Intelligent Prognostic Framework for Degradation Assessment and Remaining Useful Life Estimation of Photovoltaic Module

Nabil Laayouj*, Hicham Jamouli & Mohamed El Hail

The Laboratory of Industrial Engineering and Computer Science (LGII), National School of Applied Sciences Ibn Zohr University Agadir, Morocco
*E-mail: nabil.laayouj@edu.uiz.ac.ma

Abstract. All industrial systems and machines are subjected to degradation processes, which can be related to the operating conditions. This degradation can cause unwanted stops at any time and major maintenance work sometimes. The accurate prediction of the remaining useful life (RUL) is an important challenge in condition-based maintenance. Prognostic activity allows estimating the RUL before failure occurs and triggering actions to mitigate faults in time when needed. In this study, a new smart prognostic method for photovoltaic module health degradation was developed based on two approaches to achieve more accurate predictions: online diagnosis and data-driven prognosis. This framework of forecasting integrates the strengths of real-time monitoring in the first approach and relevant vector machine in the second. The results show that the proposed method is plausible due to its good prediction of RUL and can be effectively applied to many systems for monitoring and prognostics.

Keywords: *monitoring; photovoltaic module; prognosis; relevant vector machine; remaining useful life.*

1 Introduction

The use of photovoltaic (PV) panels for electric energy production has rapidly increased in the last few decades. Several developed countries have set up an incentive system to encourage and accelerate the deployment of energy produced from PV sources. The need for renewable energies stems from the deterioration of the environment caused by the traditional sources of energy, and they have become a possibility with the advancement of science and technology. The prognostic and health management (PHM) of these products is becoming very urgent [1] in order to improve the position of PV panels available in the market as an eco-friendly source of energy. The prognosis decision is facilitated by the ability to accurately predict their remaining useful life (RUL) (Figure 1). Several studies on the RUL of PVs have already been published [2,3]. However, these are about specific environmental conditions [4] or a specific panel type [5]. We can also find works about emission of CO2/Kwh [6] and energy consumption [7].

Received June 16th, 2016, 1st Revision December 2nd, 2016, 2nd Revision December 30th, 2016, Accepted for publication December 30th, 2016.
 ISSN: 2337-5779, DOI: 10.5614/j.eng.technol.sci.2016.48.6.9

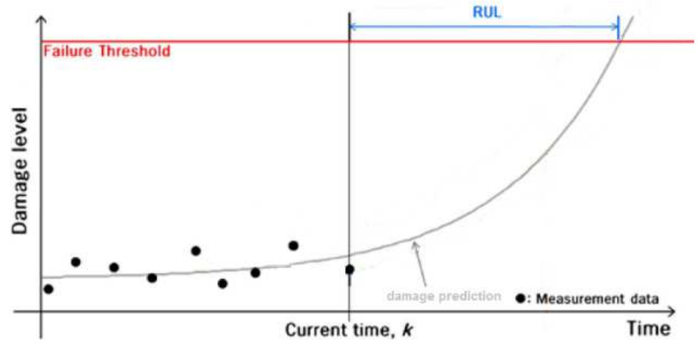

Figure 1 Illustration of RUL prediction.

In the scientific literature, prognostic techniques can be categorized into four approaches based on the usage of information, i.e. experience-based, model-based, data-driven and hybrid approaches. To estimate the parameters of the dominant reliability, the experience-based approach (or reliability-based approach) [8] uses data and knowledge gathered from experience during the system exploitation period. The works on Weibull [9] and lognormal laws [10] are of great importance. The main disadvantage of this approach is that it is difficult to have a past history of experiences representative of all systems' conditions of use.

The model-based approach [11] assumes that a model of system behaviour is available and uses this model to predict the future of the system behaviour. A number of recent developments in the model-based approach, such as Lumped parameter models [12] and functional models [13], have been reported in the literature. Thus far, model-based approaches are not suitable for many industrial applications, where physical models or mathematical models are difficult to build such that they precisely describe the mechanism of the system, which may vary under different operation conditions [14]. With respect to their accuracy, model-based approaches cannot be used for complex systems whose internal state variables are inaccessible to direct measurement using general sensors. Moreover, model-based approaches may not be the most practical approaches since the faults in question are often unique.

The data-driven approach [15] aims to collect information and data from routinely monitored components and projecting them to estimate the future evolution of degradation within the system. Compared with model-based approaches, data-driven approaches avoid developing high-level physical models of the system so they are less complex and less costly to develop. In the last few years there has been more research interest in data-driven methods for forecasting such as support vector machine (SVM) [16] and neural networks

(NN) [17]. Hybrid approaches [18] combine the data-driven and model-based approaches and includes Bayesian techniques [19]. The application of hybrid approaches is limited to cases in which sufficient knowledge of system physics-of-failure is available.

Complexity and inflexibility are the two main causes that limit the use of model-based and hybrid approaches. On the other hand, data-driven approaches can reflect an inherent relationships in which only the historical and monitoring data are required to provide an RUL estimation. At the same time, most data-driven methods are offline and therefore cannot hold online supervision. However, this results in unsatisfactory prediction accuracy. Considering the need for rapidly developed state monitoring and predicting, it is essential to propose an online RUL prediction approach.

To overcome the abovementioned difficulties, in this study a new prognostic model was developed that is a kind of mixture between two approaches for the estimation of RUL. The first approach is an online diagnosis method based on a condition-monitoring framework. It allows supervision of the critical parameters in a PV module. The second approach is a data-driven method based on the relevant vector machine (RVM) technique to predict degradation. This paper is organized as follows. Section 2 contains the new fusion prognostic framework. In Section 3, application and performance evaluation are described with a discussion of future works. Finally, Section 4 deals with the advantages and shortcomings of the suggested method.

2 Fusion Prognostic Framework for PV Module Life Forecasting

In this section, first we briefly discuss two principal components of the proposed prognostic framework: real-time monitoring and RVM prognosis. The fusion prognostic framework will then be described, so as to develop a more reliable RUL of PV module.

2.1 Real Time Monitoring (Online Diagnosis)

Before starting the prognostic process, we need to establish the current health of the system by making a diagnosis. This initial phase (diagnosis), employs pattern recognition and machine learning to detect changes in the system states [20]. During the last decade, there has been more research interest in methods of diagnosis such as condition monitoring [21]. The researchers also developed several model-based approaches for machinery diagnosis [22], such as the extended Kalman filter [23] or observer-based diagnosis [24].

Model-based diagnosis approaches require the development of an accurate mathematical model, which is not always possible, especially in the case of large-scale systems. The advantage of condition-monitoring techniques is their ability to transform high-dimensional noisy data into lower-dimensional information for diagnosis. But, the accuracy of the latter is highly dependent on the quantity and quality of the system's operational data.

Online condition monitoring determines where we are on the health curve**.** Is it ‘‘nominal”? Or are there some anomalies to have to be dealt with?

Today with new available technologies, the engineering of monitoring makes important data about a machine's condition available. However, such data are not effectively analysed due to the lack of efficient analysis methods. If we can establish a mechanism to analyse real-time data, the equipment condition can be observed and evaluated effectively and quickly, and this mechanism will be very useful in the development of a prognostic approach.

With the proposed model, we aim to develop a practical diagnostic tool that can be used to supervise the health condition of a system in real time by continuously monitoring the system's condition and using data anomaly detecting techniques. The online monitoring analyses operating data and compares it with a healthy baseline (threshold of diagnosis) to look for anomalies. When there is fault or anomaly, the diagnostic module generates an alarm and starts the process of RUL computing. The progress of diagnosis will be detailed in Section 2.3.

2.2 Relevance Vector Machine

Many data-driven approaches for prognosis have been reported in the literature, such as the auto regressive (AR) model [25], Gaussian process regression (GPR) [26], artificial neural networks (ANN) [27] and relevance vector machine (RVM) [28,29]. RVM is a Bayesian form representing a generalized linear model with a functional form identical to that of support vector machine (SVM). RVM is developed based on sparse Bayesian learning theory [30,31]. It differs from SVM in the case of solutions that provide a probabilistic interpretation of its outputs [32]. The strongest asset of RVM is its high learning ability, easy training, as well as prediction results through kernel and statistical probability learning. RVM has been successfully adopted in a variety of application fields, including renewable energy [33], mechanical fatigue [34] and battery life [35].

Consider a set of input vectors $\{x\}_{j=1...N}$ and their corresponding target vectors $\{t\}_{j=1...N}$. The goal is to learn a model of the dependency of the target vectors on the inputs in order to make an accurate prediction of t for an unseen value of x. We define function $F(x)$ in the input space to make a prediction with the learning process to identify the parameters of the function. In the context of SVM, $F(x)$ takes the form as shown in Eq. (1),

$$F(x) = \sum_{i=1}^{N} w_i K(x, x_i) + w_0 \tag{1}$$

where, $w = (w_1, w_2, ..., w_N)^T$ and $K(x, x_i)$ are weight vectors and kernel function respectively, while w_0 is bias. The output of the RVM model can be expressed as in Eq. (2),

$$t_j = F(x_j) + \varepsilon_j \tag{2}$$

where, ε_j are independent samples from some noise process (Gaussian with mean 0 and variance σ^2). The likelihood of the data set can then be written as:

$$p(t / w, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{ -\frac{1}{2\sigma^2} \|t - \Psi w\|^2 \right\} \tag{3}$$

where, Ψ is the $N \times (N+1)$ design matrix with $\Psi = [\Psi(x_1), \Psi(x_2), ..., \Psi(x_N)]^T$ and $\Psi(x_i) = [1, K(x_i, x_1), K(x_i, x_2), ..., K(x_i, x_N)]^T$. The maximum likelihood estimates of w and σ^2 from Eq. (3) will lead to severe overfitting [30]. To overcome this problem, a Bayesian perspective is adopted and constraints are superimposed on the parameters by choosing a zero-mean Gaussian prior probability distribution over w as provided by Eq. (4),

$$p(w/\alpha) = \prod_{i=0}^{N} N(w_i / 0, \alpha_i^{-1}) \tag{4}$$

where α is a vector of ($N+1$) hyper-parameters that controls how far from zero each weight is allowed to deviate [36]. The posterior distribution over all the unknowns can be computed based on Bayes' rule:

$$p(w, \alpha, \sigma^2 / t) = \frac{p\langle t | w, \alpha, \sigma^2 \rangle p(w, \alpha, \sigma^2)}{\int p\langle t | w, \alpha, \sigma^2 \rangle p(w, \alpha, \sigma^2) dw . d\alpha . d\sigma^2} \tag{5}$$

However, we cannot compute the solution of Eq. (5) directly since we cannot perform the normalizing integral in the denominator. Instead, we can decompose the posterior distribution as shown in Eq. (6),

$$p\left\langle w,\alpha,\sigma^2 \middle| t\right\rangle = p\left\langle w \middle| t,\alpha,\sigma^2\right\rangle p\left\langle \alpha,\sigma^2 \middle| t\right\rangle \tag{6}$$

For reasons of simplification, the posterior distribution of weight can be written as in Eq. (7),

$$p\left\langle w \middle| t,\alpha,\sigma^2\right\rangle = \frac{p\left\langle t \middle| w,\sigma^2\right\rangle p(w,\alpha)}{p\left\langle t \middle| \alpha,\sigma^2\right\rangle} \tag{7}$$

The posterior over the weight is then obtained from Bayes' rule in Eq. (8),

$$p\left\langle w \middle| t,\alpha,\sigma^2\right\rangle = (2\pi)^{-(N+1)/2}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(w-\mu)^T\Sigma^{-1}(w-\mu)\right\} \tag{8}$$

The posterior covariance and mean are shown in Eqs. (9) and (10),

$$\Sigma = (\Psi^T B\Psi + A)^{-1} \tag{9}$$

$$\mu = \Sigma\Psi^T Bt \tag{10}$$

where $A = diag(\alpha_1,...,\alpha_{N+1})$ and $B = \sigma^{-2}I$.

It is important to note that σ^2 is also treated as a hyper-parameter that may be estimated from the data. The sparse Bayesian learning can then be expressed as maximisation of the hyper-parameters' posterior distribution: $p\left\langle \alpha,\sigma^2 \middle| t\right\rangle \propto p\left\langle t \middle| \alpha,\sigma^2\right\rangle p(\alpha)p(\sigma^2)$. Therefore, the maximum a posteriori (MP) estimate of hyper-parameters needs only to maximize likelihood $p\left\langle t \middle| \alpha,\sigma^2\right\rangle$. The likelihood function (3) can be marginalized by integrating the weights as:

$$p\left\langle t \middle| \alpha,\sigma^2\right\rangle = (2\pi)^{-N/2}\left|B^{-1} + \Psi A^{-1}\Psi^T\right|^{-1/2}\exp\left\{-\frac{1}{2}t^T(B^{-1} + \Psi A^{-1}\Psi^T)^{-1}t\right\} \tag{11}$$

To complete the specification of this hierarchical prior, the hyper-parameters over α and σ^2 are approximated as delta functions at their most probable values α_{MP} and σ^2_{MP}. By maximizing the marginal likelihood of Eq. (11) as in [30] we can obtain the iterative re-estimation formulas as shown in Eqs. (12) and (13),

$$\alpha_i^{new} = \frac{1-\alpha_i \Sigma_{ii}}{\mu_i^2} \tag{12}$$

$$(\sigma^2)^{new} = \frac{\left\| t - \Psi \mu \right\|^2}{N - \sum_{i=0}^{N} (1-\alpha_i \Sigma_{ii})} \tag{13}$$

Having found the maximizing values α_{MP} and σ_{MP}^2, predictions for new data are then made according to Eq. (14) knowing that x_* is the new input.

$$p\left\langle t_* \middle| x_*, \alpha_{MP}, \sigma_{MP}^2 \right\rangle = \int p\left\langle t_* \middle| x_*, w, \sigma_{MP}^2 \right\rangle p\left\langle w \middle| t, \alpha_{MP}, \sigma_{MP}^2 \right\rangle dw \tag{14}$$

2.3 Description of the Proposed Methodology

The proposed model for prognosis aims to assess the machine's performance degradation and foretell the RUL through the use of diagnosis and prognosis in a manner that can take the strengths of each. To apply this method, an online monitoring process is first carried out to acquire the system's condition. The data obtained from the diagnosis can be properly managed and utilised by the RVM approach for making a prognosis. A flowchart of the combination between diagnosis and prognosis is shown in Figure 2. The suggested method consists of two steps. The role of each one can be summarized as follows.

The diagnostic routine starts with data acquisition using multi-sensors attached to the operating system. Then, embossed operating state features are extracted from the collected data (considering root mean square (RMS) as the monitoring indicator). This calculated feature (RMS) will be used to reflect the system's health and to supervise the level of anomaly using a predefined threshold of diagnosis. Next, an automatic alarm is triggered to prevent machinery performance degradation or malfunction. The alarms triggered in the diagnostic phase are dependent on the quality of feature extraction; the existence of non-robust features may lead to false or missed alarms when the monitoring tool cannot detect the existence of a system fault. The prognosis routine starts when the alarm is activated. After identifying the fault in the diagnosis step, the historical data are employed by calculating the feature. This feature can be used to represent the degradation evolution of the observed system.

The results are regarded as target vectors of degradation probability. RVM is then used for training and validation. The weights obtained from training and validation are used as predicted values of degradation. In order to evaluate the forecasting accuracy, a performance indicator is used, the mean absolute scaled error (MASE). Some data from the feature calculation unit can be used to test the performance of RVM after validation.

Definition of the MASE [37]:

A scaled error is defined as in Eq. (15),

$$Q_k = \frac{M_k - \hat{M}_k}{\frac{1}{N-1}\sum_{i=2}^{N}\left|M_i - M_{i-1}\right|} \quad , \tag{15}$$

with $N \geq 2$

where M is the real measurement, $\hat{M}$ is the predicted value and N is the number of predictions data set. The MASE in simplified form is shown in Eq. (16),

$$MASE = mean(\left|Q_k\right|) \tag{16}$$

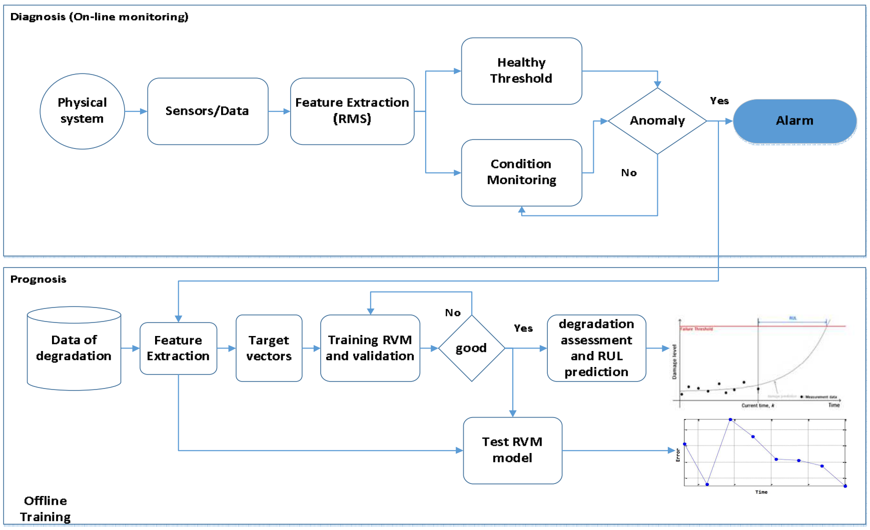

Figure 2 Flowchart of the integration of real time monitoring and prognostic (RTM-RVM).

3 Application and Performance Evaluation

The photovoltaic (PV) module consists of several PV cells that convert solar rays into direct current (DC). Degradation of a PV is the gradual deterioration of its characteristics, which may affect its performance and its ability to operate within the limits of desired criteria caused by the operating conditions. A degraded PV module may keep executing its primary function, which is to generate electricity from sunlight. However, the degraded state of the module can cause problems when the degradation exceeds a predefined critical threshold.

According to the literature, PV module performance can be degraded due to several factors such as: temperature [38], humidity [39], irradiation [40], corrosion [41], discoloration [42], and delamination [43]. The most important electrical characteristics of a PV module are the current voltage (I-V) [44] and power voltage (P-V) curves [45], short-circuit current (Isc) [46], open-circuit voltage (Voc), fill factor (FF) [47], and maximum power output (Pmax) [48].

Methods commonly applied by researchers to monitor and assess a module's electrical performance are current voltage (I-V) and power voltage (P-V) curve scanning. In general, the degradation of a photovoltaic module is assessed by measuring the power and therefore the power loss during its lifetime compared to its initial power.

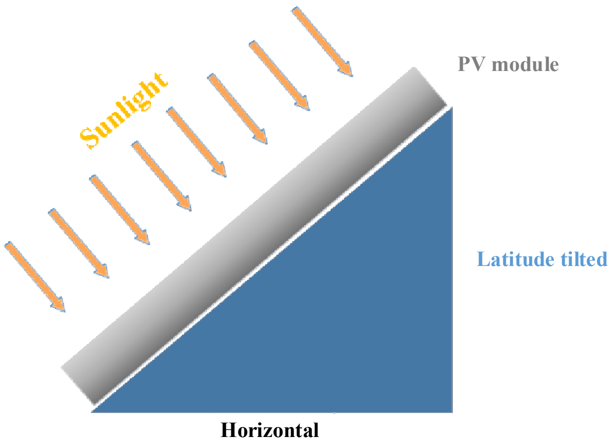

Figure 3 PV module installation.

In this study, we have analysed the performance data of one PV module collected over a period of 55 months. The PV module was mounted open-rack, open-circuit and latitude tilted as shown in Figure. 3. A temperature sensor attached to the back of the module provided the back-skin temperature. The maximum power output of the module was derived from an electrical performance test. The measurements were carried out every 50 days.

The effect of temperature on PV module performance is often neglected, but studies have shown that it cannot be ignored [49,50]. Although there is an increase in the current with temperature, the overall effect of increased temperature is a decrease in power due to the larger decrease in voltage. According to Mazer [51], the decrease of Voc with temperature is caused by the increase of the saturation current, which increases with intrinsic carrier concentration, which in turn is exponentially dependent on temperature. Figures 4 and 5 show the effect of temperature on Voc and Pmax, respectively.

Subsequently, we will evaluate the performance of the PV module by using the method from Section 2, taking into account the test measurement data.

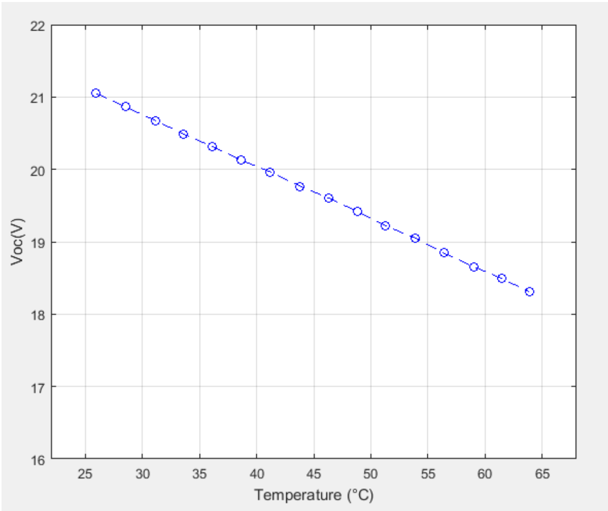

Figure 4 Temperature dependence of Voc of PV modules.

The PV module used in our study was a crystalline silicon (c-si) [52] with initial standard test condition (STC) parameters Pmax= 100.7 W, Voc = 22.84 V, Isc = 4.8A, FF = 66.8%, and η = 9.4%. STC refers to 1000 W/m^2 irradiance, 25 °C cell temperature, and AM 1.5 G spectrum. The data that were used to assess degradation and calculating the RUL were Voc and Pmax.

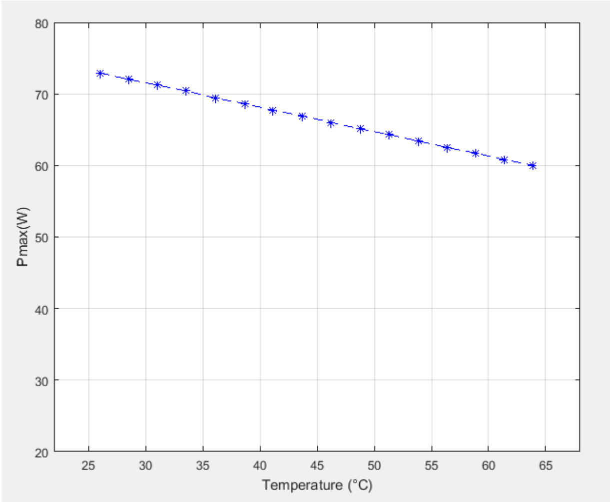

Figure 5 Temperature dependence on Pmax of PV module.

The collected data in the learning area were used to predict the degradation in the next twelve years (145 months) and to compare them with the real degradation values.

3.1 Diagnosis of PV Module

For machinery fault diagnostics and prognostics, different types of signals can be used. In this research, we used the voltage Voc data because they are readily available and the trend of voltage is closely related to PV degradation, as already indicated.

The data of voltage Voc from the PV were used for diagnosis in real time; measurements were taken in the nominal operating condition. A total of 43 measurements were collected during a period of 72 months for monitoring the health of the PV module. After measuring the voltage Voc, the health condition of the PV was checked to see if there was an anomaly or just noise in the signal. Confirmation of the existence of an abnormality was done by measuring the RMS of the difference between Vocr (open-circuit voltage reference) and Vocm (open-circuit voltage measured) measured in real time. The Vocr voltage is the value of Voc measured the first time after the installation of a new PV module (see Eq. (17)).

$$Voc_{RMS} = \left(\frac{1}{n} \sum_{t=1}^{n} (V_{ocr}(t) - V_{ocm}(t))^2 \right)^{1/2} \tag{17}$$

Figures 6 and 7 show the voltage Voc drop and the evolution of Voc_{RMS} respectively; the Voc decreases with usage and the saturation of current Isc. It was observed that during a period of 75 months the Voc_{RMS} gradually increased and at time 33 it reached the diagnosis threshold defined by the RTM, which will trigger an alarm and prognosis process. The trigger level of the alarm (diagnosis threshold) was set to $Voc_{RMS} = 1$.

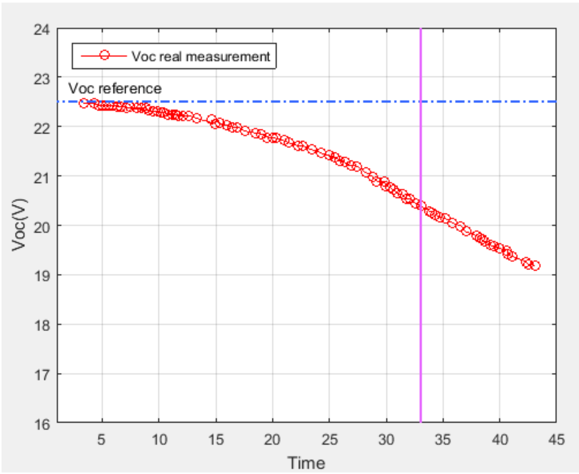

Figure 6 Real time monitoring of voltage Voc degradation.

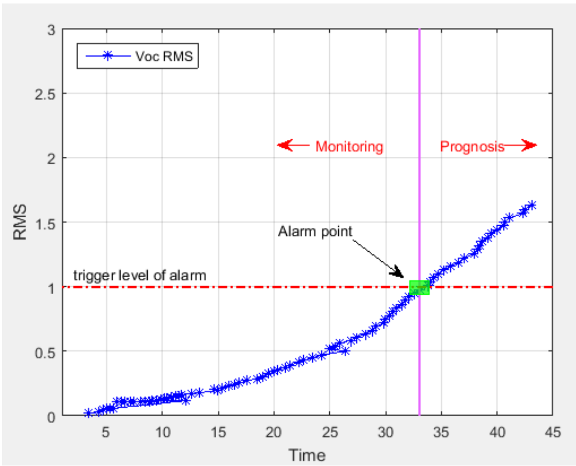

Figure 7 Online RMS evolution.

3.2 Prognosis of Degradation and RUL Estimation of PV Module

After validation of the diagnosis, the RVM model was trained by using the past data of Pmax. After the training process, which produced and saved weight and bias, the predictions results were obtained (degradation of Pmax). Two different prognostic thresholds were used, with a particular maintenance action for each. At time 70 degradation reached threshold 1 = 87 W and at time 87 degradation reached threshold 2 = 85 W, as shown in Figure 8.

Our method of prognosis was employed to make a specified number of predictions and determine the RUL of the PV module. By measuring the distance between the actual life cycle T_p and the predicted life cycle T_{rul} (at which the future degradation value hits the threshold of prognosis for the first time), we get an RUL distribution as shown in Eq. (18),

$$RUL = T_{rul} - T_p \tag{18}$$

The estimated values of degradation and corresponding 95% confidence intervals (CI) are shown in Figure 8, with additional indication of the RUL for convenience of visualization.

3.3 Results and Analysis

In Figure 8, it can be seen that the values of degradation estimated by the proposed prognostic framework were very close to the real degradation. Moreover, almost all predicted values were located within the confidence interval (CI). It can also be seen that the proposed method can forecast the RUL of a PV model over a long term (when the real degradation reaches the threshold of prognosis, the predicted value almost equals its actual value).

Then, to verify and evaluate the applicability and to further illustrate the accuracy of the proposed method, the GPR and autoregressive integrated moving average (ARIMA) model [53] were applied for the prognosis of degradation using the same data as applied in our method (Figure 9).

3.3.1 Gaussian Process Regression for Prognostic

A Gaussian Process (GP) is a collection of random variables, any finite number of which have a joint Gaussian distribution. A real GP $f(x)$ is specified by its mean function $m(x)$ and co-variance function $k(x,x')$ as defined in Eq. (19):

$$m(x) = E[f(x)],$$
$$k(x,x') = E[(f(x) - m(x))(f(x') - m(x'))], \tag{19}$$
$$f(x) \approx GP(m(x), k(x,x')).$$

The index set $X \in \Re$ is the set of possible inputs. The information about the GP and a set of training points $\{(x_i, f_i) | i = 1,...,n\}$, the posterior distribution over functions, is derived by imposing a restriction on the prior joint distribution to contain only those functions that agree with the observed data points [54]. These functions can be expressed as $y = f(x) + \varepsilon$, where ε is additive IID $N(0, \sigma_n^2)$. Once we have a posterior distribution it can be used to assess predictive values for the test data points using the equations of the predictive distribution [55] (See Eqs. (20) to (22)):

$$\text{Prior} \quad \begin{bmatrix} y \\ \mathrm{f}_{test} \end{bmatrix} \approx N\left(0, \begin{bmatrix} K(X,X) + \sigma_n^2 & K(X, X_{test}) \\ K(X_{test}, X) & K(X_{test}, X_{test}) \end{bmatrix}\right), \tag{20}$$

$$f_{test} | X, y, X_{test} \approx N(\bar{f}_{test}, \mathrm{cov}(f_{test}))$$

Since we know the values of training set y, we are interested in the conditional distribution of f_{test} given y, which is expressed as:

$$f_{test}/y = N(K(X_{test}, X)[K(X,X) + \sigma_n^2]^{-1} y, K(X_{test}, X_{test}) \\ - K(X_{test}, X)[K(X,X) + \sigma_n^2]^{-1} K(X, X_{test})) \tag{21}$$

where

$$\bar{f}_{test} \equiv E\left[\bar{f}_{test} | X, y, X_{test}\right] = K(X, X_{test})\left[K(X,X) + \sigma_n^2 I\right]^{-1} y, \tag{22}$$

$$\mathrm{cov}(f_{test}) = K(X_{test}, X_{test}) \text{-} [K(X_{test}, X) + \sigma_n^2 I]^{-1} K(X, X_{test}).$$

3.3.2 ARIMA Model for Prognosis

ARIMA is a forecasting technique, denoted as ARIMA (p,d,q). The general model, introduced by Box and Jenkins [53], is a method that allows both autoregressive (AR) and moving average (MA). It explicitly includes differencing in the formulation of the model, where p and q are the autoregressive parameter and the moving average parameter respectively, while d is the number of non-seasonal differences. The autoregressive part of the model of order p is written in Eq. (23),

$$x_t = c + \sum_{i=1}^{p} \varphi_i x_{t-i} + \varepsilon_t \tag{23}$$

where x_t is a stationary series, x_{t-i} represents lag i of x_t, the ϕ_i, $i = 1,\ldots,p$ are the parameters of the model, c is a constant and ε_t is white noise. The moving average part of the model of order q is written in Eq. (24),

$$x_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{24}$$

where the θ_i, $i = 1,\ldots,q$ are the parameters of the model, μ is the expectation of x_t, which is often assumed to be equal to zero. The $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$ are white-noise error terms. After an initial differencing step (corresponding to the integrated part of the model) the ARIMA (p,d,q) can be presented as an ARMA (p,q) process:

$$x_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{25}$$

The estimation of the ARIMA model corresponding to some learning data is done through the Box-Jenkins methods [53]. The procedure for forecasting can be summarized as follows:

1. Check stationary: If the data are not stationary, they need to be transformed into stationary data using the differencing technique.
2. Identification: Specify the appropriate number of autoregressive term p and moving average term q from the autocorrelation function (acf) and partial autocorrelation function (pacf) correlograms.
3. Forecasting: Based on the forecasting model, multi-step-ahead prediction is then conducted to forecast the final failure time.
4. Verification: If the predictions result in an unexpected trend, repeat Step 2 and Step 3 until the model fits the historical data well enough.

3.3.3 RUL Comparison and Discussion

In the implementation of the ARIMA (p,d,q) model, the appropriate number of autoregressive terms p and moving average terms q are both chosen to be 1. Parameter d is chosen to be 2 to eliminate non-stationarity.

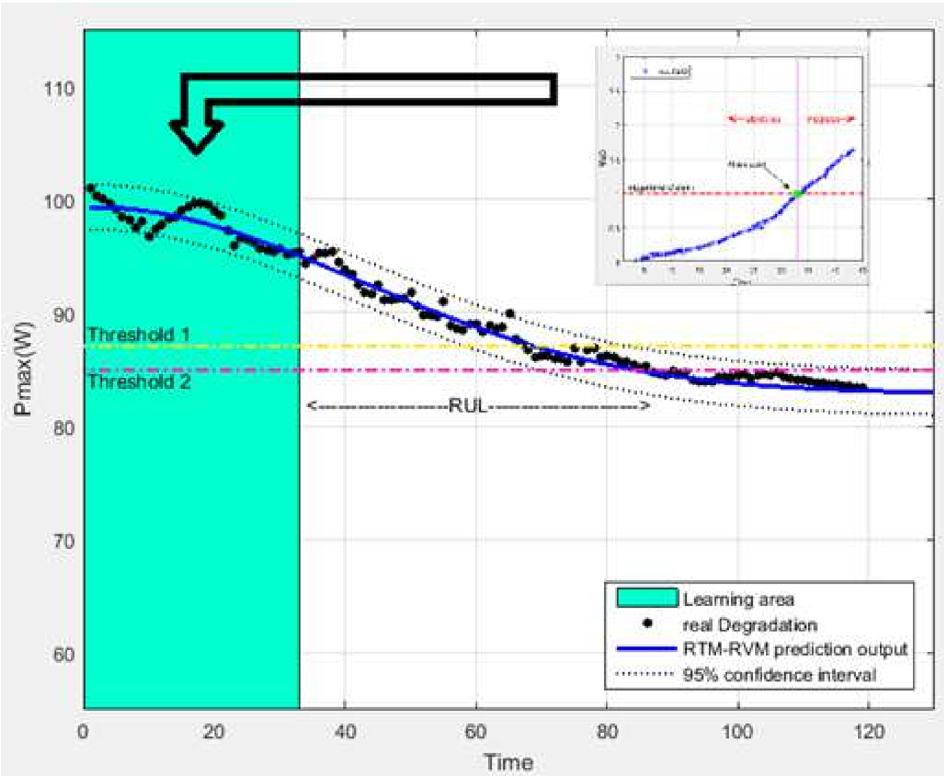

Figure 8 Degradation prediction of PV module using RTM-RVM.

Table 1 Performance Indicator Measurements

	MASE
RTM-RVM	3,832
ARIMA	10,611
GPR	4,347

The results of the application indicate that all methods can in principle come up with an RUL estimation, although the gap between the real and estimated RUL varies considerably. From the comparative results for accuracy performance presented in Figure 10 and Table 1 it can be seen that the RTM-RVM model performs significantly better than the GPR and ARIMA model via a smaller MASE value. It can be seen that the prediction values of GPR and ARIMA model are far off the real degradation values. In the prediction, we set up the same CI of 95% for all three methods. Therefore, after giving the prognostic thresholds, the RUL can be obtained for the different methods (Figure 11).

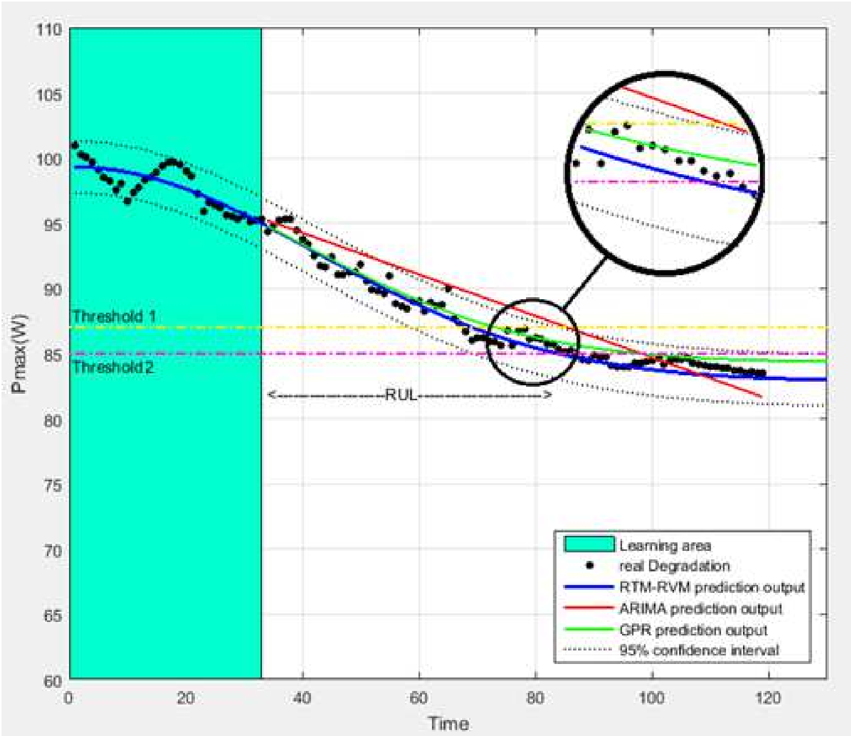

Figure 9 Comparison of prediction results with real degradation of PV module.

From Figure 9, it can also be seen that the results of RTM-RVM seem to be underestimated. In real application, an underestimated prediction is more reliable than an overestimated prediction. If the predicted RUL ends earlier than the actual value, a correct pre-maintenance can be performed. In most cases, it is more favourable to have an early prediction, instead of a late prediction.

It should be noted that the degradation thresholds (Pmax = 85W and Pmax = 87W) defined in the prognosis (Section 3.2) correspond to levels of degradation, which require the following actions:

1. In case 1 (threshold 1) the technician must check the availability of spare parts in stock.
2. In case 2 (threshold 2) the intervention will be PV module substitution or the addition of PV modules to compensate the power loss due to degradation.

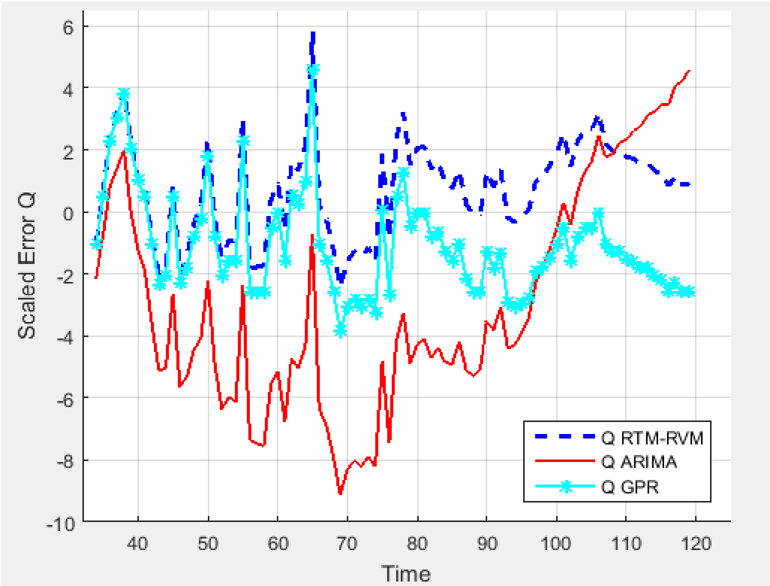

Figure 10 Scaled error Q for different prognosis methods.

The major limitations of the ARIMA model are: first, some of the traditional techniques for identifying the correct model from the class of possible models are difficult to understand. Consequently, by dint of its subjectivity, the reliability of the chosen model depends on the skill and experience of the forecaster. Second, the underlying theoretical model and structural relationships are not as distinct as some simple forecasts models, such as simple exponential smoothing.

The intrinsic ability of GPR to fit probability distribution functions (pdfs) to the data is desirable. Although GPR provides a theoretically sound framework for the prediction task it has some disadvantages concerning choosing the correct covariance function because it encodes our assumption of inter-relationships within data.

The predictions of RTM-RVM perform well and remain asymptotic to real data of degradation. Its power comes from its capacity to detect underlying trends in noisy data through the use of probabilistic kernels to account for inherent uncertainties. To always guarantee this benefit, the RTM-RVM process requires that there be sufficient points in the learning data set.

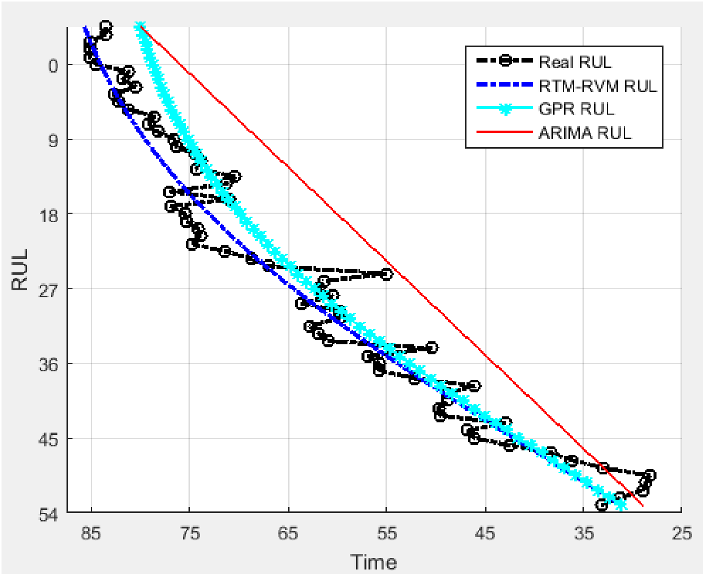

Figure 11 Comparison of real RUL and estimated RUL for different methods.

3.4 Future Work: Improving the Results

Generally, the RTM-RVM prediction accuracy seems acceptable. RVM has shown its intrinsic ability to give promising results of prognosis. However, it still suffers from the problem of uncertainties and does not take into account operating condition variation.

The obtained results are the basis of real time monitoring and historical data. Forecast accuracy can be improved by using additional information about the degradation model. In this case, hybrid approaches can be used that operate with two kinds of information (the degradation model and measurement data). According to bibliographic studies, particle filter is a hybrid method that meets our requirements and has shown its capability in multiple applications [56,57]. The degradation model that we aim to use in future work is that of output power, as mentioned in [58].

4 Conclusion

The accurate and efficient evaluation of performance degradation of PV modules is the next step after installation and commissioning to reap maximum benefit from its efficacy and to extend its lifetime. In this paper a prognostic framework was presented where a combination of real-time monitoring and the RVM algorithm was applied to estimate the RUL of a PV module device. A comparison based on prognostic performance using criteria of MASE to evaluate the performance, indicated that the proposed real-time monitoring/prognostic approach is more suitable than classical methods such as ARIMA and GPR.

The advantage of this proposed prognostic framework, outside its ability to predict the RUL, is the real-time monitoring, which triggers the prognosis process at the right time accompanied by a warning alarm. A limitation of this analysis is that data-driven prognosis does not incorporate any damage model into the computation. The use of the proposed method is not limited to PV module prognosis but can be generalized to other prognostic applications and non-repairable systems in particular.

References

[1] Sener, C. & Fthenakis, V., *Energy Policy and Financing Options to Achieve Solar Energy Grid Penetration Targets: Accounting for External Costs*, Renewable and Sustainable Energy Reviews, **32**, pp. 854-868, 2014.

[2] Sumper, A., Robledo-García, M., Villafáfila-Robles, R., Bergas-Jané, J. & Andrés-Peiró, J., *Life-cycle Assessment of A Photovoltaic System in Catalonia (Spain)*, Renewable and Sustainable Energy Reviews, **15**(8), pp. 3888-3896, 2011.

[3] Peng, J., Lu, L. & Yang, H., *Review on Life Cycle Assessment of Energy Payback and Greenhouse Gas Emission of Solar Photovoltaic Systems*, Renewable and Sustainable Energy Reviews, **19**, pp. 255-274, 2013.

[4] Jungbluth, N., *Life Cycle Assessment of Crystalline Photovoltaics in the Swiss Ecoinvent Database,* Progress in Photovoltaics: Research and Applications, **13**(5), pp. 429-446, 2005.

[5] Xiao, Q., Zhang, J., Xiao, C., Si, Z. & Tan, X., *Solar Photocatalytic Degradation of Methylene Blue in Carbon-Doped TiO_2 Nanoparticles Suspension*, Solar Energy, **82**(8), pp. 706-713, 2008.

[6] Kato, K., Murata, A. & Sakuta, K., *Energy Pay-back time and Life-cycle CO2 Emission of Residential PV Power System with Silicon PV Module*, Progress in Photovoltaics: Research and Applications, **6**(2), pp. 105-115, 1998.

[7] Alsema, E., de Wild-Scholten, M. & Fthenakis, V., *Environmental Impacts of PV Electricity Generation-a Critical Comparison of Energy Supply Options*, Paper presented at: 21st European Photovoltaic Solar Energy Conference, Dresden, Germany, 2006.

[8] Letot, C. & Dehombreux, P., *Dynamic Reliability Degradation Based Models and Maintenance Optimization*, Paper presented at: Proceedings of the 9th National Congress on Theoretical and Applied Mechanics (NCTAM2012), Brussels, Belgium, 2012.

[9] Zhang, J., Chen, W., Wang, C., Chen, X., Cheng, G., Qiu, Y. & Wu, H., *An Improved Modeling for Life Prediction of High-Power White LED Based on Weibull Right Approximation Method*, Microelectronics Reliability, **59**, pp. 49-54, 2016.

[10] Zhang, J., Liu, F., Liu, Y., Wu, H., Zhu, W., Wu, W. & Wu, L., *Life Prediction for White OLED Based on LSM under Lognormal Distribution*, Solid-State Electronics, **75**, pp. 102-106, 2012.

[11] Heng, A., Tan, A.C., Mathew. J., Montgomery, N., Banjevic, D., Jardine, A.K., *Intelligent Condition-Based Prediction of Machinery Reliability*, Mechanical Systems and Signal Processing, **23**(5), pp. 1600-1614, 2009.

[12] Marçais, J., De Dreuzy, J-R., Ginn, T., Rousseau-Gueutin, P. & Leray S., *Inferring Transit Time Distributions from Atmospheric Tracer Data: Assessment of the Predictive Capacities of Lumped Parameter Models on a 3D Crystalline Aquifer Model*, Journal of Hydrology, **525**, pp. 619-631, 2015.

[13] Berruet, P., Toguyeni, A. & Craye, E., *Structural and Functional Approach for Dependability in FMS*, Paper presented at International Conference on Systems, Man, and Cybernetics, IEEE SMC'99 Conference Proceedings, Tokyo, Japan, 1999.

[14] Pecht, M. & Jaai, R., *A Prognostics and Health Management Roadmap for Information and Electronics-rich Systems*, Microelectronics Reliability, **50**(3), pp. 317-323, 2010.

[15] Hu, C., Youn, B.D., Wang, P. & Yoon, J.T., *Ensemble of Data-Driven Prognostic Algorithms for Robust Prediction of Remaining Useful Life*, Reliability Engineering & System Safety, **103**, pp. 120-135, 2012.

[16] Pham, H.T., Yang, B.-S. & Nguyen, T.T., *Machine Performance Degradation Assessment and Remaining Useful Life Prediction using Proportional Hazard Model and Support Vector Machine*, Mechanical Systems and Signal Processing, **32**, pp. 320-330, 2012.

[17] Kutyłowska, M., *Neural Network Approach for Failure Rate Prediction*, Engineering Failure Analysis, **47**, pp. 41-48, 2015.

[18] Yan, J. & Lee, J., *A Hybrid Method for On-line Performance Assessment and Life Prediction in Drilling Operations*, Paper presented at: IEEE International Conference on Automation and Logistics, Jinan, China, 2007.

[19] Abu-Samah, A., Shahzad, M., Zamai, E. & Said, A.B., *Failure Prediction Methodology for Improved Proactive Maintenance using Bayesian Approach*, IFAC-PapersOnLine, **48**(21), pp. 844-851, 2015.

[20] Yagiz, S., Gokceoglu, C., Sezer, E. & Iplikci, S., *Application of Two Non-Linear Prediction Tools to the Estimation of Tunnel Boring Machine Performance*, Engineering Applications of Artificial Intelligence, **22**(4), pp. 808-814, 2009.

[21] Jardine, A.K., Lin, D. & Banjevic, D., *A Review on Machinery Diagnostics and Prognostics Implementing Condition-based Maintenance*, Mechanical Systems and Signal Processing, **20**(7), pp. 1483-1510, 2006.
[22] Gertler, J. & Staroswiecki, M., *Structured Fault Diagnosis in Mildly Nonlinear Systems: Parity Space and Input-output Formulations*, Paper presented at: Preprints of the 15th IFAC World Congress, 2002.
[23] Chetouani, Y., Mouhab, N., Cosmao, J. & Estel, L., *Application of Extended Kalman Filtering to Chemical Reactor Fault Detection*, Chemical engineering communications, **189**(9), pp. 1222-1241, 2002.
[24] Sotomayor, O.A. & Odloak, D., *Observer-based Fault Diagnosis in Chemical Plants*, Chemical Engineering Journal, **112**(1), pp. 93-108, 2005.
[25] Thanagasundram, S., Spurgeon, S. & Schlindwein, F.S., *A Fault Detection Tool using Analysis from an Autoregressive Model Pole Trajectory*, Journal of Sound and Vibration, **317**(3), pp. 975-993, 2008.
[26] Liu, D., Pang, J., Zhou, J., Peng, Y. & Pecht, M., *Prognostics for State of Health Estimation of Lithium-ion Batteries Based on Combination Gaussian Process Functional Regression*, Microelectronics Reliability, **53**(6), pp. 832-839, 2013.
[27] Yaïci, W. & Entchev, E., *Performance Prediction of A Solar Thermal Energy System using Artificial Neural Networks*, Applied thermal engineering, **73**(1), pp. 1348-1359, 2014.
[28] Liu, D., Zhou, J., Pan, D., Peng, Y., Peng, X., *Lithium-ion Battery Remaining Useful Life Estimation with an Optimized Relevance Vector Machine Algorithm with Incremental Learning*, Measurement, **63**, pp. 143-151, 2015.
[29] Lei, Z., *Fault Prognostic Algorithm Based on Multivariate Relevance Vector Machine and Time Series Iterative Prediction*, Procedia Engineering, **29**, pp. 678-686, 2012.
[30] Tipping, M.E., *Sparse Bayesian Learning and the Relevance Vector Machine,* The Journal of Machine Learning Research, **1**, pp. 211-244, 2001.
[31] Tipping, M.E., *Bayesian Inference: an Introduction to Principles and Practice in Machine Learning*, Advanced lectures on machine Learning: Springer, pp. 41-62, 2004.
[32] Tipping, M., *Relevance Vector Mac*hine: Google Patents, 2003.
[33] Yan, J., Liu, Y., Han, S. & Qiu, M., *Wind Power Grouping Forecasts and Its Uncertainty Analysis using Optimized Relevance Vector Machine*, Renewable and sustainable energy reviews, **27**, pp. 613-621, 2013.
[34] Zio, E. & Di Maio, F., *Fatigue Crack Growth Estimation by Relevance Vector Machine*, Expert Systems with Applications, **39**(12), pp. 10681-10692, 2012.

[35] Wang, D., Miao, Q. & Pecht, M., *Prognostics of Lithium-ion Batteries Based on Relevance Vectors and a Conditional Three-Parameter Capacity Degradation Model*, Journal of Power Sources, **239**, pp. 253-264, 2013.

[36] Schölkopf, B. & Smola, A.J., *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Cambridge, United States, 2002.

[37] Hyndman, R.J., *Another Look at Forecast-Accuracy Metrics for Intermittent Demand*, Foresight: The International Journal of Applied Forecasting, **4**(4), pp. 43-46, 2006.

[38] Cocca, M., D'Arienzo, L. & D'Orazio, L., *Effects of Different Artificial Agings on Structure and Properties of Whatman Paper Samples*, ISRN Materials Science, 2011.

[39] Miller, D.C., Carloni, J.D., Johnson, D.K., Pankow, J.W., Gjersing, E.L., To, B., Packard, C.E., Kennedy, C.E. & Kurtz, S.R., *An Investigation of the Changes in Poly (Methyl Methacrylate) Specimens after Exposure to Ultra-Violet Light, Heat, and Humidity*, Solar Energy Materials and Solar Cells, **111**, pp. 165-180, 2013.

[40] Makrides, G., Zinsser, B., Norton, M., Georghiou, G.E., Schubert, M. & Werner, J.H., *Potential of Photovoltaic Systems in Countries with High Solar Irradiation*, Renewable and Sustainable Energy Reviews, **14**(2), pp. 754-762, 2010.

[41] Kempe, M.D., *Control of Moisture Ingress into Photovoltaic Modules*, Paper presented at: Photovoltaic Specialists Conference, Conference Record of the Thirty-first IEEE, Florida, United States, 2005.

[42] Munoz, M., Alonso-Garcia, M., Vela, N. & Chenlo, F., *Early Degradation of Silicon PV Modules and Guaranty Conditions*, Solar Energy, **85**(9), pp. 2264-2274, 2011.

[43] Skoczek, A., Sample, T., Dunlop, E. & Ossenbrink, H., *Electrical Performance Results from Physical Stress Testing of Commercial PV Modules to the IEC 61215 Test Sequence,* Solar Energy Materials and Solar Cells, **92**(12), pp. 1593-1604, 2008.

[44] Berman, D. & Faiman, D., *EVA Browning and the Time-Dependence of I−V Curve Parameters on PV Modules with and without Mirror-Enhancement in a Desert Environment*, Solar energy materials and solar cells, **45**(4), pp. 401-412, 1997.

[45] Gxasheka, A., Van Dyk, E. & Meyer, E., *Evaluation of Performance Parameters of PV Modules Deployed Outdoors*, Renewable Energy, **30**(4), pp. 611-620, 2005.

[46] Ndiaye, A., Kébé, C.M., Ndiaye, P.A., Charki, A., Kobi, A. & Sambou, V., *A Novel Method for Investigating Photovoltaic Module Degradation*, Energy Procedia, **36**, pp. 1222-1231, 2013.

[47] Ndiaye, A., Kébé, C.M., Charki, A., Ndiaye, P.A., Sambou, V. & Kobi, A., *Degradation Evaluation of Crystalline-Silicon Photovoltaic Modules after a few Operation Years In A Tropical Environment*, Solar Energy, **103**, pp. 70-77, 2014.

[48] Osterwald, C., Adelstein, J., Del Cueto, J., Kroposki, B., Trudell, D. & Moriarty, T., *Comparison of Degradation Rates of Individual Modules Held at Maximum Power*, Paper presented at IEEE 4th World Conference on Photovoltaic Energy Conversion, Conference Record of the 2006, Hawaii, United States, 2006.

[49] Shaari, S., Sopian, K., Amin, N. & Kassim, M.N., *The Temperature Dependence Coefficients of Amorphous Silicon and Crystalline Photovoltaic Modules using Malaysian Field Test Investigation*, American Journal of Applied Sciences, **6**(4), pp. 586-593, 2009.

[50] Skoplaki E, Palyvos J., *On the Temperature Dependence of Photovoltaic Module Electrical Performance*: A *Review* of *Efficiency/Power Correlations*, Solar Energy, **83**(5), pp. 614-624, 2009.

[51] Mazer, J.A., Solar *Cells*: *An Introductio*n to *Crystalline Photovoltaic Technology*, Kluwer Academic Publishers, 1997.

[52] Pan, R., Kuitche, J. & Tamizhmani, G., Degradation *Analysis* of *Solar Photovoltaic Modules*: *Influence of Environmental Factor*, in 2011 Proceedings-Annual Reliability and Maintainability Symposium, IEEE Reliability Society, Florida, United States, 2011.

[53] Box, G.E., Jenkins, G.M., Reinsel, G.C. & Ljung, G.M., *Time Series Analysis*: *Forecasting and Control*, John Wiley & Sons, New Jersey, United States, 2015.

[54] Rasmussen, C.E., *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, United States, 2006.

[55] Williams, C.K. & Rasmussen C.E., *Gaussian Processes for Regression*, in Touretzky, D. S., Mozer, M. C., & Hasselmo, M. E. (eds.), Advances in Neural Information Processing Systems, **8** , pp. 514-520. MIT Press, 1996.

[56] Jouin, M., Gouriveau, R., Hissel, D., Péra, M.-C. & Zerhouni N., *Prognostics of PEM Fuel Cell in a Particle Filtering Framework*, International Journal of Hydrogen Energy, **39**(1), pp. 481-494, 2014.

[57] Zio, E. & Peloni, G., *Particle Filtering Prognostic Estimation* of the *Remaining Useful Life* of *Nonlinear Components*, Reliability Engineering & System Safety, **96**(3), pp. 403-409, 2011.

[58] Charki, A., Laronde, R. & Bigaud, D., *The Time*-variant *Degradation* of a *Photovoltaic System*, Journal of Solar Energy Engineering, **135**(2), pp. 024503, 2013.