relations between states of signals and states of the diagnosed object.

It has to be stated that in the investigation of a machine the states of an object are recognized using states of signals. Usually we do not need to identify the single state of an object: the aim is to recognize some fixed classes of the states (let's say the subclass of states *close to failure*). If only a limited number of classes is looked for then, the number of parameters (features) of signals which are used for the purpose of machine states projecting, of the machine may be limited, too.

The problem of finding the optimal lowest number of the *most sensitive* features of the signal which might enable to distinguish the requested classes of states of the investigated machine is one of the main tasks of the discussed research. Following this approach we may consider some additional *elements* in the discussed set of states:

- representation of the signal state $z_{u,t}$ by means of a limited number of features called *image of the state of signal* (marked as $\text{im}(z_{u,t})$ or $Iz_{u,t}$);
- limited representation of the state su,t which can be recognized using the image of signal which is called *image of the state of machine* and marked as $\text{im}(s_{u,t})$ or $Is_{u,t}$.

For such elements we can create a separate *n*e*t of relations* as shown in Figure 2.

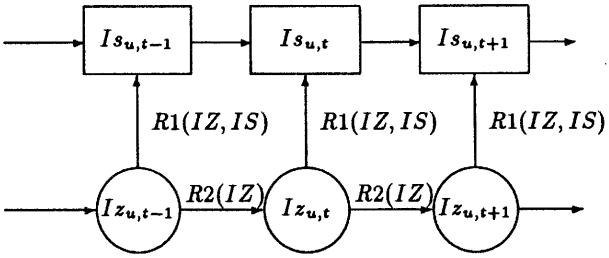


Fig. 2. Net of relations between images of states of signals and the object.

The question appears in what manner the net of relations shown in Figure 1 can be connected with the net in Figure 2. It is clear now that when using the first net we are looking for the relationship between the states of the machine and the respective states of the signal then using the second net we will intend to identify the connections between the images of the machine states and the images of the signal states. Therefore the next problem of machine diagnostics may be formulated through the contradistinction between the state of the signal and the image of such a state. The transformation T_z which connect the state $z_{u,t}$ with the image $Iz_{u,t}$ is shown in Figure 3.

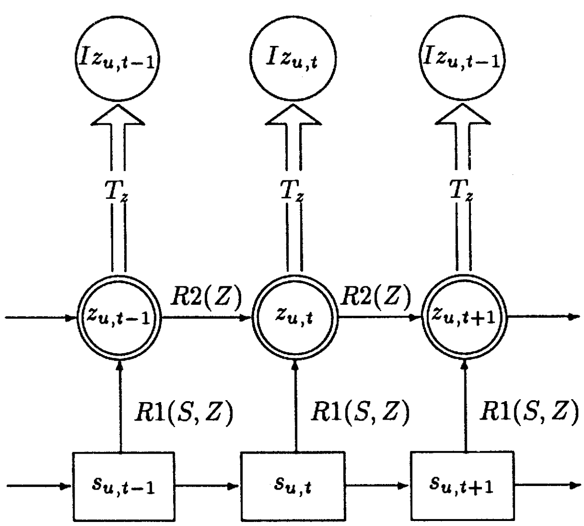


Fig. 3. Connections between states of the signals and its images in the net of *diagnostic* relations.

The *partial nets* are shown in Figures 1, 2 and 3. The *full net* of diagnostic relations is shown in Figure 4. All problems of diagnostic research of machines may be then considered using the above mentioned net (as relations between relevant states).

In opposition to relations of the $R1$ type, tying elements of two separate sets: the states of the object and the values of signals, the relations of type $R2$ associate with elements of a single set. Thus, if we intend to consider the structure of diagnostic research by means of such relations the analysis of both these two types of relations should complement each other. It is worth noting that both $R1(S, Z)$ and $R2(S)$ relations in the above presented approach are assumed to be time independent. Such an assumption may always be accepted, but it results in a significant difficulty of drawing accurate conclusions about the effect on the reasoning. In order to minimize those difficulties these relations may be considered as the processes of a higher order in which the relations (for instance of the $R2$ type) of a higher order are defined. Anyway, this approach causes an increase of complexity of generated models.

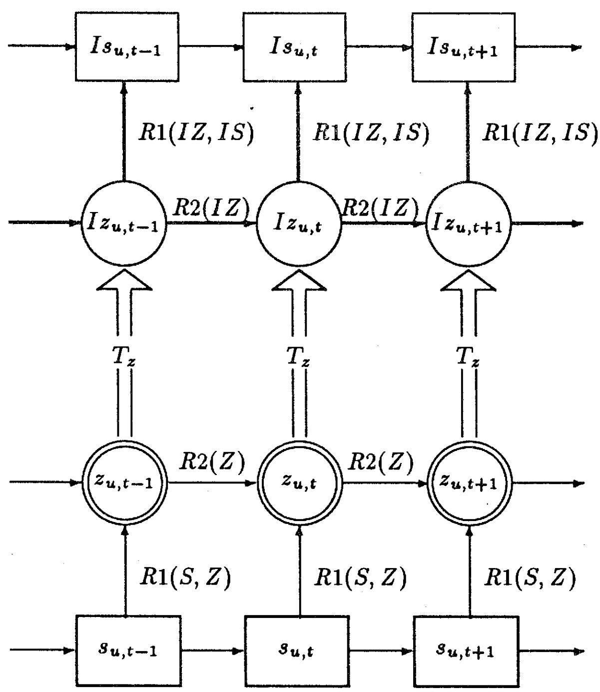


Fig. 4. Net of relations in machine research.

3. Basic Research Problems in Machine Diagnostics

The *classic* problem of machine diagnostics (the identification of the instant state of the object making use of values of the features which form the instant state of the diagnostic signal) is represented in the scheme shown in Figure 4 by means of the vertical sequences of elements. If we consider additionally the fact that the states of the investigated object and respectively the states of the diagnostic signal may vary in time, then the problems of processing diagnostic data in time domain will appear.

At first let's try to connect the *vertical* relations with the particular research problems. We may list these problems as follows:

1) The identification of the transformation T_z in order to recognize the relationship between states of signals and its images. This problem is commonly solved by the proper selection of the *sensitive* signal features (sometimes called the *diagnostic filtering*) which is described in many known works on this subject.
2) The identification of the relation of $R1$ type in order to form a base for drawing conclusions about the state of the machine using the state of diagnostic signals. As mentioned above, in practice, the relations $R1(S, Z)$ are identified by means of the relations $R1(IZ, IS)$ which may be treated as the respective models of relations. The diagnostic reasoning is the main field of solving this class of problems (see Cholewa, 1992; Cholewa and Kaźmierczak, 1992; Kaźmierczak, 1984).

This class of problems in machine diagnostics can be represented by means of the sequential scheme:

instant image of the diagnostic signal $\longmapsto$
$\longrightarrow$ *instant image of the state of the machine*

or by means of the notation:

$$\text{im}(z_{u,t}) \Rightarrow \{R1(\text{im}(S),\ \text{im}(Z))\} \Rightarrow \text{im}(s_{u,t}) \tag{9}$$

The *horizontal relations* in the scheme shown in Figure 4 relate to the separate class of research problems. Before listing these problems it is necessary to say that, in practice, only the relations between states of signals are taken into consideration. More precisely we try to identify the relations between images of these states and to treat them as appropriate *models of relations.*

In order to discuss the range of research problems concerned with the time domain the following terms have to be introduced additionally:

- *time series* Y_t : the sequence of values y_t (where y_t may relate to $s_{u,t}$ or $z_{u,t}$ or $\text{im}(z_{u,t})$ or $\text{im}(s_{u,t})$) ordered in time,
- *forecast* $\widehat{y}_{t'+la}$ of y_t: the estimate of the time series element value which is *future* in time due to the fixed *actual* element of this time series,
- *lead-time of forecast* la: the distance between the actual and forecasted element of the time series along time axis,
- *residual* values $\varepsilon_{t'+1} \in E_t$: the differences between values of time series and its forecasts.

Thus we can list the particular research problems concerned with the $R2$ type relations as follows:

1) Detection and analysis of trends in time series of diagnostic data,
2) Modelling of time series of diagnostic data (or the modelling of $R2(IZ, IS)$ relations,
3) Forecasting in time series of diagnostic data.

The above listed problems may be considered in the similar manner as the problems of the first group. For instance, in the case of forecasting one may consider the following sequence:

history of images of the signal $\longmapsto$
$\longrightarrow$ *forecast of the image of the signal* $\longmapsto$
$\longrightarrow$ *forecast of the image of the machine state*
($\vdash\!\!-\rightarrow$ *forecast of the machine state*)

or respectively the notation:

$$\mathrm{im}(z_{u,i} : i = ...t-1, t) \Rightarrow \{R2(\mathrm{im}(Z))\} \Rightarrow \mathrm{im}(z_{u,t+la}) \Rightarrow$$
$$\Rightarrow \{R1(\mathrm{im}(S),\ \mathrm{im}(Z)\} \Rightarrow \mathrm{im}(s_{u,t+la}) \qquad (10)$$

All the mentioned research problems which are to be solved in machine diagnostics have some common features from the view–point of using computer tools for the needs of assissting them. In order to introduce these features let's consider one of the listed areas: forecasting in time series of diagnostic data.

4. Forecasting as an Example of the Particular Problem in Processing of Diagnostic Data

The general division of forecasting problems which can be solved in technical diagnostics is determined by some kind of an answer which we expect to get as the result of the diagnostic forecasting. In general, two possible questions will have to be answered:

Question 1: what, in turn, (on account of the *actual* time point) is the element x_t in the analyzed time series X_t? or in diagnostic research we shall ask: when will the state marked $s_{u,t''}$ appear in the investigated chain of events $S_t(s_{u,t''} \in S_t)$?

Question 2: what will be the value of the time series element with the identifier $t' + la$, where: $la > 0$ and the element identified by t' is the last one known? or, if this question is posed in the field of diagnostics it will read: what will be the state of the investigated machine after la time units (steps) from the chosen *actual moment* t' $(s_{u,t'+la} = ?)$

The first question refers to problems of the so–called *long–time* (or long–term) *forecasting*. Usually, the solution of a problem from this area requires some forgoing assumptions which may especially be concerned with a possible form of trends behind the observed time–interval. It also requires an appropriate knowledge of the machine functioning which:

- have the same set of design features (are of the same construction),
- are exploited in similar circumstances,

and additionally, one can *a priori* assume that the rationally accepted form of trends exists in the analyzed set of diagnostic data.

The answer to the second question refers in technical diagnostics to problems of the so–called *short–term forecasting*. The considerations presented in this paper are concerned just with this kind of problems.

If we need to solve a forecasting problem in diagnostic investigations first of all we will have to find the forecasting model which will be used for this purpose. In fact, when we want to use one of the known algorithms we shall notice that there is no general rule of the choice of the proper forecasting method. Thus, we can formulate a number of criteria of such a choice, for example:

1) the ability of the considered model to represent systematic components (trends) in the analyzed time series,
2) the batch quantity of the data set which can be utilized as a *learning* set in the particular case of research,
3) the size of the sample of elements of the time series which has to be used for the needs of calculating forecasts in compliance with a fixed method, as well as to verify currently a model with a lack of exploitational time,
4) the intensity of changes of the values in the analyzed time series,
5) the complexity of calculations in the stage of:
   a) generating of forecasting model,
   b) the verification of the model,
   c) calculating forecasts by means of the model;
6) the possibility to obtain a good quality/accuracy of the forecasts,
7) the influence of sudden changes of values (jumps) on the results of forecasting by means of the considered method,
8) the dimensionality of the analyzed time series.

There is another important question how to find any rules in order to fix the hierarchy of the validity of criteria (in the particular field of machine diagnostics).

Usually, the authors of known works in the mentioned field prefer their *favorable* models (see Batko and Kaźmierczak, 1985; Cempel, 1987; Kaźmierczak, 1989) and try to prove its advantages. In general, we can state that the choice of the proper forecasting method depends on particular needs of research.

Suppose we have fixed the method of solving a particular forecasting problem and the choice is the method based on the linear models of ARMA/ARIMA class.

The basic assumption in the method of analyzing time series by means of linear modelling, as among others proposed by Box and Jenkins (1976) and developed later (for example Cadzow, 1983; Ljung, 1983; and others), is that a time series Y_t may be treated as an output (answer) of a modified linear Yule filter excited by a discrete representation of the *white–noise* process. According to such a model, the element y_t of the *output* time series Y_t can be noted as a linear combination of an *infinite* number of elements ε_t of the *input* time series E_t:

$$y_t = c_1\varepsilon_t + c_2\varepsilon_{t-1} + c_3\varepsilon_{t-2} + ... \tag{11}$$

where $c(c_1, c_2, ...)$ denotes the set of weighing coefficients.

In agreement with the modification of the Yule model proposed by Box and Jenkins the so–called *AutoRegressive Moving Average* (ARMA) model is used for the mentioned purpose. According to it the y_t can be described by a *finite* number p of the elements $\{y_{t-i} : i = 1, ...p\}$ of the time series Y_t and/or by a *finite* number q of the elements $\{\varepsilon_{t-j} : j = 1, ...q\}$ of the time series E_t, which are weighted by suitable coefficients:

$$y_t = a_1 y_{t-1} + ... + a_p y_{t-p} + \varepsilon_t - b_1\varepsilon_{t-1} - ... - b_q\varepsilon_{t-q} \tag{12}$$

where $a(a_1, a_2, ..., a_p)$: denotes the set of *autoregressive* (AR) *coefficients* of the ARMA model, and $b(b_1, b_2, ..., b_q)$ is the set of *moving average* (MA) *coefficients* of the model.

Using the *backward shift* operator $\boldsymbol{B}$, defined as:

$$By_t = y_{t-1}, \qquad B^x y_t = y_{t-x} \tag{13}$$

where x is the order of the operator $\boldsymbol{B}$ we might put:

$$a(B) = (1 - a_1 * B - a_2 * B^2 - ... - a_p * B^p) \tag{14}$$

and respectively

$$b(B) = (1 - b_1 * B - b_2 * B^2 - ... - b_q * B^q) \tag{15}$$

Then the equation (12) may be expressed by means of the following notation:

$$(1 - a_1 B - ... - a_p B^p)y_t = (1 - b_1 B - ... - b_q B^q)\varepsilon_t \tag{16}$$

or in a shorter form:

$$a(B)y_t = b(B)\varepsilon_t \quad \text{and} \quad y_t = a^{-1}(B)b(B)\varepsilon_t \quad \text{or} \quad y_t = \frac{b(B)}{a(B)}\varepsilon_t \tag{17}$$

In order to use such models for the needs of representing diagnostic data we have to note that the data usually contain a deterministic component together with a stochastic part. Box and Jenkins have proposed a form of their model which makes it possible to represent a class of non–stationary stochastic processes as well. Complying with this advanced model the deterministic component is *derived* from the original time series Y_t by means of the so–called differential operator ∇^d defined as:

$$z'_{t,0} = y'_t$$
$$z'_{t,1} = \nabla y'_t = y'_t - y'_{t-1} = (1-B)y_t \tag{18}$$
$$z'_{t,d} = \nabla^d y'_t = \nabla^{d-1} y'_t - \nabla^{d-1} y'_{t-1} = (1-B)^d y_t$$

where d is a degree of the operator ∇.

The operator ∇^d enables us to transform the time series Y_t into a new time series $Z_{t,d}$. It is possible to find such a value of d that the time series $Z_{t,d}$ represents a *stationary* stochastic process $Z(t)$ and thus can be analyzed by means of ARMA modeling. The autoregressive moving average model of the stationary stochastic process $Z(t)$ is marked by the symbol $ARIMA(p,q)$. If we analyze a nonstationary process $Y(t)$ and for this reason use the operator of the degree $d>0$ (to transform the time series Y_t into $Z_{t,d}$), this symbol will take the form $ARIMA(p,d,q)$. We may then speak of an integrated (I) autoregressive moving average model of a stochastic process.

The transformed time series $Z_{t,d}$ can be written as:

$$z'_{t,d} = a_1 z_{t-1,d} + \ldots + a_p z'_{t-p,d} + \varepsilon_t - b_1\varepsilon_{t-1} - \ldots - b_q\varepsilon_{t-q} \tag{19}$$

The degree d of the differential operator ∇ together with the autoregressive parameter p defined above as well as the moving average parameter q, form a set of parameters of the ARIMA model. The autoregressive and moving average weights are respectively called the *coefficients of the model.*

The ARMA/ARIMA method offers also the possibility to describe multidimensional time series. The exemplified notation of the ARIMA model of the L–dimensional time series Z_t and E_t $(L>1)$ can be expressed by the equation:

$$\underline{z}_{t,d} = \underline{A}_1\underline{z}_{t-1,d} + \ldots + \underline{A}_p\underline{z}_{t-p,d} + \underline{\varepsilon}_t - \underline{B}_1\underline{\varepsilon}_{t-1} - \ldots - \underline{B}_q\underline{\varepsilon}_{t-q} \tag{20}$$

where $\underline{z}'_{t,d}$, $\underline{\varepsilon}_t$ are elements of time series Z_t and E_t (given as colon matrixes $L*1$), and $\underline{A}_j$, $\underline{B}_j$ are coefficients AR and MA of the model (quadratic matrixes $L*L$).

The discussed class of models may be used as well for the needs of modelling time series with the so–called *seasonal component.* By the term *seasonal* we understand the presence of a component with a repeatability with the period Δt_S which is significantly longer in time than the time step interval of the considered time series. In this variant of the method from the basic time series Y_t with a seasonal part an auxiliary time series $Z_{t,D,S}$ in the form $\{..,z_{t-nS,D},..z_{t-S,D},z_{t,D},..\}$ is extracted for which the ARIMA model is built of the form:

$$a_S(B)\,\nabla_S^D y_t = b_S(B)\varepsilon_t \tag{21}$$

where $a_S\{a_{S,i} : i=1,2,...,P\}$ is the set of *seasonal* AR coefficients, $b_S\{b_{S,j} : j=1,2,...,Q\}$ is the set of *seasonal* MA coefficients, D is the degree of the seasonal differential operator ∇_S, and S is length of a single cycle of the *seasonal* component in the analyzed time series.

If we connect the *seasonal* model described by (21) with the *common* ARIMA model (see dependence (19)), the multiplicative ARIMA model of $(\boldsymbol{p},\boldsymbol{d},\boldsymbol{q}) * (\boldsymbol{P},\boldsymbol{D},\boldsymbol{Q})$ order will be given by the equation:

$$a(B)a_S(B)\ \nabla_S^D\ \nabla^d y_t = b_S(B)\varepsilon_t \tag{22}$$

Figure 5 shows the algorithm of building an ARMA/ARIMA model of time series in the form of the block diagram. As you can see in this diagram the algorithm can be divided into a number of groups of operations which may be described as:

1) Identification of parameters of the ARMA/ARIMA model of the time series Y_t: the parameter/parameters $\boldsymbol{p}$ (AR parameter) and/or $\boldsymbol{q}$ (MA parameter), as well as the parameter $\boldsymbol{d}$ (degree of the operator ∇) are fitted in order to describe in general the structure of trends in the analyzed time series;
2) Estimation of the coefficients of the model: the values of weights $\{a_i : i = 1,...p\}$ and/or $\{b_j : j = 1,...q\}$ are estimated in this step.
3) Testing of the model: checking of the presence of systematic components in the time series of differences between the original and the forecasted values of time series elements (called *residuals*) which are calculated according to the fixed forecasting algorithm. If a systematic component in the time series of residuals appears, then it shall be assumed that the model is inadequate to the analyzed time series. As we can see in Figure 5, a negative answer in testing the ARMA model may affect:
   i) a verification of AR and/or MA coefficients,
   ii) a verification of the parameters of the model.

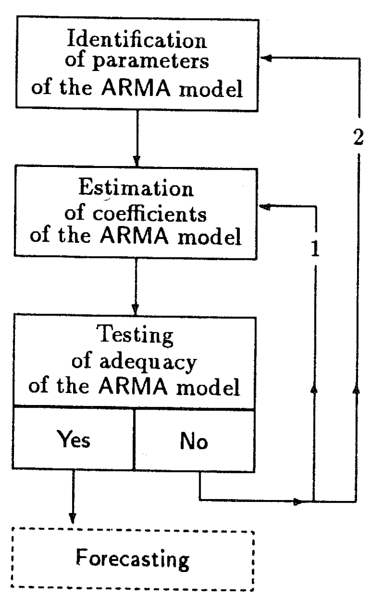


Fig. 5. Block diagram of the algorithm of generating ARMA models of time series.

A large choice of computational algorithms which can be used in all the mentioned stages of building the ARMA/ARIMA model is presented in many papers (Box and Jenkins, 1976; Cadzow, 1983; Kaźmierczak, 1989). However, if we intend to apply these algorithms for the specific needs of processing diagnostic data we will have to take into consideration some particular conditioning of such an application.

Let's now try to use the chosen method for forecasts calculations in the given time series Y_t. In order to do it, in this time series we mark out the element $\{y'_t : y'_t \in Y_t\}$ and we assume that the time point t' is the *actual* moment on the time axis. By means of a subset of elements of the time series $\{y_t : t \in [(t'-M);t']\} \subset Y_t$ we have constructed a model for the analyzed time series. By the term to *construct a model* or *to generate a model* we should understand the fixing of the set of the model parameters. The algorithm of forecast calculation $\widehat{y}_{t'+la}$ with the lead-time la by means of the model we can show as the following sequence of operations:

- we assume that the subset of elements of the time series Y_t is given and the *actual* element y'_t is the last one in this subset (in the direction of increasing time),
- making use of the model we calculate the forecast of the elements of the analyzed time series in the time point $t'+1$,
- using the forecast $\widehat{y}_{t'+1}$ of the $(t'+1)$–th element of Y_t we are then able to calculate the forecast $\widehat{y}_{t'+2}$ of the $(t'+2)$–th element.

When the forecast of the element of Y_t with the identifier $t'' = t' + la$ has been calculated in compliance with the above presented routine, the iterative procedure is finished. The algorithm does not make any limits to the lead–time la of the calculated forecasts. Anyway, we should remember that for higher values of la we can compute *forecasts from forecasts*. Because of that, the likelihood (accuracy) of the forecasts is more and more limited.

Using the above presented example we are able to mention the areas which can be supported by computer tools as well as the range of the assistance in each mentioned case. Let's name the successive steps as follows:

a) choice of the proper forecasting model,
b) necessary calculations in every steps of the algorithm of fitting the model to the given set of data (time series),
c) calculations of forecasts with the fixed lead–time,
d) estimation of accuracy (quality) of forecasts.

It is well observed that the mentioned steps should require various computer tools. For the steps (b), (c) and (d) the computer supporting consists of arithmetic calculations. In the case (a) the problem is more complicated. We have here the example of the task which can be supported by more advanced computer tools, namely the methods from the area of knowledge engineering or artificial intelligence *(expert systems)*. In particular, such classes of research problems in machine diagnostics like multicriterial choice of algorithms or various cases of drawing conclusions are typical cases of applying expert systems for the purpose of computer aiding of diagnostic data processing.

We have shown above the example of application of a general method in the particular field of machine diagnostics. It must be stated, that such an application usually considering needs some specific requests of the field which may affect the manner of using the chosen method. In the discussed case of modelling time series of diagnostic data some particular aspects of the meaning of the term time *ought* to be considered.

5. Time Domain in Machine Diagnostics and Monitoring

The machine states and values of state symptoms vary in the domain of time. Thus time–ordered sequences of data which describe such variability may be treated as representations of particular processes. The time sequence of states of an object may be considered as a representation of the *diagnosed process* and the sequence of values of diagnostic symptoms – as a projection of the *diagnostic signal.*

It is a specific feature of machine diagnostics that one can make the question: can the *natural* meaning of the term *time* be used directly in this area of problems? Usually, the commonly accepted meaning of this term is treated in diagnostics as the sets of time: we will talk in these cases about *real time.* But the processes considered in machine diagnostics are not necessarily continuous in the real time. It is connected for instance with the problems of breaks in the machine functioning. It has also to be taken into consideration that in processing of diagnostic signals the *instant* values of the features of signals are estimated not in points but in intervals of time. Thus, it may be stated that both the diagnosed process and the diagnostic signals occur in some *other* time domain, because the considered intervals between changing states and features are significantly longer than intervals in the domain of signal processing.

Because of the aforesaid reasons it was proposed to distinguish in the field of diagnostic research a number of *separate* sets (domains) of time (Cholewa and Kaźmierczak, 1992):

- real time $\{\nu\}$,
- micro–time $\{\tau\}$,
- macro–time $\{t\}$.

The following mapping is possible:

$$\nu \rightarrow t \bigwedge(\nu, t) \rightarrow \tau \tag{23}$$

or simply

$$\nu \rightarrow (t, \tau) \tag{24}$$

so that for

$$\nu_1 \rightarrow (t_1, \tau_1)$$

$$\nu_2 \rightarrow (t_2, \tau_2) \tag{25}$$

we have

$$t_1 > t_2 \Leftrightarrow \nu_1 > \nu_2 \tag{26}$$

$$(t_1 = t_2 \bigwedge \tau_1 > \tau_2) \Leftrightarrow \nu_1 > \nu_2 \tag{27}$$

$$\nu_1, \nu_2 \quad \text{are incomparable for} \quad t_1 \neq t_2 \tag{28}$$

In terms of the micro–time τ and the macro–time t we can discuss here the problem of *instantaneous* states of processes considered in machine diagnostics. In practice the *instant* states of diagnosed processes and instant values of diagnostic signals are identified in the micro–time interval $[\tau, \tau + \Delta\tau]$. The term *instant* used in this case is relative in some sense because the values of signal features are measured and estimated at a fixed point t of macro–time and in a closed interval of the micro–time τ.

The variable which orders the elements of the discussed sequence in time has not been necessarily given in *real* units of time (seconds, hours, years etc.). A value, a like for instance number of kilometers passed by a vehicle since the moment of its leaving the assembly belt (or since its last repair or maintenance) or a number of hours of operation (the so– called *motor–hours*), may also be treated as a basis to fix the succession of elements in time series. In the second case we have to take into consideration the fact that the time–indexing value is not necessarily continuous in the sense of the real time ν axis (for instance some exploitational breaks may appear in the functioning of any particular machine or industrial installation). Thus, we should state that the existence of any indexing value connected with the leak of time is a sufficient condition to use the expression *time series* to describe a set of discrete data.

The important requirement for time series is the demand of a constant *time–step* between successive elements of the sequence. This demand should be understood in such a way that the value of a time–indexing function $f(t)$ for every element of the time series should differ from the value $f(t)$ for the directly preceding element, as well as the directly following one with a constant value, so that:

$$\text{dist}(t_i, t_{i+1}) = f(t+1) - f(t) = f(t) - f(t-1) = \ldots = \text{const} \tag{29}$$

where $f(i)$ is the function which orders the elements of time series, and $\text{dist}(t_i, t_{i+1})$ is the distance between two successive elements along the axis of the t time.

It is interesting to search effects of connecting the above formulated general requirement with specific needs of diagnostics as a particular field of implementing some general methodology. If the analyzed time series is a projection of process in the domain of macro–time the requirement of the constant time step will have some specific meaning (Kaźmierczak, 1989).

Apart from the problem of time–step, the results of primary research may cause that in a particular diagnostic experiment the states of objects are identified by more than one symptom. The number of symptoms taken into account in the experiment refers certainly to the dimensionality of the time series. Such an approach makes it possible to use single– or multidimensional algorithms for the needs of analyzing diagnostic data. On the other hand, it has been proved in investigations that, because of the specific properties of diagnostic investigations, in many cases the variability of multiplied symptoms in time can be considered separately i.e. using the proper number of single–dimensional time series.

In diagnostic investigations of machines we have usually at disposal values of features of signals which enable us to formulate conclusions about states of the machine. If we measure such values along the axis of exploitational time they may be ordered in the form of time series and then processed as time series. Of course, the specific nature of diagnostic data provokes some troubles in such ordering of the data. There are two main reasons for such troubles:

1) breaks in the functioning of machines of which may have a different origin and have to be treated in a different manner when the time series are formed and analyzed,
2) the number of values which are used for the needs of state recognition of the diagnosed object.

The first one of the mentioned factors affects the value of time step in time series, the second one refers to the *dimensionality* of the time series.

The necessity of adapting general methods and algorithm to particular requests of the field of application opens other possibilities of using the advanced computer tools for the needs of tasks supporting. This area of problems may be also effectively supported using tools as expert systems.

6. Practical Solutions of Computer–Aided Data Processing in Machine Diagnostics and Monitoring

The results of research which have been carried out in the Chair of Fundamentals of Machine Design of the Silesian Technical University in Gliwice have proved the possibility to support effectively a number of particular tasks in the field of machine diagnostics by means of computer tools.

At the first step the computer system was developed for the needs of processing diagnostic signals. The system, called *Programmable Signal Analyzer PAS6* (Cholewa, 1988) has been designed in the *friendly–for–user* formula. It offers the user a very rich range of calculating routines which performs basic as well as more specialized analysis of random signals. The system disposes the own data base, graphic driver and convenient manner of management based on multilevel menus. The utilization of the PAS system in many practical cases of diagnostic research has given the authors good experience in computer–aided diagnostic signal processing.

According to the approach introduced in this paper the PAS system should be treated as a tool of performing tasks connected with the *vertical* part of the net of relations in diagnostic research. Because of that, in the following steps of the research works two main problems have been undertaken:

- development of the tool of supporting the time–domain analysis of diagnostic data,
- development of the tool of supporting the drawing conclusions in machine diagnostics.

As the result of the first mentioned effort, the system of processing time series of diagnostic symptoms was developed (Kaźmierczak, 1989). The system is now in a prototype version and enables user to carry out following analysis:

1) detection of trends in time series,
2) fitting the ARMA/ARIMA linear models of time series,
3) calculating forecasts in time series using the fitted models.

The system is developed as a perspective tool of computer–aided data processing from machine monitoring. At present the special attention is paid to problems of data *management* in monitoring systems (data bases, formats of input data for the calculating routines, visualization of data in various steps of the processing). In order to present possibilities of the discussed system let's show some results of its functioning.

Figure 6 shows an example of visualization of a single–dimensional time series of the diagnostic symptom. The figure shows clearly how some problems of visualization of this kind of data has been solved. In particular the axis of time in the diagram represents the *exploitational time*, i.e. *pure* time of functioning.

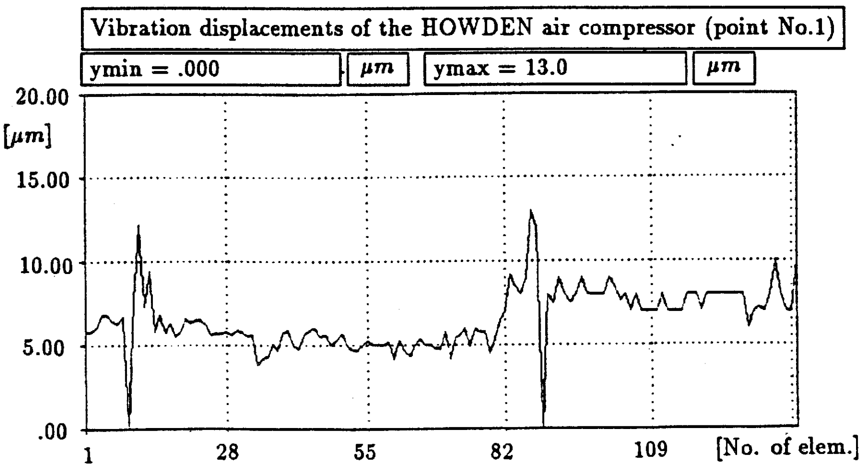


Fig. 6. Example of time series of a diagnostic symptom.

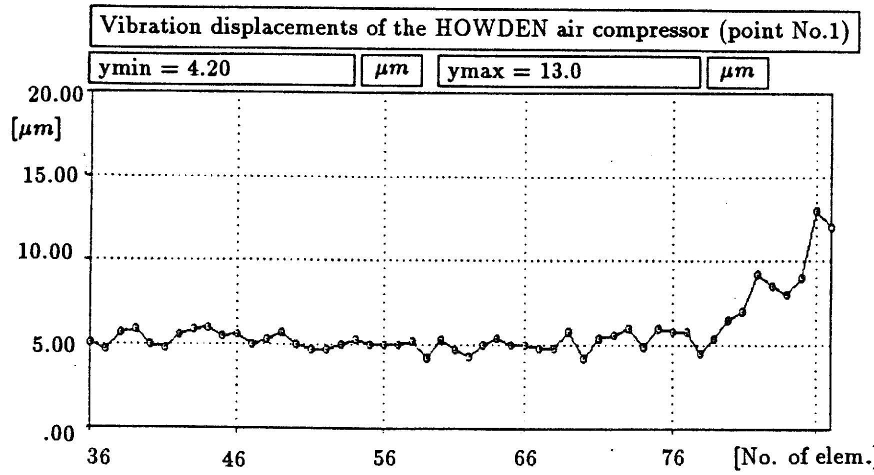


Fig. 7. Subset of elements of the time series plotted in Figure 6.

The breaks are marked in the figure by additional *zero* elements of the series. The tools of visualization of such data should offer additionally a number of possibilities like the possibility of zooming selected parts of the time series or the possibility of scrolling the diagram when the new *actual* element is added into the time series. Figure 7 shows the selected subset of elements of the time series presented in Figure 6 which will be used for the needs of illustrating results of processing the time series in the system of machine monitoring.

In analysis of time series of diagnostic symptoms as results of machine monitoring we are the most often looking for some tendencies (*trends*) in the time history of the investigated values. There is a number of algorithms which make it possible to detect trends in time series (Batko and Kaźmierczak, 1985; Cholewa and Kaźmierczak, 1992). For the needs of data processing in systems of machine monitoring these algorithms seem to be the most useful which offer quantitative measures of presence as well as intensity of trends in the analyzed set of data. The quantitative values are a better base for drawing conclusions about changes of states of the investigated object, especially when we want to support the process of drawing conclusions using computer tools (like the so-called *expert systems*). Figure 8 illustrates results of detecting trends in the time series shown in Figure 7 using the exemplified quantitative measure (*Spearman's rank correlation coefficient*). Comparing the results with the *original* time series we may state that the measure of trend enables the user to conclude about the growing-up tendency in time series far earlier than it can be observed in the diagram shown in Figure 7.

Apart from the analyzing of trends in the *historically existing* part of time series we may also want to predict such trends in future: beyond the *actual* point of time. We may formulate the problem of such prediction (*forecasting*) in the terms introduced above. Forecasts of diagnostic symptoms may be in turn the basis for conclusions concerning future states of the investigated machine. Anyway, the methods of reasoning go beyond the methods discussed in this paper. In known research works (Batko and Kaźmierczak, 1985; Cempel, 1987; Kaźmierczak, 1989) which discuss the problems of diagnostic forecasting the stage of drawing conclusions has been commonly dealt with in the same way as in *classic* diagnostics, i.e. by comparing the forecasts with a fixed threshold/limiting value. Figure 9 shows an example which illustrates the results of performing the forecasting task. The twin diagram compares values of the *original* time series (in this case: values of vibration displacements measured in a chosen point on the shell of the air compressor) with the forecasts calculated by means of a fixed model. In Figure 10 the values of residuals corresponding to Figure 9 are presented.

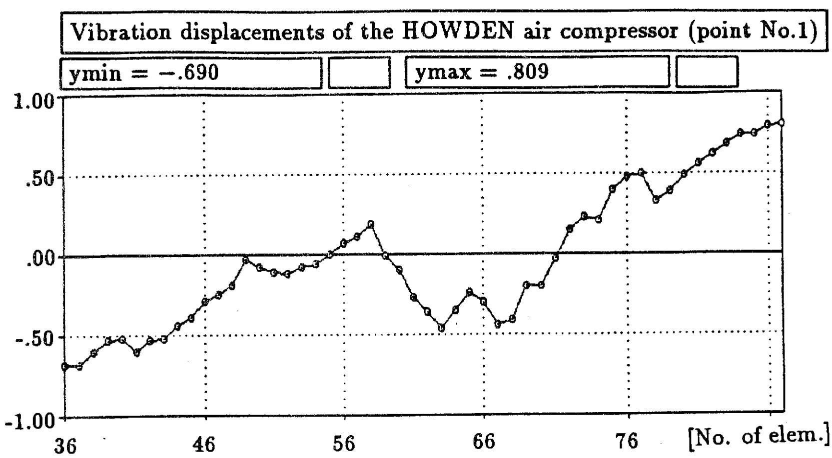


Fig. 8. Values of the trend measure for the subset of time series shown in Figure 7.

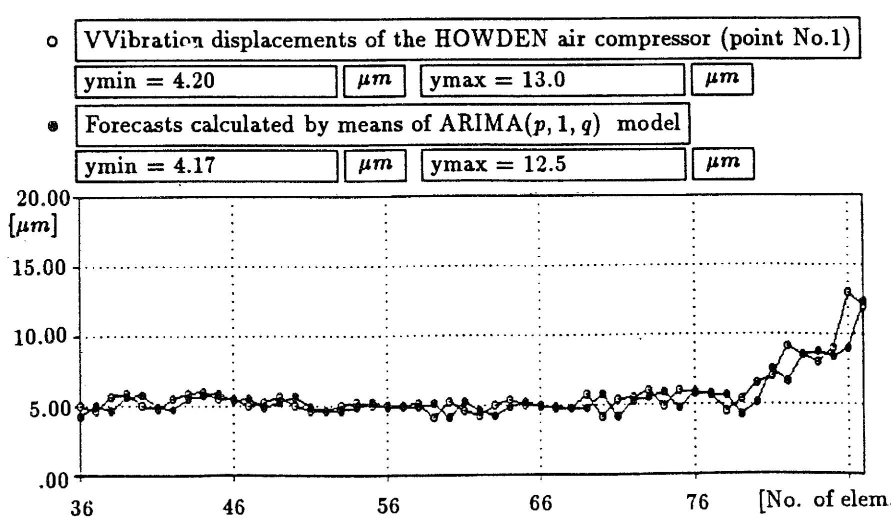
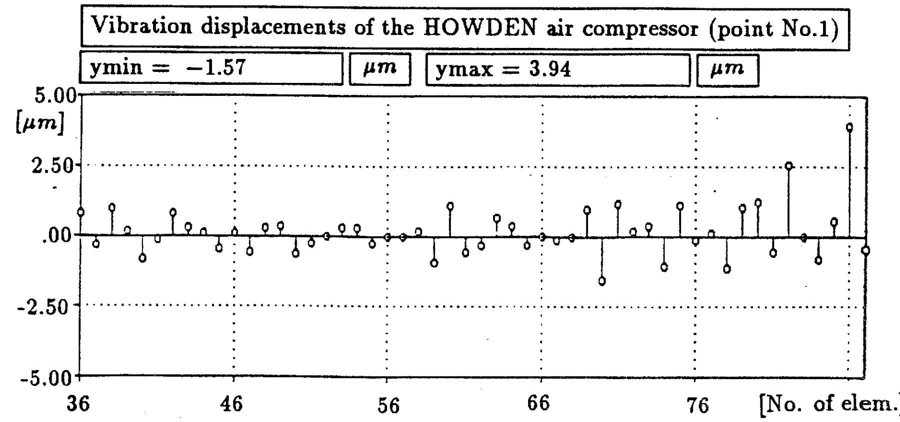


Fig. 9. Example of the forecasting task of type 1 in diagnostics (A).



Sum of residuals SUM= 9.0395 μm
Average value of residuals MER= .1738 μm
Standard deviation of residuals SD= .9166 μm

Fig. 10. Example of the forecasting task of type 1 in diagnostics (B).

In the Chair of Fundamentals of Machine Design in Gliwice the works concerning the advanced tools of computer–aided processing of diagnostic data have been carried out, too. In particular, the shell expert system called VV_SHELL has been also developed by Cholewa (1992, 1988). The system is meant first of all as a tool supporting diagnostic data processing as well as of other kind of engineering tasks (for example in the field of CAD).

There is a number of possible solutions of programs which belong to the family of *expert systems*. The most significant feature of all these programs is a manner of representing, storing and processing knowledge which is used in a particular system. In the VV_SHELL the knowledge is stored in the form of *frames*. The frame describes a selected object or a group of objects and its structure contains a number of fields called *slots*. Every slot contains *facets* which are storing *descriptors* of the object/objects covered by the frame. The values of the features of the object or the connections between this object and other ones may be treated as the descriptors. The frames, slots and facets are identified by means of names so that in one frame two slots with the same name and in one slot two facets with the same name can not exist.

The VV_SHELL system has some particular features of the above mentioned structure. In every slot the facets of special destination may appear which are marked as *value*, *if_added*, *if_needed*, *if_removed*. Value of the facet *value* is understood as the value of the slot. If such a facet doesn't exist then the facet of the name *if_needed* will be looked for which should contain the description of the proper subroutine (the so–called *demon*). The value which is given back as the result of running this subroutine is treated as the value of the slot. The facets *if_added* and *if_removed* should contain the subroutines which are run for the purpose of adding something to the contents of the facet value or removing the contents of it, respectively.

For every frame the list of higher-order (superior) frames may be determined. The list of frames which are directly superior to the considered frame is contained in the slot with the name ako in this frame. The lack of such a slot means that for the considered frame no superior frames have been determined. The contents of the facets of the superior frames in the VV_SHELL system are inherited. If the sought facet doesn't exist in the slot of a given frame or if the demon contained in it can not be run then the search will be continued in the superior frame.

In order to use the VV_SHELL for a particular purpose, i.e. for carrying out calculations, the only thing to do is to complement the knowledge base. The complementation of the knowledge base in the VV_SHELL system consists of recording the proper frames. The knowledge base in the system can be divided into two parts which deal with two types of frames:

1) frames which contain information about the subject of activity (referring to decisive tables),
2) *organizing* frames which utilize information from the first part for the needs of controlling the run of action (for instance – the run of calculation).

The VV_SHELL is at present tested and applied for some limited tasks (see Wyczółkowski and Kaźmierczak, 1992).

7. Conclusions

This paper should play a double role according to intention of the author. First, it introduces the general concept of the structure of machine diagnostic research and shows some possibilities of supporting particular tasks in this research by means of computer tools. Secondly the paper presents an overview of works which were carried out in the Chair of Fundamentals of Machine Design of the Silesian University of Technology. This part of paper shows by means of practical examples what areas of the diagnostics and machine monitoring may be effectively supported by computer tools. On the other hand, if the monitoring system is intended to be used in industrial practice the problem will appear how to make this system to be *friendly-for-user*. There is a range of questions, for instance, concerned with the visualization of results of data processing or with the supporting of drawing conclusions by the user. The last problem may be solved using the tools from the field of *artificial intelligence* which is proved in research works carried out in the Chair of Fundamentals of Machine Design (the project KBN PB-132/9/91).

References

Batko W. and Kaźmierczak J. (1985): *Forecasting in technical diagnostics.* — Proc. VII School *DIAGNOSTICS'85*, Rydzyna, Poland, Sept., pp.31–42.

Box G.E.P. and Jenkins G.M. (1976): *Time Series Analysis – Forecasting and Control.* — San Francisco: Holden-Day.

Cadzow J.A. (1983): *ARMA time series modelling: An effective method.* — IEEE Trans. Aerospace and Electronic Systems, v.AES–19, No.1, pp.49–58.

Cempel Cz. (1987): *Simple condition forecasting techniques in vibroacoustical diagnostics.* — Mechanical Systems and Signal Processing, v.1, No.1, pp.75–82.

Cholewa W. (Ed.) (1988): *Programmable signal analyser PAS6.* — Zeszyt IMiPKM Nr 109, Politechnika Śląska, Gliwice (in Polish).

Cholewa W. (Ed.) (1988): *Expert system VV_SHELL - User's manual.* — Raport Nr G–547/RMT–4/91, Instytut Mechaniki i Podstaw Konstrukcji Maszyn Politechniki Śląskiej, Gliwice (in Polish).

Cholewa W. (1992): *Diagnostic reasoning.* — Proc. 1st Inter. Scientific Conf. *Achievements in the Mechanical and Material Engineering*, Gliwice, Poland, Sept., pp.19–28.

Cholewa W. and Kaźmierczak J. (1992): *Technical Diagnostics of Machines. Processing of Signal Featurs.* — Gliwice: Silesian University of Technology Press (in Polish).

Kaźmierczak J. (1984): *The representation of the functioning of an electric arc furnace in the space of aconstic signals.* — Archives of Acoustics, v.19, No.3, pp.201–208 (in Polish).

Kaźmierczak J. (1989): *The use of linear models of stochastic processes for the needs of forecasting in machine diagnostics.* — Zeszyty Naukowe Politechniki Śląskiej, v.1037, series: Mechanics No.95, Gliwice (in Polish).

Ljung L. (1983): *Theory and Practice of Recursive Identification.* — Cambridge: MIT Press.

Wyczółkowski R. and Kaźmierczak J. (1992): *An application of an expert system for the needs of aiding the choice of algorithms in engineering calculations.* — Proc. Inter. Scientific Conf. *Computer Integrated Manufacturing*, Gliwice–Zakopane, Poland, pp.425–434.

Received February 12, 1993
Revised June 15, 1993