Fast and Robust Online Dynamic System Identification

Andrzej Latocha(✉)

Faculty of Electrical Engineering, Automatics,
Computer Science and Biomedical Engineering,
AGH University of Science and Technology, Krakow, Poland
alatocha@agh.edu.pl

Abstract. A new method is proposed for black-box linear model identification of a dynamic system embedded at a nearly Gaussian noise. The Gaussian process can highlight areas of the output spaces where the prediction quality is poor, due to the lack of data or its complexity, by indicating the higher variance of the predicted mean; the input spaces in which we can reconstruct data represent the expected values. This paper proposed a new approach for the online system identification for non-zero initial conditions in the moving window.

Keywords: Identification · Filtration · Estimation · Modeling · Control

1 Introduction

To solve many of the problems in the design, implementation, and operation of automatic control systems, relatively precise mathematical models for the static and dynamic behavior of industrial processes are required. If the underlying physical laws are not known or are only partially known, or if significant parameters are not known precisely enough, one has to perform an experimental modeling, which is called process or system identification. There are different ways to identify systems when the input and output of the system are known. In real systems, the signals are always more or less subject to interference with noise. The expected values can only be estimated. Preferably, identification algorithms should be fast algorithms that allow for the identification of dynamic systems around the operating point in real time. There are different methods that model different situations with respect to the noise. This study was based on identification methods using least squares estimation (LSE) [1,2] for the black-box model (Fig. 1). The work is of use in the field of control systems engineering. The innovation in the paper is a fast system for the high-precision identification of linear dynamics that is independent of the initial conditions.

J.M. Kościelny et al. (eds.), *Advanced Solutions in Diagnostics and Fault Tolerant Control*, Advances in Intelligent Systems and Computing 635,
DOI 10.1007/978-3-319-64474-5_18

2 Problem Formulation

Describing the process using equations with an acceptable error margin is very difficult due to the complexity of the system structure and the noise distortion. These data [1] have been obtained from an experiment. For identification, we assume that the structure of the black-box system being tested (Fig. 1) will be approximated by the parametric ARMAX model, which will have the same dynamic properties in terms of system input/output.

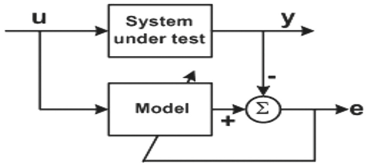

Fig. 1. Output error model

For this reason, we define a cost function that minimizes the error between the tested system and the approximating model (i.e., the model used to approximate the system dynamics). The test signal satisfies the assumptions [1] of zero initial conditions, and the system is embedded in nearly Gaussian noise [1–3]. The test signal u(k) is permanently changeable, and the sampling sequences are equal lengths of time [1]. For the above assumptions, we can solve problems fast with robust identification of linear dynamic systems.

3 Least Squares Estimator (LSE)

Assuming the parametric model (1),

$$y(t) = \varphi^T(t)\theta \tag{1}$$

where $y(t)$ is the measure of the output value, $\varphi(t)$ stands for the n-dimensional vector of the data samples, and the θ represents the n-dimensional vector of the unknown coefficient. For the measurement data, we can write the Eq. (2) as follows:

$$Y = \Phi^T\theta. \tag{2}$$

Due to account noise, and the inaccuracy of the model, it is better to use an overly large number of samples, as additional data improves the accuracy of the estimation. For $N \gg n$, the system is overdetermined, and there is no exact solution. For oversized samples, the data matrix will not be a square matrix. In this case, the samples matrix can be replaced by a pseudo square matrix. Taking into account the inaccuracy of samples (3) [1–3],

$$\varepsilon(k) = y(k) - \varphi^T(k)\theta,\ k \in \mathbb{N},\ k > 0 \tag{3}$$

The least squares error (LSE) estimator $\hat{\theta}$ is defined as a vector that minimizes the cost function (4):

$$V(\theta) = \frac{1}{2}\sum_{t=1}^{N}\varepsilon^2(k) = \frac{1}{2}\varepsilon^T\varepsilon = \frac{1}{2}||\varepsilon||^2 \tag{4}$$

where $||\,\mathbf{.}\,||$ is the Euclidean vector norm. For the positive definite matrix $\Phi^T\Phi$, the cost function (4) has a minimum:

$$minV(\theta) = V(\hat{\theta}) = \frac{1}{2}[Y^TY - Y^T\Phi(\Phi^T\Phi)^{-1}\Phi^TY], \tag{5}$$

$$E = Y - \Phi\theta, \tag{6}$$

$$0 = \frac{dV}{d\theta} = -Y^T\Phi + \theta^T(\Phi^T\Phi), \tag{7}$$

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^TY. \tag{8}$$

Equation (8) in the field of control systems engineering can be considered to represent good or bad numerical task conditioning (9) for computing:

$$(\Phi^T\Phi)(\Phi^T\Phi)^{-1} = \tilde{I} \approx I. \tag{9}$$

Perturbations outside the main diagonal show poor conditions for the numeric task. For significant perturbations outside the main diagonal obtained numerically, a pseudo-square matrix can be close to the losing row, although it is reversible. Equation (9) returns the predictive indices, probability of good identification results for LSE (8). If the matrix $\Phi^T\Phi$ is known as the Gramian matrix of Φ, which possesses several correct properties, such as being a positive semi-definite matrix, the matrix Φ^TY is known as the moment matrix. Finally, $\hat{\theta}$ is the coefficient vector of the least-squares hyperplane, expressed as (8). For this reason, we consider an equation in the field of discrete time on the moving window. Systems can be described by the autoregressive moving average model with exogenous inputs (ARMAX) (10):

$$y(k) = z^{-n}\frac{B(z^{-1})}{A(z^{-1})}u(k) + \frac{C(z^{-1})}{A(z^{-1})}\varepsilon(k) \tag{10}$$

The model does not require a preliminary asumption of the system stability, as shown in Sect. 4, as this is a contribution of the new identification algorithm. A necessary and sufficient condition to identify the system is satisfy the controllability condition in the sense limited input and output, bounded input generates a bounded signal as an output over limited time range (b.i.b.o.), and a sufficient amount of data. We assume the following:

$$\varepsilon = \frac{C(z^{-1})}{A(z^{-1})}\varepsilon(k) \tag{11}$$

where $y(k)$, $u(k)$, and $\varepsilon(k)$ are a series of discrete data equally distant in time. By describing the system using a difference equation, the following equation is obtained:

$$y(k) + a_1 y(k-1) + ... + a_n y(k-n) + \varepsilon = b_1 u(k-1) + ... + b_m u(k-m);$$
$$k \gg n,\ n \geq m;\ m, n \in \mathbb{N};\ u \in \mathbb{R};\ y \in \mathbb{R};\ \varepsilon \in \mathbb{R} \quad (12)$$

where the linearization error $\varepsilon = 0$, and $b_0, b_1, ..., b_m; a_0, a_1, ..., a_n$ are search coefficients. By applying the discrete Z-transform, the zero initial condition and $\varepsilon = 0$ are obtained as follows:

$$\hat{G}(z) = \frac{Y(z)}{U(z)} = \frac{\hat{b}_1 z^{m-1} + \hat{b}_2 z^{m-2} + ... + \hat{b}_{m-1} z + \hat{b}_m}{z^n + \hat{a}_1 z^{n-1} + \hat{a}_2 z^{n-2} + ... + \hat{a}_{n-1} z + \hat{a}_n}, \quad (13)$$

The discrete transfer function, from the definition of the discrete "z" operator, requires the assumption that the signal does not grow faster than the exponential function (14)

$$Z[f^*(t)] = Z[f(kT)] = F(z),\ F(z) = \sum_{k=-\infty}^{\infty} f(kT) z^{-k}$$
$$k \in \mathbb{N},\ T \in \mathbb{R},\ f(k) < k!,\ f(k) < e^{ak^2};\ a > 0,\ a \in \mathbb{R}. \quad (14)$$

4 Non-zero Initial Condition

A linear system without noise fulfills the principle of causality and can be identified by the LSE in any state. For the non-zero initial condition, we have a non-continuous function. The problem appears when the system is exposed to noise, because such a system does not fulfill the principle of causality. The goal is satisfy zero initial condition on $u(k)$ signal for Eq. (12), zero initial condition is being arbitrarily imposed with regard to the input signal, an output error is added to the noise. For this reason, the discontinuity on the input is modeled as a nonlinearity $f(.)$ (Fig. 2):

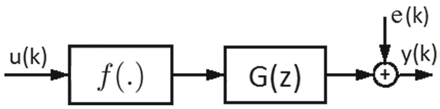

Fig. 2. The system structure of the identification model.

A nonlinear function $f(.)$ (Fig. 2) is estimated by the proposed delay function (15), which optimally carries out the $\tilde{u}(k)$ (16) signal from the zero initial condition to its actual state on range of data samples used to identification and has an insignificant impact on the dynamics of the system (18). The function is use to satisfy the zero initial condition on the input signal $\tilde{u}(k)$ for the computation algorithm (LSE) of error Eq. (24). A strong nonlinear function, [12,13], corrects the discrete input value in Eqs. (8) and (12), by imposing the zero initial condition for the optimal first initials of samples by delay time (15). The discrete output signal in Eqs. (8) and (12), are unchanged, and the original value is retained (as the modification of Eqs. (8) and (12) breaks the principle of causation through the delay time of the input signal (15).

$$h(z) = \frac{1}{z^{\eta}};\ \eta\epsilon\mathbb{N},\ \eta > 1 \tag{15}$$

A proposed function (15) is defined as the zero input initial reconstructor (ZIIR).

$$\tilde{u}(k) = Z^{-1}[h(z)U(z)];\ k\epsilon\mathbb{N};\ k > 0;\ \hat{u}\epsilon\mathbb{R} \tag{16}$$

where $u[1, ..., j)$ is arbitrary assuming the imposition of the zero initial condition:

$$u[1, ..., j) \equiv 0 \tag{17}$$

$$k = j,\ j + 1, ..., N;\ N\epsilon\mathbb{N};\ j\epsilon\mathbb{N}, \tag{18}$$

$$N - j \gg n, \tag{19}$$

$$y(k) = z^{-n}\frac{B(z^{-1})}{A(z^{-1})}\tilde{u}(k) + \frac{C(z^{-1})}{A(z^{-1})}\varepsilon(k), \tag{20}$$

$$\varepsilon = \frac{C(z^{-1})}{A(z^{-1})}\varepsilon(k), \tag{21}$$

where (21) includes an equation error.

$$y(k)+a_1y(k-1)+...+a_ny(k_i-n)+\varepsilon = b_1\tilde{u}(k-1)+...+b_m\tilde{u}(k-m);\ y\epsilon\mathbb{R},\ \varepsilon \in \mathbb{R} \tag{22}$$

$$\hat{\theta}_i = (\tilde{\Phi}_i^T\tilde{\Phi}_i)^{-1}\tilde{\Phi}_i^T Y_i. \tag{23}$$

By applying the discrete Z-transform, the following is obtained:

$$\hat{G}_i(z) = \frac{\hat{b}_1 z^{m-1} + \hat{b}_2 z^{m-2} + ... + \hat{b}_{m-1}z + \hat{b}_m}{z^n + \hat{a}_1 z^{n-1} + \hat{a}_2 z^{n-2} + ... + \hat{a}_{n-1}z + \hat{a}_n} \tag{24}$$

$$\hat{y}(k) = Z^{-1}[\hat{G}(z)\tilde{u}(z)] \tag{25}$$

The mean squar error (MSE) is based on the window (27):

$$e_i = \frac{1}{N-j}\sum_{k=0}^{N-j}(y_{j+k} - Ey_{j+k})^2 \tag{26}$$

$$\hat{e}_i = \frac{1}{N-j}\sum_{k=0}^{N-j}(y_{j+k} - E\hat{y}_{j+k})^2 \tag{27}$$

The optimality function (15) can be calculated as follows:

$$\eta = f(inf(e_i(1(t))), inf(e_i(\delta(t)))), \tag{28}$$

and the optimal identification we obtain for the minimum of error (29) is

$$\hat{G}(z) = \underset{\hat{G}_i(z)}{arg} \ inf(e_i). \tag{29}$$

5 Numerical Experiments

5.1 Example System Identification

A discrete example system is described by Eq. (30), where the sampling discretization step $\triangle t = 0.1[s]$.

$$G(z) = \frac{-0.3832z^2 - 0.2338z + 0.06683}{z^3 - 1.127z^2 + 0.494z - 0.1129} \tag{30}$$

Using the relationship (20) can identify the model of the system for different cases as demonstrated below.

System Without Noise. A plant that is not subject to noise is identified by the LSE in any state by a minimum number of samples (Fig. 3). Oversizing data in relation to the system dimensions is a result of numerical errors.

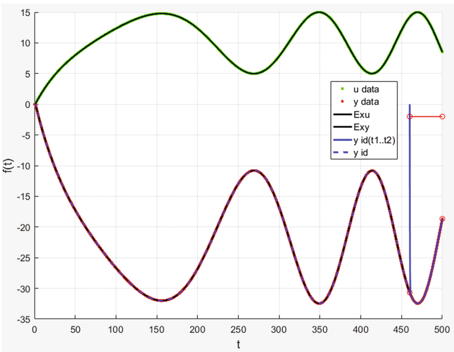

Fig. 3. Response-identified model without noise.

System with Unit Distorted Input. If unit noise is introduced into the system on the input (e.g. $u(N-5)=0$, for $k>N-5$ samples), the principle of causality for the dynamic system will not be satisfied. The results of such a disturbance are presented in Table 1; the quality of these results depends on the number of data gathered before disturbance. Here: the $e_{1(t)}$ is the MSE (26), of the response-identified model for the step signal, the $e_{\delta(t)}$ is the MSE (26) of the response-identified model for the impulse signal, the $e_{u(1..N)}$ is the MSE (26) distorted u signal by function $h(z)$, and the $e_{(j..N)}$ is the MSE (26) response-identified model.

By comparing the results of Tables 1 and 2, it can be seen that the disturbance of zero at the first position on the input signal for a large number of samples fulfills the principle of causality and response-identified model for the identification window.

Table 1. Distorted input $u(N-5)=0$ for $k>N-5$ samples.

k	$e_{(j..N)}$	$e_{1(t)}$	$e_{\delta(t)}$	$e_{u(1..N)}$
$N-10$	396.8864	219.9079	93.5970	0.0151
$N-50$	423.0187	219.9079	93.5970	0.0151
$N-100$	493.6943	219.9079	93.5970	0.0151
$N-1000$	1451.8	219.9079	93.5970	0.0151
$N-3000$	3.30e+10	220.3036	93.5978	0.0151

Table 2. Distorted input $u(N-j)=0$ on the first position input samples.

k	$e_{(j..N)}$	$e_{1(t)}$	$e_{\delta(t)}$	$e_{u(1..N)}$
$N-10$	114,1275	0,0142	3,5951	0.0156
$N-50$	29.7687	0.0218	5.6054	0.0203
$N-100$	19.6988	0.0218	5.5959	0.0270
$N-1000$	0.5552	4.89e-4	0.1163	0.0076
$N-3000$	7.336	4.90e-4	0.1164	0.0305

System with Noise on Output. The next experiment identifies the system (30) embedded in Gaussian noise in the output. It was assumed that the output system signal was exposed to a noise of covariance (32) (Fig. 4).

$$ErrCov = \frac{1}{N}\sum_{i=1}^{N}(Ex_i - x_i)^2 \tag{31}$$

$$y_{ErrCov} = 0.0494 \tag{32}$$

Table 3. Response error of the identified model with noise on the output.

k	$N-100$	$N-1000$	$_{opt(N-2698)}$	$N-3000$
$e_{u(1..N)}$	0.0807	0.0235	0.0238	0.0928
$e_{(j..N)}$	1089.6	2.3990	0.6006	5.4661
$e_{(j..N)}(MSIT)$	6.7094	2.5969	0.4274	0.5190
$e_{1(t)}$	346.5430	22.1021	9.4401	53.5393
$e_{1(t)}(MSIT)$	3.2970	35.0409	1.5530	2.3846
$e_{\delta(t)}$	105.5208	54.2538	41.0381	70.1437
$e_{\delta(t)}(MSIT)$	194.5097	63.7391	170.8675	120.9573

Table 3 shows the comparison error (26) of the response-identified model by the Matlab System Identification Toolbox (MSIT) and the proposed algorithm.

Table 4 displays the dependence of the estimated model coefficients on the horizon of the data and the dependence of the noise on the output.

Table 4. Dependence of the estimated coefficients on the discrete samples.

coeff\k	Model	$N-1000$	$_{opt(N-2698)}$	$N-3000$
a_1	−1.1269	−0.3604	−0.3267	−0.3572
a_2	0.4940	−0.3185	−0.2953	−0.3442
a_3	−0.1129	−0.2609	−0.223	−0.2614
b_1	−0.3832	−0.2139	−0.2843	−0.0875
b_2	−0.2328	0.0207	−0.0998	−0.0485
b_3	0.0668	0.0623	0.0034	0.0559

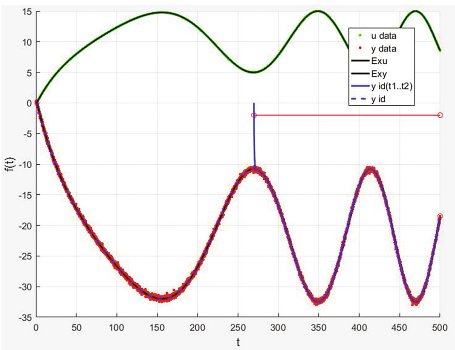

Fig. 4. Response-identified model with noise on the output.

System with Noise on Input and Output. The next experiment identifies the system (30) embedded in Gaussian noise on the input and output. It was assumed that the input and output system signals were exposed to a noise of covariance (33) (Fig. 6 and Tables 5 and 6).

$$u_{ErrCov} = 0.0494, \ y_{ErrCov} = 0.0494. \tag{33}$$

Fig. 5 shows the time constants of the identified system based on proposed algorithm and MSIT.

Table 5. Response error of the identified model with noise on the input and output.

k	$N-100$	$N-1000$	$_{opt(N-2713)}$	$N-3000$
$e_{u(1..N)}$	0.0832	0.0415	0.0651	0.1526
$e_{(j..N)}$	1136.6	2.8476	0.7436	11.6981
$e_{(j..N)}(MSIT)$	76.2574	3.5377	2.6045	4.2952
$e_{1(t)}$	358.3644	28.4760	12.9394	101.9493
$e_{1(t)}(MSIT)$	42.2193	54.8410	55.5994	54.7985
$e_{\delta(t)}$	108.5663	58.6032	46.8358	46.8358
$e_{\delta(t)}(MSIT)$	68.2047	90.8851	92.9625	92.0620

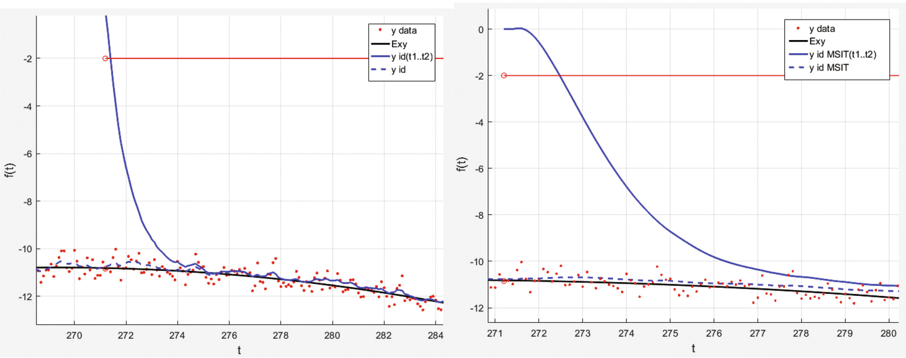

Fig. 5. The time constants of the response-identified model based on proposed algorithm and MSIT.

5.2 Laboratory System Distillation Column

Identification Laboratory Subsystem of Distillation Column. Identification of the non-Gaussian distribution of noise is done using registered data from the measurement level point $u = L175$ to the measurement level point $y = L176$, where: $\eta = 3$, $k = 6000$ samples, and the discretization step $\triangle t = 0.1[s]$ (Figs. 7, 8, 9 and 10).

Table 6. Dependence of the estimated coefficients on the discrete samples.

coeff\k	Model	$N - 1000$	$opt(N-2713)$	$N - 3000$
a_1	−1.1269	−0.5878	−0.5854	−0.5950
a_2	0.4940	−0.0964	−0.0679	−0.1212
a_3	−0.1129	−0.2650	−0.2074	−0.2642
b_1	−0.3832	−0.1129	−0.1455	−0.0306
b_2	−0.2328	0.0020	−0.0797	−0.0123
b_3	0.0668	0.0045	0.0765	0.0011

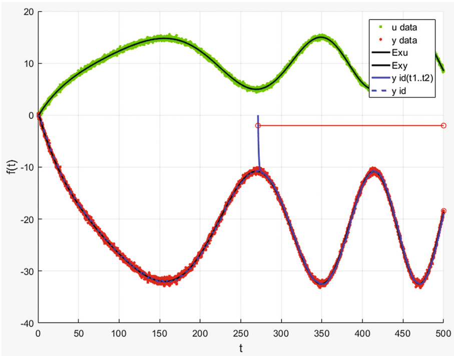

Fig. 6. Response-identified model with noise on the input and output.

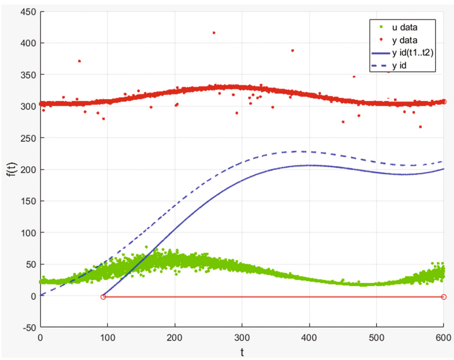

Fig. 7. Direct identification at the technology operating point.

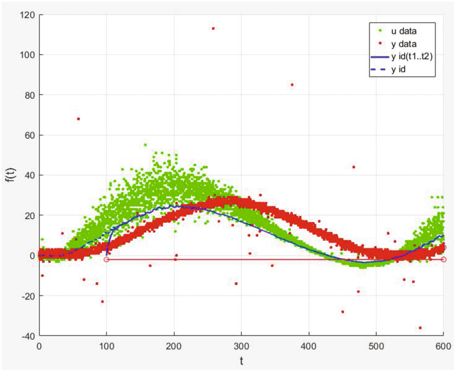

Fig. 8. Identification by biased operating points to the zero initial condition.

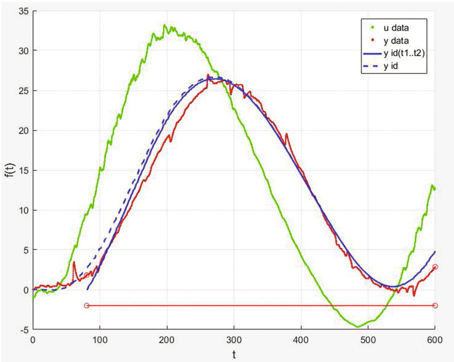

Fig. 9. Identification by biased to zero initial condition with optimal filtering.

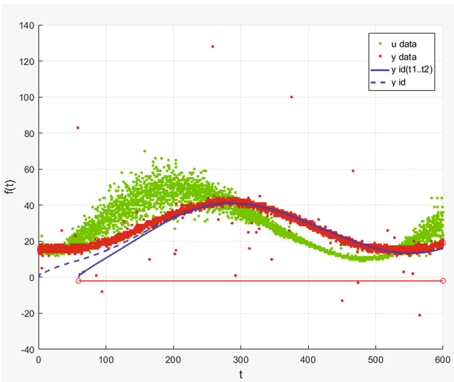

Fig. 10. Identification by biased to operating point with optimal filtering.

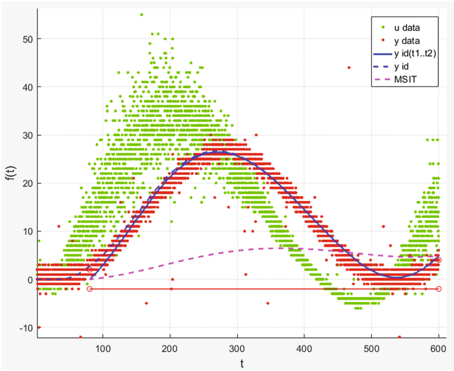

Fig. 11. Comparison response identified models between the proposed algorithm with optimal filtering and MSIT, by biased to zero initial condition (deviation model).

Comparison of the Proposed Algorithm and the Matlab System Identification Toolbox (MSIT). A comparison of the identified model transfer function response between the proposed algorithm and MSIT (Fig. 11).

$$G_{F1}(z) = \frac{-0.0001766z^2 + 2.856e - 05z + 0.0001745}{z^3 - 2.953z^2 + 2.928z - 0.9745} \tag{34}$$

$$G_{MSIT\ F1}(z) = \frac{0.0001183z^2 - 0.0002365z + 0.0001183}{z^3 - 3z^2 + 3z - 0.9998} \tag{35}$$

where the error of identification (27) is obtained as

$$e_{F1} = 8.3425, \; e_{MSIT\ F1} = 140.7838 \tag{36}$$

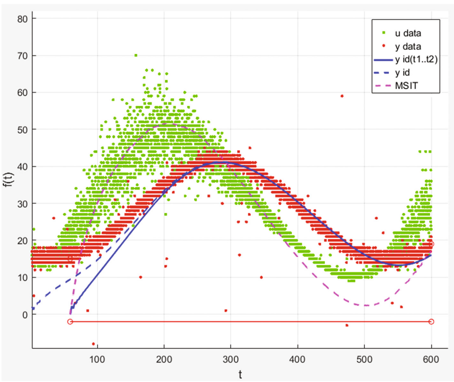

Fig. 12. Comparison of the response-identified models between the proposed algorithm with optimal filtering and MSIT, by biased to operating point.

A comparison of the identified model transfer function laboratory subsystems between the proposed algorithm and MSIT around the operating point is presented in (Fig. 12). If the operating point is biased to the neighborhood of the zero initial condition and optimal filters are used, the proposed algorithm produces acceptable results. The state matrix changes very little, but a coefficient of the control matrix changes, which has an impact on the system.

$$G_{F2}(z) = \frac{2.671e-05z^2 - 6.632e-07z + 5.943e-07}{z^3 - 2.954z^2 + 2.93z - 0.976} \tag{37}$$

$$G_{MSIT\,F2}(z) = \frac{-0.3832z^2 - 0.2338z + 0.06683}{z^3 - 1.127z^2 + 0.494z - 0.1129} \tag{38}$$

where the error of identification (27) is obtained as

$$e_{F2} = 14.5630,\ e_{MSIT\,F2} = 153.6657 \tag{39}$$

6 Conclusion

The proposed algorithm gives exact and repeatable results for systems embedded at a nearly Gaussian noise on the input and output. The results of identification are independent of the initial conditions. The algorithm allows for the correct of the time constants in the identified model through the modulation of the function (15). In the literature there is a lot of theoretical proposals of concepts that are based on the mathematics of dynamic systems [4–11]. The problem appears when these concepts are used in control systems engineering, which requires more generalized assumptions: limited precision of data representation and perturbed Gaussian distribution on the input and output, as shown in Sect. 5.2. The study demonstrated that the proposed algorithm is an innovation in the fields of control systems engineering and applied mathematics. It returns acceptable quality indices for online real-system identification, is independent of the system state and preliminary parametrization [10,11], and can be used for a wide range of test signals for the assumptions given in [1]. Direct calculation of the identification results for the window data range allows for robust identification of the optimal model of LSE. A new achievement of the presented algorithm is the ability to identify unstable systems, which satisfies the controllability condition in the b.i.b.o. sense. The proposed algorithm also opens up a new horizon of possibilities in process diagnostics, enabling high-precision faults detection and reconstruction of damaged data using mathematical models and linear regression.

References

1. Soderstrom, T., Stoica, P.: System Identification. Prentice-Hall, Hemel Hempstead (1989)
2. Keesman, K.J.: System Identification. An Introduction. Springer, London (2011). doi:10.1007/978-0-85729-522-4

3. Isermann, R., Munchhof, M.: Identification of Dynamic Systems. An Introduction with Applications. Springer, Heidelberg (2011). doi:10.1007/978-3-540-78879-9
4. Niedźwiecki, M., Meller, M., Pietrzak, P.: System identification based approach to dynamic weighing revisited. Mechan. Syst. Sig. Process. **80**, 582–599 (2016). Elsevier
5. Sastry, S.: Nonlinear Systems Analysis, Stability, and Control. Springer Science+Business Media, New York (1999). doi:10.1007/978-1-4757-3108-8
6. Greblicki, W., Pawlak, M.: Nonparametric System Identification. Cambridge University Press, New York (2008)
7. Donga, S., Liua, T., Wanga, W., Baob, J., Caoc, Y.: Identification of discrete-time output error model for industrial processes with time delay subject to load disturbance. J. Process Control **50**, 40–55 (2017). Elsevier
8. Liu, T., Yao, K., Gao, F.: Identification and autotuning of temperature-control system with application to injection molding. IEEE Trans. Control Syst. Technol. **17**(6), 1282–1294 (2009)
9. Aljanaideh, K.F., Bernstein, D.S.: Closed-loop identification of unstable systems using noncausal FIR models. Int. J. Control **90**(2), 184–201 (2017)
10. Wang, W., Garnier, H.: System Identification, Environmental Modelling, and Control System Design. Springer, London (2012)
11. Ljung, L.: System Identification. Theory for the User. PTR Prentice Hall Inc, Boston (1987)
12. Brooder, F.E.: Strongly nonlinear parabolic boundary-value problems. Am. J. Math. **86**, 339–357 (1964)
13. Lions, J.L.: Sur certaines equations paraboliques nonlineaires. Bull. Sm. Math. Frunce **93**, 155–175 (1965)