10.24425/acs.2023.146281

Archives of Control Sciences
Volume 33(LXIX), 2023
No. 2*, pages 391–411*

Discrete identification of continuous non-linear and non-stationary dynamical systems that is insensitive to noise correlation and measurement outliers

Janusz KOZŁOWSKI and Zdzisław KOWALCZUK

The paper uses specific parameter estimation methods to identify the coefficients of continuous-time models represented by linear and non-linear ordinary differential equations. The necessary approximation of such systems in discrete time in the form of utility models is achieved by the use of properly tuned 'integrating filters' of the FIR type. The resulting discrete-time descriptions retain the original continuous parameterization and can be identified, for example, by the classical least squares procedure. Since in the presence of correlated noise, the estimated parameter values are burdened with an unavoidable systematic error (manifested by asymptotic bias of the estimates), in order to significantly improve the identification consistency, the method of instrumental variables is used here. In our research we use an estimation algorithm based on the least absolute values (LA) criterion of the least sum of absolute values, which is optimal in identifying linear and non-linear systems in the case of sporadic measurement errors. In the paper, we propose a procedure for determining the instrumental variable for a continuous model with non-linearity (related to the Wienerian system) in order to remove the evaluation bias, and a recursive sub-optimal version of the LA estmator. This algorithm is given in a simple (LA) version and in an instrumental variable version (IV-LA), which is robust to outliers, removes evaluation bias, and is suited to the task of identifying processes with non-linear dynamics (semi-Wienerian/NLID). In conclusion, the effectiveness of the proposed algorithmic solutions has been demonstrated by numerical simulations of the mechanical system, which is an essential part of the suspension system of a wheeled vehicle.

Key words: instrumental variable, non-linear continuous-time models, optimization, supervisory or security systems, system identification

J. Kozłowski and Z. Kowalczuk are with Gdansk University of Technology, WETI, Narutowicza 11/12, 80-952 Gdańsk, e-mail: {jk23, kowa}@eti.pg.edu.pl, ORCIDs: JK: 000-0002-8608-7754; ZK: 0000-0001-9174-546X.

Received 06.03.2023. Revised 24.04.2023.

1. Introduction

In supervisory systems, dedicated methods of identifying an entire process or detecting some changes are usually implemented in order to obtain appropriate diagnostic information about the evolution of the monitored process. The dynamics of a process can conveniently be modeled using a simple input-output description or a more complex state-space representation. The usage of easy-to-use discrete representations of the observed process in the form of difference equations simplifies the problem of computer implementation of identification schemes for obvious reasons. This is because in such mathematical models the regression data takes the simple form of delayed samples of the recorded inputs and outputs. On the other hand, these representations are not physically motivated and not intuitive, and the parameter values of these mathematical models do not have a physical scale (unit) and strongly depend on the sampling frequency used [1–3].

Whereas in the case of mathematical descriptions of the continuous-time nature, the modeling parameters can usually be assigned specific physical units. The original parameterization of the model is also not affected by other data processing conditions (e.g. sampling period). Moreover, thanks to the intuitive and interpretable physical parameters, the initial verification of the estimation results is possible and very easy [4, 5].

It should be realized that all identification algorithms are in fact numerical procedures and therefore the estimation of parameters that shape continuous models must be based on sampled input-output data. Consequently, the operation of approximating continuous-time concepts (quantities, like derivatives appearing in differential equations) in discrete time should be performed [6–16].

Some time ago, some hopes were pinned on the so-called "delta" approach, which uses finite differences to imitate the original differentials, which at certain computational costs can be used in the design of both identification systems and control systems [17]. However, it has also been proven analytically and numerically that this approach can lead to significant inaccuracies in the numerical mechanization of descriptions in continuous time [18].

Moreover, in practice, an 'integral' approach can easily be applied, which overcomes the noise attenuation problem of delta-based modeling [19]. Indeed, a properly crafted low-pass multiple-integration operator eliminates the detrimental additive noise to a large extent, but the regression variables themselves computed as multiple integrals of the inputs and outputs inevitably diverge to infinity. The integrated free response of the system (due to unknown non-zero initial conditions) must also not be underestimated.

Fortunately, an effective remedy for the above problems seems to be the use of dedicated low-pass operators, such as the Poisson IIR filter [4] or the Sagara FIR filter [1]. Note that with such IIR or FIR filters, the additive noise (characteristic of measurement signals) is eliminated, the values of the regressors (calculated as filtered derivatives of the input and output signals) are limited, and the properly filtered transient response of the system (in effect tending to zero) does not affect on the ultimate accuracy of such identification (and modeling).

In practical industrial automation systems, the result of processing measurement data is usually affected by noise and various system disturbances. The emergence of this type of parasitic phenomena often results in a significant deterioration in the quality of the controlled industrial process. In general, high-frequency additive noise (e.g. from quantization in AD converters) can be effectively eliminated with properly tuned low-pass anti-aliasing prefilters. In turn, in the case of systematic errors, the problem is usually solved by proper calibration of measuring instruments and the use of dedicated compensation techniques. Importantly, the results of diagnostic procedures that rely on classical least squares (LS) methods can turn out to be very skewed, especially when the measurements are contaminated with sporadic errors called outliers. A radical improvement in the estimation quality can be easily obtained by implementing an identification method resistant to occasional errors in the sense of the least sum of absolute values (LA).

Reliable modeling and effective identification of dynamic systems in continuous time play a key role in the supervision and control of various industrial processes [20–23]. We will discuss such problems [24] regarding the estimation of the parameters of models of physical systems in the following sections of this article.

In Section 2, mechanizations of the discrete-time equivalents of linear and non-linear differential equations are presented. In particular, an efficient method of representing the specific (NLID) non-linearity present in the differential equation of a certain practical system (car suspension) is proposed. Section 3 shows the various parameter estimation procedures along with a discussion of their asymptotic properties and robustness (insensitivity) to outliers in the processed data. Our novelties, introduced to the system identification practice, are recursive algorithms optimal in the sense of the least sum of absolute values (basic LA and IV-LA) and an innovative method of generating instrumental variables for the analyzed non-linear dynamic objects (called semi-Wienerian systems).

As a consequence of the proposed methodology, Section 4 discusses the results of several numerical tests that demonstrate the effectiveness of the described estimation procedures. At the end of the work, in Section 5, the originality of the research results is emphasized and the prospects for further research in the field of robust identification of practical processes are outlined.

2. Continuous-time modeling

Let the dynamics of the supervised industrial system (the object of parameter estimation) be expressed by the ordinary differential equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y = b_0 u, \tag{1}$$

where n denotes the system order, while $u = u(t)$ and $y = y(t)$ stand for the input and output signals, respectively. The description (1) can be subject to any initial conditions, and the parameterization coefficients $\{a_{n-1}, \ldots, a_0, b_0\}$ are assumed to be unknown. Since the practical identification of continuous-time models must be based on the processing of sampled data, an appropriate discrete-time approximation of (1) must be used.

2.1. Discrete approximation of differential equations

There are many practical approaches to the numerical mechanization of differential equations, such as the solution using the *delta* operator for direct evaluation of derivatives, the classical approximation method based on multiple integration of both sides of the differential equation, and the use of suitably tuned low-pass *matching* (or *state variable*) filters [1,4].

Unfortunately, due to the high-pass nature of the delta operator, additive noise is amplified, which can significantly distort the evaluated continuous quantities (derivatives). Moreover, in the case of pure integration, regression data represented by multiple integrals of signals (inputs and outputs) inevitably tend to infinity. In addition, the free response of the integrator also affects modeling accuracy and usually cannot be ignored. The above indicated problems can effectively be eliminated by employing low-pass matching filters used in the discrete-time approximation synthesis for the model (1).

Among others things, a finite-horizon integrating filter of the FIR type deserves attention [1]. The resulting operator, referred to as the 'linear integral filter' (LIF), takes the form of a multiple integral of the signal (or its i-th derivative) in a finite time horizon $[t-\tau,\ t]$

$$J_i^n x(t) = \int_{t-\tau}^{t} \int_{t_1-\tau}^{t_1} \cdots \int_{t_{n-1}-\tau}^{t_{n-1}} x^{(i)}(t_n)\,\mathrm{d}t_n \cdots \mathrm{d}t_2 \mathrm{d}t_1\,, \tag{2}$$

where the length of the integration horizon (τ) can easily be selected so as to obtain a proper bandwidth of (2).

Applying (2) to both sides of the model (1), the following "integral" equation with the original parameterization of the underlying continuous-time system is obtained

$$J_n^n y + a_{n-1} J_{n-1}^n y + \ldots + a_0 J_0^n y = b_0 J_0^n u. \tag{3}$$

It is important for modeling consistency that with (2) the effect of the "integrated" free response of the system becomes irrelevant after a finite time ($n\tau$). The integrals used above can easily be calculated based on sampled data. With the aid of the bilinear (Tustin's) operator the discrete-time mechanization of (2) can be shown as

$$J_i^n x(t)\,\Big|_{t=kT} \approx I_i^n x(k), \tag{4}$$

$$I_i^n = \left[\frac{T}{2}\left(1+q^{-1}\right)\right]^{n-i}\left(1-q^{-1}\right)^i\left(1+q^{-1}+\ldots+q^{-L+1}\right)^n, \tag{5}$$

where T stands for the sampling time, q^{-1} symbolizes the unit delay operator, L is the "numerical" length of the integration horizon ($\tau = LT$), and for brevity, the time moment $t = kT$ is represented by the index k. As a result of the above assumptions, the discrete equivalent of the original continuous-time model takes the convenient form of a regression equation for the *converted output*

$$I_n^n y(k) = \chi(k) = \phi^{\mathrm{T}}(k)\,\theta + e(k), \tag{6}$$

$$\phi(k) = \left[-I_{n-1}^n y \quad \ldots \quad -I_0^n y \quad I_0^n u\right]^{\mathrm{T}}, \tag{7}$$

$$\theta = [a_{n-1} \quad \ldots \quad a_0 \quad b_0]^{\mathrm{T}}, \tag{8}$$

where the reference term $\chi(k)$ represents the converted output of the discrete-time counterpart of the original continuous-time model.

The above model takes into account a residual or equation error $e(k)$, which is a stochastic component and includes both system disturbances and other inaccuracies of this type of modeling. It is worth noting that the model (6)–(8) retains original continuous parameterization, while the "integral" regressors (7) are numerically well-conditioned (bounded). According to the rule proposed by [1], the horizon L (or $\tau = LT$) should be selected so that the frequency bandwidth of the filter (2) matches as closely as possible the frequency band of the identified system (1). The LIF operator with a too narrow frequency band (large τ and L horizons) simply falsifies the system dynamics (3). On the other hand, when this bandwidth is too large (small τ and L), the broadband noise strongly affects the accuracy of the estimation (due to bias).

2.2. Non-linear continuous-time models

Consider now a continuous-time differential equation model

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y^{(1)} + f(y) = b_0 u \tag{9}$$

with the non-linear term $f(y)$ representing an unknown function such that $f(0) = 0$. Here we can apply a method suitable for many practical cases, which is to use the series expansion as an approximation to the non-linear term, provided that $f(y)$ is at least r times differentiable,

$$f(y) = c_1 y + c_2 y^2 + \ldots + c_r y^r. \tag{10}$$

Now, using the FIR integrating filters (5), the regression equivalent of the non-linear system (9) can be

$$I_n^n y(k) = \chi(k) = \phi^{\mathrm{T}}(k)\,\theta + e(k), \tag{11}$$

$$\phi(k) = \left[-I^n_{n-1}y \quad \ldots \quad -I^n_1 y \quad -I^n_0 y \quad \ldots \quad -I^n_0 y^r \quad I^n_0 u\right]^{\mathrm{T}}, \tag{12}$$

$$\theta = [a_{n-1} \quad \ldots \quad a_1 \quad c_1 \quad \ldots \quad c_r \quad b_0]^{\mathrm{T}}, \tag{13}$$

where, as before, $\chi(k)$ is the converted output.

Once a regression representation of the non-linear system is obtained, typical identification procedures can be applied that will allow us to appropriately evaluate the unknown but physically motivated coefficients.

3. Estimation procedures

As shown, the dynamics of a supervised system can be expressed using the regression model (6)–(8) or (11)–(13). Now the system parameters can be effectively estimated with appropriate identification procedures. Three identification methods are considered below: the least squares algorithm, the instrumental variable method, and the least absolute values procedure.

3.1. Least squares method

The classical weighted LS estimation scheme results from the minimization of the quadratic loss index

$$V_{\mathrm{LS}} = \sum_{\ell=1}^{k} \lambda^{k-\ell} e^2(\ell) = \sum_{\ell=1}^{k} \lambda^{k-\ell} \left[\chi(\ell) - \phi^{\mathrm{T}}(\ell)\,\theta\right]^2, \tag{14}$$

where the weighting factor λ constrained to unity ($0 \ll \lambda \leqslant 1$) controls the rate of exponential forgetting when tracking variable parameters of a non-stationary system.

The above LS criterion can be analytically minimized by nullifying its gradient

$$\nabla_\theta V_{\mathrm{LS}} = -2 \sum_{\ell=1}^{k} \lambda^{k-\ell} \phi(\ell)\,[\chi(\ell) - \phi^{\mathrm{T}}(\ell)\,\theta] = \mathbf{0}. \tag{15}$$

As a consequence, the weighted (WLS) estimator takes the following algebraic form

$$\hat{\theta}(k) = \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \phi(\ell)\,\phi^{\mathrm{T}}(\ell)\right]^{-1} \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \phi(\ell)\,\chi(\ell)\right]. \tag{16}$$

It is easy to verify that the LS procedure (using $\lambda = 1$) generates consistent estimates, when the regression data $\phi(k)$ and the residual error $e(k)$ are uncorrelated: $\mathrm{E}\{\phi(k)\,e(k)\} = \mathbf{0}$. By definition, this condition is met when $e(k)$ takes the form of zero-mean white noise (a sequence of independent random variables).

Most often, the residual error turns out to be correlated, which makes the LS estimators asymptotically biased. The consistency of it can be significantly improved by using the technique of instrumental variables (IV).

3.2. Instrumental variable method

The idea behind the IV method stems from the analysis of the rearrangement (15) that leads to minimization of the LS index (14). Since consistency problems are attributed to the non-zero correlation $\mathrm{E}\{\phi(k)\, e(k)\} \neq \mathbf{0}$, one can replace the original regressors $\phi(k)$ in (15) with a suitably selected instrument $\xi(k)$, provided $\xi(k)$ and $e(k)$ are uncorrelated: $\mathrm{E}\{\xi(k)\; e(k)\} = \mathbf{0}$. In this way the algebraic IV algorithm can be written down as

$$\hat{\theta}(k) = \left[\sum_{\ell=1}^{k} \lambda^{k-\ell}\, \xi(\ell)\, \phi^{\mathrm{T}}(\ell)\right]^{-1} \left[\sum_{\ell=1}^{k} \lambda^{k-\ell}\, \xi(\ell)\, \chi(\ell)\right]. \tag{17}$$

To avoid the non-recommended inversion of the matrix present in (16) and (17), the well-known matrix inversion lemma [25] can be used. The resulting recursive form of both algorithms, including prediction error estimation, covariance matrix update and final correction of the estimation vector, can be presented as

$$\varepsilon(k) = \chi(k) - \phi^{\mathrm{T}}(k)\, \hat{\theta}(k-1), \tag{18}$$

$$P(k) = \frac{1}{\lambda}\left[P(k-1) - \frac{P(k-1)\, \psi(k)\, \phi^{\mathrm{T}}(k)\, P(k-1)}{\lambda + \phi^{\mathrm{T}}(k)\, P(k-1)\, \psi(k)}\right], \tag{19}$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)\, \psi(k)\, \varepsilon(k) \tag{20}$$

with $\psi(k) = \phi(k)$ and $\psi(k) = \xi(k)$ for the LS and IV estimation schemes, respectively. Typically, the initial diagonal covariance matrix is set as $P(0) = \mathrm{diag}[10^5 \ldots 10^5]$ at the start-up of these estimators.

According to its purpose, the instrumental variable $\xi(k)$ should match the original regressors $\phi(k)$ as much as possible, while keeping $\xi(k)$ and $e(k)$ uncorrelated: $\mathrm{E}\{\xi(k)\, e(k)\} = \mathbf{0}$. Among the various implementations of such variables, it is preferable to use a method based on simple deterministic filtering [1,26]. The desired noise-free output of the linear system (1) can be obtained from

$$\hat{y}(k) = \frac{\hat{b}_0}{\rho^n + \hat{a}_{n-1}\rho^{n-1} + \ldots + \hat{a}_1\rho + \hat{a}_0} u(k) \tag{21}$$

with the differentiation operator $\rho = d/dt$ (being the time counterpart of the frequency Laplace variable 's') implemented, for instance, using the well-known bilinear (Tustin's) formula: $\rho = [2(1-q^{-1})]/[T(1+q^{-1})]$. Note that this indirect transformation [12] maps the left-half s-plane onto the unit disc in the z-plane, so the resulting discrete-time filter (21) remains stable.

Note that in the case of identifying the non-linear model (9), processing based on the filter (21) used to generate the instrumental variable cannot be directly applied. Thus, we propose an instrumental variable based on the use of the Taylor series scheme to numerically approximate the non-linear system [27]. This novel solution can be used to evaluate the noise-free output of the dynamic system as

$$\begin{aligned} \hat{y}^{(n)}(k-1) &= -\hat{a}_{n-1}\hat{y}^{(n-1)}(k-1) + \ldots - \hat{a}_1\hat{y}^{(1)}(k-1) \\ &\qquad - \hat{c}_1\hat{y}(k-1) + \ldots - \hat{c}_r\hat{y}^r(k-1) + \hat{b}_0 u(k-1), \\ \hat{y}^{(n-1)}(k) &= \hat{y}^{(n-1)}(k-1) + T\,\hat{y}^{(n)}(k-1), \\ \hat{y}^{(n-2)}(k) &= \hat{y}^{(n-2)}(k-1) + T\,\hat{y}^{(n-1)}(k-1) + \frac{T^2}{2!}\hat{y}^{(n)}(k-1), \\ &\ \ \vdots \\ \hat{y}(k) &= \hat{y}(k-1) + T\,\hat{y}^{(1)}(k-1) + \ldots + \frac{T^n}{n!}\hat{y}^{(n)}(k-1), \end{aligned} \tag{22}$$

where the estimates of the parameters a_i, b_i and c_i of the model (11)–(13) are used in the above processing.

Now, the instrumental variable $\xi(k)$ can be determined in a similar fashion as $\phi(k)$, provided the noise-free output, (21) or (22), replaces $y(k)$ contaminated with noise:

$$\xi(k) = \left[-I^n_{n-1}\hat{y} \quad \ldots \quad -I^n_0\hat{y} \quad I^n_0 u\right]^{\mathrm{T}}, \tag{23}$$

$$\xi(k) = \left[-I^n_{n-1}\hat{y} \quad \ldots \quad -I^n_1\hat{y} \quad -I^n_0\hat{y} \quad \ldots \quad -I^n_0\hat{y}^r \quad I^n_0 u\right]^{\mathrm{T}} \tag{24}$$

for linear (6)–(8) and non-linear (11)–(13) systems, respectively. Of fundamental importance is that the processing (22) is completely deterministic, and therefore the estimate of $y(k)$ will always be noise-free (neglecting the effects of quantization). It is thus obvious that the instrumental variable obtained from (24) is fundamentally (completely) uncorrelated with the residual (noise) process $e(k)$ occurring in the non-linear model (11).

The described IV procedure effectively solves the problem of consistent identification in the presence of correlated noise. Both LS and IV schemes resulting from the minimization of quadratic criteria are completely ineffective for reliable estimation of parameters in the case of large measurement errors. Therefore, to overcome the problem of sporadic outliers, we consider an algorithm that penalizes/weights errors less drastically, namely based on their absolute values rather than square values.

3.3. Least absolute values method

The weighted least absolute value (LA) algorithm quoted here is based on minimizing the following (non-differentiable) loss index [6, 28]

$$V_{\mathrm{LA}} = \sum_{\ell=1}^{k} \lambda^{k-\ell} \, |e(\ell)| = \sum_{\ell=1}^{k} \lambda^{k-\ell} \, |\chi(\ell) - \phi^{\mathrm{T}}(\ell)\, \theta|, \tag{25}$$

where, as before, λ stands for the weighting factor.

Thus the estimation process becomes resistant to outliers that may appear in the processed measurement data. Assuming that an estimate of the residual error $e(k)$ is available (e.g. from a running LS algorithm), the criterion (25) can be converted to

$$V_{\mathrm{LA}} \approx \sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{e^2(\ell)}{|\hat{e}(\ell)|} = \sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{[\chi(\ell) - \phi^{\mathrm{T}}(\ell)\, \theta]^2}{|\hat{e}(\ell)|}\,. \tag{26}$$

Now, with the residual error estimate known, nulling the gradient of (26) leads to the following estimator

$$\hat{\theta}(k) = \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{\phi(\ell)\, \phi^{\mathrm{T}}(\ell)}{|\hat{e}(\ell)|} \right]^{-1} \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{\phi(\ell)\, \chi(\ell)}{|\hat{e}(\ell)|} \right]. \tag{27}$$

This result can be further improved by using the idea of successive approximations. Since the current residual error $e(k)$ can be estimated simply from the recent estimate of the θ parameter vector, the complete iterative identification scheme ($p = 0, 1, \ldots$) takes the following form

$$\hat{e}^{/p/}(\ell) = \chi(\ell) - \phi^{\mathrm{T}}(k)\, \hat{\theta}^{/p/} \tag{28}$$

$$\hat{\theta}^{/p+1/} = \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{\phi(\ell)\, \phi^{\mathrm{T}}(\ell)}{|\hat{e}^{/p/}(\ell)|} \right]^{-1} \left[\sum_{\ell=1}^{k} \lambda^{k-\ell} \frac{\phi(\ell)\, \chi(\ell)}{|\hat{e}^{/p/}(\ell)|} \right], \tag{29}$$

where the simplest method is to use the LS estimate

$$\hat{\theta}^{/0/} = \hat{\theta}_{\mathrm{LS}} \tag{30}$$

to initiate the iteration loop (with $p = 0$).

The iterative implementation presented above ends when the improvement in minimization drops below the assumed threshold ($\Delta_{\min}$)

$$\left| V_{\mathrm{LA}}(\hat{\theta}^{/p/}) - V_{\mathrm{LA}}(\hat{\theta}^{/p+1/}) \right| < \Delta_{\min}\,. \tag{31}$$

This condition is mathematically justified, because the key factor here is the decreasing sequence of successively calculated values of the base LA criterion [29].

Assuming that the above processing is limited to only one iteration and using the already mentioned matrix inversion lemma, the recursive version of the iteration LA scheme (including prediction error estimation, covariance matrix update, and final estimation vector correction) can be roughly approximated as [5]

$$\varepsilon(k) = \chi(k) - \phi^{\mathrm{T}}(k)\,\hat{\theta}(k-1), \tag{32}$$

$$P(k) = \frac{1}{\lambda}\left[P(k-1) - \frac{P(k-1)\,\psi(k)\,\phi^{\mathrm{T}}(k)\,P(k-1)}{\lambda|\varepsilon(k)| + \phi^{\mathrm{T}}(k)\,P(k-1)\,\psi(k)}\right], \tag{33}$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)\,\psi(k)\,\mathrm{sign}\,[\varepsilon(k)] \tag{34}$$

with $\psi(k) = \phi(k)$ and $\psi(k) = \xi(k)$ used for the LA estimation routines in the straight and IV versions, respectively, and the converted output $\chi(k)$ given by (6) or (11). We recommend starting the procedure (32)–(34) with LS (auxiliary) results (18)–(20) for covariance matrices and parameter estimates.

The proposed initiation is indeed crude, but due to the weighting mechanism ($\lambda < 1$) contained in the LA estimation algorithm, such forced start-up data will be gradually eliminated from the finite memory of the estimator. Note that the effective number of observations, also referred to as the memory length of the weighted estimator, equals $\Gamma = 1/(1-\lambda)$.

All three LS, IV and LA estimation algorithms discussed above will be verified in a numerical simulation study reported below.

4. Numerical simulations

Consider the mechanical system shown in Fig. 1. Such a structure usually constitutes the essential part of the wheeled vehicle suspension system.

Given that $u(t)$ represents the external force, the reaction of the damper is proportional to the velocity (with a coefficient B), and the classical spring (K) follows the linear Hooke's rule, the resulting continuous-time model can be described as follows

$$m\,\ddot{y}(t) + B\,\dot{y}(t) + K\,y(t) = u(t). \tag{35}$$

Modern manufacturers of cars and wheeled vehicles prefer to use the so-called progressive springs as the basis of automotive suspension mechanisms, which can be shown as a specific non-linear model as follows

$$m\,\ddot{y}(t) + B\,\dot{y}(t) + K\,y^3(t) = u(t). \tag{36}$$

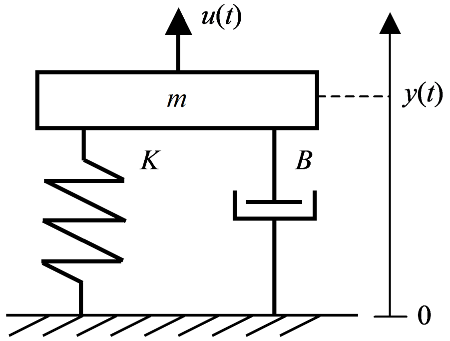

Figure 1: Suspension: a non-linear dynamical system

In order to use the idea of realizable discrete-time integrating filters, the corresponding counterpart of the non-linear model (36) should be presented in the necessary monic form (9) as follows

$$\ddot{y}(t) + a_1\dot{y}(t) + c_3 y^3(t) = b_0 u(t), \tag{37}$$

where $a_1 = B/m$, $c_3 = K/m$, and $b_0 = 1/m$.

It should be emphasized that the considered non-linear description (37) clearly differs from the known (classical) representations of non-linear systems, such as the Wiener and Hammerstein models. This is because these models are based on a simple series structure with cascaded linear-dynamic and non-linear-static blocks. In contrast, a non-linear block is used in (37) to generate feedback, so it directly influences the dynamics of the modeled system (see block diagram in Fig. 2 based on a canonical regulatory structure or a general approach to simulating

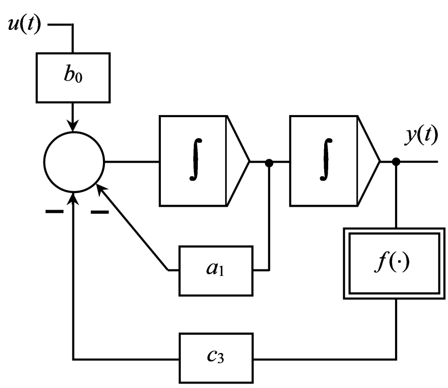

Figure 2: NLID: non-linear-in-dynamics model of the suspension with the non-linear function $f(y) = y^3$ in the loop

continuous-time systems). Such a system with non-linearity in dynamics (as opposed to Hammerstein and Wiener models with an isolated linear dynamic part) will be referred to as a *non-linear-in-dynamics* (NLID) system. Note that shifting the output beyond the non-linear element creates a new system that could be called *semi-Wienerian*.

The appropriate regression model of the non-linear dynamics of the mechanical suspension can therefore be presented as (11)–(13), in which $n = 2$ and $r = 3$

$$I_2^2 y(k) = \chi(k) = \phi^{\mathrm{T}}(k)\,\theta + e(k), \tag{38}$$

$$\phi(k) = \left[-I_1^2 y \quad -I_0^2 y^3 \quad I_0^2 u\right]^{\mathrm{T}}, \tag{39}$$

$$\theta = [a_1 \;\; c_3 \;\; b_0]^{\mathrm{T}}. \tag{40}$$

The instrumental variable procedure which makes our identification process insensitive to correlated noise requires an appropriate instrument (24)

$$\xi(k) = \left[-I_1^2 \hat{y} \quad -I_0^2 \hat{y}^3 \quad I_0^2 u\right]^{\mathrm{T}}, \tag{41}$$

where the noise-free output for the dynamical system (37) of the order $n = 2$, has the form of (22)

$$\begin{aligned} \hat{y}^{(2)}(k-1) &= -\hat{a}_1 \hat{y}^{(1)}(k-1) - \hat{c}_1 \hat{y}^3(k-1) + \hat{b}_0 u(k-1), \\ \hat{y}^{(1)}(k) &= \hat{y}^{(1)}(k-1) + T\,\hat{y}^{(2)}(k-1), \\ \hat{y}(k) &= \hat{y}(k-1) + T\,\hat{y}^{(1)}(k-1) + \frac{T^2}{2!}\hat{y}^{(2)}(k-1). \end{aligned} \tag{42}$$

A series of computer simulations of the identification procedures were carried out for the non-linear process under consideration using a convenient persistently exciting (periodic) input signal

$$u(t) = \sum_{i=1}^{N} \sin \omega_i t, \tag{43}$$

where $N = 5$, and the angular frequencies [rad/s] were assumed as: $\omega_1 = 0.7$, $\omega_2 = 1.1$, $\omega_3 = 1.4$, $\omega_4 = 1.7$, $\omega_5 = 2.1$.

In the first numerical experiment, the non-linear model was identified in the presence of additive correlated noise. The coefficients were assumed as $a_1 = 3.5$, $c_3 = 5$, and $b_0 = 4$, while the residual error $e(k)$ was implemented in the form of a zero-mean correlated process such that the eventual noise-to-signal ratio of the respective standard deviations was $N/S = 10\%$. The discrete-time counterpart of the identified model was obtained using finite-horizon integrating filters with

$L = 50$. The sampling time was set to $T = 0.02$ s, and for the simulated stationary dynamics, the forgetting mechanism was turned off ($\lambda = 1$). Simulation time of the exercised LS and IV recursive estimation schemes (18)–(20) was limited to 200 s (10000 samples).

From the estimation runs shown in Figs. 3 and 4, it is clear that the IV scheme, which uses the innovative procedure (22) to generate instrumental variables, radically improves the consistency of estimation. The LS method, in turn, shows an evident asymptotic bias.

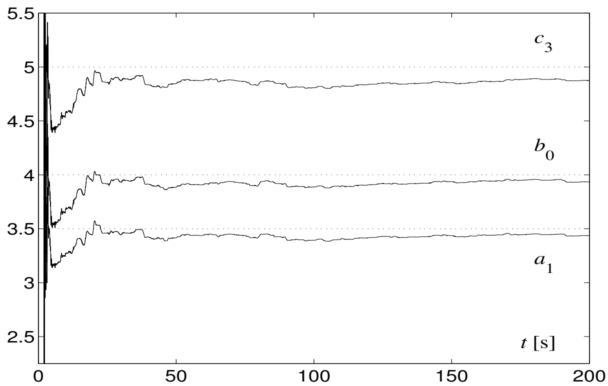

Figure 3: Parameter estimation of the non-linear and stationary suspension system using the LS method

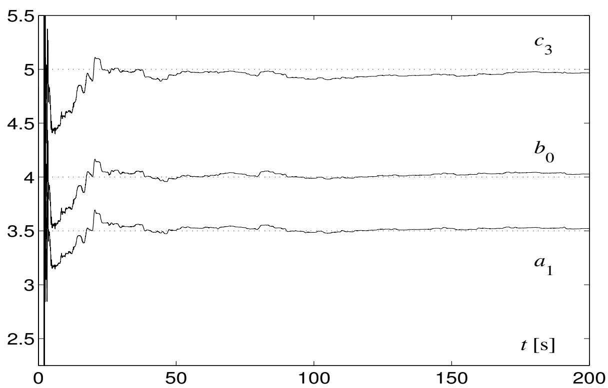

Figure 4: Parameter estimation of the non-linear and stationary suspension system using the IV method

The same model was then identified using the resistant-to-outliers LA procedure. This time, however, the simulated variable parameter a_1 was changing gradually (3.5 . . . 2.8) in a restricted time interval [50 s , 150 s], while the other parameters were kept constant ($c_3 = 5$, $b_0 = 4$). In this test the residual error $e(k)$ was represented by a normally-distributed zero-mean white noise sequence with variance selected so that the already mentioned noise-to-signal ratio was $N/S = 1\%$. Measurement faults were represented by missing data ($y(k) = 0$) introduced at $t \in [95\ \text{s},\ 96\ \text{s}]$. Again, the integration horizon was set at $L = 50$ and the sampling period was set to $T = 0.02$ s. Yet, the forgetting mechanism with $\lambda = 0.99$ was activated for tracking the time-variant parameter a_1. Simulation time of the recursive LS (18)–(20) and LA (32)–(34) schemes was limited to 200 s (10000 samples).

It is evident from Figs. 5 and 6, that the weighted LA procedure allows for reliable tracking of variable parameters regardless of occasional faults. On the other hand, the LS algorithm with pronounced sensitivity to such errors simply fails in this case.

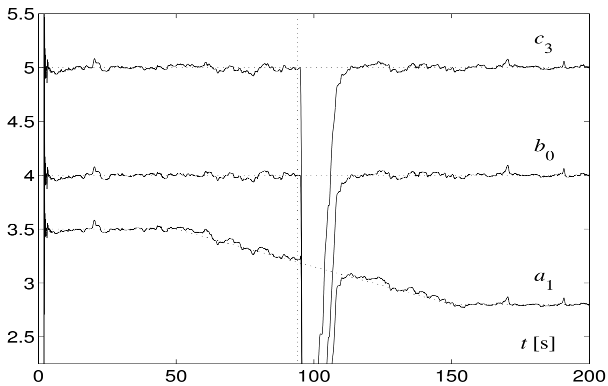

Figure 5: Tracking variable parameters of the non-linear and non-stationary suspension system using the LS method; vertical dashed line shows the appearance of measurement faults

Ultimately, the most challenging case was considered, in which the non-linear model was identified provided the two parasitic phenomena occur simultaneously. Namely, the zero-mean correlated residual error $e(k)$, such that $N/S = 10\%$, represented a considerable disturbance, while missing data ($y(k) = 0$) introduced at $t \in [95\ \text{s},\ 96\ \text{s}]$ simulated sporadic measurement faults. Once more, the integration horizon was set to $L = 50$, the sampling period was set to $T = 0.02$ s, and the forgetting mechanism $\lambda = 0.999$ was implemented in the exercised procedures.

It is apparent from Figs. 7 and 8 that the tested IV-LA, henceforth marked as IV_{abs}, the IV in the sense of absolute values: (32)–(34) with $\psi(k) = \xi(k)$, significantly suppresses the asymptotic bias of estimates, regardless of sporadic outliers corrupting the measurement data. Contrary to this, the classical IV_{sqr}, the IV in the square sense: (18)–(20) with $\psi(k) = \xi(k)$, inherits its drawback

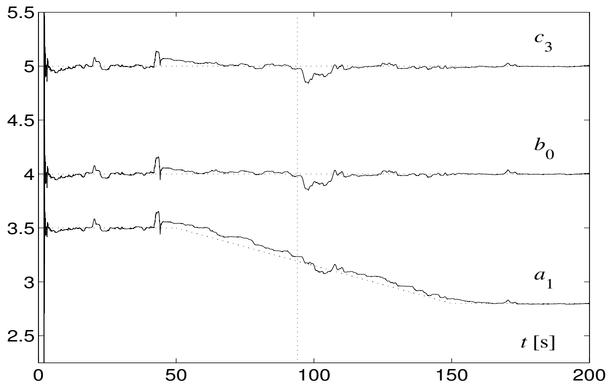

Figure 6: Tracking variable parameters of the non-linear and non-stationary suspension system using the LA method; vertical dashed line shows the appearance of measurement faults

(i.e. sensitivity to faults in data) typical for estimation algorithms derived from minimization of square indices.

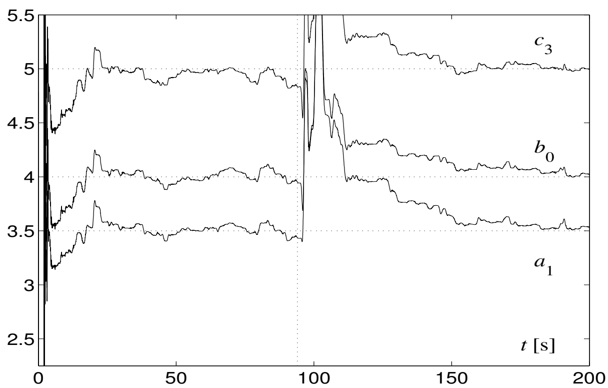

Figure 7: Parameter estimation of the non-linear and stationary suspension system using the IV_{sqr} method; vertical dashed line shows the appearance of measurement faults

Proper selection of the weighting factor λ is critical when implementing estimators. In general, the weighting factor (λ) used in the weighted estimator can be related to the so-called memory length of the algorithm (Section 3): $\Gamma = 1/(1 - \lambda)$. This means that the parameters of the identified model are evaluated based on the measurement data (i.e. the converted output $\chi(\ell)$ and the regression vector $\phi(\ell)$) obtained in the current time interval $[k - \Gamma, k]$. As a consequence, for λ close to unity, the weighted algorithm is very "conservative",

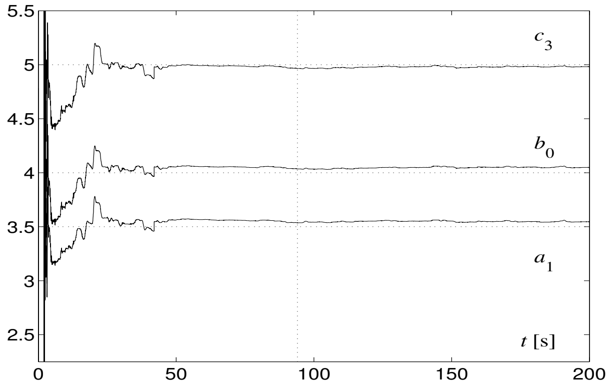

Figure 8: Parameter estimation of the non-linear and stationary suspension system using the IV_{abs} method; vertical dashed line shows the appearance of measurement faults

resulting in a delayed response to an actual parameter change. In contrast, with smaller values of λ the weighted algorithm reacts much faster to changes in the parameters of the identified process, but at the expense of the evidently increased variance of the estimation process.

More specifically, the choice of λ is mostly intuitive, based on unique prior knowledge of the underlying system dynamics and predictions of parameter evolution in the supervised system (e.g. sudden or gradual parameter changes). A known remedy for the problem of proper tuning of the λ coefficient is the idea of parallel estimation (e.g. the use of a "battery" of 3 competing estimators working in parallel with differently tuned weighting factors: $\lambda_1 < \lambda_2 < \lambda_3$). In such a solution [2], the on-line decision-making mechanism evaluates the performance quality indicator (such as the local mean square residual error $e(k)$ used in the regression model) and proposes the estimation results generated by the 'winning' procedure (i.e. the one with the smallest index).

5. Summary and further study

The article addresses the problem of parametric estimation in order to obtain identification procedures resistant to correlated noise and occasional outliers in the analysis of non-stationary and non-linear models in continuous time. The presented numerical simulations show that the IV scheme used successfully solves the problem of correlated noise and therefore can be recommended as a good tool for suppressing the asymptotic bias (systematic error) of the obtained estimates. In addition, an innovative method of generating instrumental variables adapted to non-linear dynamics was presented, which proved its effectiveness and practical usefulness in experiments. The non-trivial problem of parametric identification resistant to measurement errors is solved by the applied algorithm that takes into account penalization according to the sum of absolute values. On the other hand, the weighting mechanism makes it possible to track the time-varying parameters of non-stationary systems.

Our original contribution to the field of identifying continuous-time models can be summarized as follows: The most important achievements include the procedure for determining the instrumental variable for a continuous model with non-linearity (NLID) and thus removing the evaluation bias (systematic error) for such semi-Wienerian systems, and the suboptimal estimation algorithm in the sense of the minimum sum of absolute deviations (LA) in the recursive version, which actually implements minimization of the objective function (absolute values) only approximately, but it can be performed on-line and maintains the system's robustness to sporadic outliers.

The fundamental direction of this research was initiated and described in our conference article [24]. This study has been consistently and further developed in this article. In particular, the difficult problem of such a synergistic implementation of an estimation procedure that is both insensitive to occasional outliers in the data and consistent (asymptotically convergent) in the presence of correlated noise appears to have been innovatively and successfully solved here as well as illustrated by new working examples.

Let us remember that the LA algorithm itself (applied even in the simple version) brings the feature of insensitivity to outliers (deviations), and in the form enriched with an instrumental variable (IV-LA or IV_{abs}) additionally removes the evaluation error and is suitable for the task of identifying processes with non-linear dynamics (semi-Wienerian/NLID). It is worth noting that the developed IV_{abs} method with the applied weighting mechanism (λ) is also suitable for tracking variable parameters of non-stationary systems (both linear and semi-Wienerian/NLID).

Finally, the proposed method of instrumental variable in the sense of LA (marked as IV_{abs}) was also verified in the numerical study of the task of identifying a practical non-linear system (vehicle suspension with "progressive" springs).

Since the presented approach takes into account only one non-linear static element $f(y)$, which is an intercept (free term) in the differential equation, an interesting direction for further research seems to be the search for an effective discrete approximation model for other (more complex) types of continuous-time systems (e.g. with non-linear functions corrupting derivatives).

In view of the above, further research in this area may focus on the following issues.

1. *Identification of non-linear MIMO systems*: Given that any system with multiple inputs and multiple outputs can simply be written as a set of ordinary

differential equations, the appropriate SISO to MIMO generalization need not actually be problematic. In the literature, one can find dedicated practical approaches to modeling [30] and identification [31] of linear MIMO systems.

2. *Identification of non-linear models with input delay*: The simultaneous identification of both system parameters and input delay is definitely a non-trivial matter. For a delay system represented by a linear differential equation, efficient estimation schemes are available [32].

3. *Identification of non-linear distributed-parameter objects*: Since the dynamics of a system of distributed parameters is by definition represented by a partial differential equation, the necessary discrete-time approximation seems to be quite a challenge. Interesting results in this regard, including linear systems with distributed parameters, were presented by, for example, Sagara [33].

It is particularly important that the developed methods of identifying non-linear systems are resistant to (destructive) measurement errors, thanks to which they can be directly used in reliable diagnostics of wheeled vehicle components. Some useful techniques and results in the diagnosis of vehicle dynamics, taking into account the well-known quarter-car model, can be found in the work [16]. In the available literature [15] you can also find many other interesting and practical studies in the wide field of identification of non-linear systems.

Finally, it is worth noting that there are specific modern methods (based on AI, including neural networks) that can be effectively used to identify [34–36] the parameters of discrete- or continuous-time dynamical systems.

The authors declare that there is no competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

References

[1] S. Sagara and Z. Zhao: Numerical integration approach to on-line identification of continuous-time systems. *Automatica*, **26**(1), (1990), 63–74. DOI: 10.1016/0005-1098(90)90158-E.

[2] Z. Kowalczuk: Competitive identification for self-tuning control: robust estimation design and simulation experiments. *Automatica*, **28**(1), (1992), 193–201. DOI: 10.1016/0005-1098(92)90021-7.

[3] Z. Kowalczuk and J. Kozłowski: Continuous-time approaches to identification of continuous-time systems. *Automatica*, **36**(8), (2000), 1229–1236. DOI: 10.1016/S0005-1098(00)00033-9.

[4] S. Sagara, Z. Yang and K. Wada: Identification of continuous systems using digital low-pass filters. *International Journal of Systems Science*, **22**(7), (1991), 1159–1176. DOI: 10.1080/00207729108910693.

[5] Z. Kowalczuk and J. Kozłowski: Non-quadratic quality criteria in parameter estimation of continuous-time models. *IET Control Theory and Applications*, **5**(13), (2011), 1494–1508. DOI: 10.1049/iet-cta.2010.0310.

[6] E. Schlossmacher: An iterative technique for absolute deviations curve fitting. *Journal of the American Statistical Association*, **68**(344), (1973), 857–865. DOI: 10.1080/01621459.1973.10481436.

[7] P. Young: Parameter estimation for continuous-time models – a survey. *Automatica*, **17**(1), (1981), 23–39. DOI: 10.1016/0005-1098(81)90082-0.

[8] H. Unbehauen and G. Rao: *Identification of Continuous Systems* North Holland, Amsterdam, Netherlands, 1987.

[9] R. Middleton and G. Goodwin: *Digital Control and Estimation. A Unified Approach*. Prentice-Hall, Upper Saddle River, NJ, USA, 1990.

[10] J. Schoukens: Modeling of continuous time systems using a discrete time representation. *Automatica*, **26**(3), (1990), 579–583. DOI: 10.1016/0005-1098(90)90029-H.

[11] H. Unbehauen and G. Rao: Continuous-time approaches to system identification – a survey. *Automatica*, **26**(1), (1990), 23–35. DOI: 10.1016/0005-1098(90)90155-B.

[12] Z. Kowalczuk: Discrete approximation of continuous-time systems – a survey. *IEE Proceedings G (Circuits, Devices and Systems)*, **140**(4), (1993), 264–278. DOI: 10.1049/ip-g-2.1993.0045.

[13] R. Johansson: Identification of continuous-time models. *IEEE Transactions on Signal Processing*, **42**(4), (1994), 887–897. DOI: 10.1109/ 78.285652.

[14] Z. Kowalczuk: Discrete-time realization of on-line continuous-time estimation algorithms. *IASTED Journal on Control and Computers*, **23**(2), (1995), 33–37.

[15] J. Schoukens and L. Ljung: Nonlinear system identification: a user-oriented road map. *IEEE Control Systems Magazine*, **39**(6), (2019), 28–99. DOI: 10.1109/MCS.2019.2938121.

[16] J. Kozłowski and Z. Kowalczuk: Intelligent monitoring the vertical dynamics of wheeled inspection vehicles. *IFAC-PapersOnLine*, **52**(8), (2019), 251–256. DOI: 10.1016/j.ifacol.2019.08.079.

[17] P. Suchomski and Z. Kowalczuk: Analytical design of stable delta-domain generalized predictive control. *Optimal Control Applications and Methods*, **23**(5), (2002), 239–273. DOI: 10.1002/oca.712.

[18] T. Söderström, H. Fan, B. Carlsson and S. Bigi: Least squares parameter estimation of continuous-time arx models from discrete-time data. *IEEE Transactions on Automatic Control*, **42**(5), (1997), 659–673. DOI: 10.1109/9.580871.

[19] Y. Chao, C. Chen and H. Huang: Recursive parameter estimation of transfer function matrix models via simpson's integrating rules. *International Journal of Systems Science*, **18**(5), (1987), 901–911. DOI: 10.1080/00207728708964017.

[20] K. Inoue, K. Kumamaru, Y. Nakahashi, H. Nakamura and M. Uchida: A quick identification method of continuous-time nonlinear systems and its application to power plant control. *Proceedings of the 10th IFAC Symposium on System Identification*, **1** Copenhagen, Denmark, (1994), 319–324. DOI: 10.1016/S1474-6670(17)47729-9.

[21] W. Byrski, M. Drapała and J. Byrski: An adaptive identification method based on the modulating functions technique and exact state observers for modeling and simulation of a nonlinear MISO glass melting process. *International Journal of Applied Mathematics and Computer Science*, **29**(4), (2019), 739–757. DOI: 10.2478/amcs-2019-0055.

[22] W. Byrski and M. Drapała: On-line process identification using the Modulating Functions Method and non-asymptotic state estimation. *Archives of Control Sciences*, **32**(3), (2022), 535–555. DOI: 10.24425/acs.2022.142845.

[23] M. Drapała and W. Byrski: Online continuous-time adaptive predictive control of the technological glass conditioning process. *Archives of Control Sciences*, **32**(4), (2022), 755–782. DOI: 10.24425/acs.2022.143670.

[24] J. Kozłowski and Z. Kowalczuk: Resistant to correlated noise and outliers discrete identification of continuous non-linear non-stationary dynamics objects. *Intelligent and Safe Computer Systems in Control and Diagnostics*, **545** of LNNS: Lecture Notes in Networks and Systems, Springer Nature AG, Cham, Switzerland, 2023, 317–327. DOI: 10.1007/978-3-031-16159-9 26.

[25] L. Ljung: *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, NJ, USA, 1987.

[26] T. Söderström and P. Stoica: Comparison of some instrumental variable methods – consistency and accuracy aspects. *Automatica*, **17**(1), (1981), 101–115. DOI: 10.1016/0005-1098(81)90087-X.

[27] J. Craig: *Introduction to Robotics: Mechanics and Control*. Pearson Education, Cranbury, NJ, USA, 2014.

[28] K. Janiszowski: Towards estimation in the sense of the least sum of absolute errors. *IFAC Proceedings Volumes*, **31**(20), (1998), 605–610. DOI: 10.1016/S1474-6670(17)41862-3.

[29] J. Kozłowski and Z. Kowalczuk: Robust to measurement faults, parameter estimation algorithms in problems of systems diagnostics. *Intelligent Extraction of Information for Diagnostic Purposes*, Pomorskie Wydawnictwo Naukowo-Techniczne, Gdańsk, Poland, (2007), 221–240.

[30] Z. Kowalczuk: On discretization of continuous-time state-space models: A stable normal approach. *IEEE Transactions, Circuits and Systems*, **38**(1), (1991), 1460–1477. DOI: 10.1109/31.108500.

[31] S. Sagara and Z. Zhao: Recursive identification of transfer function matrix in continuous systems via linear integral filter. *International Journal of Control*, **50**(2), (1989), 457–477. DOI: 10.1080/00207178908953377.

[32] Z. Zhao and S. Sagara: Consistent estimation of time delay in continuous-time systems. *Transactions of the Society of Instrument and Control Engineers*, **27**(1), (1991), 64–69. DOI: 10.9746/sicetr1965.27.64.

[33] S. Sagara and Z. Zhao: Identification of system parameters in distributed parameter systems. *Proceedings of the 11th IFAC World Congress*, IFAC, Tallinn, Estonia, (1990), 471–476. DOI: 10.1016/S1474-6670(17)51960-6.

[34] D. Goldberg: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, USA, 1989.

[35] M. Willis, G. Montague, C. Di Massimo, M. Tham and A. Morris: Artificial neural networks in process estimation and control. *Automatica*, **28**(6), (1992), 1181–1188. DOI: 10.1016/0005-1098(92)90059-O.

[36] D. Uciński and M. Patan: Sensor network design for the estimation of spatially distributed processes. *Internarional Journal of Applied Mathematics and Computer Science*, **20**(3), (2010), 459–481. DOI: 10.2478/v10006-010-0034-2.