Data Driven, Additive Fault Detection for Wind Turbines

Dušan Krokavec*, Anna Filasová

Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University of Košice
Letná 9, 042 00 Košice, Slovakia
e-mail: {dusan.krokavec, anna.filasova}@tuke.sk
*Corresponding author

Abstract: This paper covers an approach for fault residuals construction, when carrying out the data driven techniques. The method is based on the system input and output data and is built on a difference between data of a "normal" and a "faulty" system behavior. The research evaluates the performance of the structures, based on the subspace identification technique, to highlight the acceptable usefulness of the resulting data driven approach. Following the same lines, it delineates the connection to input/output data and provides an approach to computation of the parameterized data matrices. The principle is applied directly in analysis of the data generated by the wind turbine model.

Keywords: fault detection; fault residuals; data driven techniques; wind turbine models

1 Introduction

Beginning in the 70s of the last century, the analytical redundancy approach, instead of hardware redundancy, forces the concept of fault detection filters, constructed on the observations derived from the system measurements using Kalman filters [1] and system state observers [2]. The principal focus points on the fault detection and diagnosis were oriented on the chemical and petrochemical processes [3], the first applications of the data-driven methods are also linked to this industry area [4]. Thus, in the main context, the fault detection and isolation (FDI) part of such diagnostic systems have to give an efficient solution with small fault detection time delays.

Many models based FDI techniques have been proposed including the above mentioned Kalman filters [5], unknown input observers [6] and H_∞/H_2 observer-based residual filter schemes [7]. The underlying idea behind the model based FDI is to use a mathematical model of the system as the source of redundant information and to produce the state estimate-based fault residuals by using the systems measured outputs. For all representations based on system models the fundamental question is the system observability. The disadvantage of this approach is the need for an accurate model of the process.

An alternative approach known as data driven FDI (DD FDI) uses a collection of measured data to discover patterns related to the normal and faulty system behavior. The data driven FDIs prevailingly reflect results in machine learning, computational intelligence and data mining. The most popular DD FDI techniques are data feature-based prediction techniques using the nonlinear time series based on the projection-based techniques [8], prediction principle [9], the correlation analysis and Bayesian inference principle [10] and the fault feature extraction and classification using computational intelligence principle [11]. The additive fault estimators can also be considered as data driven procedures, usually applicable in the case of poor analytical knowledge of the system dynamics and external disturbances [12] [13].

The wind turbine technology is one of the appropriate ways for increased use of renewable energy. Thus, in order to improve a wind turbine behavior and reliability, faults prevention in the parts of the wind turbine and fault tolerant control stay relevant objectives. A way to ensure these tasks consists in introducing advanced fault detection, isolation, and accommodation schemes [14-15], where FDIs are designed to estimate the filter speed by using a bank of state observers and bank of unknown input observers, respectively. To improve diagnosis of faults the wind farm level can be applied, when a wind turbine is considered in comparison to another turbine of the wind farm [16].

The starting data-driven solution in the wind turbine diagnosis relies on Takagi–Sugeno (T-S) fuzzy models that are derived from a clustering c-means algorithm, followed by an identification procedure [17]. To achieve maximization of the wind power extraction, T-S fuzzy models of reduced dimension are proposed [18]. The power of the approach can be amplified when combining the wind power extraction T-S models by DD FDI and the weighting two or more metrics related with analytical FDI [19] [20]. In all schemes, a properly large data set should be available as a-priori knowledge to train the structure in the fault-free case.

In this paper a basic ground is taken on the data driven approaches, based on subspace identification methods, for additive fault detection of wind turbines. Since the system state-space description projects the system state into output variables, it is a natural way for constructive data driven approaches to suppress the system model description and to identify only the parameterized matrices of system data. This leads to formulas for the matrix pseudo inverse operations on the predetermined data matrices that seems to be basic. Of course, this application of the orthogonal complement of non-square data matrices for identification of parameterized matrices of the system is not new and the presented way seems to be only a new adaptation.

The paper is organized as follows. Ensuing introduction in Sect. 1 and continuing by the system description using lagged variables in Sect. 2, constructions of data matrix null spaces are given in Sect. 3. The fault residuals generation is analyzed in Sect. 4 and suppression of noise effects is explained in Sect. 5. In the task relation, Sect. 6 describes substantial details in a reference model of wind turbines and applicability of the proposed method using benchmark model parameters is in Sect. 7. Finally, in Sect. 8, some prioritized concluding remarks are presented.

For sake of convenience, throughout this paper used notations reflect usual conventionality so that $\mathbf{x}^T$, $\mathbf{X}^T$ mean transpose of the vector $\mathbf{x}$ and the matrix $\mathbf{X}$, the notation $\mathbf{X}^{\ominus 1}$ denotes the Moore-Penrose pseudo inverse of a non-square matrix, $\mathbf{X}^{\perp}$ reflects the orthogonal complement of a non-square matrix $\mathbf{X}$, $\mathbf{I}_n$ is the n^{th} order identity matrix and $\mathbb{R}^{n \times r}$ signifies the set *nxr* real matrices.

2 Lagged Variables

Considering that the task is focuses on developing the FDI method in the data-driven fashion, the linear discrete-time model of the square system is used in the standard state space representation given by

$$\mathbf{q}(i+1) = \mathbf{F}\mathbf{q}(i) + \mathbf{G}\mathbf{u}(i) + \mathbf{v}(i) \tag{1}$$

$$\mathbf{y}(i) = \mathbf{C}\mathbf{q}(i) + \mathbf{w}(i) \tag{2}$$

where $\mathbf{q}(i) \in \mathbb{R}^n$ is the state vector, $\mathbf{u}(i) \in \mathbb{R}^{r_u}$, $\mathbf{y}(i) \in \mathbb{R}^m$ is the input and the output vector, respectively, while $\mathbf{F} \in \mathbb{R}^{n \times n}$, $\mathbf{G} \in \mathbb{R}^{n \times r_u}$, $\mathbf{C} \in \mathbb{R}^{m \times n}$. The system is corrupted by the stochastic disturbances $\mathbf{v}(i) \in \mathbb{R}^n$, $\mathbf{w}(i) \in \mathbb{R}^n$, which are zero means and normally distributed white noise.

Taking (1), (2) the following can be performed for the square system ($r_u = m$)

$$\begin{aligned}
\mathbf{y}(i) &= \mathbf{C}\mathbf{q}(i) + \mathbf{w}(i) \\
\mathbf{y}(i+1) &= \mathbf{C}\mathbf{F}\mathbf{q}(i) + \mathbf{C}\mathbf{G}\mathbf{u}(i) + \mathbf{C}\mathbf{v}(i) + \mathbf{w}(i) \\
&\ \vdots \\
\mathbf{y}(i+s-1) &= \begin{matrix} \mathbf{C}\mathbf{F}^{s-1}\mathbf{q}(i) + \left[\mathbf{C}\mathbf{F}^{s-2}\mathbf{G} \ \cdots \ \mathbf{C}\mathbf{G} \ \ \mathbf{0}\right]\mathbf{u}_o(i) + \\ + \left[\mathbf{C}\mathbf{F}^{s-2} \ \cdots \ \mathbf{C} \ \ \mathbf{0}\right]\mathbf{v}_o(i) + \mathbf{w}(i+s-1) \end{matrix}
\end{aligned} \tag{3}$$

writing with

$$\mathbf{u}_o(i) = \begin{bmatrix} \mathbf{u}(i) \\ \vdots \\ \mathbf{u}(i+s-1) \end{bmatrix}, \quad \mathbf{v}_o(i) = \begin{bmatrix} \mathbf{v}(i) \\ \vdots \\ \mathbf{v}(i+s-1) \end{bmatrix}, \quad \mathbf{w}_o(i) = \begin{bmatrix} \mathbf{w}(i) \\ \vdots \\ \mathbf{w}(i+s-1) \end{bmatrix} \tag{4}$$

which can be rewritten as:

$$\mathbf{y}_o(i) = \mathbf{P}\mathbf{q}(i) + \mathbf{R}\mathbf{u}_o(i) + \mathbf{S}\mathbf{v}_o(i) + \mathbf{w}_o(i) \tag{5}$$

where,

$$\mathbf{y}_o(i)=\begin{bmatrix}\mathbf{u}(i)\\ \vdots\\ \mathbf{u}(i+s-1)\end{bmatrix},\quad \mathbf{R}=\begin{bmatrix}\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{CG} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathbf{CF}^{s-2}\mathbf{G} & \mathbf{CF}^{s-3}\mathbf{G} & \cdots & \mathbf{CG} & \mathbf{0}\end{bmatrix} \tag{6}$$

$$\mathbf{P}=\begin{bmatrix}\mathbf{C}\\ \mathbf{CF}\\ \vdots\\ \mathbf{CF}^{s-1}\end{bmatrix},\quad \mathbf{S}=\begin{bmatrix}\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathbf{CF}^{s-2} & \mathbf{CF}^{s-3} & \cdots & \mathbf{C} & \mathbf{0}\end{bmatrix} \tag{7}$$

and $\mathbf{y}_o(i),\mathbf{u}_o(i),\mathbf{w}_o(i)\in\mathbb{R}^{sm}$, $\mathbf{v}_o(i)\in\mathbb{R}^{sn}$, $\mathbf{P}\in\mathbb{R}^{sn\times n}$, $\mathbf{R}\in\mathbb{R}^{sm\times sm}$, $\mathbf{S}\in\mathbb{R}^{sm\times sn}$.

Thus, applying the input and output data sets, the related data structures can be investigated, defined as:

$$\mathbf{y}_g(i)=\begin{bmatrix}\mathbf{y}(i-r)\\ \vdots\\ \mathbf{y}(i-1)\end{bmatrix},\quad \mathbf{u}_g(i)=\begin{bmatrix}\mathbf{u}(i-r)\\ \vdots\\ \mathbf{u}(i-1)\end{bmatrix},\quad \mathbf{z}_g(i)=\begin{bmatrix}\mathbf{y}_g(i)\\ \mathbf{u}_g(i)\end{bmatrix},\quad \mathbf{z}_o(i)=\begin{bmatrix}\mathbf{y}_o(i)\\ \mathbf{u}_o(i)\end{bmatrix} \tag{8}$$

where $\mathbf{z}_g(i)\in\mathbb{R}^{2rm}$, $\mathbf{z}_o(i)\in\mathbb{R}^{2sn}$ and $s\geq r>n$

The structural relations (5) can be parameterized such that:

$$\mathbf{Y}_{o,p}=\mathbf{PQ}_{o,p}+\mathbf{RU}_{o,p}+\mathbf{D}_{o,p} \tag{9}$$

where,

$$\mathbf{Y}_{o,p}=\left[\mathbf{y}_o(i)\ \cdots\ \mathbf{y}_o(i+p-1)\right],\quad \mathbf{Q}_{o,p}=\left[\mathbf{q}(i)\ \cdots\ \mathbf{q}(i+p-1)\right] \tag{10}$$

$$\mathbf{U}_{o,p}=\left[\mathbf{u}_o(i)\ \cdots\ \mathbf{u}_o(i+p-1)\right],\quad \mathbf{V}_{o,p}=\left[\mathbf{v}_o(i)\ \cdots\ \mathbf{v}_o(i+p-1)\right] \tag{11}$$

$$\mathbf{D}_{o,p}=\left[\mathbf{SV}_{o,p}\ \mathbf{W}_{o,p}\right],\quad \mathbf{W}_{o,p}=\left[\mathbf{w}_o(i)\cdots\mathbf{w}_o(i+p-1)\right] \tag{12}$$

and $\mathbf{Q}_{o,p}\in\mathbb{R}^{sm\times p}$, $\mathbf{Y}_{o,p},\mathbf{U}_{o,p}\in\mathbb{R}^{sm\times pm}$, $\mathbf{D}_{o,p}\in\mathbb{R}^{sm\times 2p}$, $p\gg s\geq r>n$

When the new samples $\mathbf{y}_o(i)$ and $\mathbf{u}_o(i)$ are available, updated matrices are formulated by appending the new columns on the right and deleting the first columns in $\mathbf{Y}_{o,p}$, $\mathbf{U}_{o,p}$

Corollary 1 The subspace method aided data-driven fault detection can be detailed for systems with the additive faults, including the actuator faults. This application requests to alter the state equation (1) as:

$$\mathbf{q}(i+1)=\mathbf{Fq}(i)+\mathbf{Gu}(i)+\mathbf{Hf}(i)+\mathbf{v}(i),\quad \mathbf{f}(i)\in\mathbb{R}^{r_f},\mathbf{H}\in\mathbb{R}^{n\times r_f} \tag{13}$$

$$f_o(i)=\begin{bmatrix}\mathbf{f}(i)\\ \vdots\\ f(i+s-1)\end{bmatrix},\quad \mathbf{R}_f=\begin{bmatrix}\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \mathbf{CH} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathbf{CF}^{s-2}\mathbf{H} & \mathbf{CF}^{s-3}\mathbf{H} & \cdots & \mathbf{CH} & \mathbf{0}\end{bmatrix} \tag{14}$$

and modifies the matrix $\mathbf{D}_o$ as follows, when applying in (9),

$$\mathbf{D}_{o,p}=\left[\mathbf{R}_f\mathbf{f}_o(i)+\mathbf{Sv}_o(i)+\mathbf{w}_o(i)\ \cdots\ \mathbf{R}_f\mathbf{f}_o(i+p-1)+\mathbf{Sv}_o(i+p-1)+\mathbf{w}_o(i+p-1)\right] \tag{15}$$

3 Construction of Data Matrix Null Spaces

The presented method is built on non-square matrices, which gives the possibility to eliminate the column redundance by principle of the matrix null space.

Lemma 1 *The orthogonal projector onto the kernel of a non-square matrix* $\mathbf{L} \in \mathbb{R}^{a \times b}, b > a,$ *(an orthogonal complement of* $\mathbf{L}$ *) takes the form:*

$$\mathbf{L}^{\perp} = \mathbf{I}_b - \mathbf{L}^{\ominus 1}\mathbf{L} \cong \mathbf{0} \tag{16}$$

Where,

$$\mathbf{L}^{\ominus 1} = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1} \tag{17}$$

is the Moore-Penrose pseudo-inverse of $\mathbf{L}$ *, where,* $\mathbf{L}^{\ominus 1} \in \mathbb{R}^{b \times a}$

Remark 1 Noting that for a non-square matrix it yields also that $\mathbf{L} = \mathbf{L}$, then multiplying this equality from the left-hand side by the identity matrix $\mathbf{I}_a$ it yields:

$$\mathbf{L} = \mathbf{I}_a\mathbf{L} = \mathbf{L}\mathbf{I}_b = \mathbf{L}\mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L} \tag{18}$$

which implies,

$$\mathbf{I}_b = \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L} = \mathbf{L}^{\ominus 1}\mathbf{L} \tag{19}$$

Rewriting the equality (19) as:

$$\mathbf{L} = \mathbf{L}\mathbf{L}^{\ominus 1}\mathbf{L} \tag{20}$$

it can be derived that:

$$\mathbf{L}(\mathbf{I}_b - \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}) = \mathbf{L}(\mathbf{I}_b - \mathbf{L}^{\ominus 1}\mathbf{L}) = \mathbf{L}\mathbf{L}^{\perp} \cong \mathbf{0} \tag{21}$$

which defines (16).

Lemma 2 *Singular value decomposition (SVD) of a non-square matrix* $\mathbf{L} \in \mathbb{R}^{a \times b}, b > a,$ *constructs the matrix relation:*

$$\mathbf{M}^T\mathbf{L}\mathbf{N} = \mathbf{\Sigma} \tag{22}$$

where,

$$\mathbf{M} = \begin{bmatrix}\mathbf{m}_1 & \cdots & \mathbf{m}_a\end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix}\mathbf{n}_1 & \cdots & \mathbf{n}_b\end{bmatrix}, \quad \mathbf{M} \in \mathbb{R}^{a \times a}, \mathbf{N} \in \mathbb{R}^{b \times b} \tag{23}$$

are the ortthogonal matrices of the left and the right singular vectors of $\mathbf{L}$ *and*

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \mathbf{0}_{a,b-a} \\ & & & \sigma_a \end{bmatrix} \tag{24}$$

where σ_l, $l = 1, \ldots, a$ *are the singular variables of* $\mathbf{L}$ *, generally ordered in such a way that* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_a \geq 0$.

Remark 2 Since (22) can be rewritten as:

$$\mathbf{M}^T\mathbf{L}\left[\mathbf{N}_1\ \mathbf{N}_2\right] = \mathbf{\Sigma}, \quad \mathbf{N}_1 \in \mathbb{R}^{b\times a}, \mathbf{N}_2 \in \mathbb{R}^{b\times(b-a)} \tag{25}$$

the structure of $\mathbf{\Sigma}$ implies that:

$$\mathbf{L}\mathbf{N}_2 = \mathbf{0}_{a,b-a} \tag{26}$$

that is, the orthogonal complement of $\mathbf{L}$ is $\mathbf{N}_2$

Remark 3 To given $\mathbf{L}$ it can be constructed the square matrix $\mathbf{L}\mathbf{L}^T$ and for the eigenvalue structure of $\mathbf{L}\mathbf{L}^T$ it yields:

$$\mathbf{L}\mathbf{L}^T\mathbf{s}_l = \rho_l\mathbf{s}_l, \quad \mathbf{u}_l^T\mathbf{L}\mathbf{L}^T = \rho_l\mathbf{u}_l^T, \quad l = 1,\dots,a \tag{27}$$

where ρ_l it the l^{th} eigenvalue of $\mathbf{L}\mathbf{L}^T$, $\mathbf{s}_l$ it the l^{th} right eigenvector of $\mathbf{L}\mathbf{L}^T$ and $\mathbf{u}_l^T$ it the l^{th} left eigenvector of $\mathbf{L}\mathbf{L}^T$. Multiplying the left side of (27) by $\mathbf{u}_h^T$ it has to be:

$$\mathbf{u}_h^T\mathbf{L}\mathbf{L}^T\mathbf{s}_l = \rho_h\mathbf{u}_h^T\mathbf{s}_l = \rho_l\mathbf{u}_h^T\mathbf{s}_l \tag{28}$$

which can be satisfied only if:

$$\mathbf{u}_h^T\mathbf{s}_l = \begin{cases} 1, & h = l, \\ 0, & h \neq l. \end{cases} \quad h,l = 1,\dots,a \tag{29}$$

that is the left and the right eigenvectors are orthonormal. Consequently (28) can be written as:

$$\rho_h\mathbf{u}_h^T\mathbf{s}_l = \mathbf{u}_h^T\mathbf{L}\mathbf{L}^T\mathbf{s}_l = \mathbf{u}_h^T\mathbf{L}\mathbf{L}^T\mathbf{s}_l\frac{\left\|\mathbf{L}^T\mathbf{s}_l\right\|}{\left\|\mathbf{L}^T\mathbf{s}_l\right\|} = \mathbf{u}_h^T\mathbf{L}\mathbf{v}_l\sqrt{\rho_l} = \mathbf{u}_h^T\mathbf{L}\mathbf{v}_l\sigma_l \tag{30}$$

which can be satisfied only if:

$$\mathbf{u}_h^T\mathbf{s}_l = \begin{cases} 1, & h = l, \\ 0, & h \neq l, \\ 0, & l > a, \end{cases} \quad h,l = 1,\dots,a \tag{31}$$

and where,

$$\mathbf{v}_l = \frac{\mathbf{L}^T\mathbf{s}_l}{\left\|\mathbf{L}^T\mathbf{s}_l\right\|}, \quad \sqrt{\rho_l} = \left\|\mathbf{L}^T\mathbf{s}_l\right\|, \quad \sqrt{\rho_l} = \sigma_l \tag{32}$$

Then (30) defines (22).

To find the fault residual filter construction and structure, the following section is especially offered.

4 Fault Residuals

A general fault diagnosis can be performed using residual signal that represents a deviation from standard operating conditions and can be generated by comparing, for example, a model output with the actual system output.

The data-driven fault residuals must be generated from the system input and output data using the lagged variables.

Rewriting (9) as:

$$\mathbf{Y}_{o,p} - \mathbf{R}\mathbf{U}_{o,p} = \mathbf{P}\mathbf{Q}_{o,p} + \mathbf{D}_{o,p} \tag{33}$$

and multiplying (33) from the left-hand side by $\mathbf{P}^{\perp}$ (constructed by (16) to eliminate the effect of $\mathbf{Q}_{o,p}$), it yields:

$$\mathbf{P}^{\perp}(\mathbf{Y}_{o,p} - \mathbf{R}\mathbf{U}_{o,p}) = \mathbf{P}^{\perp}\mathbf{D}_{o,p} \tag{34}$$

and the fault residuals can be defined as:

$$\Omega_{o,p} = \mathbf{P}^{\perp}\mathbf{Y}_{o,p} - \mathbf{P}^{\perp}\mathbf{R}\mathbf{U}_{o,p} \tag{35}$$

$$\omega_o(i) = \mathbf{P}^{\perp}\mathbf{y}_o(i) - \mathbf{P}^{\perp}\mathbf{R}\mathbf{u}_o(i) \tag{36}$$

Thus, the residual generation can be achieved if $\mathbf{P}^{\perp}$, $\mathbf{P}^{\perp}\mathbf{R}$ can be acquired from the system input and output data.

To use the residual vector $\omega_o(i)$ at the time instant step *i*, the statistic test has to be applied, based on the following covariance matrix computed in the fault-free case.

$$\mathbf{\Xi}_o = \tfrac{1}{p-1}\sum_{j=1}^{p}\mathbf{\Omega}_{o,p}(j)\mathbf{\Omega}_{o,p}^T(j) \tag{37}$$

to give the possibility of generating the fault occurrence threshold as:

$$t_\omega = \boldsymbol{\omega}_o^T(i)\,\mathbf{\Xi}_o^{-1}\boldsymbol{\omega}_o(i) \tag{38}$$

when defining the detection logic:

$$\begin{cases} t_\omega < t_h, & no\ \ fault \\ t_\omega \geq t_h, & a\ \ fault \end{cases} \tag{39}$$

where $t_h \in \mathbb{R}_+$ is a defined threshold.

The presented approach takes advantage of the actual system input and output data sets to generate a discrepancy (residuals) that are indicative as a potential additive fault occurrence.

It can be underlined that the data-driven technique addresses mostly anticipated fault conditions only.

5 Suppression of Noise Effects

When the system is corrupted by the system and measurement noise, both taking Gaussian white noise properties, the instrumental variable $\mathbf{z}_o(i)$ can be exploited to suppress noise effects. Since in this case the past data $\mathbf{z}_g(i)$ are uncorrelated with the future noise realizations then, using (12):

$$\lim_{p\to\infty}\frac{1}{p}\mathbf{D}_{o,p}\mathbf{Z}_{g,p}^T = \lim_{p\to\infty}\frac{1}{p}\left(\mathbf{S}\mathbf{V}_{o,p}\mathbf{Z}_{g,p}^T + \mathbf{W}_{o,p}\mathbf{Z}_{g,p}^T\right) = \mathbf{0} \tag{40}$$

where,

$$\mathbf{Z}_{g,p} = \begin{bmatrix}\mathbf{Y}_{g,p}\\ \mathbf{U}_{g,p}\end{bmatrix} = \begin{bmatrix}\mathbf{y}_g(i-1) & \cdots & \mathbf{y}_g(i-p)\\ \mathbf{u}_g(i-1) & \cdots & \mathbf{y}_g(i-p)\end{bmatrix} \tag{41}$$

The following remark shows the computation of the orthogonal complement $\mathbf{P}^{\perp}$

Remark 4 Constructing according to (16) the orthogonal complement $\mathbf{U}_{o,p}^{\perp}$ then it yields for the auxiliary variable Φ

$$\Phi = \lim_{p\to\infty}\frac{1}{p}\left(\mathbf{Y}_{o,p} - \mathbf{R}\mathbf{U}_{o,p}\right)\mathbf{U}_{o,p}^{\perp}\mathbf{Z}_{g,p}^T = \lim_{p\to\infty}\frac{1}{p}\mathbf{Y}_{o,p}\mathbf{U}_{o,p}^{\perp}\mathbf{Z}_{g,p}^T = \lim_{p\to\infty}\frac{1}{p}\mathbf{P}\mathbf{Q}_{o,p}\mathbf{U}_{o,p}^{\perp}\mathbf{Z}_{g,p}^T \tag{42}$$

and the result of SVD to Φ, when exploiting the property (25), means:

$$\mathbf{\Gamma}^T\mathbf{\Phi}\left[\mathbf{\Psi}_1 \ \ \mathbf{\Psi}_2\right] = \mathbf{\Sigma}_{\Phi} \tag{43}$$

which defines that $\mathbf{P}^{\perp}$ can be computed as:

$$\mathbf{P}^{\perp} = \mathbf{\Psi}_2 \tag{44}$$

Finally, since (40) implies for (34) that:

$$\mathbf{0} = \mathbf{P}^{\perp}\lim_{p\to\infty}\frac{1}{p}\left(\mathbf{Y}_{o,p}\mathbf{Z}_{g,p}^T - \mathbf{R}\mathbf{U}_{o,p}\mathbf{Z}_{g,p}^T\right) \tag{45}$$

(45) means that:

$$\mathbf{P}^{\perp}\lim_{p\to\infty}\frac{1}{p}\mathbf{Y}_{o,p}\mathbf{Z}_{g,p}^T = \mathbf{P}^{\perp}\mathbf{R}\lim_{p\to\infty}\frac{1}{p}\mathbf{U}_{o,p}\mathbf{Z}_{g,p}^T \tag{46}$$

and using the notations:

$$\mathbf{\Phi}_{YZ} = \lim_{p\to\infty}\frac{1}{p}\mathbf{Y}_{o,p}\mathbf{Z}_{g,p}^T, \quad \mathbf{\Phi}_{UZ} = \lim_{p\to\infty}\frac{1}{p}\mathbf{U}_{o,p}\mathbf{Z}_{g,p}^T \tag{47}$$

then, with the pseudoinverse of the matrix $\mathbf{\Phi}_{UZ}$ it can be computed that:

$$\mathbf{P}^{\perp}\mathbf{R} = \mathbf{P}^{\perp}\mathbf{\Phi}_{YZ}\mathbf{\Phi}_{UZ}^{-1} \tag{48}$$

Thus, all the parameters are now known to construct (35) when computing the fault residuals from the system input and system output data.

Note, in practice, the limit to infinity in (40), (42), (47) is replaced by a substantially large value of *p*.

6 Reference Wind Turbine Model

The proposed reference linear model can be given by the model of the three-blade horizontal wind turbine [21].

In this scheme the aerodynamic torque is given as [22]:

$$\tau_r(t) = \tfrac{1}{2}\rho\pi r^3 C_q(\lambda(t),\beta(t))v^2(t) \tag{49}$$

where $v(t)$ is the actual speed of wind [m.s^{-1}], r is the rotor radius [m], $C_q(*)$ is the torque coefficient and ρ is the air density [kg.m^{-3}]. Specifically, $C_q(*)$ depends on the tip speed ratio $\lambda(t)$ [rad] between the tangential speed of the tip of a blade and the actual speed of wind and the blades pitch angle $\beta(t)$ [rad], where:

$$C_q(\lambda(t),\beta(t)) = \frac{C_p(\lambda(t),\beta(t))}{\beta(t)}, \quad \lambda(t) = \frac{r\omega_r(t)}{v(t)} \tag{50}$$

whilst $\omega_r(t)$ is the rotor angular speed [rad.s^{-1}] and $C_p(\lambda(t),\beta(t))$ is a rotor torque coefficient.

The turbine speed is controlled to track the reference trajectory $\omega_{rt}(t)$, to maintain the tip-speed ratio at its optimal value λ_o, computed from the relation:

$$\omega_{rt}(t) = \lambda(t)\frac{v(t)}{r} \tag{51}$$

Thus, using $\lambda(t)$ it yields:

$$\tau_r(t) = \frac{1}{2\lambda^2(t)}\rho\pi r^5 C_q(\lambda(t),\beta(t))\omega_r^2(t) \tag{52}$$

The two-mass drive train equations take the forms:

$$J_r\frac{d\omega_r(t)}{dt} = \tau_r(t) - (b_{dt}+b_r)\omega_r(t) - k_{dt}\vartheta(t) + \frac{b_{dt}}{N_g}\omega_g(t) \tag{53}$$

$$J_g\frac{d\omega_g(t)}{dt} = \frac{\mu_{dt}k_{dt}}{N_g}\vartheta(t) + \frac{\mu_{dt}b_{dt}}{N_g}\omega_r(t) - \left(\frac{\mu_{dt}b_{dt}}{N_g^2} + b_g\right)\omega_g(t) - \tau_g(t) \tag{54}$$

$$\frac{d\vartheta(t)}{dt} = \omega_r(t) - \frac{1}{N_g}\omega_g(t) \tag{55}$$

where $\omega_r(t)$ is the rotor angular speed [rad.s^{-1}], $\omega_g(t)$ is the generator rotating speed [rad.s^{-1}], $\vartheta(t)$ is the torsion angle [rad], $\tau_g(t)$ is the aerodynamic torque [kg.m^2.s^{-2}], J_r is rotor inertia moment [kg.m^2], N_g is gear ration, b_{dt} is the torsion damping coefficient [N.m.s.rad^{-1}], b_r is rotor external damping [N.m.s.rad^{-1}], k_{dt} is the torsion stiffness [N.m.rad^{-1}], b_g is the generator external damping [N.m.s.rad^{-1}] and μ_{dt} is efficiency od drive train.

The hydraulic pitch is modeled by the differential equation of the form [23]

$$\frac{d^2\beta(t)}{dt^2} + 2\varsigma\omega_n \frac{d\beta(t)}{dt} + \omega_n^2\beta(t) = \omega_n^2\beta_r(t) \tag{56}$$

while $\beta(t)$ is the measured pitch angle [rad], $\beta_r(t)$ is the reference pitch angle [rad], ω_n is the pitch filter natural frequency [rad.s^{-1}], ς is the damping factor and the generator and converter dynamics are modeled by the first-order differential equation [24]:

$$\frac{d\tau_g(t)}{dt} + \alpha\tau_g(t) = \alpha\tau_{gr}(t) \tag{57}$$

where $\tau_g(t)$ is the generator torque [Nm], $\tau_{gr}(t)$ is the generator torque reference [Nm] and α is the generator and converter time-constant parameter [s^{-1}].

Considering the continuou-tine system state-space description of the reference wind turbine model of the following form:

$$\frac{d\mathbf{q}(t)}{dt} = \mathbf{A}\mathbf{q}(t) + \mathbf{B}\mathbf{u}(i), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{q}(t) \tag{58}$$

then the used vector variables and the reference model matrix parameters are defined as follows:

$$\mathbf{u}^T(t) = \begin{bmatrix} \tau_{gr}(t) & \beta_r(t) \end{bmatrix}, \quad \mathbf{y}^T(t) = \begin{bmatrix} \omega_r(t) & \omega_g(t) \end{bmatrix} \tag{59}$$

$$\mathbf{q}^T(t) = \begin{bmatrix} \omega_r(t) & \omega_g(t) & \vartheta(t) & \dot{\beta}(t) & \beta(t) & \tau_g(t) \end{bmatrix} \tag{60}$$

$$\mathbf{B}^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & \omega_n & 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{61}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & \frac{b_{dt}}{J_r N_g} & -\frac{k_{dt}}{J_r} & 0 & 0 & 0 \\ \frac{\mu_{dt} b_{dt}}{J_g N_g} & -\frac{\mu_{dt} b_{dt}}{J_g N_g^2} - \frac{b_g}{J_g} & \frac{\mu_{dt} k_{dt}}{J_g N_g} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{N_g} & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\varsigma\omega_n & -\omega_n^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha \end{bmatrix} \tag{62}$$

$$a_{11}(t) = \frac{1}{2J_r\lambda^2(t)}\rho\pi r^5 C_q(\lambda(t), \beta(t))\omega_r(t) - \frac{b_{dt} + b_r}{J_r} \tag{63}$$

In simulation this scheme is accommodated by accordingly adding the vectors of the system noise and the measurement noise.

7 Simulations

The parameters used in the reference model are [21]

$\rho = 1.225\,\text{kg.m}^{-3}$	$r = 57.5\,\text{m}$	$N_\text{g} = 95$	$\alpha = 501\text{s}^{-1}$
$b_\text{dt} = 775.49\ \text{N.m.s.rad}^{-1}$	$b_\text{r} = 7.11\,\text{N.m.s.rad}^{-1}$	$b_\text{g} = 45.6\,\text{N.m.s.rad}^{-1}$	$\xi = 0.6$
$k_\text{dt} = 2.7\times10^6\ \text{N.m.rad}^{-1}$	$J_\text{g} = 390\,\text{kg.m}^2$	$J_\text{r} = 55\times10^6\ \text{kg.m}^2$	$\mu_\text{dt} = 0.97$
$\omega_\text{n} = 11.11\text{rad.s}^{-1}$			

and, consequently, the derived matrix elements are

$a_{12} = 1.4840\times10^{-2}$	$a_{13} = -49.0909$	$a_{21} = 0.0203$
$a_{23} = 7.0688\times10^{4}$	$a_{22} = -0.1171$	$a_{32} = -0.0105$
$a_{44} = -13.3320$	$a_{45} = -123.4321$	$a_{66} = -50.0$

Finally, supposing the constant speed wind profile, then the matrix element a_{11} is defined as:

$\beta = 0.05 \quad \lambda_\text{s} = 2.4 \quad \omega_\text{r} = 1.25 \quad C_\text{q} = 0.25$

$$a_{11} = 1.0499$$

To obtain parameters of the discrete-time system model (1), (2), the continuos-time representation was converted to the discrete-time form by the standard Matlab function using the sampling period $t_s = 0.02\,s$ and the resulted equations were used for generating the input and output data to construct the instrumental variable in the simulation.

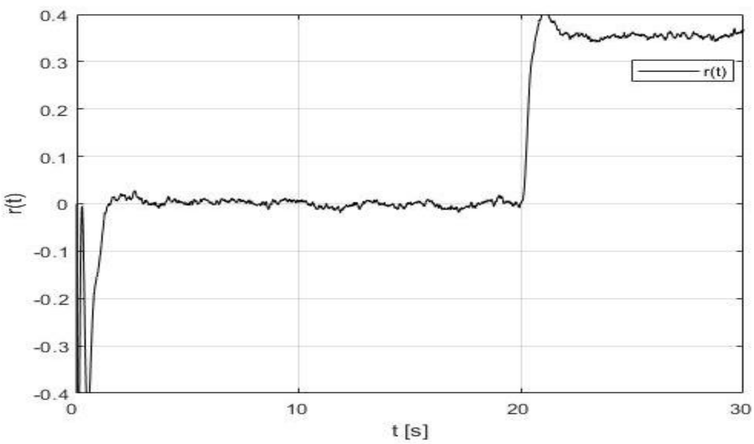

Figure 1

Time response of the residual filter on the first actuator fault

As the result Figure 1 presents the fault residual response reflecting a step-like fault in the gain loss of the first actuator at the time $t_f = 20\,s$. The FDI was trained according to the algorithm discussed in the paper and the result is analyzed in the term of detection threshold. In order to generate a suitable training data set with the proportion balance of the faulty-free and the faulty event, there were set both interval parameters as $r = s = 8$ and the value $p = 12$.

The system model was extended by the system noise input vector:

$$\mathbf{V}^T = \begin{bmatrix} 0.693 & 0.397 & 0.872 & 0 & 0 & 0 \end{bmatrix}$$

while the system noise $v(i)$ obeys normal distribution with zero mean and standard deviation $\sigma = 0.3$.

These examples illustrate the power that can be invoked through the prescribed method properties.

Conclusions

In the paper, a benchmark model is used for building and simulation testing a fault detection scheme, based on the subspace method data-driven principle, applicable to wind turbine benchmarks. To construct the reference residual from fault-free data, the discrete-time system model parameters are used for the prediction data set generation. Exploited data-driven methodology has sufficient adaptability in correspondence to different system operating points. Simulation case indicates that the residual performance is satisfactory.

Compared with the model-based fault detection method, prior knowledge of complex system model is not needed and only the parameterized matrices, need to be identified. Modifications of this framework to obtain hybrid solutions, combining neural networks and fuzzy inference system approaches, are the subject of future research in this algorithmic field, since the hybrid input-output representations are challenging research works in the wind turbine diagnosis. It is worthwhile to point out here, that the proposed method can be simply adjusted to disturbances or noise [23], which would be welcome applications in the wider industrial systems.

Acknowledgement

The work presented in this paper was supported by VEGA, the Grant Agency of Ministry of Education and Academy of Science of Slovak Republic, under Grant No. 1/0483/21. This support is very gratefully acknowledged.

References

[1] R. K. Mehra and J. Peschon, “An innovations approach to fault detection and diagnosis in dynamic systems," Automatica, 1971, Vol. 7, pp. 637-640

[2] A. S. Willsky, “A survey of design methods for failure detection in dynamic systems," Automatica, 1976, Vol. 12, pp. 601-611

[3] D. M. Himmelblau, Fault Detection and Diagnosis in Chemical and Petrochemical Processes, 1978, Amsterdam, Elsevier

[4] V. Venkatasubramanian, R. Vaidyanathan, and Y. Yamamoto, "Process fault detection and diagnosis using neural networks - I. Steady-state processes," Computers & Chemical Engineering, 1990, Vol. 14(7), pp. 699-712

[5] Z. Wang, Y. Wang, and Z. Ji, Advances in Fault Detection and Diagnosis Using Filtering Analysis, 2022, Singapore, Springer Nature

[6] D. Du, S. Xu, and V. Cocquempot, Observer-Based Fault Diagnosis and Fault-Tolerant Control for Switched Systems, 2021, Singapore, Springer Nature

[7] S. X. Ding, Advanced Methods for Fault Diagnosis and Fault-tolerant Control, 2021, Berlin, Springer Nature

[8] Z. Chen, Data-Driven Fault Detection for Industrial Processes. Canonical Correlation Analysis and Projection Based Methods, 2017, Wiesbaden, Springer Nature

[9] J. Zhao, W. Wang, and C. Sheng, Data-Driven Prediction for Industrial Processes and Their Applications, 2018, Cham, Springer Nature

[10] J. Wang, J. Zhou, and X. Chen, Data-Driven Fault Detection and Reasoning for Industrial Monitoring, 2022, Singapore, Springer Nature

[11] G. Niu, Data-Driven Technology for Engineering Systems Health Management. Design Approach, Feature Construction, Fault Diagnosis, Prognosis, Fusion and Decision, 2017, Singapore, Springer Nature

[12] Q. Jia, W. Chen, Y. Zhang, and Y. Jiang, "Robust fault reconstruction in discrete-time Lipschitz nonlinear systems via Euler-approximate proportional integral observers," Mathematical Problems in Engineering, 2015, Vol. 2015, pp. 1-14

[13] D. Krokavec and A. Filasová, "Descriptor approach in design of discrete-time fault estimators," WSEAS Transactions on Systems and Control, 2018, Vol. 13, pp. 481-490

[14] W. Chen, S. X. Ding, A. Haghani, A. Naik, A. Q. Khan, and S. Yin, "Observer-based FDI schemes for wind turbine benchmark," IFAC Proceedings Volumes, 2011, Vol. 44(1), pp. 7073-7078

[15] P. F. Odgaard and J. Stoustrup, "Fault tolerant control of wind turbines using unknown input observers," IFAC Proceedings Volumes, 2012, Vol. 45(20), pp. 313-318

[16] P. F. Odgaard and J. Stoustrup, "Fault tolerant wind farm control - a benchmark model," Proceedings of the IEEE International Conference on Control Applications CCA 2013, Hyderabad, India, 2013, pp. 412-417

[17] S. Simani, S. Farsoni, and P. Castaldi, "Wind turbine simulator fault diagnosis via fuzzy modelling and identification techniques," Sustainable Energy, Grids and Networks, 2015, Vol. 1(1), pp. 45-52

[18] X. Liu, Z. Gao, and M. Z. Q. Chen, "Takagi-Sugeno fuzzy model based fault estimation and signal compensation with application to wind turbines," IEEE Transactions on Industrial Electronics, 2017, Vol. 64(7), pp. 5678-5689

[19] D. Krokavec and A. Filasová, "One approach to design the fuzzy fault detection filters for Takagi-Sugeno models," Advanced and Intelligent Computations in Diagnosis and Control, 2016, Cham, Springer, pp. 19-33

[20] D. Krokavec and A. Filasová, "Model-based fault diagnosis of wind turbines built on Takagi–Sugeno fuzzy observers," Proceedings of the 10th International Scientific Symposium Elektroenergetika 2019, Stara Lesna, Slovakia, 2019, pp. 377-382

[21] P. F. Odgaard, J. Stoustrup, and M. Kinnaert, "Fault-tolerant control of wind turbines. A benchmark model,“ IEEE Transactions on Control Systems Technology, 2013, Vol. 21(4), pp. 1168-1182

[22] J. F. Manwell, J. G. McGowan, and A. L. Rogers, Wind Energy Explained. Theory, Design and Application, 2009, New York, John Wiley & Sons

[23] H. Sanchez, T. Escobet, V. Puig, and P. F. Odgaard, "Fault diagnosis of an advanced wind turbine benchmark using interval-based ARRs and observers," IEEE Transactions on Industrial Electronics, 2015, Vol. 62(6), pp. 3783-3793

[24] S. Simani, S. Farsoni, and P. Castaldi, "Data–driven techniques for the fault diagnosis of a wind turbine benchmark," International Journal of Applied Mathematics and Computer Science, 2018, Vol. 28(2), pp. 247-268

[25] D. Krokavec and A. Filasová, "Data driven approaches in fault detection of wind turbines," Proceedings of the 11th International Scientific Symposium Elektroenergetika 2022, Stara Lesna, Slovakia, pp. 126-130