

Support Vector Machines for Fault Detection in Wind Turbines

Nassim Laouti, Nida Sheibat-Othman, Sami Othman*

**Université de Lyon, Université Lyon 1, CNRS, CPE Lyon, UMR 5007, Laboratoire d'Automatique et de Génie des Procédés (LAGEP), F-69616 Villeurbanne, France
(*Tel. 33.4.72.43.18.88; e-mail: othman@lagep.univ-lyon1.fr)*

Abstract: Support Vector Machines (SVM) are used for fault detection and isolation in a variable speed horizontal-axis wind turbine composed of three blades and a full converter. The SVM approach is data based and is therefore robust to process knowledge. Moreover, it is based on structural risk minimization which enhances generalization and it allows accounting for process non linearity by using flexible Kernels. In this work, a radial basis function was used as Kernel. Different parts of the process were investigated including actuators, sensors and process faults. With duplicated sensors, we could detect sensor faults in blade pitch positions, generator and rotor speeds rapidly. Fixed value fault were detected in 2 sample periods and offset faults could be detected for $\Delta\beta \geq 0.5^\circ$ with a detection time that depends on the offset level. The converter torque fault (an actuator) could be detected within two sample periods. Faults in the actuators of the pitch systems could not be detected. Faults in the process concerning friction in the drive train could be detected only for very high offset ($\Delta\eta_{dt} \geq 50\%$).

Keywords: Fault detection, Support vector machines, Wind turbines.

1. INTRODUCTION

Methods used for fault diagnosis can be classified as model based or data based. Model based methods require a comprehensive model of the system. Success of data based methods is conditioned by the significance of historical data and the mathematical method used to detect the patterns in data. For industrial systems where an important amount of data is stored regularly and process model is not available, the use of statistical methods is preferred.

Among statistical methods for fault detection and diagnosis appear artificial neural networks (Schlechtingen and Ferreira Santos 2011), principal component analysis (Sun et al. 2005) and more recently support vector machines (SVM). SVM are based on structural risk minimization principle based on the statistical learning theory introduced in 1964 by Vapnik and Chervonenkis. Only recently, SVM were introduced as machine learning algorithms for classifying data from two different classes (Boser et al. 1992, Vapnik in 1995). Basically, a binary support vector classifier constructs a separating hyperplane. The hyperplane should have the maximum margin which is the width up to which the boundary can be extended on both sides before it hits any data point. These contact points are called the support vectors. In order to allow classifying non linearly separable sets, a nonlinear Kernel function can be used. The main differences between SVM and many other statistical methods are therefore: first, the structural risk minimization (training by traditional classifiers usually minimizes only the empirical risk) that improves the ability of generalization even with a reduced number of samples and avoids over-fitting in view of good parameter tuning; Second, SVM use nonlinear Kernels which allows separation of non linearly separable data.

SVM have been extensively used to solve classification problems in many domains ranging from face, object and text detection and categorization, information and image retrieval and so on. Their use for fault detection started in 1999 and was found to improve the detection accuracy. Widodo and Yang (2007) presented a review about the use of SVM for fault detection. They reported 37 papers in academic journals on this subject. Nowadays, the number of journal papers using SVM for fault detection has almost doubled. The concerned domains are in majority restricted to mechanical machinery as for instance roller bearings, gear box, power transmission system, induction motors, turbo pump rotor but are also extended to other domains such as electro-mechanical machinery, semi-conductors, refrigeration system, sheet metal stamping, air conditioning systems, and few chemical processes such as the Tennessee Eastman benchmark.

In this work, SVM are used for fault detection in a wind turbine that is used to generate electrical energy from the wind energy. A specific kind of turbines was simulated and controlled by Odgaard *et al.* (2009). The proposed benchmark is used for fault detection in this work. Even though the wind turbine functionality might be similar to rotating machinery, it encloses a number of difficulties ranging from a high variability in the wind speed, aggression by the environment, measurement difficulties besides the fact that wind turbines are supposed to run continuously for several years.

With the widespread use of wind turbines as renewable energy systems, control and supervision should be included in the system design. Fault detection of wind turbines allows reducing of maintenance costs. Indeed, online supervision of

all parts of the system allows early detection of faults which avoids degradation of the material and other side effects. Also, online supervision can suggest the best maintenance time as a function of the wind speed in order to ensure high performance. Fault detection is also interesting for control reconfiguration in order to ensure optimal power in case of partial fault. Note however that only few works treat this subject (Amirat *et al.* 2009, Hameed *et al.* 2009).

In the first part of this work, basic hints about SVM classification are given. Thereafter, the wind turbine is described and the locations and types of faults are defined. Then SVM learning is presented showing the different tuning levels. Finally, SVM validation is considered through simulation results using a real wind sequence.

2. SVM Classification

Consider N training vectors $x_i \in \mathfrak{R}^p$ characterized by a set of p descriptive variables $x_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\}$ and by the class label $y_i \in \{-1, +1\}$. For nonlinearly separable data x , the data can be mapped by some nonlinear function $\phi(x)$ into a high-dimensional feature space where linear classification becomes possible. Rather than fitting nonlinear curves to the data, SVM handle this by using a kernel function $K(x_i, x) = \langle \phi(x_i), \phi(x) \rangle$ to map the data into a different space where a hyperplane can be used to do the separation. The optimization problem is given by:

$$\min_{w,b} \frac{1}{2} \|w\|^2$$

Subject to:

$$y_i f(x_i) = y_i (\langle w, \phi(x_i) \rangle + b) \geq 1, \quad i = 1, \dots, N \quad (1)$$

Using the Lagrange function, the optimization problem is solved giving the following decision function:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x) + b \quad (2)$$

With the property:

$$w = \sum_{i=1}^N \alpha_i y_i \phi(x_i)$$

Where b is the bias term (a scalar) and $\alpha_i \geq 0$ are the Lagrange multipliers. It is important then to define a threshold for $f(x)$ (usually 0) to allow decision making. A slack variable can be introduced into eq. 1 to relax the margin constraints and allow misclassification of a controlled part of data.

The Gaussian kernel (a Radial Basis Function) with the variance σ is used in this work for data mapping:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (3)$$

3. Wind turbine description

A horizontal axis variable speed turbine composed of three blades is considered in this work (Odgaard *et al.* 2009). The

system contains a full converter coupled to a generator that allows converting the mechanical energy to electrical energy. A drive train is used to increase the rotational speed from the rotor to the generator.

The system is equipped with duplicated sensors to measure the three pitch positions ($\beta_{k,mi}$, $k=1, 2, 3$, $i=1, 2$) and the speeds of the generator and rotor ($\omega_{g,mi}$, $\omega_{r,mi}$, $i=1, 2$). This gives a total of ten sensors all subject to two kinds of faults: fixed value and offset (see table 1). Twenty faults are therefore to be detected with a detection time (T_D) that is less than 10 times the sampling time ($T_s=0.01s$).

As a function of the wind speed, a control system allows controlling the aerodynamics of the turbine to get the optimal power. The actuators are the three pitch systems and the converter. They allow respectively pitching the blades and setting the generator torque to control the rotational speed of the generator and the rotor. These actuators are also subject to fault. The converter system that sets the generator torque might have an offset that should be detected rapidly ($T_D < 5T_s$). The three pitching systems might have a change in the dynamics that can be due to abrupt change in the hydraulic system (5a) or to high air content in the oil at a slower rate (5b). In this case, the total number of actuator faults is seven. Finally, a system fault might occur in the driving train due to friction changes with time that might break down the train. The total number of faults to be supervised is therefore 28. However, it can be seen that some faults are similar. For instance, a fault of the sensor measuring blade position $\beta_{1,mi}$ is similar to those in the sensors measuring positions of blades 2 and 3 under the same conditions. By this way, it can be seen that we have ten different kinds of faults to be considered distinctly as classified in table 1. The process has other sensors, measuring for instance the wind speed, that are not supervised in this work.

The benchmark allows simulating the wind turbine control under normal operation (zone II: power optimization and zone III: constant power production). Fault detection will be studied using the closed-loop simulation in these zones with a real measured sequence of wind of 4400s.

The model of the turbine is given in Odgaard *et al.* (2009). It is nonlinear and the measurements are noisy. Note also the switching control structure.

Let us recall the pitch system and converter models that we will explicitly refer to in the fault scenarios. The pitch system is hydraulic and can be modelled by a second order transfer function:

$$\frac{\beta_{k,m}(s)}{\beta_k^d(s)} = \frac{w_n^2}{s^2 + 2\zeta w_n s + w_n^2} \quad (4)$$

Where $\beta_{k,m}(s)$ and $\beta_k^d(s)$ are the measured and desired positions of pitch $k=1, 2, 3$ and $[w_n, \zeta] = [11.11, 0.6]$ are the model parameters.

The converter dynamics can be modeled by a first order transfer function:

$$\frac{\tau_g^m(s)}{\tau_g^d(s)} = \frac{1}{\tau_g s + 1} \quad (5)$$

Where τ_g^m and τ_g^d are the real and desired generator torques, and $\tau=0.02s$. The real torque being non measured, it is calculated from the measured generator speed $\omega_{g,mi}$.

4. SVM for fault detection in wind turbines

Fault detection by SVM is developed in two parts. First of all, a set of measurement data with and without fault is used to learn models for detection of each fault (using the given wind sequence as an input). The obtained models are then validated in a new fault scenario.

4.1 SVM Learning

The key step in learning a new model for fault detection by SVM is the definition of the vector x to be used for classification. This vector should contain the most pertinent information on the behavior of the system. It should not be limited to the measurement output. It can include the inputs, the set-points, combination of those or variation of the outputs with time. In order to build a useful vector, one should carefully observe the process outputs for each fault and propose a combination that ensures a sufficiently high impact of the considered fault in x . Using some statistical analysis such as principal component analysis or partial least square can be useful for pretreatment.

Different vectors were proposed for the different kinds of faults, but the Kernel used for learning all the faults is Gaussian (with different values of variance). Most of the data is filtered (filtered data is noted with a hat, ^) using a first order filter with a time constant τ in order to reduce the sensitivity to process disturbances or measurement noise.

Detection/isolation of fault sensors of pitch position

For the 6 sensors measuring the pitch positions ($\beta_{k,mi}$, $k=1, 2, 3, i=1, 2$), the variance σ is regulated at 10.

Type 1a:

For faults of type 1a, the following vector is used for detection and isolation:

$$x = \begin{bmatrix} |\hat{\beta}_{k,m1}(t_j) - \hat{\beta}_{k,m2}(t_j)| \\ |\beta_{k,m1}(t_j) - \beta_{k,m1}(t_{j-1})| \\ |\beta_{k,m2}(t_j) - \beta_{k,m2}(t_{j-1})| \end{bmatrix} \quad (6)$$

Where t_j and t_{j-1} are the time instance j and $j-1$ respectively and $\hat{\beta}$ is filtered using $\tau=0.06s$. Note that absolute values are used in x . When $|\beta_{k,mi}(t_j) - \beta_{k,mi}(t_{j-1})| = 0$, this term is replaced by a large constant value (5000) in order to enhance distinguishability between the fixed value fault and normal case (no fault) where these values oscillate between 1×10^{-2}

and 2.

Type 1b:

This fault is detected and isolated in two steps. First of all, the fault is detected using the following equation:

$$x = \begin{bmatrix} |\hat{\beta}_{k,m1}(t_j) - \hat{\beta}_{k,m2}(t_j)| \\ |\hat{\beta}_{k,m1}(t_j) - \hat{\beta}_{k,m1}(t_{j-1})| \\ |\hat{\beta}_{k,m2}(t_j) - \hat{\beta}_{k,m2}(t_{j-1})| \end{bmatrix} \quad (7)$$

The second and third lines in 7 are important in order to exclude faults of type 1a. In a second step, if a fault of type b is detected, for isolation between sensors 1 and 2, the following vector is used:

$$x = \begin{bmatrix} |\hat{\beta}_{k,r}(t_j) - \hat{\beta}_{k,m1}(t_j)| \\ |\hat{\beta}_{k,r}(t_j) - \hat{\beta}_{k,m2}(t_j)| \end{bmatrix} \quad (8)$$

Where $\beta_{k,r}$ is the desired value of the pitch angle β_k and $\hat{\beta}$ is filtered using $\tau=0.08s$.

Detection/isolation of fault sensors of generator and rotor speeds

For sensor faults of the speeds of the generator and rotor ($\omega_{g,mi}, \omega_{r,mi}, i=1, 2$), the Gaussian variance is tuned at $\sigma=15$ in order to increase the ability of detection. Note however that very high variance values might lead to false alarms.

Types 2a and 3a:

For faults of type a, the following vector is used for detection and isolation:

$$x = \begin{bmatrix} |\hat{\omega}_{p,m1}(t_j) - \hat{\omega}_{p,m2}(t_j)| \\ |\omega_{p,m1}(t_j) - \omega_{p,m1}(t_{j-1})| \\ |\omega_{p,m2}(t_j) - \omega_{p,m2}(t_{j-1})| \end{bmatrix}, p = g, r \quad (9)$$

$\hat{\omega}_g$ is obtained using a filter with $\tau=0.02s$ and $\hat{\omega}_r$ using $\tau=0.6s$.

Types 2b and 3b:

For the detection of faults of type b (excluding faults of type a), the following vector is used:

$$x = \begin{bmatrix} |\hat{\omega}_{p,m1}(t_j) - \hat{\omega}_{p,m2}(t_j)| \\ |\hat{\omega}_{p,m1}(t_j) - \hat{\omega}_{p,m1}(t_{j-1})| \\ |\hat{\omega}_{p,m2}(t_j) - \hat{\omega}_{p,m2}(t_{j-1})| \end{bmatrix}, p = g, r \quad (10)$$

In a second step, isolation between sensors 1 and 2 in case of fault of type b is done using the following vector:

$$x = \begin{bmatrix} \frac{\hat{\tau}_g^m \times \hat{\omega}_{p,m1}(t_j)}{\hat{P}_g^m} \\ \frac{\hat{\tau}_g^m \times \hat{\omega}_{p,m2}(t_j)}{\hat{P}_g^m} \end{bmatrix}, p = g, r \quad (11)$$

Where, P_g^m is the measured power of the generator. The

measurements are filtered with $\tau=0.06s$ for the estimation of faults of $\hat{\omega}_g$ and using $\tau=0.6s$ for $\hat{\omega}_r$.

Detection/isolation of the converter actuator fault and system fault

For faults (4a and 6), the following vector is used:

$$x = \begin{bmatrix} |\omega_{g,m1}(t_j) - \omega_{g,m2}(t_i)| \\ \|\tau_g^d(t_j) - \tau_g^m(t_j)\| \\ \left| \lambda_2 \times |\omega_g^d(t_j) - (\omega_{g,m1}(t_j) + \omega_{g,m2}(t_i)) / 2| \right| \end{bmatrix} \quad (12)$$

Where ω_g^d is the desired generator speed, calculated from the desired generator torque τ_g^d obtained by the controller (P_r / τ_g^d , with P_r the desired power). The factor $\lambda_2 = 10^{-10} \times v_{wind}^6$ in the 3rd component of x is used to take into account the wind speed and for normalization. Note that τ_g^d is also filtered using a first order filter with a time constant $\tau=0.02s$. The objective of this filter is to take into account the dynamic of the control system (time necessary for τ_g^m to attain τ_g^d , (see eq. 5) and not to reject measurement noise or disturbances. The variance corresponding to x in 12 is $\sigma=10$ for fault type 4a and $\sigma=200$ for 6.

Detection/isolation of the pitch position actuator fault

For the detection of faults 5a and 5b, the following vector is used with $\sigma=10$:

$$x = \begin{bmatrix} |\omega_{g,m1}(t_j) - \omega_{g,m2}(t_i)| \\ |\beta_{k,m1}(t_j) - \beta_{k,m2}(t_j)| \\ |\beta_{k,m1}(t_j) - \beta_{k,m1}(t_{j-1})| \\ \|\beta_{k,m2}(t_j) - \beta_{k,m2}(t_{j-1})\| \end{bmatrix} \quad (9)$$

Once the learning vectors are defined for each fault, different fault scenarios are simulated and each sample is attributed $y=+/-1$ (with or without fault) for each kind of faults. About six scenarios were considered for each fault with different amplitudes. The SVM learning algorithm uses the outputs (x) and the corresponding y values to identify α_i and the support vectors (x_i) to be used in eq. 2 for decision making. Identification depends on the Kernel type and tuning parameters: σ and the slack variable. Note that the same “model” (x_i and α_i) is used for all faults of type 1a, another “model” for all faults of type 1b and so on. Ten “models” were therefore developed.

4.2 SVM validation

Let us consider the following scenario that we simulate using the wind sequence given in Odgaard *et al.* (2009):

1. Fault type 1a, $\beta_{1,m1} = -3^\circ$ (fixed value, stuck) occurring between 100s and 200s.
2. Fault type 1b, $\beta_{2,m2} = 5 \times \beta_{2,m2}$ (gain factor) on 3200-

3300s.

3. Fault type 1a, $\beta_{3,m1} = 7^\circ$ on 900-1000s.
4. Fault type 2a, $\omega_{r,m1} = 2 \text{ rad.s}^{-1}$ on 1200-1300s.
5. Faults type 2b and 3b, $\omega_{r,m2} = 0.5 \times \omega_{r,m2}$ and $\omega_{g,m1} = 1.5 \times \omega_{g,m1}$ on 1700-1800s.
6. Fault type 4a, $\tau_g = \tau_g - 1000 \text{ Nm}$ on 4200-4300s.
7. Fault type 6, $\eta_{dt} = 0.22 \times \eta_{dt}$ on 300-500s.
8. Fault type 5a, parameters in pitch actuator 2 (w_n, ζ) abruptly changed from [11.11, 0.6] to [5.73, 0.45] from 3200 and 3300s.
9. Fault type 5b, parameters in pitch actuator 3 (w_n, ζ) changed slowly (with a linear function) from [11.11, 0.6] to [3.42, 0.9] over 30s, remained constant during 40s, and then decreased again over 30s from 3400 and 3500s.

Fixed value faults of the pitch position could be detected in the required time (see Table 1) in both controller zones easily (faults n° 1 and 3).

Fig. 1 shows the estimation results of fault of type 1b (gain factor) for the same sensor (pitch position). Only faults with an offset $\Delta\beta \geq 2^\circ$ can be detected during the required detection time ($<10 T_s$). For an offset of 1.5, the detection time is about $10 T_s$ depending on the control phase. If a gain factor is applied rather than an offset, oscillations might increase if the order of β is high but the detection capacity remains equivalent. Of course if β is close to 0, the gain factor does not introduce a fault.

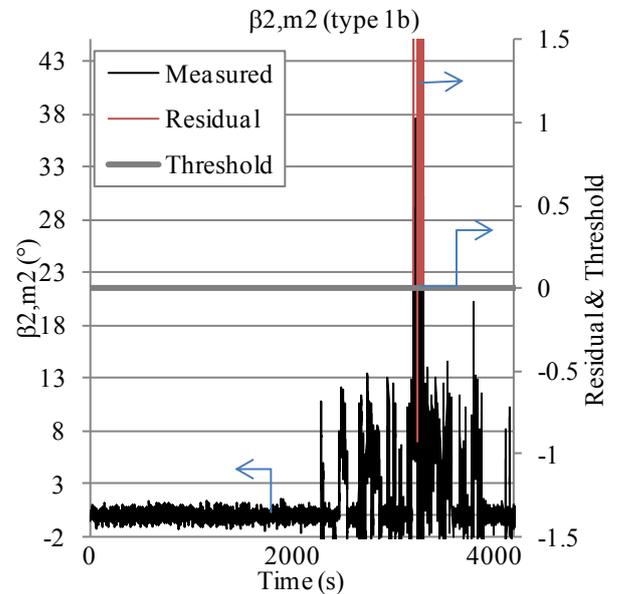


Fig. 1. Fault detection and isolation of pitch position (fault n°2, type 1b).

Fig. 2 indicates the occurrence of a fixed value fault in sensor $\omega_{r,m1}$ (the rotor speed). It can be seen that it is achieved instantaneously without difficulty.

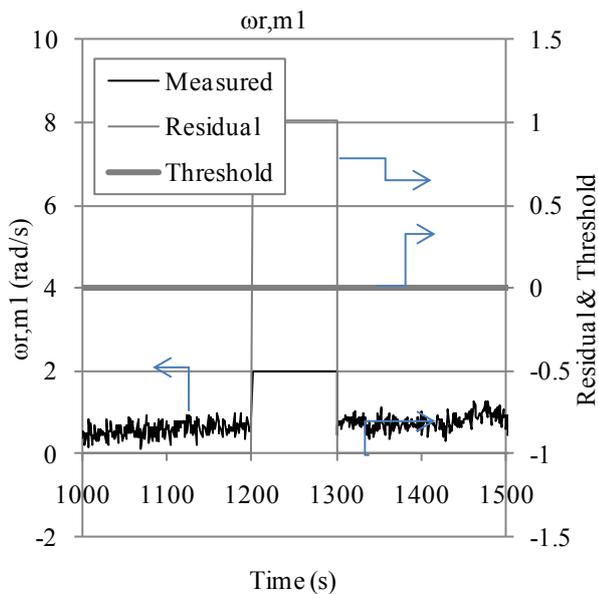


Fig. 2. Fault detection and isolation of the rotor speed (fault n°3, type 2a).

Faults in the rotor speed sensor could be detected rapidly for $\Delta\omega_{r,m2} > 0.8 \text{ rad}\cdot\text{s}^{-1}$ as shown by Fig. 3 where a gain factor error is applied. The objective is however to detect 10% error in this sensor, therefore approximately $\Delta\omega_r > 0.13$ which was not possible with the obtained model. Probably, this can be achieved by introducing more data in the learning step and adapting an adequate filtering method.

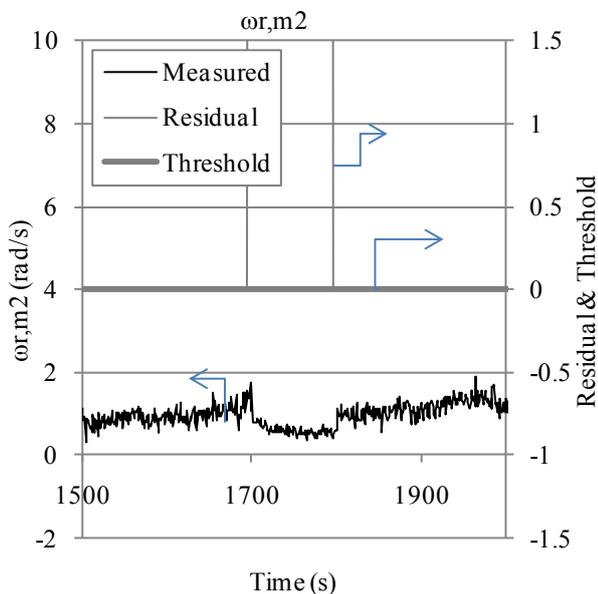


Fig. 3. Fault detection and isolation of the rotor speed (faults n°5, type 2b).

Fig. 4 shows the fault detection results of the sensor of the generator speed $\omega_{g,m1}$. In case of a gain factor error leading to $\omega_g > 50 \text{ rad}\cdot\text{s}^{-1}$, an important change in the residual occurs rapidly.

Concerning the estimation of the rotator torque speed (an actuator), it could be detected as required in terms of fault

level and rapidity (Fig. 5). Note that x uses the desired torque value that is compared to the measured one with 2 sample periods delay. This fault could be detected in both of the controller zones.

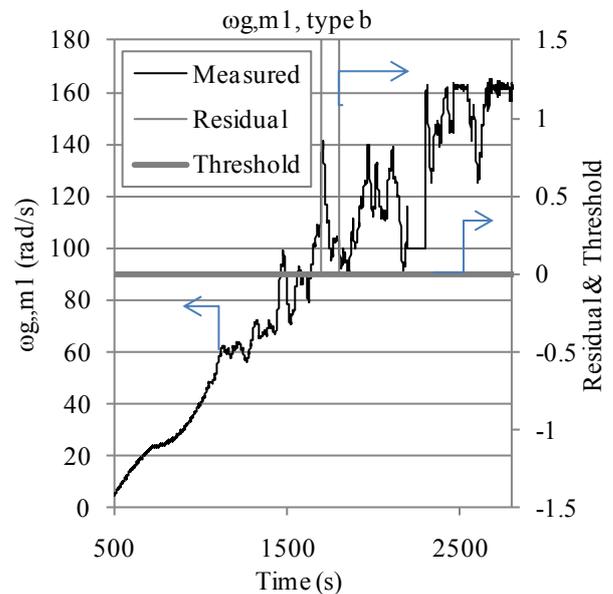


Fig. 4. Fault detection and isolation of the generator speed (faults n°5, type 3b).

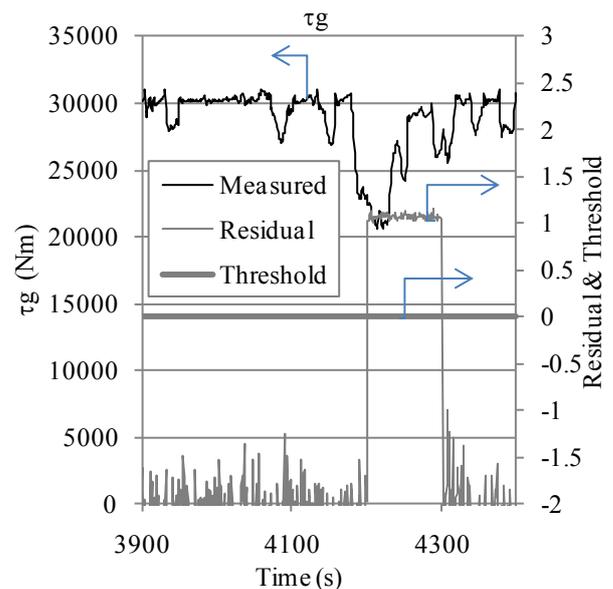


Fig. 5. Fault detection and isolation of the convertor torque (fault n°6, type 4a).

Concerning the actuators of the pitch positions, their faults could not be detected by the proposed vector x . Further investigation of this vector and parameter tuning should be done in order to extract eventual hidden information about these actuators among the measurements.

Finally, Fig. 6 shows fault detection of the system consisting of the drive train friction. This error is modelled by changing the values of the model parameter $\eta_{dt} = 0.22 \times \eta_{dt}$. However, the error could be detected only with much higher fault level

in this parameter, $\eta_{dt} = 0.95 \times \eta_{dt}$.

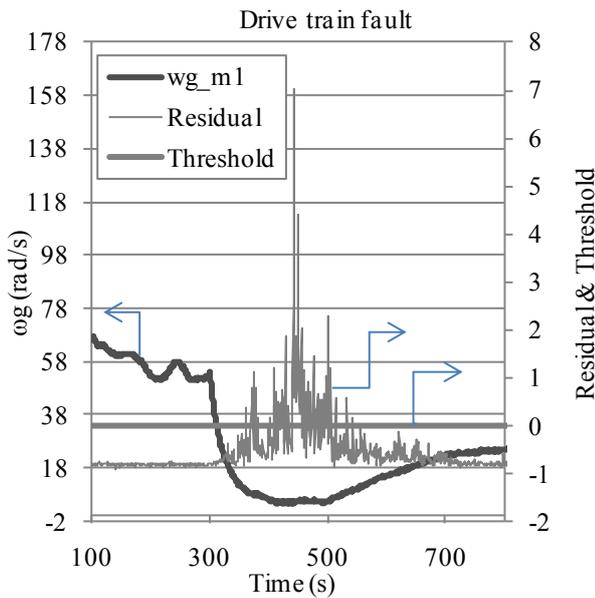


Fig. 6. Fault detection and isolation of the system (fault n°7, type 6).

Table 1. Fault detection results

N°	Fault	Fault level to detect	Fault detected & isolated	n	n ^{des}	
1a)	FV $\Delta\beta_{k,mi}$		Yes	2	<10	
1b)	Offset (or GF) $\Delta\beta_{k,mi}$	$1.2 \times \beta$ ($\Delta\beta \approx 1$)	Yes if $\Delta\beta \geq 0.5$	10 if $\Delta\beta = 1.5$		
2a)	FV $\Delta\omega_{r,mi}$		yes	2		
2b)	GF $\Delta\omega_{r,mi}$	$0.1 \times \omega_r$	Yes if $\geq 0.2 \times \omega_r$	67		
3a)	FV $\Delta\omega_{g,mi}$		Yes	2		
3b)	GF $\Delta\omega_{g,mi}$	$0.1 \times \omega_g$ ($\cong 10 \text{ rad.s}^{-1}$)	Yes if $\geq 0.1 \times \omega_g$	2		
4a)	Offset $\Delta\tau_g$	$\Delta\tau_g \geq 100$	Yes	2		<5
5a)	Abrupt $\Delta\beta_k$	$\Delta w_n \geq 0.5$ $\Delta\zeta \geq 0.25$	No	-		<8
5b)	Slow $\Delta\beta_k$	$\Delta w_n \geq 0.25$ $\Delta\zeta \geq 0.5$	No	-		<600
6)	Drive train, $\Delta\omega_r$, $\Delta\omega_g$	$\Delta\eta_{dt} \geq 5\%$	Yes if $\Delta\eta_{dt} \geq 50\%$	2835		Free

FV: Fixed value, GF: Gain factor, n: n° of sample periods and n^{des} the desired n° of sample periods for detection (note

that the sampling period is 0.01s).

5. CONCLUSIONS

The wind energy is profitable if the technology of the turbines is optimized and online supervised. In view of the large number of components in the system, high number of frequent but noisy measurements besides the system disturbances, a good statistical method should be used for fault detection and isolation. The SVM is found to be a good method for pattern recognition. A “model” is learned to detect all the sensors, actuators and system faults. Defining the input vector of the model as well as parameter tuning are primordial in order to detect and isolate the faults. A compromise between sensitivity to noise and fault detection is to be determined.

Most of the requirements for fault detection were realized. Faults of type 1a, 2a, 3a, 4a, 1b and 3b could be detected without further constraints. Faults n° 2b and 6 could be detected only with higher error levels than required. Finally, faults n° 5a and 5b could not be detected. However, further investigation might be necessary in order to improve the quality of the estimations mainly by improving the input data vector x.

REFERENCES

- Amirat, Y., Benbouzid, M.E.H., Al-Ahmar, E., Bensaker, B., and Turri, S. (2009). A brief status on condition monitoring and fault diagnosis in wind energy conversion systems. *Renewable and sustainable energy reviews*, 13, 2629-2639.
- Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the 5th Workshop on Computational Learning Theory*, 144-152.
- Hameed, Z., Hong, Y.S., Cho, Y.M., Ahn, S.H., and Song, C.K. (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and sustainable energy reviews*, 13, 1-39.
- Odgaard, P.F., Stoustrup, J., and Kinnaert, M. (2009). Fault tolerant control of wind turbines- a benchmark model. *7th IFAC symposium on fault detection, supervision and safety of technical processes*, 155-160.
- Schlechtingen, M., and Ferreira Santos, I. (2011). Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical Systems and Signal Processing*.
- Sun, X., Marquez, H.J, Chen, T., and Riaz, M. (2005). An improved PCA method with application to boiler leak detection. *ISA Transactions*, 44, 379-397.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*, Springer, New York.
- Vapnik, V. and Chervonenkis, A. (1964). A note on one class of perceptrons. *Automation and Remote control*, 25.
- Widodo, A., and Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 21, 2560-2574.