PIOTR LICHOTA [*], *MACIEJ LASEK* [*]

MAXIMUM LIKELIHOOD ESTIMATION FOR IDENTIFICATION OF AIRCRAFT AERODYNAMIC DERIVATIVES

This article investigates identification of aircraft aerodynamic derivatives. The identification is performed on the basis of the parameters stored by Flight Data Recorder. The problem is solved in time domain by Quad-M Method. Aircraft dynamics is described by a parametric model that is defined in Body-Fixed-Coordinate System. Identification of the aerodynamic derivatives is obtained by Maximum Likelihood Estimation. For finding cost function minimum, Lavenberg-Marquardt Algorithm is used. Additional effects due to process noise are included in the state-space representation. The impact of initial values on the solution is discussed. The presented method was implemented in Matlab R2009b environment.

1. Introduction

The development of civil aviation is associated with permanent problem-solving in safety and economy. Determining changes in aerodynamic characteristics is one of the areas that could be used for both improving safety and lowering costs.

At present, there is a lack of effective methods allowing for determining changes in aerodynamic derivatives due to e.g. maintenance. In order to find out the values of various aerodynamic derivatives, an identification experiment could be performed. Such an approach is possible, however, because of valid regulations, is highly unprofitable. Identification experiment costs would be higher than possible profits (e.g. due to maintenance program changes). Therefore, in civil aircraft, this approach is used only for flight simulators – due to certification regulations.

A possible solution is to identify the aerodynamic characteristics on the basis of parameters that are registered during each scheduled flight by flight

---

[*] *The Institute of Aeronautics and Applied Mechanics, Warsaw University of Technology, Warsaw, Poland; E-mail: plichota@meil.pw.edu.pl*

data recorders (FDR). The use of FDR is obligatory, so there would be no need for an additional certification process in contrast to the approach which requires specially mounted sensors that register data during scheduled flight.

A separate issue is the safety level. Identification of aerodynamic derivatives on the basis of FDR recordings concerns a specific aircraft in its particular condition. Therefore, it provides information about system dynamics that can be used for investigating aircraft accidents in order to raise the safety level of civil aviation.

Summarizing, the use of FDR for the purpose of aircraft aerodynamic characteristics identification may lead to benefits in safety and economy.

2. Quad-M method and basic assumptions

When dealing with dynamic objects, one often uses the Quad-M method. In this method, one may distinguish four main stages, which for aircraft aerodynamics identification on the basis of FDR recordings are [4], [7]:
▶ Manoeuvre – scheduled flight
▶ Measurement – FDR recording
▶ Model – rigid body model
▶ Method – identification method
Except of this stages, one should consider an additional one – validation.

In the analyzed case, first two stages are determined by flight plan and the installed FDR capability, and therefore they can be treated as user-independent. However, the following assumptions are made:
▶ exogenous input signal
▶ statistically independent measurements errors
▶ no process noise
▶ input signal provides full description of system dynamics
The form of the model depends on the considered problem, and for identification of aerodynamic characteristics it can be presented in the following way:
▶ rigid body with 6 DOF and a plane of symmetry
▶ constant mass and uniform mass distribution
▶ constant gravity
▶ limited deflections of flight control surfaces and deflection rates
▶ flight control through deflections of flight control surfaces and engine thrust
▶ linear influence of motion parameters changes on aerodynamic forces and moments
▶ quasi-stationary flow
▶ flat earth approximation

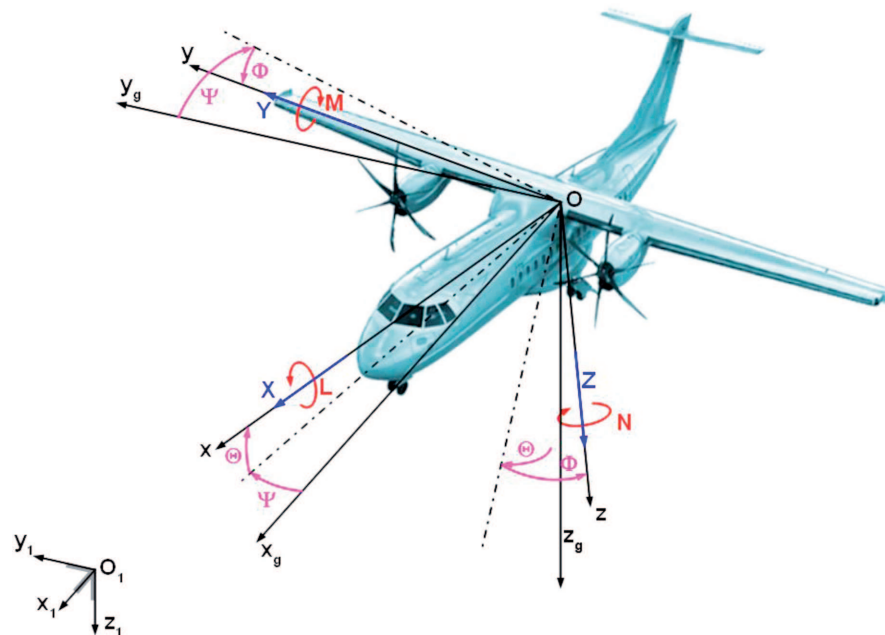Usually, the model is described in the Body-Fixed-Coordinate system Oxyz [2].

Fig. 1. Coordinate systems

Oxyz coordinate system is right-handed and orthogonal. Its origin is fixed to an arbitrary point of the aircraft, but usually it is located at centre of gravity of an aircraft (CG). Ox-axis lies in symmetry plane (xz), is parallel to the mean aerodynamic chord and directed forward. Oy-axis is perpendicular to the symmetry plane and directed towards right wing. Oz-axis lies in symmetry plane and is directed downward.

Fig. 1. shows dependencies between Oxyz coordinate system and $Ox_g y_g z_g$ frame, which is carried by the aircraft, but is parallel to non-rotating rectangular earth-fixed frame ($O_1 x_1 y_1 z_1$).

3. Mathematical model

Dynamic systems can be represented in state-space and described by: state, output and measurement equations [3]. These are: vector-valued-functions of state **x** (linear and angular velocities components), input **u** (control surfaces deflections), output **y** (linear and angular velocities and accelerations components), system parameters **Θ** (dimensional derivatives), measurement noise **v** (corresponding to output vector) and time t.

State equation:

$$\dot{\mathbf{x}}=f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{\Theta}) \tag{1}$$

Output Equation:

$$\mathbf{y} =g(\mathbf{x}(t), \mathbf{u}(t), \mathbf{\Theta}) \tag{2}$$

Measurement equation in particular time point t_k:

$$\mathbf{z}(t_k)= \mathbf{y}(t_k) + \mathbf{v}(t_k) \tag{3}$$

The state equation is based on Newton's Second Law of Motion. The forces acting on aircraft in flight can be defined as longitudinal X, lateral Y, and vertical Z force. Moments about those axis are rolling moment L, pitching moment M and yawing moment N, respectively.

For Oxyz coordinate system, whose origin is located at CG, one can write momentum $\mathbf{\Pi}$ and angular momentum $\mathbf{K_O}$ equations as:

$$\begin{aligned}
\frac{\tilde{\delta}\mathbf{\Pi}}{dt}+\mathbf{\Omega} \times \mathbf{\Pi} &= \mathbf{F} \\
\frac{\tilde{\delta}\mathbf{K_O}}{dt}+\mathbf{\Omega}\times\mathbf{K_O} &=\mathbf{M_O}
\end{aligned} \tag{4}$$

where:

$\mathbf{F} =[\ X\ \ Y\ \ Z\ ]^T$ – force vector,
$\mathbf{M_O}=\left[\ L\ \ M\ \ N\ \right]^T$ – moment vector,
$\mathbf{\Omega} =[\ P\ \ Q\ \ R\ ]^T$ – angular momentum,
$\tilde{\delta}$– local derivative.

Aerodynamic derivatives are defined with respect to linear velocity $\mathbf{V}= [\ U\ \ V\ \ W\ ]^T$, angular momentum $\mathbf{\Omega}$ and deflections of flight surfaces δ.

The subscripts in flight control deflections stand for: ailerons A, elevator H, rudder V and thrust Th. Dots above letters denote derivatives with respect to time.

Aerodynamic derivatives with respect to j- parameter are defined as follows:

$$X_j = \frac{1}{m}\frac{\partial X}{\partial j}$$
$$Y_j = \frac{1}{m}\frac{\partial Y}{\partial j}$$
$$Z_j = \frac{1}{m}\frac{\partial Z}{\partial j}$$
$$L_j = \frac{1}{I_{xx}}\frac{\partial L}{\partial j} \tag{5}$$
$$M_j = \frac{1}{I_{yy}}\frac{\partial M}{\partial j}$$
$$N_j = \frac{1}{I_{zz}}\frac{\partial N}{\partial j}$$

where I_{kl} denotes the moment of inertia around the k-axis when the objects are rotated around the l-axis and m stands for mass.

Applying kinematic equations and using small disturbances theory, one obtains the equations of motion. The trim condition is straight symmetric flight with wing levels. 0 denotes the trimmed state.

The equations of motion can be reduced to simultaneous homogeneous differential equation by applying primed derivatives and ignoring second order effects.

$$L'_j = L_j + \frac{I_{xz}}{I_{zz}}N_j$$
$$N'_j = \frac{I_{xz}}{I_{xx}}L_j + N_j \tag{6}$$

To simplify state equation the following transformation can be performed:

$$\tilde{M}_U = M_U + M_{\dot{W}}Z_U$$

$$\tilde{M}_W = M_W + M_{\dot{W}}Z_W$$

$$\tilde{M}_Q = M_Q + M_{\dot{W}}(Z_Q + U_0) \tag{7}$$

$$\tilde{M}_{\delta H} = M_{\delta H} + M_{\dot{W}}Z_{\delta H}$$

$$\tilde{M}_{\delta Th} = M_{\delta Th} + M_{\dot{W}}Z_{\delta Th}$$

Defining state $\mathbf{x}$ as a vector of linear velocity $\mathbf{V}$ and angular momentum $\mathbf{\Omega}$ components increments, pitch $\mathbf{\Theta}$ and roll $\mathbf{\Phi}$ angles increments, and ap-

plying process noise (despite earlier assumption), one obtains the following dependencies:

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} + \mathbf{w} \tag{8}$$

where:
State vector:

$$\mathbf{x} = [\ \Delta U \quad \Delta V \quad \Delta W \quad \Delta P \quad \Delta Q \quad \Delta R \quad \Delta\Theta \quad \Delta\Phi \ ]^{\mathrm{T}} \tag{9}$$

Input vector:

$$\mathbf{u} = [\ \Delta\delta_{\mathrm{H}} \quad \Delta\delta_{\mathrm{V}} \quad \Delta\delta_{\mathrm{A}} \quad \Delta\delta_{\mathrm{Th}} \ ]^{\mathrm{T}} \tag{10}$$

Process noise:

$$\mathbf{w} = [\ w_{\dot{U}} \quad w_{\mathrm{U}} \quad w_{\dot{V}} \quad w_{\mathrm{V}} \quad w_{\dot{W}} \quad w_{\mathrm{W}} \quad w_{\dot{P}} \quad w_{\mathrm{P}} \quad \dots \ ]^{\mathrm{T}} \tag{11}$$

State matrix:

$$\mathbf{A} = \begin{bmatrix} X_{\mathrm{U}} & 0 & X_{\mathrm{W}} & 0 & X_{\mathrm{Q}} & 0 & -g & 0 \\ 0 & Y_{\mathrm{V}} & 0 & Y_{\mathrm{P}} & 0 & Y_{\mathrm{R}}+U_0 & 0 & g \\ Z_{\mathrm{U}} & 0 & Z_{\mathrm{W}} & 0 & Z_{\mathrm{Q}}+U_0 & 0 & 0 & 0 \\ 0 & L'_{\mathrm{V}} & 0 & L'_{\mathrm{P}} & 0 & L'_{\mathrm{R}} & 0 & 0 \\ \tilde{M}_{\mathrm{U}} & 0 & \tilde{M}_{\mathrm{W}} & 0 & \tilde{M}_{\mathrm{Q}} & 0 & 0 & 0 \\ 0 & N'_{\mathrm{V}} & 0 & N'_{\mathrm{P}} & 0 & N'_{\mathrm{R}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{12}$$

Input matrix:

$$\mathbf{B} = \begin{bmatrix} X_{\delta\mathrm{H}} & 0 & Z_{\delta\mathrm{H}} & 0 & \tilde{M}_{\delta\mathrm{H}} & 0 & 0 & 0 \\ 0 & Y_{\delta\mathrm{V}} & 0 & L'_{\delta\mathrm{V}} & 0 & N'_{\delta\mathrm{V}} & 0 & 0 \\ 0 & Y_{\delta\mathrm{A}} & 0 & L'_{\delta\mathrm{A}} & 0 & N'_{\delta\mathrm{A}} & 0 & 0 \\ X_{\delta\mathrm{Th}} & 0 & Z_{\delta\mathrm{Th}} & 0 & \tilde{M}_{\delta\mathrm{Th}} & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \tag{13}$$

The components of process noise (11) correspond to the terms in right hand side of output equation (2) i.e. to linear and angular components of velocities.

4. Identification

Model identification is the third step in the Quad-M Method. This term refers to parametric identification, i.e. means determining aircraft mathematical model parameters for which the best fits between the model and
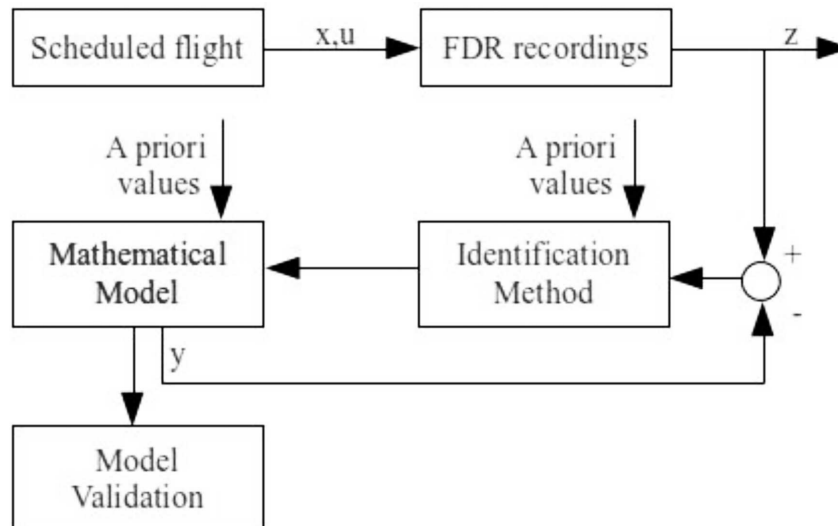
Fig. 2. Identification scheme

the real object is achieved. The identification is based on input and output measurements and makes it possible to find the state vector function. Fig. 2. presents general scheme for aircraft identification on the basis of FDR.

The identification can be carried out in both frequency and time domain. Due to physical nature of the process, we selected time domain identification. The methods presented below concern analysis in the this domain.

One can distinguish four main classes of identification methods
▶ Equation Error Methods (EEM)
▶ Filter Error Methods (FEM)
▶ Output Error Methods (OEM)
▶ Artificial Neural Networks (ANN)

Basic differences between particular methods are as follows: Equation Error Methods (e.g. Least Squares Method) minimize quadratic error in state equation, Filter Error Methods (e.g. Kalman) minimize filter error, whereas Output Error Method minimize the difference between the model response and measurements (e.g. Maximum Likelihood Estimation). Due to a completely different approach, Artificial Neural Networks are a separate class of identification methods.

Each identification method can be used in one of two modes: online (parameters are used for identification during data gathering) or offline (parameters are used after data gathering process).

For the presented aerodynamic derivatives identification, there is no need to process the data from FDR in online mode. Furthermore – the model ob-

tained from offline identification is more accurate than that identified online. Therefore, identification in offline mode was selected.

5. Maximum Likelihood Estimation

For aircraft identification, one can use the Maximum Likelihood Estimation [5]. This method belongs to OEM group, and consist in determining such a set of identified parameters $\hat{\Theta}$ (dimensional derivatives) that maximizes log-likelihood $L(\mathbf{z}|\Theta, \mathbf{R})$ for measurements $\mathbf{z}$ when Θ, and the covariance matrix $\mathbf{R}$.

$$\hat{\Theta} = \arg\{L(\mathbf{z}|\Theta, \mathbf{R})\} \tag{14}$$

Assumptions for the method are as follows [6]:
► the model is identifiable
► the support of probability density function does not depend on parameters
► random variables are independent and identically distributed
► the set of parameters contains an open neighbourhood in which
   ► log-likelihood function is twice continuously differentiable with respect to the identified parameters
   ► third derivatives of log-likelihood exist and are absolutely bounded
► the Fisher Information Matrix is a negative defined one.
The idea of Maximum Likelihood Estimation is based on searching such parameters' estimates that maximize the conditional probability density function $p(\mathbf{z}|\Theta, \mathbf{R})$. The measurements are assumed to be normally distributed, however different multivariate distribution is also acceptable.

Referring to identification: for non-degenerate case and N particular discrete time points, conditional probability density function for n-dimensional Multivariate Normal Distribution is given in the following form:

$$p(\mathbf{z}(t_k)|\Theta, \mathbf{R}) = \frac{1}{((2\pi)^n |\mathbf{R}|)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}[\mathbf{z}(t_k) - \mathbf{y}(t_k)]^T \mathbf{R}^{-1}[\mathbf{z}(t_k) - \mathbf{y}(t_k)]\right) \tag{15}$$

Due to independence of variables, there is:

$$p(\mathbf{z}|\Theta, \mathbf{R}) = \frac{1}{((2\pi)^n |\mathbf{R}|)^{\frac{N}{2}}} \exp\left(-\frac{1}{2}\sum_{k=1}^{N}[\mathbf{z}(t_k) - \mathbf{y}(t_k)]^T \mathbf{R}^{-1}[\mathbf{z}(t_k) - \mathbf{y}(t_k)]\right) \tag{16}$$

The quest for probability density maximum function can be replaced by minimizing the negative log-likelihood function, which corresponds to (16).

$$L(\mathbf{z}|\Theta, \mathbf{R}) = \frac{1}{2}\sum_{k=1}^{N}[\mathbf{z}(t_k) - \mathbf{y}(t_k)]^T \mathbf{R}^{-1}[\mathbf{z}(t_k) - \mathbf{y}(t_k)] + \frac{N}{2}\ln|\mathbf{R}| + \frac{nN}{2}\ln(2\pi) \tag{17}$$

Maximum Likelihood Estimator of the covariance matrix from N discrete time points is obtained from:

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{k=1}^{N} \left[\mathbf{z}(t_k) - \mathbf{y}(t_k) \right] \left[\mathbf{z}(t_k) - \mathbf{y}(t_k) \right]^T \tag{18}$$

The above estimator makes it possible to simplify the cost function J due to constant value of some terms in the log-likelihood function.

$$J(\mathbf{\Theta}, \mathbf{R}) = |\mathbf{R}| \tag{19}$$

One can find minimum of the cost function given by (19) by applying any optimization algorithm.

6. Levenberg-Marquardt Algorithm

For finding minimum of the cost function, Levenberg-Marquardt Algorithm can be used. This method interpolate between the Gauss-Newton Algorithm and the Steepest Descent in order to exploit their best features.

The Gauss-Newton Algorithm uses linear approximation for finding cost function minimum. Linearization is done by using Taylor series expansion:

$$(\mathbf{\nabla}_{\mathbf{\Theta}} J)_i \approx (\mathbf{\nabla}_{\mathbf{\Theta}} J)_{i-1} + \left(\mathbf{\nabla}_{\mathbf{\Theta}}^2 J \right)_{i-1} \Delta \mathbf{\Theta} \tag{20}$$

where (neglecting the higher-order terms):

Gradient vector

$$\mathbf{\nabla}_{\mathbf{\Theta}} J = - \sum_{k=1}^{N} \left[\frac{\partial \mathbf{y}(t_k)}{\partial \mathbf{\Theta}} \right]^T \mathbf{R}^{-1} \left[\mathbf{z}(t_k) - \mathbf{y}(t_k) \right] \tag{21}$$

Fisher Information Matrix

$$\mathbf{\nabla}_{\mathbf{\Theta}}^2 J = - \sum_{k=1}^{N} \left[\frac{\partial \mathbf{y}(t_k)}{\partial \mathbf{\Theta}} \right]^T \mathbf{R}^{-1} \frac{\partial \mathbf{y}(t_k)}{\partial \mathbf{\Theta}} \tag{22}$$

Therefore, the parameter update can be calculated from expression:

$$\mathbf{\Theta}_i = \mathbf{\Theta}_{i-1} \left(\mathbf{\nabla}_{\mathbf{\Theta}}^2 J \right)_{i-1}^{-1} (\mathbf{\nabla}_{\mathbf{\Theta}} J)_{i-1} \tag{23}$$

The Steepest Descent method uses the negative gradient direction calculated in successive steps and the step-size parameter in order to find minimum of the cost function. In this method, the parameter update is described by the following formula:

$$\mathbf{\Theta}_i = \mathbf{\Theta}_{i-1} \lambda (\mathbf{\nabla}_{\mathbf{\Theta}} J)_{i-1} \tag{24}$$

Combining (23) and (24) leads to:

$$\boldsymbol{\Theta}_i = \boldsymbol{\Theta}_{i-1} \left(\boldsymbol{\nabla}^2_{\boldsymbol{\Theta}} J + \lambda \mathbf{I} \right)^{-1}_{i-1} (\boldsymbol{\nabla}_{\boldsymbol{\Theta}} J)_{i-1} \tag{25}$$

Step-size λ determines the impact of the Gauss-Newton Algorithm and the Steepest Descent on the parameter update: when it tends to 0, the Gauss-Newton Algorithm is dominant, when to infinity the Steepest Descent is prevailing.

When Levenberg-Marquardt Algorithm is used for finding cost function minimum, an additional variable is defined – the reduction factor ν. In each iteration, the cost function is calculated for an unchanged value of step-size parameter λ and for the step-size parameter reduced by ν. Comparison of cost functions (in successive steps) makes it possible to specify whether the step-size parameter should be decreased, increased or remains unchanged.

The scheme of the procedure is as follows. When the cost function calculated for a reduced λ is smaller than the cost function from previous iteration, λ is reduced (the influence of Gauss-Newton Algorithm is increased). If the above condition is not satisfied, but the cost function evaluated for an unchanged λ is smaller than the cost function from previous iteration, λ remains unchanged. If one of these conditions is satisfied, the parameters are updated according to (25). Otherwise λ is increased (higher impact of Steepest Descent) and the cost functions are compared again.

7. Results

Presented method was implemented in MATLAB environment. The program was tested with the data obtained from [1].

Fig. 3. and Fig. 4. present verification of the identified parameters: red lines denote flight parameters obtained through identification, whereas blue lines denote measured data. For lateral-directional identification, ailerons deflection, rudder deflection and side slip angle was also measured, but there was no need for identification of these parameters (therefore are not presented).

The presented graphs refer to two cases, i.e. initial values of aerodynamic derivatives far from or near to the solution.

8. Conclusions

Maximum Likelihood Estimation with Levenberg-Marquardt Algorithm can be used for identification of aircraft aerodynamic characteristics. Therefore, it can be used for obtaining aerodynamic derivatives on the basis of flight data recorders.
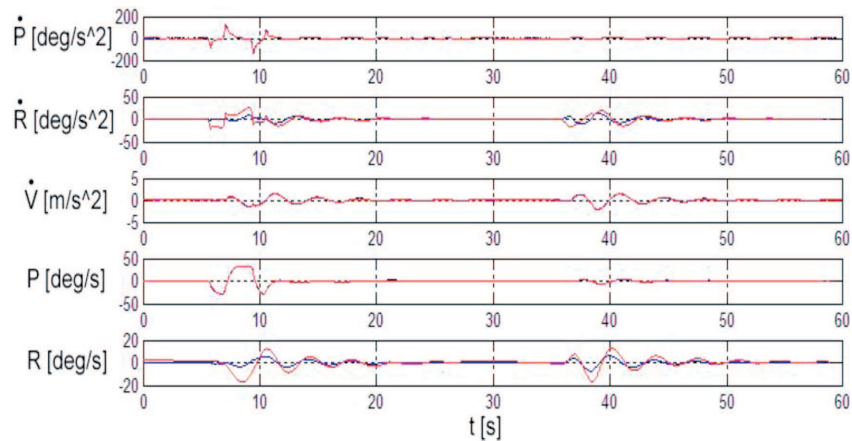
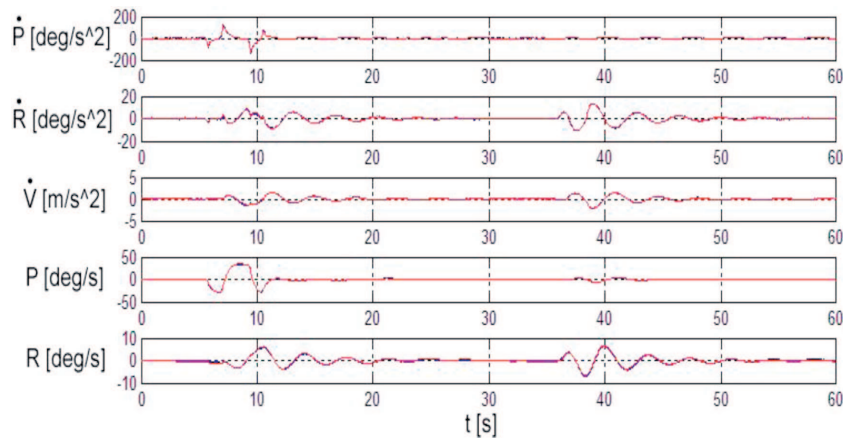Fig. 3. Initial values far from the optimum

Fig. 4. Initial values near the optimum

However, special attention must be paid to defining initial values used for identification. When these values significantly differ from the optimum, identification error, i.e. the difference between identified and measured data is larger and identification is slower than in the near-optimum case. For this reason, additional studies are necessary – to determine initial values of model's parameters.

All drawings in this article are by Piotr Lichota.

PIOTR LICHOTA, MACIEJ LASEK

REFERENCES

[1] Jategaonkar R.: Flight Vehicle System Identification: A Time Domain Methodology, Reston, Progress in Astronautics and Aeronautics, 2006.

[2] Manerowski J.: Identyfikacja modeli ruchu sterowalnych obiektów latających, Warsaw, Askon, 2003.

[3] Stevens B. L., Lewis F. L.: Aircraft Control and Simulation, 2nd ed, New York, Wiley-Interscience, 2003.

[4] Goszczyński J.: Practical aspects of the identification of the aerodynamic characteristics, Journal of Theoretical and Applied Mechanics, 2006, Vol. 44, No. 1, pp. 31-50.

[5] Iliff K. W.: Estimation of Aerodynamic Characteristics from Dynamic Flight Test Data. Stability Parameters, 1978, AGARD-CP-235, Paper 15.

[6] Iliff, K. W., Maine, R. E.: Practical aspects of using a maximum likelihood estimator, 1975, AGARD Methods for Aircraft State and Parameter Identification.

[7] Jategaonkar R., Fischenberg D., von Gruenhagen W.: Aerodynamic Modeling and System Identification from Flight Data – recent Applications at DLR.

[8] Morelli E.A.: Flight Test Maneuvers for Efficient Aerodynamic Modeling, 2011, AIAA Atmospheric Flight Mechanics Conference.

[9] Ratnayake N. A., Waggoner E. R., Taylor B. R.: Lateral-Directional Parameter Estimation on the X-48B Aircraft Using an Abstracted, Multi-Objective Effector Model, 2011, AIAA Applied Aerodynamics Conference.

Identyfikacja pochodnych aerodynamicznych Metodą Największej Wiarygodności

S t r e s z c z e n i e

Artykuł zawiera informacje na temat identyfikacji pochodnych aerodynamicznych. Estymacja opiera się o parametry zapisywane przez Pokładowy Rejestrator Lotu. Zagadnienie jest rozważane w dziedzinie czasu przy użyciu podejścia Quad-M. Do opisu dynamiki samolotu wykorzystano model parametryczny zdefiniowany w układzie sztywno związanym z samolotem. Do identyfikacji wykorzystano Metodę Największej Wiarygodności. Do znalezienia minimum funkcji celu użyto algorytm Levenberga-Marquardta. W modelu uwzględniono wpływ dodatkowych czynników reprezentowany przez szum przetwarzania. Omówiono wpływ wartości początkowych na rozwiązanie. Prezentowane wyniki uzyskano w środowisku Matlab R2009b.