

## A COMPARATIVE AND EXPERIMENTAL STUDY ON GRADIENT AND GENETIC OPTIMIZATION ALGORITHMS FOR PARAMETER IDENTIFICATION OF LINEAR MIMO MODELS OF A DRILLING VESSEL

STANISŁAW BAŃKA<sup>a</sup>, MICHAŁ BRASEL<sup>a</sup>, PAWEŁ DWORAK<sup>a,\*</sup>, KRZYSZTOF JAROSZEWSKI<sup>a</sup>

<sup>a</sup>Faculty of Electrical Engineering  
West Pomeranian University of Technology, 26 Kwietnia 10, 71-126 Szczecin, Poland  
e-mail: {michal.brasel, pawel.dworak, krzysztof.jaroszewski}@zut.edu.pl

The paper presents algorithms for parameter identification of linear vessel models being in force for the current operating point of a ship. Advantages and disadvantages of gradient and genetic algorithms in identifying the model parameters are discussed. The study is supported by presentation of identification results for a nonlinear model of a drilling vessel.

**Keywords:** MIMO dynamic plant, identification, nonlinear system.

### 1. Introduction

Parameter identification of dynamic models of controlled plants has been the subject of studies since the very beginning of automatic control engineering. Although a large variety of identification methods have been developed over a period of decades, new, more accurate and reliable methods are still sought for (Nelles, 2001; Ljung, 2001; Al Seyab and Cao, 2008; Billings, 2013; Orjuela *et al.*, 2013; Mzyk, 2013). The problem is of particular importance for parameter identification of vessels. The synthesis of a ship's control systems often requires dynamic plant models that are both multi-input multi-output (MIMO), nonlinear and time-varying for ships operating under various environmental conditions (Fossen, 2011).

In accomplishing advanced control systems for sea vessels, which represent usually nonlinear and MIMO control plants, a fundamental and difficult problem to resolve is identification of varying ship parameter values in linear (linearized) models for a multi-degree of freedom (DOF) motion. Such values are needed for the synthesis and realization of control algorithms for varying environment circumstances. Parameter variations of linear models (or the lack of precise knowledge of

their “true” values) in nonlinear plant description can be induced by changes in the direction and velocity of the sea current, as well as the direction and state of sea waves. Usually, changes of parameters in linear models are frequently caused by changing in the ship's operating points when the vessel course (yaw angle) and/or position over the drilling point on the sea bottom (or the ship velocity along the demanded trajectory) must be changed, as well as due to the changes in the mass distribution (moment of inertia), the vessel displacement and dipping during its operation.

Identification of parameters can be done on the basis of archived data of vessel motions measured during a passive or active experiment. Active experiments, when linear ship models are used, should be performed after each important change in the operating point of the controlled plant. Usually, it may be performed on the basis of recorded measurements of the position and course of an appropriately equipped vessel (or from simulation of its known nonlinear model). One of the experimental trials whose results are used to identify the parameters of the ship model is the Kempf zigzag test.

The Kempf test usually is carried out for ships equipped with a helm and a steering gear where the primary objective of control is course keeping. For identification purposes, the so-called Nomoto model

\*Corresponding author

(Fossen, 1994) is then considered for two degrees of freedom of the ship's motion (sway and yaw), which after its linearization and adoption of certain simplifications is usually given in the form of transfer function models relating the yaw rate to the rudder angle and sway velocity to the rudder angle (Aström and Källström, 1976; 1981; Velasco *et al.*, 2013; Padilla and Yuz, 2013). Frequently, identification of ship models boils down to determining the hydrodynamic coefficients or hydrodynamic derivatives occurring in differential equations of the ship's model (Yoon and Rhee, 2003; Herrero and González, 2012; Zhang and Zou, 2011; Tran Khanh *et al.*, 2013; Skjetne *et al.*, 2004). It happens that identification is confined only to a selected degree of freedom of the ship's motion (Yoon *et al.*, 2004; 2007; Zhang *et al.*, 2010; Revestido *et al.*, 2011). Also, parameter identification of linear ship models described in state space may be found in 3DOF: sway roll yaw (Dai *et al.*, 2012), and in 2DOF: sway, yaw (Wang *et al.*, 2011).

In the case under study, the Kempf test was carried out for a multivariable model of a drilling ship provided with a main engine and four azimuth thrusters. There was identified comprehensively (3DOF: surge, sway, yaw) a discrete linearized MIMO model of such a ship. The multivariable ship model to be identified is described in state space, and its parameters after identification may be employed directly for synthesis and adaptation of an MIMO controller.

Because of the multivariable character of the plant model to be identified and in view of the occurrence of non-stationary stochastic disturbances affecting the nonlinear plant, it is necessary to use very advanced identification algorithms with proper simplifications introduced to allow for the nature of stochastic disturbances, which are usually non-stationary and are not white noise in sea conditions. The disturbances originating from the sea environment generally do not satisfy such assumptions. This is particularly true for disturbances induced by waving and wind gusts. However, it is possible to model them (in fully developed wind waving) using additional dynamic models stimulated by Gaussian white noise type signals. Unfortunately, in real conditions, the parameters of such additional "whitening" models should also be identified. Also the obvious correlation between wave and wind, as we can see in the works of Barth and Eecen (2006), Bredmose *et al.* (2012), Anil Ari Guner *et al.* (2013), or Kondo *et al.* (1972), seems to be region specific, and the stochastic models of such relations have to be developed for site-specific implementations. That is mainly why estimation algorithms known from the literature are commonly developed under the assumption that stochastic disturbances affecting the controlled plant are stationary and independent (white) (Aström and Källström, 1976; Dai *et al.*, 2012; Yoon *et al.*, 2007; Zhang *et al.*, 2010;

Revestido *et al.*, 2011).

All these problems increase the level of complication of the ship model and practically exclude the possibility of adopting identification algorithms performed in an on-line mode. However, passive experiments performed with a feedback loop created by control systems do not assure obtaining unbiased results of parameter estimation. Hence, acquisition of truly useful measurements, e.g., for implementation of adaptive control systems, is possible only by conducting active experiments under conditions being as close as possible to those real out at sea.

All these problems render watercrafts very complex controlled plants in terms of their identification. Many classical identification algorithms give unsatisfactory results and suffer from numeric sensitivity and long computational times. That is way many authors try, e.g., to combine the classical identification methods with neural networks (Mahfouz and Haddara, 2003) or ant (Dai *et al.*, 2010) and swarm intelligence algorithms (Chen *et al.*, 2008). In this paper, a nonlinear model of a drilling vessel is used as an example to present an algorithm to identify the parameters of linear models of watercrafts being in force for their current operating point. Also, the advantages and disadvantages of the gradient and genetic versions of the algorithm to identify the model parameters are discussed.

The organization of this paper is as follows. In Section 2, a nonlinear ship model is presented. Then, in Section 3 a description of the popular active maneuvering tests useful for parameter identification is presented. In Section 4, a gradient based identification algorithm is derived. An example of identification parameter of a linear discrete-time model of a drilling vessel is presented in Section 5. We end the paper in Section 6 with conclusions.

## 2. Ship's model

To synthesize a ship's control system, models of vessel motions in a horizontal plane with three degrees of freedom are most commonly used (Fossen, 1994; Wise and English, 1975), and therefore this type of models (after linearization) will be considered in this paper.

All of these models describing the motions of a rigid body in 6-, 4-, 3- and 1DOF have the form of ordinary differential equations. They can be derived from the Newton–Euler and Lagrange equations, or by using the principles of the conservation of the momentum in the linear motion and the conservation of the angular momentum in the rotational motion (Lewis, 1989; Fossen, 1994; 2011).

These equations describe most commonly low frequency (LF) components of motion and velocity caused by the impact of average marine disturbances (wind, wave and current)  $d(t)$  and the ship's propulsion  $u(t)$ .

In the structure of these equations, we can find a part describing the kinematics of the ship's motion in the adopted reference systems and a dynamic part, which takes into account the whole possible impact of hydrodynamic (and hydrostatic) forces on the hull of the ship (along with the added mass of the accompanying water) and operation of the main engine, the rudder and the propellers and tunnel thrusters (if any).

Motion components induced by high frequency (HF) components of marine disturbances stemming from the gust of wind and sea waves are usually modeled by using separate models applicable only for HF motions (usually linear ones). On the other hand, it is a common practice to take into account HF effects produced by the gust of wind and sea waves in the models describing the ship's LF motions by regarding them as additional (non-measured) stochastic disturbances  $\mathbf{w}(t)$ .

The mathematical description of the LF motion of the ship is most commonly given in the form of nonlinear state space and output equations,

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), V_c, \psi_c, V_w, \psi_w, \mathbf{w}(t)), \\ \mathbf{y}(t) &= \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{v}(t), \end{aligned} \quad (1)$$

where  $\mathbf{x}(t) \in \mathbb{R}^n$  is the state vector describing changes in the ship's position and velocity in 3DOF,  $\mathbf{y}(t)$  is the output vector,  $\mathbf{u}(t)$  is the vector of manipulated variables (force components  $F_x(t)$ ,  $F_y(t)$  and torque  $M_z(t)$  produced by the main engine and azimuth thrusters), and  $V_c$ ,  $\Psi_c$ ,  $V_w$ ,  $\Psi_w$  are velocities and directions of the sea current and wind, respectively, while  $\mathbf{w}(t) \in \mathbb{R}^p$  represents the HF part of disturbances generated by the sea wave and the gusts of the wind. The disturbances are usually regarded as measurement errors that burden the output  $\mathbf{y}(t)$ .

The last signals  $\mathbf{w}(t)$  and  $\mathbf{v}(t)$  have a stochastic nature. In the paper, we assume that these disturbances are white Gaussian noise with zero mean values and (known) covariance matrices  $\mathbf{W} \geq 0$  and  $\mathbf{V} > 0$ . This assumption is not necessarily true, especially when it comes to the input disturbances  $\mathbf{w}(t)$ .

In the structure of the state equations in (1), generally two parts can be seen. The first part describes the kinematics and the second one the dynamic properties of the ship's model (see Eqn. (37) in the example; Section 5). But the final form of (1) depends on the number of degrees of freedom (DOF) and the number of manipulated variables, as well as on which components of the state vector  $\mathbf{x}(t)$  are associated with the controlled output signals  $\mathbf{y}(t)$ .

The needed nominal values for the state vector  $\mathbf{x}_o$  and components of forces and torque  $\mathbf{u}_o$  that allow the ship to be kept over the drilling point at known (constant) values  $V_c$ ,  $\Psi_c$ ,  $V_w$  and  $\Psi_w$  can be determined from a

system of nonlinear algebraic equations,

$$\begin{aligned} \mathbf{0} &= \mathbf{f}(\mathbf{x}_o, \mathbf{u}_o, V_c, \Psi_c, V_w, \Psi_w) \\ \mathbf{y}_o &= \mathbf{g}(\mathbf{x}_o, \mathbf{u}_o), \end{aligned} \quad (2)$$

which is valid for the steady state.

The knowledge of nominal values  $\mathbf{u}_o$  and  $\mathbf{x}_o$  at the values of  $V_c$ ,  $\Psi_c$ ,  $V_w$ ,  $\Psi_w$  corresponding to them enables linearization of equations (1) to be made. Then the state equations for the linear model, valid for limited deviations of all signals from their nominal values, assume the form

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \bar{\mathbf{A}}\tilde{\mathbf{x}}(t) + \bar{\mathbf{B}}\tilde{\mathbf{u}}(t) + \bar{\mathbf{G}}_{V_c}\tilde{v}_c(t) + \bar{\mathbf{G}}_{\psi_c}\tilde{\psi}_c(t) \\ &+ \bar{\mathbf{G}}_{V_w}\tilde{v}_w(t) + \bar{\mathbf{G}}_{\psi_w}\tilde{\psi}_w(t) + \bar{\mathbf{G}}_w\mathbf{w}(t), \\ \tilde{\mathbf{y}}(t) &= \bar{\mathbf{C}}\tilde{\mathbf{x}}(t) + \bar{\mathbf{D}}\tilde{\mathbf{u}}(t), \end{aligned} \quad (3)$$

where  $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \mathbf{x}_o$ ,  $\tilde{\mathbf{u}}(t) = \mathbf{u}(t) - \mathbf{u}_o$ ,  $\tilde{v}_c(t) = v_c(t) - V_c$ ,  $\tilde{\psi}_c(t) = \psi_c(t) - \Psi_c$ ,  $\tilde{v}_w(t) = v_w(t) - V_w$  and  $\tilde{\psi}_w(t) = \psi_w(t) - \Psi_w$ .

In the case of ships, the nature of operating points defined by (2) is unstable or is on the border of stability. Furthermore, forcing angular motions about the vertical axis of the vessel, which moves in relation to water as a result of the occurrence of a sea current, usually also causes the loss of the ship's position.

To prevent this, it is necessary to provide additionally acting manipulated variables  $\tilde{\mathbf{u}}(t)$  (in the form of deliberately introduced changes in longitudinal and transverse components of forces developed by the main engine and thrusters) added to the previously established nominal values  $\mathbf{u}_o$ , which will permit maintaining the conditions applicable to linear models. This can be accomplished in a closed-loop control system or by using a specially designed manual control system.

### 3. Circular and Kempf maneuvering tests

In the case of conventional ships (e.g., commercial ships), their basic dynamic properties can be described by using simple SISO (or MISO) models with a small number of degrees of freedom of the ship's motion (1DOF, 2DOF). Here, maneuvering trials called *circulation tests* may be applied for this purpose. The tests are carried out on the ship under certain conditions assumed for the wind and waves. Usually the tests are performed for different states of loading (ballasting) and different initial translational velocities of the vessel with appropriately selected rudder deflection. After making one-step change in the rudder position, it remains unchanged during each circulation test. Measurements of the transverse velocity and the resulting curved trajectory are most commonly used to evaluate the characteristics of the maneuvering properties of the vessel, such as stability and maneuverability rates of the ship (Fossen, 2011), less frequently for evaluation of parameter values of their

linear models. This is due to the fact that the circulation maneuvers are conducted under conditions not satisfying the assumptions of linearity of the model corresponding to the desired operating conditions (operating points) of the vessel. To detect possible asymmetry characteristics of maneuvering the trials are performed for rudder deflections on both the right and on the left side.

In the case of parameter identification of linear models useful for implementation of precise control of the motion and/or yaw angle and position of highly specialized naval units (e.g., drilling vessels) with more degrees of motion freedom (3DOF or more), equipped with additional propellers (bow thrusters) and the corresponding systems for measurement of position and course (yaw angle) components, we should use tests called *Kempf zigzag tests*. The tests are performed as follows: the vessel that moves initially with the assumed constant translational velocity and the assumed course (yaw angle), according to the defined operating points of the vessel in a specified regime of its operation (positioning or moving along a trajectory) at the specific configuration of the marine environment, makes periodic rudder deflections to the left and right sides to enforce snake-like moving of the vessel with a possible constant translational velocity. For conventional ships, the periodical deflections of the rudder to the right and left sides usually equal about  $\pm 15^\circ$  or  $\pm 20^\circ$ . However, for vessels equipped with additional propellers (or bow thrusters), extortion of the course is done by changing the moment of turning the ship around the vertical axis. Under the influence of changes in the torque, the ship changes its course when accompanied frequently by progressive decreases in the velocity components of the vessel with respect to water. If changes in the yaw angle will reach an adopted (preset) value, the turning moment (and/or rudder position) is reversed.

During Kempf tests made for more specialized vessels, e.g., drilling ships, the vessel should be maintained at a predetermined position over the drilling point at the sea bottom and perform temporary deviations from the adopted nominal value of the yaw angle only. This requires the designation of nominal values for longitudinal and transverse components of the forces and rotating torque for offsetting the impact of hydrodynamic forces coming from the sea current and the wind. Moreover, to prevent the departing of the ship from the drilling point, application of appropriate methods should also provide conditions of an active experiment that comply with those for which linearization of the model was done. The deviations of manipulated variables  $\tilde{\mathbf{u}}(t)$  and outputs  $\tilde{\mathbf{y}}(t)$  from their nominal values measured with a sampling period  $T_p$  during the tests created measurement sequences

$$\mathbf{U}^{N-1} = [\tilde{\mathbf{u}}_0, \tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_{N-1}]$$

and

$$\mathbf{Y}^N = [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_N],$$

used as input data to estimate unknown parameters of the ship's discrete model applicable for the ship's operation point.

#### 4. Derivation of a gradient-based identification algorithm

It is assumed that the controlled plant to be identified is modeled by a linear dynamic system of the  $n$ -th order with  $m$  inputs and  $l$  outputs given by discrete-time state space equations

$$\begin{aligned}\tilde{\mathbf{x}}(k+1) &= \mathbf{A}(\mathbf{a})\tilde{\mathbf{x}}(k) + \mathbf{B}(\mathbf{b})\tilde{\mathbf{u}}(k) + \mathbf{G}_w(\mathbf{g})\mathbf{w}(k), \\ \mathbf{y}(k) &= \mathbf{C}(\mathbf{c})\tilde{\mathbf{x}}(k) + \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}(k) + \mathbf{v}(k),\end{aligned}\quad (4)$$

where  $\mathbf{w}(k) \in \mathbb{R}^p$  and  $\mathbf{v}(k) \in \mathbb{R}^l$  are zero-mean mutually uncorrelated discrete white noise signals with (known) covariance matrices  $\mathbf{W}$  and  $\mathbf{V}$ , acting on the controlled plant. The noises may be described by normal probability densities  $\mathbf{f}_w(\mathbf{w}) = \mathcal{N}_w(\mathbf{0}, \mathbf{W})$  and  $\mathbf{f}_v(\mathbf{v}) = \mathcal{N}_v(\mathbf{0}, \mathbf{V})$ . It is also assumed that the disturbances  $\mathbf{w}(k)$  represent the combined effect of all (unmeasured) stochastic disturbances coming from rapidly varying wind gusts and waves, while  $\mathbf{v}(k)$  is regarded as measurement errors that affect the plant outputs.

**Remark 1.** In Eqns. (4) the matrices  $\mathbf{G}_{V_c}$ ,  $\mathbf{G}_{\psi_c}$ ,  $\mathbf{G}_{V_w}$  and  $\mathbf{G}_{\psi_w}$  are omitted, since the signals  $V_c$ ,  $\Psi_c$ ,  $V_w$ ,  $\Psi_w$  in the process of identification take on constant values by assumption, and their effect manifests itself only in the nominal values  $\mathbf{u}_o$  and  $\mathbf{x}_o$  defined by the algebraic equations (2).

The unknown model parameters are the respective entries of matrices  $\mathbf{A}(\mathbf{a}) \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B}(\mathbf{b}) \in \mathbb{R}^{n \times m}$ ,  $\mathbf{C}(\mathbf{c}) \in \mathbb{R}^{l \times n}$ ,  $\mathbf{D}(\mathbf{d}) \in \mathbb{R}^{l \times m}$  and  $\mathbf{G}_w(\mathbf{g}) \in \mathbb{R}^{n \times p}$  forming the vector

$$\boldsymbol{\theta} \triangleq [\mathbf{a}^T, \mathbf{b}^T, \mathbf{c}^T, \mathbf{d}^T, \mathbf{g}^T]^T \quad (5)$$

which is to be identified (estimated). The models of the plant to be identified (4) can assume either a natural form obtained through linearization of the nonlinear plant description (1) and discretization of its linear model (3), where the state variables  $\tilde{\mathbf{x}}(k)$  have a physical interpretation, or—in the absence of *a priori* knowledge about the structure of the model—take a selected canonical form that contains the least possible number of unknown parameters.

Because of the occurring stochastic disturbances  $\mathbf{w}(k)$ , it is assumed that the initial state  $\tilde{\mathbf{x}}(0) = \tilde{\mathbf{x}}_0$  of the plant model can be a random variable with a known mean value  $\bar{\mathbf{x}}_0$  and a covariance matrix  $\mathbf{P}_0$  described

by a normal probability density function  $f_x(\tilde{\mathbf{x}}_0) = \mathcal{N}_x(\tilde{\mathbf{x}}_0, \mathbf{P}_0)$ .

Subject to the assumptions of normality (and independence) of disturbances  $\mathbf{w}(k)$  and  $\mathbf{v}(k)$  and the random initial state vector  $\tilde{\mathbf{x}}(0)$ , there exists a total multivariable conditional probability density  $f^N(\mathbf{Y}^N | \mathbf{U}^{N-1}, \boldsymbol{\theta}, \tilde{\mathbf{x}}_0, \mathbf{P}_0, \mathbf{W}, \mathbf{V})$  that creates, as in, e.g., the works of Aström and Källström (1976) or Bańka (2007), a chain of normal conditionally independent probability densities

$$\begin{aligned} f^N(\mathbf{Y}^N | \mathbf{U}^{N-1}, \boldsymbol{\theta}, \cdot) &= f_N(\tilde{\mathbf{y}}_N | \mathbf{Y}^{N-1}; \mathbf{U}^{N-1}, \boldsymbol{\theta}, \cdot) \\ &\times f_{N-1}(\tilde{\mathbf{y}}_{N-1} | \mathbf{Y}^{N-2}; \mathbf{U}^{N-2}, \boldsymbol{\theta}, \cdot) \\ &\times f_k(\tilde{\mathbf{y}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot) \times \dots \times f_1(\tilde{\mathbf{y}}_1 | \tilde{\mathbf{u}}_0, \boldsymbol{\theta}, \cdot) \\ &= \prod_{k=1}^N f_k(\tilde{\mathbf{y}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot) \end{aligned} \quad (6)$$

where the dot means the values of  $\tilde{\mathbf{x}}_0, \mathbf{P}_0, \mathbf{W}$  and  $\mathbf{V}$  given *a priori*. They are described by two conditional moments:

- the conditional expected value (one-step expectation),

$$\begin{aligned} \hat{\mathbf{y}}_{k|k-1}(\boldsymbol{\theta}) &\triangleq E \{ \tilde{\mathbf{y}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \} \\ &= E \left\{ \mathbf{C}(\mathbf{c})\tilde{\mathbf{x}}(k) + \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}(k) \right. \\ &\quad \left. + \mathbf{v}(k) | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \right\} \\ &= \mathbf{C}(\mathbf{c})E \{ \tilde{\mathbf{x}}(k) | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \} + \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}(k) \\ &\quad + E \{ \mathbf{v}(k) | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \} \\ &= \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k|k-1}(\boldsymbol{\theta}) + \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_k, \end{aligned} \quad (7)$$

where

$$\hat{\mathbf{x}}_{k|k-1}(\boldsymbol{\theta}) \triangleq E \{ \tilde{\mathbf{x}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \};$$

- the conditional covariance matrix,

$$\begin{aligned} \text{cov } \tilde{\mathbf{y}}_k, \tilde{\mathbf{y}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \} &= \mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}(\boldsymbol{\theta})\mathbf{C}^T(\mathbf{c}) + \mathbf{V}, \end{aligned} \quad (8)$$

where

$$\mathbf{P}_{k|k-1}(\boldsymbol{\theta}) \triangleq \text{cov} \{ \tilde{\mathbf{x}}_k, \tilde{\mathbf{x}}_k | \mathbf{Y}^{k-1}; \mathbf{U}^{k-1}, \boldsymbol{\theta}, \cdot \}.$$

Adopting the maximum likelihood (ML) approach defined by the relationship

$$L_N(\boldsymbol{\theta}; \mathbf{Y}^N, \mathbf{U}^{N-1}, \cdot) = f^N(\mathbf{Y}^N; \mathbf{U}^{N-1}, \boldsymbol{\theta}, \cdot) \quad (9)$$

as a method of estimation, and applying the logarithm to both the sides, and neglecting insignificant constant

quantities, we arrive at the problem of minimizing the equivalent criterion

$$\begin{aligned} J_N(\boldsymbol{\theta}) &= \frac{1}{2} \sum_{k=0}^{N-1} \left( \log \det [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}(\boldsymbol{\theta})\mathbf{C}^T(\mathbf{c}) + \mathbf{V}] \right. \\ &\quad \left. + [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k+1|k}(\boldsymbol{\theta}) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_{k+1}]^T \right. \\ &\quad \left. \times [\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}(\boldsymbol{\theta})\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \right. \\ &\quad \left. \times [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k+1|k}(\boldsymbol{\theta}) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_{k+1}] \right), \end{aligned} \quad (10)$$

with

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k}(\boldsymbol{\theta}) &= \mathbf{A}(\mathbf{a})\hat{\mathbf{x}}_{k|k-1}(\boldsymbol{\theta}) + \mathbf{B}(\mathbf{b})\tilde{\mathbf{u}}_k \\ &\quad + \mathbf{A}(\mathbf{a})\mathbf{K}_k(\boldsymbol{\theta}) [\tilde{\mathbf{y}}_k - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k|k-1}(\boldsymbol{\theta}) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_k], \end{aligned} \quad (11)$$

$\hat{\mathbf{x}}_{0|-1} = \tilde{\mathbf{x}}_0$  and

$$\begin{aligned} \mathbf{P}_{k+1|k}(\boldsymbol{\theta}) &= \mathbf{A}(\mathbf{a}) [\mathbf{I}_n - \mathbf{K}_k(\boldsymbol{\theta})\mathbf{C}(\mathbf{c})] \mathbf{P}_{k|k-1}(\boldsymbol{\theta})\mathbf{A}^T(\mathbf{a}) \\ &\quad + \mathbf{G}_w(\mathbf{g})\mathbf{W}\mathbf{G}_w^T(\mathbf{g}), \\ \mathbf{P}_{0|-1} &= \mathbf{P}_0, \end{aligned} \quad (12)$$

where

$$\begin{aligned} \mathbf{K}_k(\boldsymbol{\theta}) &= \mathbf{P}_{k|k-1}(\boldsymbol{\theta})\mathbf{C}^T(\mathbf{c}) \\ &\quad \times [\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}(\boldsymbol{\theta})\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \end{aligned} \quad (13)$$

is the weight matrix of the discrete (optimal) Kalman filter.

Reaching the minimum of the criterion (10) is subject to the following conditions:

- necessary,

$$\frac{\partial J_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0},$$

- and sufficient,

$$\mathbf{H}(\boldsymbol{\theta}) = \left[ \frac{\partial^2 J_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right] > \mathbf{0},$$

for  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^* \in \Xi$ .

The above criterion consists of two components. The first, logarithmic, is characteristic of the maximum likelihood methods. The second is an (optimally weighted) sum of squared errors of the one-step prediction of outputs

$$\begin{aligned} \hat{\mathbf{e}}_{k+1|k}(\boldsymbol{\theta}) &\triangleq \tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}_{k+1|k}(\boldsymbol{\theta}) \\ &= \tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k+1|k}(\boldsymbol{\theta}) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_{k+1} \end{aligned} \quad (14)$$

described by the criterion (performance index)

$$\begin{aligned}
 J'_N(\theta) &= \frac{1}{2} \sum_{k=0}^{N-1} [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k+1|k}(\theta) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_{k+1}]^T \\
 &\quad \times [\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}(\theta)\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \\
 &\quad \times [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c})\hat{\mathbf{x}}_{k+1|k}(\theta) - \mathbf{D}(\mathbf{d})\tilde{\mathbf{u}}_{k+1}].
 \end{aligned} \tag{15}$$

It should be noted that in the full version of the ML criterion given by Eqn. (10), at low values of the covariance matrices  $\mathbf{W}$  and  $\mathbf{V}$ , the first (logarithmic) component of the criterion can also take negative values, which may make the reaching of the minimum impossible. Therefore, in practice, the minimization is performed only with respect to the second component described by Eqn. (15).

If the model structure is fully consistent with the structure of the plant on which the measurements of  $\mathbf{U}^{N-1}$  and  $\mathbf{Y}^{N-1}$  are taken, and all the adopted assumptions about disturbances are satisfied, then minimizing the criterion (10)—provided that the global minimum is reachable—leads, as  $N \rightarrow \infty$ , to asymptotically consistent and unbiased estimates  $\hat{\theta}$  of the unknown parameters of the (discrete-time) plant (Bańka, 2007). In practice, this is usually not the case, since (non-random) components of the discrepancy between the (continuous) plant and its discrete-time model come into play, resulting from linearization and discretization errors and the impact of unmeasured external disturbances acting on the plant.

If the deviations in the manipulated variables and plant outputs from their nominal values during the measurements are too large and/or measurements are taken with an improper (too large) sampling period  $T_p$ , then the result of identification may be a vector  $\theta^*$  that defines under the circumstances the best, in terms of the criterion (15), predictive linear model of the plant, comprising a discrete Kalman filter in its structure.

Minimization of the criteria (10) (or, more frequently, (15)) with respect to the unknown (constant) values  $\theta$  represents a difficult problem of static optimization that can be solved by and large only numerically, i.e., by using gradient-based methods for an iterative search for the minimum of a scalar function with respect to the vector argument, or by employing genetic methods, if the ranges of the values of unknown parameters  $\theta$  are known.

This issue can be effectively solved by replacing the static optimization problem with a formally defined dynamic optimization one through varying all unknown parameters of the model  $\theta$  by means of additional difference equations  $\mathbf{a}(k+1) = \mathbf{a}(k)$ ,  $\mathbf{b}(k+1) = \mathbf{b}(k)$ ,  $\mathbf{c}(k+1) = \mathbf{c}(k)$ ,  $\mathbf{d}(k+1) = \mathbf{d}(k)$ ,  $\mathbf{g}(k+1) = \mathbf{g}(k)$  with unknown initial conditions  $\mathbf{a}_0$ ,  $\mathbf{b}_0$ ,  $\mathbf{c}_0$ ,  $\mathbf{d}_0$  and  $\mathbf{g}_0$

(Bańka, 2007). It will allow us to derive a relation between increments in the criterion (15)  $\Delta J'_N$  and variations in the initial conditions  $\Delta \mathbf{a}_0$ ,  $\Delta \mathbf{b}_0$ ,  $\Delta \mathbf{c}_0$ ,  $\Delta \mathbf{d}_0$  and  $\Delta \mathbf{g}_0$ .

Then, rewriting the criterion (15) in the form (16) and regarding Eqns. (11) and (12) as constraint equations (equality constraints)

$$\begin{aligned}
 \hat{\mathbf{x}}_{k+1|k} &= \mathbf{A}(\mathbf{a}_k)\hat{\mathbf{x}}_{k|k-1} + \mathbf{B}(\mathbf{b}_k)\tilde{\mathbf{u}}_k + \mathbf{A}(\mathbf{a}_k)\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}_k) \\
 &\quad \times [\mathbf{C}(\mathbf{c}_k)\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}_k) + \mathbf{V}]^{-1} \\
 &\quad \times [\tilde{\mathbf{y}}_k - \mathbf{C}(\mathbf{c}_k)\hat{\mathbf{x}}_{k|k-1} - \mathbf{D}(\mathbf{d}_k)\tilde{\mathbf{u}}_k],
 \end{aligned} \tag{17}$$

$\hat{\mathbf{x}}_{0|-1} = \bar{\mathbf{x}}_0$  known, and

$$\begin{aligned}
 \mathbf{P}_{k+1|k} &= \mathbf{A}(\mathbf{a}_k)\mathbf{P}_{k|k-1}\mathbf{A}^T(\mathbf{a}_k) - \mathbf{A}(\mathbf{a}_k)\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}_k) \\
 &\quad \times [\mathbf{C}(\mathbf{c}_k)\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}_k) + \mathbf{V}]^{-1}\mathbf{C}(\mathbf{c}_k) \\
 &\quad \times \mathbf{P}_{k|k-1}\mathbf{A}^T(\mathbf{a}_k) + \mathbf{G}_w(\mathbf{g}_k)\mathbf{W}\mathbf{G}_w^T(\mathbf{g}_k),
 \end{aligned} \tag{18}$$

$$\mathbf{P}_{0|-1} = \mathbf{P}_0 \text{ known,}$$

and

$$\begin{aligned}
 \mathbf{a}_{k+1} &= \mathbf{a}_k, & \mathbf{a}_0 & \text{freely assumed,} \\
 \mathbf{b}_{k+1} &= \mathbf{b}_k, & \mathbf{b}_0 & \text{freely assumed,} \\
 \mathbf{c}_{k+1} &= \mathbf{c}_k, & \mathbf{c}_0 & \text{freely assumed,} \\
 \mathbf{d}_{k+1} &= \mathbf{d}_k, & \mathbf{d}_0 & \text{freely assumed,} \\
 \mathbf{g}_{k+1} &= \mathbf{g}_k, & \mathbf{g}_0 & \text{freely assumed,}
 \end{aligned} \tag{19}$$

we obtain a discrete version of the Lagrange problem with (partially) unknown initial conditions, fixed terminal time  $N$  and free endpoints.

To solve the problem, the vector-matrix Lagrange multipliers are introduced:

$$\begin{aligned}
 \psi_{k+1} & \text{ adjoint to } \hat{\mathbf{x}}_{k+1|k}, \\
 \Psi_{k+1} & \text{ adjoint to } \mathbf{P}_{k+1|k}, \\
 \alpha_{k+1} & \text{ adjoint to } \mathbf{a}_{k+1}, \\
 \beta_{k+1} & \text{ adjoint to } \mathbf{b}_{k+1}, \\
 \gamma_{k+1} & \text{ adjoint to } \mathbf{c}_{k+1}, \\
 \delta_{k+1} & \text{ adjoint to } \mathbf{d}_{k+1}, \\
 \eta_{k+1} & \text{ adjoint to } \mathbf{g}_{k+1},
 \end{aligned} \tag{20}$$

and the Hamiltonian is defined as in (21), into which the right-hand sides of the constraint equations (17)–(19) are to be inserted.

After substituting and expressing the criterion (16) through the Hamiltonian (21), we obtain (22), where the following conditions apply for terminal values of the conjugate variables  $\psi_N = \mathbf{0}$ ,  $\Psi_N = \mathbf{0}$ ,  $\alpha_N = \mathbf{0}$ ,  $\beta_N = \mathbf{0}$ ,  $\gamma_N = \mathbf{0}$ ,  $\delta_N = \mathbf{0}$ ,  $\eta_N = \mathbf{0}$  due to the non-occurrence of constraints imposed on terminal values

$$J'_N(\theta) = \frac{1}{2} \sum_{k=0}^{N-1} [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c}_{k+1})\hat{\mathbf{x}}_{k+1|k} - \mathbf{D}(\mathbf{d}_{k+1})\tilde{\mathbf{u}}_{k+1}]^T [\mathbf{C}(\mathbf{c}_{k+1})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}_{k+1}) + \mathbf{V}]^{-1} \quad (16)$$

$$\times [\tilde{\mathbf{y}}_{k+1} - \mathbf{C}(\mathbf{c}_{k+1})\hat{\mathbf{x}}_{k+1|k} - \mathbf{D}(\mathbf{d}_{k+1})\tilde{\mathbf{u}}_{k+1}],$$

$$H_k = H_k(\hat{\mathbf{x}}_{k|k-1}, \mathbf{a}_k, \mathbf{b}_k, \mathbf{c}_k, \mathbf{d}_k, \mathbf{g}_k, \psi_k, \alpha_k, \beta_k, \gamma_k, \delta_k, \eta_k) \quad (21)$$

$$\triangleq \frac{1}{2} [\mathbf{y}_{k+1} - \mathbf{C}(\mathbf{c}_{k+1})\hat{\mathbf{x}}_{k+1|k} - \mathbf{D}(\mathbf{d}_{k+1})\tilde{\mathbf{u}}_{k+1}]^T [\mathbf{C}(\mathbf{c}_k)\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}_k) + \mathbf{V}]^{-1}$$

$$\times [\mathbf{y}_{k+1} - \mathbf{C}(\mathbf{c}_{k+1})\hat{\mathbf{x}}_{k+1|k} - \mathbf{D}(\mathbf{d}_{k+1})\tilde{\mathbf{u}}_{k+1}] + \psi_{k+1}^T \hat{\mathbf{x}}_{k+1|k} + \text{tr} [\Psi_{k+1}\mathbf{P}_{k+1|k}]$$

$$+ \alpha_{k+1}^T \mathbf{a}_{k+1} + \beta_{k+1}^T \mathbf{b}_{k+1} + \gamma_{k+1}^T \mathbf{c}_{k+1} + \delta_{k+1}^T \mathbf{d}_{k+1} + \eta_{k+1}^T \mathbf{g}_{k+1},$$

$$J'_N = \sum_{k=0}^{N-1} (H_k - \psi_{k+1}^T \hat{\mathbf{x}}_{k+1|k} - \text{tr} [\Psi_{k+1}\mathbf{P}_{k+1|k}]) - \alpha_{k+1}^T \mathbf{a}_{k+1} - \beta_{k+1}^T \mathbf{b}_{k+1} \quad (22)$$

$$- \gamma_{k+1}^T \mathbf{c}_{k+1} - \delta_{k+1}^T \mathbf{d}_{k+1} - \eta_{k+1}^T \mathbf{g}_{k+1})$$

$$\equiv \sum_{k=0}^{N-1} (H_k - \psi_k^T \hat{\mathbf{x}}_{k|k-1} - \text{tr} [\Psi_k\mathbf{P}_{k|k-1}]) - \alpha_k^T \mathbf{a}_k - \beta_k^T \mathbf{b}_k - \gamma_k^T \mathbf{c}_k - \delta_k^T \mathbf{d}_k - \eta_k^T \mathbf{g}_k) + \psi_0^T \hat{\mathbf{x}}_{0|-1}$$

$$+ \text{tr} [\Psi_0\mathbf{P}_{0|-1}] + \alpha_0^T \mathbf{a}_0 + \beta_0^T \mathbf{b}_0 + \gamma_0^T \mathbf{c}_0 + \delta_0^T \mathbf{d}_0 + \eta_0^T \mathbf{g}_0 - \psi_N^T \hat{\mathbf{x}}_{N|N-1} - \text{tr} [\Psi_N\mathbf{P}_{N|N-1}]$$

$$- \alpha_N^T \mathbf{a}_N - \beta_N^T \mathbf{b}_N - \gamma_N^T \mathbf{c}_N - \delta_N^T \mathbf{d}_N - \eta_N^T \mathbf{g}_N,$$

$$\Delta J'_N = \sum_{k=0}^{N-1} \left( \left[ \frac{\partial H_k}{\partial \hat{\mathbf{x}}_{k|k-1}} - \psi_k \right]^T \Delta \hat{\mathbf{x}}_{k|k-1} + \text{tr} \left[ \frac{\partial H_k}{\partial \mathbf{P}_{k|k-1}} - \Psi_k \right]^T \Delta \mathbf{P}_{k|k-1} + \left[ \frac{\partial H_k}{\partial \mathbf{a}_k} - \alpha_k \right]^T \Delta \mathbf{a}_k \quad (23)$$

$$+ \left[ \frac{\partial H_k}{\partial \mathbf{b}_k} - \beta_k \right]^T \Delta \mathbf{b}_k + \left[ \frac{\partial H_k}{\partial \mathbf{c}_k} - \gamma_k \right]^T \Delta \mathbf{c}_k + \left[ \frac{\partial H_k}{\partial \mathbf{d}_k} - \delta_k \right]^T \Delta \mathbf{d}_k + \left[ \frac{\partial H_k}{\partial \mathbf{g}_k} - \eta_k \right]^T \Delta \mathbf{g}_k \right)$$

$$+ \psi_0^T \Delta \hat{\mathbf{x}}_{0|-1} + \Psi_0^T \Delta \mathbf{P}_{0|-1} + \alpha_0^T \Delta \mathbf{a}_0 + \beta_0^T \Delta \mathbf{b}_0 + \gamma_0^T \Delta \mathbf{c}_0 + \delta_0^T \Delta \mathbf{d}_0 + \eta_0^T \Delta \mathbf{g}_0$$

of the variables  $\hat{\mathbf{x}}_{N|N-1}$ ,  $\mathbf{P}_{N|N-1}$  and  $\mathbf{a}_N$ ,  $\mathbf{b}_N$ ,  $\mathbf{c}_N$ ,  $\mathbf{d}_N$ ,  $\mathbf{g}_N$ .

$$\frac{\partial H_k}{\partial \mathbf{g}_k} = \eta_k. \quad (24)$$

Assuming that the Hamiltonian (21) is differentiable, after performing variations in the loss function  $J'_N$  with respect to all the variables occurring there, we get a total differential (23), from which it is inferred that the following equalities will be satisfied for all  $k = 0, 1, \dots, N - 1$ :

$$\frac{\partial H_k}{\partial \hat{\mathbf{x}}_{k|k-1}} = \psi_k, \quad \frac{\partial H_k}{\partial \mathbf{P}_{k|k-1}} = \Psi_k,$$

$$\frac{\partial H_k}{\partial \mathbf{a}_k} = \alpha_k, \quad \frac{\partial H_k}{\partial \mathbf{b}_k} = \beta_k,$$

$$\frac{\partial H_k}{\partial \mathbf{c}_k} = \gamma_k, \quad \frac{\partial H_k}{\partial \mathbf{d}_k} = \delta_k,$$

At the known values  $\hat{\mathbf{x}}_{0|-1} = \bar{\mathbf{x}}_0$  and  $\mathbf{P}_{0|-1} = \mathbf{P}_0$ , the criterion increments  $\Delta J'_N$  caused by variations in  $\Delta \mathbf{a}_0$ ,  $\Delta \mathbf{b}_0$ ,  $\Delta \mathbf{c}_0$ ,  $\Delta \mathbf{d}_0$  and  $\Delta \mathbf{g}_0$  will be determined simply as

$$\Delta J'_N = \alpha_0^T \Delta \mathbf{a}_0 + \beta_0^T \Delta \mathbf{b}_0 + \gamma_0^T \Delta \mathbf{c}_0 \quad (25)$$

$$+ \delta_0^T \Delta \mathbf{d}_0 + \eta_0^T \Delta \mathbf{g}_0.$$

This means that the components of the loss function gradient vector (16) regarded with respect to the sought-for parameter values  $\theta^T \triangleq [\mathbf{a}^T, \mathbf{b}^T, \mathbf{c}^T, \mathbf{d}^T, \mathbf{g}^T]$  will equal the

initial values of conjugate variables

$$\begin{aligned} \frac{\partial J'_N(\boldsymbol{\theta})}{\partial \mathbf{a}} &= \boldsymbol{\alpha}_0, & \frac{\partial J'_N(\boldsymbol{\theta})}{\partial \mathbf{b}} &= \boldsymbol{\beta}_0, \\ \frac{\partial J'_N(\boldsymbol{\theta})}{\partial \mathbf{c}} &= \boldsymbol{\gamma}_0, & \frac{\partial J'_N(\boldsymbol{\theta})}{\partial \mathbf{d}} &= \boldsymbol{\delta}_0, \\ \frac{\partial J'_N(\boldsymbol{\theta})}{\partial \mathbf{g}} &= \boldsymbol{\eta}_0. \end{aligned} \tag{26}$$

Determination of the derivatives (24) leads to the conjugate equations (27)–(31), where  $\partial \mathbf{P}_{k+1|k} / \partial c_{i,j}$ ,  $(\boldsymbol{\delta}_k)_{i,j}$  and  $(\boldsymbol{\eta}_k)_{i,j}$  are described by (32)–(34), respectively (Bańka, 2007).

In the above relationships, the derivatives

$$\begin{aligned} \frac{\partial \mathbf{A}^T(\mathbf{a})}{\partial a_{i,j}}, & \quad \frac{\partial \mathbf{B}^T(\mathbf{b})}{\partial b_{i,j}}, & \quad \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{i,j}}, \\ \frac{\partial \mathbf{D}^T(\mathbf{d})}{\partial d_{i,j}}, & \quad \frac{\partial \mathbf{G}_w^T(\mathbf{g})}{\partial g_{j,i}} \end{aligned} \tag{35}$$

are matrices filled with zeros excepted for one entry equal to 1 that lies at the intersection of the  $i$ -th row and the  $j$ -th column.

On the basis of the above discussion and the obtained relationships, one may formulate the identification algorithm as Algorithm 1 (Bańka, 2007) for plants being subject to unmeasured stochastic disturbances of the white Gaussian noise type.

The calculations are to be continued until the successive values of the criterion (16) differ by less than a given preset value  $\varepsilon_J > 0$ , that is,

$$J_N^{(i)} - J_N^{(i+1)} < \varepsilon_J, \tag{36}$$

where  $\varepsilon_J$  is a small positive number, or are to be stopped after performing a preset number of iterations  $i = i_{\max}$ . Since there is no guarantee that the global minimum is achievable or the computations will stop on the border of the permissible domain  $\Xi$ , the computations should be repeated for different values of starting parameters  $\boldsymbol{\theta}^{(0)} \in \Xi$ .

### 5. Example: Parameter identification of the “Wimpey Sealab” drilling vessel

According to the procedure described at the beginning of the paper, an active experiment was conducted on the controlled plant, i.e., the Kempf zigzag test on the nonlinear model of the ship “Wimpey Sealab” given by equations (Wise and English, 1975; Bańka et al., 2010;

#### Algorithm 1. Identification procedure.

**Step 1.** Read in *a priori* data  $\bar{\mathbf{x}}_0$ ,  $\mathbf{P}_0$ ,  $\mathbf{W}$  and  $\mathbf{V}$  and results of measured inputs  $\mathbf{U}^{N-1} = [\tilde{\mathbf{u}}_0, \tilde{\mathbf{u}}_1, \dots, \tilde{\mathbf{u}}_{N-1}]$  and outputs  $\mathbf{Y}^N = [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_N]$  of the plant taken from within the range of observations  $[0, N]$ .

**Step 2.** In iteration  $i := 0$  choose the starting values for the parameters  $\boldsymbol{\theta}^{(i)} \in \Xi$  (5).

**Step 3.** For  $k = 0, 1, 2, \dots, N - 1$  solve the constraint equations (18) and (17) in the normal course of time, memorizing the sequences of evaluated values as arrays

$$[\mathbf{P}_{1|0}^{(i)}, \mathbf{P}_{2|1}^{(i)}, \mathbf{P}_{3|2}^{(i)}, \dots, \mathbf{P}_{N|N-1}^{(i)}]$$

and

$$[\hat{\mathbf{x}}_{1|0}^{(i)}, \hat{\mathbf{x}}_{2|1}^{(i)}, \hat{\mathbf{x}}_{3|2}^{(i)}, \dots, \hat{\mathbf{x}}_{N|N-1}^{(i)}].$$

Compute the entries of the Kalman filter weight matrix (gain matrix) (13) and the values of the squared errors of the one-step predictions of outputs (14) and store them as arrays

$$[\mathbf{K}_1^{(i)}, \mathbf{K}_2^{(i)}, \mathbf{K}_3^{(i)}, \dots, \mathbf{K}_N^{(i)}]$$

and

$$[\hat{\mathbf{e}}_{1|0}^{(i)}, \hat{\mathbf{e}}_{2|1}^{(i)}, \hat{\mathbf{e}}_{3|2}^{(i)}, \dots, \hat{\mathbf{e}}_{N|N-1}^{(i)}].$$

Determine and store the value of the loss function  $J_N^{(i)}$  computed in accordance with (16).

**Step 4.** In utilizing the values computed and memorized in Step 3 solve the adjoint equations (27)–(34) backwards in time for  $k = N - 1, N - 2, \dots, 0$  and determine the gradient vector by substitution

$$\mathbf{g}_\theta^{(i)} = \begin{bmatrix} \boldsymbol{\alpha}_0^{(i)} \\ \boldsymbol{\beta}_0^{(i)} \\ \boldsymbol{\gamma}_0^{(i)} \\ \boldsymbol{\delta}_0^{(i)} \\ \boldsymbol{\eta}_0^{(i)} \end{bmatrix}$$

**Step 5.** Determine the direction the search for a minimum  $\boldsymbol{\zeta}^{(i)}$  according to one of the selected gradient methods and minimize the loss function  $J_N^{(i)}$  along the direction  $\boldsymbol{\zeta}^{(i)}$  so that

$$J_N^{(i)}(\boldsymbol{\theta}^{(i)} + d^{(i)} \boldsymbol{\zeta}^{(i)}) \rightarrow \min.$$

Store the last length of the step  $d_{\text{opt}}^{(i)} = d^{(i)}$  and compute

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} + d_{\text{opt}}^{(i)} \boldsymbol{\zeta}^{(i)}.$$

Store the obtained (best) values of the model parameters  $\boldsymbol{\theta}^{(i+1)}$  and the achieved criterion value

$$J_N^{(i+1)} < J_N^{(i)}.$$

**Step 6.** At  $i := i + 1$  go to Step 1.

$$\begin{aligned} \psi_k &= \frac{\partial H_k}{\partial \hat{\mathbf{x}}_{k|k-1}} \\ &= [\mathbf{I}_n - \mathbf{C}^T(\mathbf{c})\mathbf{K}_k^T] \mathbf{A}^T(\mathbf{a}) [\psi_{k+1} - \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k}], \quad \psi_N = \mathbf{0}, \end{aligned} \quad (27)$$

$$\begin{aligned} \Psi_k &= \frac{\partial H_k}{\partial \mathbf{P}_{k|k-1}} \\ &= \psi_k \hat{\mathbf{e}}_{k|k-1}^T [\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \mathbf{C}(\mathbf{c}) \\ &\quad + [\mathbf{I}_n - \mathbf{C}^T(\mathbf{c})\mathbf{K}_k^T] \mathbf{A}^T(\mathbf{a}) \left[ \psi_{k+1} - \frac{1}{2} \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} [\hat{\mathbf{e}}_{k+1|k} \hat{\mathbf{e}}_{k+1|k}^T] \right] \\ &\quad \times [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \mathbf{C}(\mathbf{c}) \mathbf{A}(\mathbf{a}) [\mathbf{I}_n - \mathbf{K}_k \mathbf{C}(\mathbf{c})], \quad \Psi_N = \mathbf{0}, \end{aligned} \quad (28)$$

$$\begin{aligned} (\alpha_k)_{ij} &= \frac{\partial H_k}{\partial a_{i,j}} \\ &= [\hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_k \hat{\mathbf{e}}_{k|k-1}]^T \frac{\partial \mathbf{A}(\mathbf{a})}{\partial a_{i,j}} [\psi_{k+1} - \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k}] \\ &\quad - \frac{1}{2} \hat{\mathbf{e}}_{k+1|k}^T [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \mathbf{C}(\mathbf{c}) \left[ \frac{\partial \mathbf{A}(\mathbf{a})}{\partial a_{i,j}} [\mathbf{I}_n - \mathbf{K}_k \mathbf{C}(\mathbf{c})] \mathbf{P}_{k|k-1} \mathbf{A}^T(\mathbf{a}) \right. \\ &\quad \left. + \mathbf{A}(\mathbf{a}) [\mathbf{I}_n - \mathbf{K}_k \mathbf{C}(\mathbf{c})] \mathbf{P}_{k|k-1} \frac{\partial \mathbf{A}^T(\mathbf{a})}{\partial a_{j,i}} \right] \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k} \\ &\quad + \text{tr} \left[ \left[ \mathbf{A}(\mathbf{a}) \mathbf{P}_{k|k-1} [\mathbf{I}_n - \mathbf{C}^T(\mathbf{c})\mathbf{K}_k^T] \frac{\partial \mathbf{A}^T(\mathbf{a})}{\partial a_{j,i}} + \frac{\partial \mathbf{A}(\mathbf{a})}{\partial a_{i,j}} [\mathbf{I}_n - \mathbf{C}^T(\mathbf{c})\mathbf{K}_k^T] \mathbf{P}_{k|k-1} \mathbf{A}^T(\mathbf{a}) \right] \Psi_{k+1} \right] \\ &\quad + (\alpha_{k+1})_{ij}, \quad (\alpha_N)_{ij} = 0, \end{aligned} \quad (29)$$

$$\begin{aligned} (\beta_k)_{ij} &= \frac{\partial H_k}{\partial b_{i,j}} \\ &= \tilde{\mathbf{u}}_k^T \frac{\partial \mathbf{B}^T(\mathbf{b})}{\partial b_{i,j}} [\psi_{k+1} - \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k}] + (\beta_{k+1})_{ij}, \quad (\beta_N)_{ij} = 0, \end{aligned} \quad (30)$$

$$\begin{aligned} (\gamma_k)_{ij} &= \frac{\partial H_k}{\partial c_{i,j}} \\ &= \left[ \hat{\mathbf{x}}_{k+1|k}^T \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{j,i}} + \frac{1}{2} \hat{\mathbf{e}}_{k+1|k}^T [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \tilde{\mathbf{C}}(\mathbf{c}) \right] \\ &\quad \times \left[ \frac{\partial \mathbf{P}_{k+1|k}}{\partial c_{i,j}} \right] \mathbf{C}^T(\mathbf{c}) [\mathbf{C}(\mathbf{c})\mathbf{P}_{k+1|k}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k} \\ &\quad - \left[ \hat{\mathbf{x}}_{k|k-1}^T \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{j,i}} \mathbf{K}_k^T + \hat{\mathbf{e}}_{k|k-1}^T [\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \right. \\ &\quad \left. \times \left[ \left[ \frac{\partial \mathbf{C}(\mathbf{c})}{\partial c_{i,j}} \mathbf{P}_{k|k-1} \mathbf{C}^T(\mathbf{c}) + \mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1} \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{j,i}} \right] \mathbf{K}_k^T - \frac{\partial \mathbf{C}(\mathbf{c})}{\partial c_{i,j}} \mathbf{P}_{k|k-1} \right] \mathbf{A}^T(\mathbf{a}) \right] \psi_{k+1} \\ &\quad + \text{tr} \left[ \left[ \frac{\partial \mathbf{P}_{k+1|k}}{\partial c_{i,j}} \right]^T \Psi_{k+1} \right] + (\gamma_{k+1})_{ij}, \quad (\gamma_N)_{ij} = 0. \end{aligned} \quad (31)$$

$$\frac{\partial \mathbf{P}_{k+1|k}}{\partial c_{i,j}} = \mathbf{A}(\mathbf{a}) \left[ -\mathbf{P}_{k|k-1} \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{i,j}} + \mathbf{K}_k \left[ \frac{\partial \mathbf{C}(\mathbf{c})}{\partial c_{i,j}} \mathbf{P}_{k|k-1} \mathbf{C}^T(\mathbf{c}) + \mathbf{C}(\mathbf{c}) \mathbf{P}_{k|k-1} \frac{\partial \mathbf{C}^T(\mathbf{c})}{\partial c_{j,i}} \right] \mathbf{K}_k^T \right. \\ \left. - \frac{\partial \mathbf{C}(\mathbf{c})}{\partial c_{i,j}} \mathbf{P}_{k|k-1} \right] \mathbf{A}^T(\mathbf{a}), \quad (32)$$

$$(\delta_k)_{ij} = \frac{\partial H_k}{\partial d_{j,i}} = -\tilde{\mathbf{u}}_{k+1}^T \frac{\partial \mathbf{D}^T(\mathbf{d})}{\partial d_{j,i}} [\mathbf{C}(\mathbf{c}) \mathbf{P}_{k+1|k} \mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \hat{\mathbf{e}}_{k+1|k} \\ - \tilde{\mathbf{u}}_k^T \frac{\partial \mathbf{D}^T(\mathbf{d})}{\partial d_{j,i}} \mathbf{K}_k^T \mathbf{A}^T(\mathbf{a}) \psi_{k+1} + (\delta_{k+1})_{ij}, \quad (\delta_N)_{ij} = 0. \quad (33)$$

2013; Dworak and Brasel, 2013)

$$\begin{aligned} \dot{x}_1 &= x_4 \cos x_3 - x_5 \sin x_3 + V_c \cos \Psi_c, \\ \dot{x}_2 &= x_4 \sin x_3 + x_5 \cos x_3 + V_c \sin \Psi_c, \\ \dot{x}_3 &= x_6, \\ \dot{x}_4 &= 0.088x_5^2 - 0.132x_4V_s + 0.958x_5x_6 \\ &\quad + 0.958u_1, \\ \dot{x}_5 &= -1.4x_5V_s - 0.978x_5^3/V_s - 0.543x_4x_6 \\ &\quad + 0.037x_6|x_6| + 0.544u_2, \\ \dot{x}_6 &= \frac{1}{a}(0.258x_5V_s - 0.764x_4x_5 \\ &\quad - 0.162x_6|x_6| + u_3), \\ y_1 &= x_1, \quad y_2 = x_2, \quad y_3 = x_3, \end{aligned} \quad (37)$$

where the state variables  $x_1, \dots, x_3$  represent the ship position and course angle over the drilling point and  $x_4, \dots, x_6$  her longitudinal, transversal and angular velocities (see Fig. 1),  $V_s = \sqrt{x_4^2(t) + x_5^2(t)}$  is the translational velocity of the ship measured with respect to water. The coefficient  $a = k_{zz}^2 + 0.0431$  describes the ship's inertia moment together with water associated with the angle motion of the ship around its vertical axis. Here  $k_{zz}^2$  is the square of the relative inertia radius referenced to the ship's length  $L_{pp}$ , and  $V_c$  and  $\Psi_c$  are, respectively, the velocity and direction of the sea current as indicated in Fig. 1. All the signals appearing in (38) are dimensionless, i.e., related to the ship's dimensions and displacement together with the dimensionless time  $t = t_r/\sqrt{L_{pp}/g} \approx 0.32 t_r$  (Wise and English, 1975).

The Kempf test was carried out using the MATLAB/Simulink environment on the developed nonlinear simulation ship model with animated ship's motions for a selected operating point over the drilling point. It is defined by the nominal values of the course angle  $x_{30} = 45^\circ$  at  $V_c = 2$  knots and  $\Psi_c = 180^\circ$  and  $V_w = 0$ , under the assumption that the radius of gyration of the vessel is constant during the test and equals  $k_{zz} = 1/4$ .

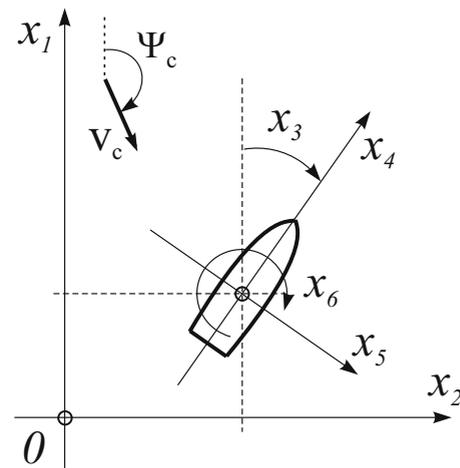


Fig. 1. Ship's co-ordinate systems.

At small course deviations  $\tilde{x}_3(t) = x_3(t) - x_{30}$  caused by torque variations, for example, in the form of a “square wave” with an appropriately selected amplitude  $\pm \Delta M_z$  (e.g., equaling 20% of the maximum possible torque to be generated by the propellers), it can be assumed that the longitudinal and transverse components of the forces developed by the main engine and the thrusters will depend directly on the value of this deviation in accordance with the following trigonometric relationships:

$$\begin{aligned} \tilde{u}_1(t) &= \tilde{F}_x(t) = (1 - \cos(\tilde{x}_3(t))) \Delta F_x, \\ \tilde{u}_2(t) &= \tilde{F}_y(t) = -\sin(\tilde{x}_3(t)) \Delta F_y, \end{aligned} \quad (38)$$

where  $\Delta F_x$  and  $\Delta F_y$  are experimentally chosen values of the amplitudes of these forces (representing, e.g., 1% of their maximum value possible to be generated by the main engine or thrusters, respectively).

The nominal values for the remaining state coordinates  $\mathbf{x}_o$  and components of manipulated variables  $\mathbf{u}_o$  required to keep the ship's position over the drilling

$$\begin{aligned}
 (\boldsymbol{\eta}_k)_{ij} &= \frac{\partial H_k}{\partial g_{i,j}} \\
 &= \frac{1}{2} \hat{\mathbf{e}}_{k+1|k}^T [\mathbf{C}(\mathbf{c}) \mathbf{P}_{k+1|k} \mathbf{C}^T(\mathbf{c}) + \mathbf{V}]^{-1} \left[ \mathbf{C}(\mathbf{c}) \left[ \frac{\partial \mathbf{G}_w(\mathbf{g})}{\partial g_{i,j}} \mathbf{W} \mathbf{G}_w^T(\mathbf{g}) + \mathbf{G}_w(\mathbf{g}) \mathbf{W} \frac{\partial \mathbf{G}_w^T(\mathbf{g})}{\partial g_{i,j}} \right] \mathbf{C}^T(\mathbf{c}) \right] \\
 &\quad \times \left[ \mathbf{C}(\mathbf{c}) \mathbf{P}_{k+1|k} \mathbf{C}^T(\mathbf{c}) + \mathbf{V} \right]^{-1} \hat{\mathbf{e}}_{k+1|k}^T + \text{tr} \left[ \left[ \frac{\partial \mathbf{G}_w(\mathbf{g})}{\partial g_{i,j}} \mathbf{W} \mathbf{G}_w^T(\mathbf{g}) + \mathbf{G}_w(\mathbf{g}) \mathbf{W} \frac{\partial \mathbf{G}_w^T(\mathbf{g})}{\partial g_{i,j}} \right] \boldsymbol{\Psi}_{k+1} \right] \quad (34) \\
 &+ (\boldsymbol{\eta}_{k+1})_{ij}, \quad (\boldsymbol{\eta}_N)_{ij} = 0.
 \end{aligned}$$

point, according to (2), are equal to

$$\begin{aligned}
 u_{10} &= 5.8816 \cdot 10^{-5} & (F_{x0} = 3.2715 \text{ kN}), & \quad (39) \\
 u_{20} &= -2.8041 \cdot 10^{-3} & (F_{y0} = -15.5970 \text{ kN}), \\
 u_{30} &= -2.2791 \cdot 10^{-4} & (M_{z0} = -1.1979 \text{ MNm}).
 \end{aligned}$$

In order for the vessel to make only angular motions about the vertical axis during the zigzag test and to keep the position over the drilling point, the control policy defined by (38) was employed with the following experimentally chosen amplitudes:

$$\Delta u_3 = 3.9000 \cdot 10^{-4} \quad (\Delta M_z = 2049.80 \text{ kNm})$$

and

$$\begin{aligned}
 \Delta u_1 = \Delta u_2 &= 4.1183 \cdot 10^{-4} \\
 (\Delta F_x = \Delta F_y &= 22.907 \text{ kN})
 \end{aligned}$$

for the respective manipulated variables in the form of a “square wave” produced manually for the torque  $\hat{u}_3(t) \triangleq M_z(t)$  and automatically generated forces  $\hat{u}_1(t) \triangleq F_x(t)$  and  $\hat{u}_2(t) \triangleq F_y(t)$ , counteracting the ship’s departure from the drilling point.

**Remark 2.** Providing similar conditions on a real ship under sea operating conditions is, of course, more difficult and requires full compensation of all forces and moments acting on the ship’s hull at a specific configuration of impacts (being constant as far as possible) produced by the marine environment. Provision of such conditions is generally possible only on a board of fully equipped vessels with an automatic control system switched on. Then, for the sake of the active nature of the experiment, it should be remembered that the angular motions in the Kempf zigzag test should be forced by means of an additional torque rotating the vessel, generated independently of the control system.

It is assumed that during the simulation the ship’s model is acted upon by some additional white noise  $\mathbf{w}(t)$ , replacing the total impact of (rapidly varying) forces coming from the wind and waves with zero mean values

and variances defined by the covariance matrix

$$\begin{aligned}
 \mathbf{W} &= \begin{bmatrix} 8.6484 \cdot 10^{-10} & 0 & 0 \\ 0 & 1.9658 \cdot 10^{-6} & 0 \\ 0 & 0 & 1.2986 \cdot 10^{-8} \end{bmatrix}.
 \end{aligned}$$

The variances correspond to standard deviations of randomly generated forces, the longitudinal and the transverse one, and those of the torque rotating the vessel, approximately equal to 50% of their nominal values  $\mathbf{u}_o$  given in (39).

Also, it was assumed that measurements of the position and the course angle are affected by random errors  $\mathbf{v}(t)$  with zero mean values and variances defined by the covariance matrix

$$\begin{aligned}
 \mathbf{V} &= \begin{bmatrix} 1.0000 \cdot 10^{-4} & 0 & 0 \\ 0 & 1.0000 \cdot 10^{-4} & 0 \\ 0 & 0 & 3.0000 \cdot 10^{-4} \end{bmatrix},
 \end{aligned}$$

which correspond to standard deviations of errors in measurements of the position and course angle equal about  $\pm 1$  m and  $\pm 1$  deg. The noise signals  $\mathbf{w}(t)$  and  $\mathbf{v}(t)$  were generated during the tests by generators of pseudorandom numbers and added directly to the manipulated variables  $\mathbf{u}(t)$  acting on the plant and  $\mathbf{v}(t)$  to the plant outputs  $\mathbf{y}(t)$ .

After subtracting the nominal values  $\mathbf{u}_o$  and  $\mathbf{y}_o$  from the obtained signals  $\mathbf{u}(t)$  and  $\mathbf{y}(t)$  and sampling them with a period of  $T_p = 0.1$  ( $\approx 0.32$  s), measurement sequences  $\mathbf{U}^{N-1}$  and  $\mathbf{Y}^N$  were created, used as input data for the algorithm of parameter identification of the discrete-time linear ship model together with the adopted *a priori* mean value and covariance matrix for the initial state of the linear model, namely,  $\bar{\mathbf{x}}(0) = \mathbf{0}$  and  $\mathbf{P}_0 = \text{diag}\{2.5 \cdot 10^{-6}, 2.5 \cdot 10^{-6}, 3.05 \cdot 10^{-4}, 1.06 \cdot 10^{-8}, 2.0 \cdot 10^{-7}, 7.0 \cdot 10^{-7}\}$ .

The structure of the linear model was established during linearization of the nonlinear state equations (37), which describe LF motions made by the “Wimpey Sealab” drilling vessel. As a result, we obtained linear ship models

defined by the difference state and output equations (4), in a natural form with the matrices

$$\mathbf{A} = \bar{\mathbf{A}}T_p + \mathbf{I}_n = \begin{bmatrix} 1 & 0 & 0 & a_{14} & a_{15} & 0 \\ 0 & 1 & a_{23} & a_{24} & a_{25} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 0 & a_{64} & a_{65} & 1 \end{bmatrix}, \tag{40}$$

$$\mathbf{B} = \bar{\mathbf{B}}T_p = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ b_{31} & 0 & 0 \\ 0 & b_{52} & 0 \\ 0 & 0 & b_{63} \end{bmatrix},$$

and the output matrix

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

In the natural form (40), there occur  $l_p = 17$  parameters in total to be identified. Among them, there are 14 varying parameters  $a_{ij}$  and  $b_{63}$ , with the latter (and also  $a_{64}$  and  $a_{65}$ ) being dependent on the value of the adopted assumed ship radius of gyration  $k_{zz}$ . During the tests,  $k_{zz} \in [0.25, 0.5]$  was adopted. Except for the overall structure of the model (40), no other assumptions have been made, e.g., correlation between particular parameters. The only limitation applies to the ranges of parameters values. They were determined by the linearization of the nonlinear ship model (37) carried out over the entire range of the course angle  $x_{30} \in [-\pi, \pi]$  rd, at different values of the sea current velocity  $V_c \in [0.01, 3.5]$  knots and at the angle  $\Psi_c = \pi$ . The values are summarized in Table 1. On the other hand, the canonical observable Luenberger form for the same linear model includes altogether  $l_p = 23$  parameters to be identified.

The intervals given in Table 1 form subspaces of admissible values  $\Xi$  in  $l_p$ -dimensional parameter spaces. Due to limitations of parameter values, they may not be convex. This is indicated by the fact that taking terminal (upper or lower) values for the parameters  $\theta$  creates models that have more unstable poles than all natural models obtained for the studied ranges of the ship’s course angle  $x_{30}$  and the sea current velocity  $V_c$ .

Also, it is worth noting that some of the parameters take their values from within very “tight” intervals. It is therefore to be expected that, if we have to do with strongly varying dynamic properties of the vessel over the studied range of the course angle  $x_{30}$  and the ship’s velocity  $V_s - V_c$ , the “hypersurfaces” generated by the criterion  $J'_N(\theta)$  may have narrow “valleys” with very

Table 1. Constraints on the parameters of the natural model.

Item	Parameter
1	$-0.1 \leq a_{14} \leq 0.1$
2	$-0.1 \leq a_{15} \leq 0.1$
3	$0 \leq a_{23} \leq 0.0068$
4	$-0.1 \leq a_{24} \leq 0.1$
5	$-0.1 \leq a_{25} \leq 0.1$
6	$0.99 \leq a_{44} \leq 1$
7	$-0.0014 \leq a_{45} \leq 0.0014$
8	$-0.007 \leq a_{46} \leq 0.007$
9	$-0.0034 \leq a_{54} \leq 0.0034$
10	$0.9679 \leq a_{55} \leq 0.9998$
11	$-0.0037 \leq a_{56} \leq 0.0037$
12	$-0.0495 \leq a_{64} \leq 0.0495$
13	$-0.0324 \leq a_{65} \leq 0.0324$
14	$0.05 \leq b_{41} \leq 0.1$
15	$0.02 \leq b_{52} \leq 0.06$
16	$0.35 \leq b_{63} \leq 1$

steep “slopes”, which may impede the search for the minimum.

The identification of the model parameters was carried out by means of two methods: using the gradient-based algorithm presented in the paper and with a genetic algorithm. Both methods allow correct results to be obtained, but they inherently differ in their implementation and in the way the calculations are conducted.

**Results produced by the gradient-based algorithm.**

On the basis of the performed calculations and the obtained results, it can be concluded that the parameter identification of linear models of the drilling ship “Wimpey Sealab” proved to be a difficult optimization problem that requires lengthy and numerically sensitive numerical calculations to be carried out. The main reason is that the nonlinear ship model is intrinsically dimensionless, where all the signals assume very small values within the range of real numbers  $\ll 1$ .

As a result, during the calculations, warnings were very often generated related to the inability to reverse the ill-conditioned weighting matrix of the criterion  $[\mathbf{C}(\mathbf{c})\mathbf{P}_{k|k-1}(\theta)\mathbf{C}^T(\mathbf{c}) + \mathbf{V}]$ , which occurs repeatedly in relationships needed to calculate the objective function (criterion)  $J'_N(\theta)$  and the gradient. This is confirmed by the fact that, after setting  $\mathbf{W} = \mathbf{0}$  and  $\mathbf{P}_0 = \mathbf{0}$  (that is, after ignoring the available *a priori* information about the disturbances  $\mathbf{w}(k)$ ) in the developed identification algorithm, the calculations run quickly, delivering Markovian estimates with the criterion weighting matrix that takes into account only statistical data about plant output measurement errors contained in the covariance matrix  $\mathbf{V}$  of the measurement noise  $\mathbf{v}(k)$ . Unfortunately, parameter estimates  $\theta$  obtained in such a way are much worse than those obtained by the maximum likelihood

method, i.e., by using all available *a priori* information.

Even worse results from the same measurement data  $U^{N-1}$  and  $Y^N$  are obtained by replacing the weighting matrix  $[C(c)P_{k|k-1}(\theta)C^T(c)+V]$  in the criterion (15) by the identity matrix  $I_l$ , which is possible after setting  $W = 0$ ,  $P_0 = 0$  and  $V = I_l$  in the algorithm. In doing so, the discussed estimation method (ML) is reduced to the ordinary least-squares method.

The numerical sensitivity intensifies the need for using advanced optimization methods, such as second order gradient-based techniques with a variable metric (e.g., the Davidon–Fletcher–Powell (DFP) method or the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method with calculating a Hessian of large dimensions  $l_p \times l_p$ ). The use of genetic methods, especially when, in the case under study, the ranges of the sought-for parameters are known, greatly facilitates the solution to this problem.

**Results produced by the genetic algorithm.** The criterion (15) was optimized using the genetic method implemented in the Matlab environment with the Optimization Toolbox for the same (identical) measurement data obtained from the Kempf zigzag test and with adopted constraints conditioned by physical meaning that are imposed on the parameter values of the natural model to be identified (Table 1).

In order to have a possibility to compare other optimization methods, the same fitness function in all cases was employed.

In the processes of optimization, a roulette selection function with two elite individuals was used. Moreover, the scaling of the fitness function was conducted by using a range function. After preliminary comparative research, the initial population was established to ten individuals. Two different approaches were applied to prepare the initial population. Apart from constraints, in one case six from ten individuals were set in advance. In the second case, all individuals were randomly chosen. As the difference in the value of the fitness function of these individuals is rather small, we suggest that using an initial population without in-advance prepared individuals is a good approach, as good as with in advance prepared ones. The convergence of so prepared calculations allowed us to stop them before one hundred generations.

Illustrative examples of parameter identification results obtained for the model given in the natural form (40) are presented in Table 2 for the following vectors:  $\theta_{start\_grad}$ , initial values for the gradient-based algorithm;  $\hat{\theta}_{best\_grad}$ , the “best” final result of computations delivered by the gradient-based algorithm;  $\theta_{theoretic}$ , the set of parameters describing the model yielded directly after linearization and discretization for a given nominal operating point;  $\hat{\theta}_{best\_gen}$  the best result yielded by the genetic algorithm.

In making the calculations with due account of all *a*

Table 2. Identification results.

Parameters	$\hat{\theta}_{best\_grad}$	$\theta_{theoretic}$	$\hat{\theta}_{best\_gen}$
$a_{14}$	0.045513	0.070711	0.079061
$a_{15}$	-0.080083	-0.070711	-0.074403
$a_{23}$	0.003517	0.003379	0.003162
$a_{24}$	0.051545	0.070711	0.019711
$a_{25}$	0.096987	0.070711	0.088827
$a_{44}$	0.993691	0.999331	0.995490
$a_{45}$	-0.001047	-0.000197	-0.000566
$a_{46}$	-0.003056	-0.002289	-0.002290
$a_{54}$	-0.001202	0.001539	-0.001887
$a_{55}$	0.994031	0.988772	0.999461
$a_{56}$	-0.000909	-0.001297	-0.001644
$a_{64}$	0.049026	0.013160	0.045848
$a_{65}$	0.023308	-0.004904	-0.004075
$b_{41}$	0.050158	0.095800	0.063409
$b_{52}$	0.060000	0.054300	0.027839
$b_{63}$	0.816091	0.946970	0.886175
$J'_N$	1914.1	1922.7	1920.7

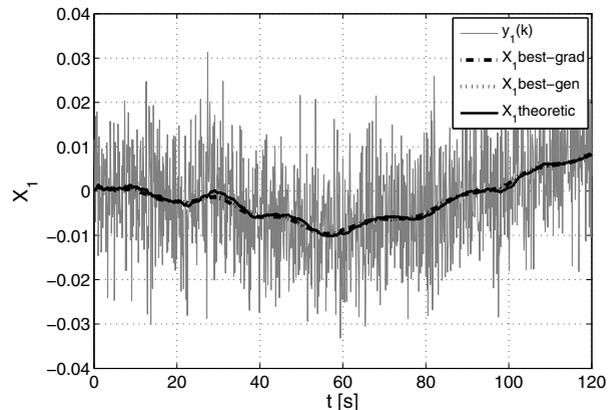


Fig. 2. Time-dependent behavior of the first component of the state vector  $\hat{x}_{k+1|k}$  against the background of  $\tilde{y}_1(k)$ .

*priori* data that minimize the objective function (15), the best values of the parameters  $\hat{\theta}_{best\_grad}$  and  $\hat{\theta}_{best\_gen}$  were obtained, for which the time curves for components of the state vector  $\hat{x}_{k+1|k}$  are presented in Figs. 2–7. The corresponding curves of the one-step output prediction errors  $\hat{e}_{k+1|k}(\hat{\theta}_{best\_grad})$ , neglected for their low legibility, are sufficiently white and have means close to zero.

When comparing the identification results of the ship’s model parameters delivered by the gradient-based method and the genetic one, it may be stated that the obtained results are similar. The time behaviors of state variables estimated by the Kalman filter obtained by means of both the methods largely coincide with each other—the curves for best-gen, best-grad and the theoretical solution in Figs. 2–7 largely overlap.

It should be noted that the best individual from

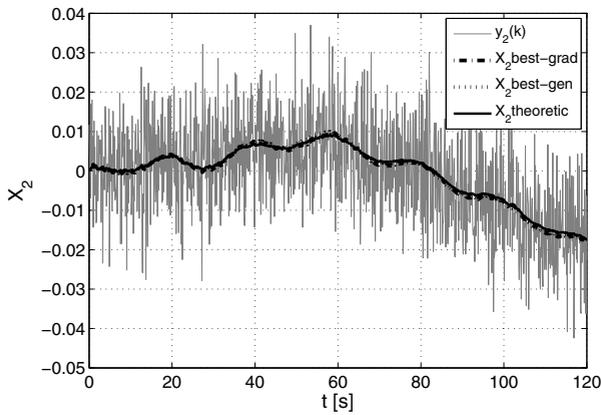


Fig. 3. Time-dependent behavior of the second component of the state vector  $\hat{x}_{k+1|k}$  against the background of  $\tilde{y}_2(k)$ .

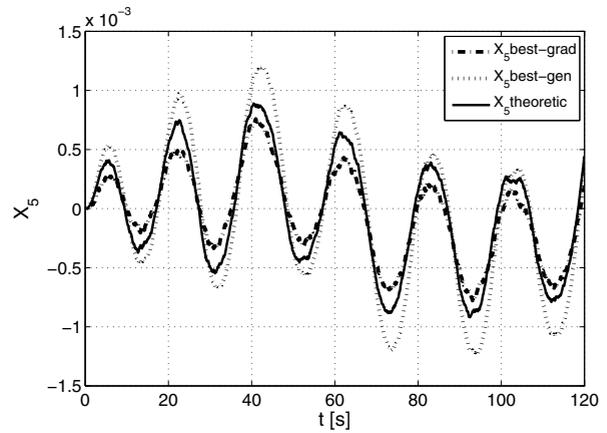


Fig. 6. Time-dependent behavior of the fifth component of the state vector.

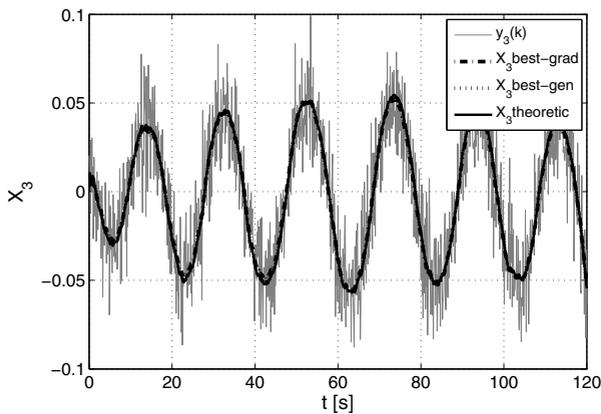


Fig. 4. Time-dependent behavior of the third component of the state vector  $\hat{x}_{k+1|k}$  against the background of  $\tilde{y}_3(k)$ .

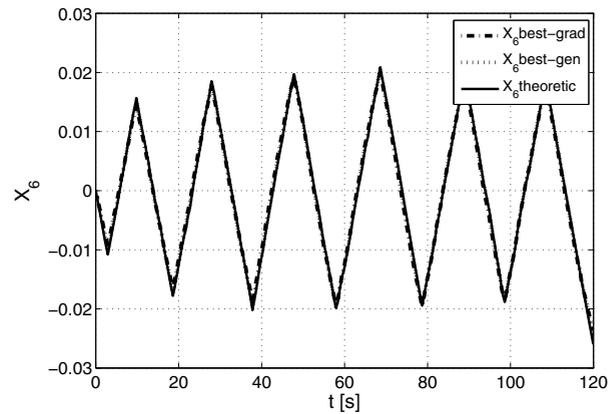


Fig. 7. Time-dependent behavior of the sixth component of the state vector.

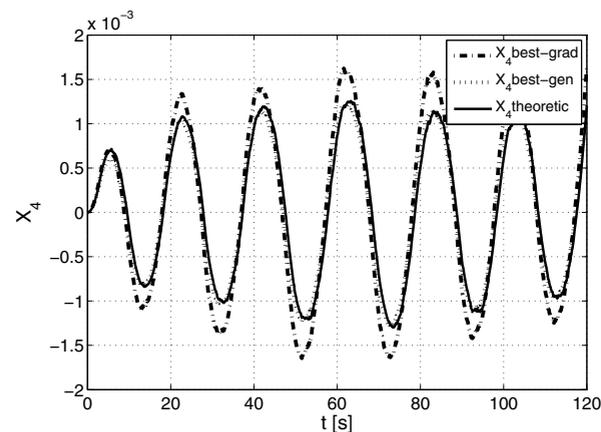


Fig. 5. Time-dependent behavior of the fourth component of the state vector.

the genetic algorithm is “more” fitted to the value of the nominal operating point (only one element differs in

sign), but the value of the criterion for this individual is greater than that of the best solution produced by the gradient-based algorithm. In both the cases, the criterion values are similar and are less than the value calculated for parameters in the nominal operating point of the ship.

The presented results also confirm the possibility of a practical implementation of the algorithm. The computations were conducted on a PC-class computer with an i3 processor, 4 GB RAM. The average time of computation was around 5 minutes inclusive of plotting each point of the chart flow on the screen.

In addition, the analysis of changes experienced by the value of the criterion  $J'_N$  in the process of minimization (Fig. 8) justifies the conclusion that it is possible to further reduce the number of iterations/generations without a significant deterioration in the value of the criterion by which the average computation times will be reduced.

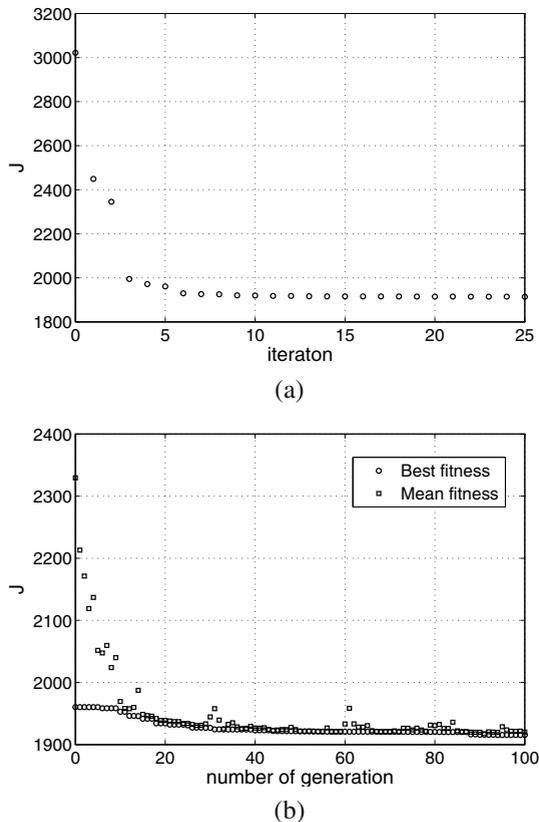


Fig. 8. Value of the criterion  $J'_N$  after each iteration/generation delivered by the respective algorithm: gradient-based (a), genetic (b).

## 6. Concluding remarks

On the basis of the conducted calculations and the results obtained it may be concluded that identification of parameters of linear models of the ship “Wimpey Sealab” proved to be a difficult optimization problem that requires lengthy and sensitive numerical calculations to be carried out. Both the methods tested produced correct, comparable results.

However, if we take into account the complexity of calculations, the encountered numerical problems and the obtained, in consequence, computation times, it seems that the genetic algorithm is a better choice here. Regardless of how the initial individuals are selected, the criterion achieves its acceptable value after about 50 generations, and the obtained accuracy and computation times allow the use of the algorithm in practice.

Finally, taking into account all pros and cons of both the methods, it seems to us that the best solution would be a combination—the genetic algorithm for starting the optimization procedure and the gradient one for finding the best solution in the designated area. Such an algorithm may have a chance to be successfully used in practice.

## Acknowledgment

This article was inspired by the late Professor Stanisław Bańka and prepared with the use of jointly developed material and results of experiments. Its completion and final editing is a form of gratitude to Professor Bańka for the long-term, fruitful cooperation.

## References

- Al Seyab, R. and Cao, Y. (2008). Nonlinear system identification for predictive control using continuous time recurrent neural networks and automatic differentiation, *Journal of Process Control* **18**(6): 568–581.
- Anil Ari Guner, H., Yuksel, Y. and Cevik, E.O. (2013). Estimation of wave parameters based on nearshore wind-wave correlations, *Ocean Engineering* **63**: 52–62.
- Aström, K. and Källström, C. (1976). Identification of ship steering dynamics, *Automatica* **12**(1): 9–22.
- Aström, K. and Källström, C. (1981). Experiences of system identification applied to ship steering, *Automatica* **17**(1): 187–198.
- Bańka, S. (2007). *Multivariable Control Systems: A Polynomial Approach*, Monographs of the Committee of Automation and Robotics, Polish Academy of Sciences, Szczecin University of Technology Press, Szczecin, (in Polish).
- Bańka, S., Dworak, P., Brasel, M. and Latawiec, K.J. (2010). Switched structure of linear MIMO controllers for positioning of a drillship on a sea surface, *Proceedings of the 15th International Conference on Methods and Models in Automation and Robotics, MMAR 2010, Międzyzdroje, Poland*, pp. 249–254.
- Bańka, S., Dworak, P. and Jaroszewski, K. (2013). Linear adaptive structure for control of a nonlinear MIMO dynamic plant, *International Journal of Applied Mathematics and Computer Science* **23**(1): 47–63, DOI: 10.2478/amcs-2013-0005.
- Barth, S. and Eecen, P. (2006). Description of the relation of wind, wave and current characteristics at the offshore wind farm Egmond aan Zee, *Technical report*, NoordzeeWind, IJmuiden.
- Billings, S. (2013). *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-temporal Domains*, Wiley, Chichester.
- Bredmose, H., Larsen, S., Matha, D., Rettenmeier, A., Marino, E. and Seattran, L. (2012). Collation of offshore wind-wave dynamics, *Technical report*, Marine Renewables Infrastructure Network.
- Chen, W., Yan, K., Shi, G. and Lu, H. (2008). A new swarm intelligence and parameter identification of hydrodynamic forces acting on axis-symmetric body exiting water obliquely, *Journal of Ship Mechanics* **12**(2): 204–210.
- Dai, Y., Li, Y. and Song, J. (2012). Parameter identification of ship lateral motions using evolution particle swarm optimization, *Proceedings of the 5th International Joint Conference on Computational Sciences and Optimization, CSO 2012, Harbin, China*, pp. 797–801.

- Dai, Y., Zhao, X. and Liu, L. (2010). Parameter identification of ship longitudinal motions using continuous ant colony algorithm with period searching, *Journal of Ship Mechanics* **14**(8): 872–878.
- Dworak, P. and Brasel, M. (2013). Improving quality of regulation of a nonlinear MIMO dynamic plant, *Elektronika Ir Elektrotehnika* **19**(7): 3–6.
- Fossen, T.I. (1994). *Guidance and Control of Ocean Vehicles*, John Wiley and Sons, Chichester.
- Fossen, T.I. (2011). *Handbook of Marine Craft Hydrodynamics and Motion Control*, John Wiley and Sons, Chichester.
- Herrero, E. and González, F.V. (2012). Two-step identification of non-linear manoeuvring models of marine vessels, *Ocean Engineering* **53**: 72–82.
- Kondo, J., Fujinawa, Y. and Naito, G. (1972). Wave-induced wind fluctuation over the sea, *Journal of Fluid Mechanics* **51**(4): 751–771.
- Lewis, E.V. (1989). *Principles of Naval Architecture, Volume III: Motions in Waves and Controllability*, Society of Naval Architects and Marine Engineers, Alexandria, VA.
- Ljung, L. (2001). *System Identification. Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ.
- Mahfouz, A. and Haddara, M. (2003). Effect of the damping and excitation on the identification of the hydrodynamic parameters for an underwater robotic vehicle, *Ocean Engineering* **30**(8): 1005–1025.
- Mzyk, G. (2013). Nonparametric instrumental variables for identification of block-oriented systems, *International Journal of Applied Mathematics and Computer Science* **23**(3): 521–537, DOI: 10.2478/amcs-2013-0040.
- Nelles, O. (2001). *Nonlinear System Identification*, Springer-Verlag, Berlin/Heidelberg.
- Orjuela, R., Marx, B., Ragot, J. and Maquin, D. (2013). Nonlinear system identification using heterogeneous multiple models, *International Journal of Applied Mathematics and Computer Science* **23**(1): 103–115, DOI: 10.2478/amcs-2013-0009.
- Padilla, A. and Yuz, J. (2013). Continuous-time system identification of a ship on a river, *Proceedings of the 52nd Annual Conference on Decision and Control, CDC 2013, Firenze, Italy*, pp. 4553–4558.
- Revestido, E., Velasco, F., Zamanillo, I., Lopez, E. and Moyano, E. (2011). Parameter estimation of ship linear maneuvering models, *Proceedings of IEEE OCEANS 2011, Santander, Spain*, pp. 1–8.
- Skjetne, R., Smogeli, Ø. and Fossen, T. (2004). A nonlinear ship manoeuvring model: Identification and adaptive control with experiments for a model ship, *Modeling, Identification and Control* **25**(1): 3–27.
- Tran Khanh, T., Ouahsine, A., Naceur, H. and Wassifi, K.E. (2013). Assessment of ship manoeuvrability by using a coupling between a nonlinear transient manoeuvring model and mathematical programming techniques, *Journal of Hydrodynamics B* **25**(5): 788–804.
- Velasco, F., Herrero, E., Lopez, L. and Moyano, E. (2013). Identification for a heading autopilot of an autonomous in-scale fast ferry, *IEEE Journal of Oceanic Engineering* **38**(2): 263–274.
- Wang, Y., Meng, H., Liu, Z.-L. and Yu, R.-T. (2011). On-line of modeling ship manoeuvring motion by using least squares support vector machines, *Proceedings of the IEEE 18th International Conference on Industrial Engineering and Engineering Management, Changchun, China*, pp. 1712–1716.
- Wise, D.A. and English, J.W. (1975). Tank and wind tunnel tests for a drill-ship with dynamic position control, *Offshore Technology Conference, Dallas, TX, USA*, pp. 103–118.
- Yoon, H. and Rhee, K. (2003). Identification of hydrodynamic coefficients in ship maneuvering equations of motion by estimation-before-modeling technique, *Ocean Engineering* **30**(18): 2379–2404.
- Yoon, H., Son, N.S. and Lee, C.M. (2004). Estimation of roll related hydrodynamic coefficients through the free running model tests, *Proceedings of IEEE TECHNO-OCEAN'04, Kobe, Japan*, pp. 1086–1092.
- Yoon, H., Son, N. and Lee, G. (2007). Estimation of the roll hydrodynamic moment model of a ship by using the system identification method and the free running model test, *IEEE Journal of Oceanic Engineering* **32**(4): 798–806.
- Zhang, H.-G. and Zou, Z.-J. (2011). Identification of abkowitz model for ship manoeuvring motion using  $\varepsilon$ -support vector regression, *Journal of Hydrodynamics B* **23**(3): 353–360.
- Zhang, Z., Li, X., Yang, L. and Wu, G.Y. (2010). Parameter identification method for ship swaying motion differential equations, *Proceedings of the 3rd International Joint Conference on Computational Science and Optimization, CSO 2010, Huangshan, Anhui, China*, pp. 205–208.



**Stanisław Bańka** received the M.Sc., Ph.D. and D.Sc. degrees in control engineering in 1969, 1976 and 1992, respectively. In the years 1996–2002 he was the dean of the Electrical Engineering Faculty at the Technical University of Szczecin and in 2003–2005 the director of the Institute of Control Engineering of that university. From 2009 a professor, then the head of the Department of Control Engineering and Robotics at the West Pomeranian University of Technology at Szczecin. A member of the Committee on Automatic and Robotics of the Polish Academy of Sciences (from 2003). An author of over 80 journal papers and three books. Main scientific interests: multivariable control systems, adaptive and optimal control, estimation and parameter identification, applications of informatics in automation. Professor Bańka passed away in 2014.



**Michał Brasel** received the M.Sc. degree in control engineering in 2009. Currently he is a Ph.D. student at the Department of Control Engineering and Robotics, West Pomeranian University of Technology in Szczecin. His main scientific interests concentrate on multivariable control systems, nonlinear and optimal control, and its applications in mechanical and electrical engineering.



**Krzysztof Jaroszewski** holds the M.Sc. (2001) and Ph.D. (2007) degrees, both in control engineering. Currently an assistant professor at the Department of Control Engineering and Robotics, West Pomeranian University of Technology in Szczecin. Main scientific interests: artificial intelligence, especially neural networks, control systems and visualization, industrial diagnostics.



**Paweł Dworak** received the M.Sc. and Ph.D degrees in control engineering in 1999 and 2005, respectively. Currently an assistant professor at the Department of Control Engineering and Robotics, West Pomeranian University of Technology in Szczecin. Main scientific interests: multivariable control systems, adaptive control, process data acquisition and visualization, industrial applications of modern control algorithms.

Received: 6 October 2014

Revised: 28 April 2015

Re-revised: 7 September 2015