Diagnosis of a battery energy storage system based on principal component analysis

Edison Banguero[a, b, *], Antonio Correcher[a], Ángel Pérez-Navarro[c], Emilio García[a], and Andrés Aristizabal[d]

[a]*Instituto de Automática e Informática Industrial-ai2, Universitat Politècnica de València, CP 46022 Valencia, Spain.*
[b]*Universidad Tecnológica del Chocó, A.A 292, Calle 22 No 18B-10, Quibdó, Colombia.*
[c]*Instituto Universitario de Ingeniería Energética—IUIIE, Universitat Politècnica de València, CP 46022 Valencia, Spain.*
[d]*Engineering Department, Universidad de Bogotá Jorge Tadeo Lozano, Carrera 4 # 22-61, 34185 Bogotá D.C., Colombia.*

**Correspondence: edbanpa@doctor.upv.es; Tel.: +34-69-796-6373*

Abstract

This paper proposes the use of principal component analysis (PCA) for the state of health (SOH) diagnosis of a battery energy storage system (BESS) that is operating in a renewable energy laboratory located in Chocó, Colombia. The presented methodology allows the detection of false alarms during the operation of the BESS. The principal component analysis model is applied to a parameter set associated to the capacity, internal resistance and open circuit voltage of a battery energy storage system. The parameters are identified from experimental data collected daily. The PCA model retains the first 5 components that collect 80.25% of the total variability. During the test under real operation contidions, PCA diagnosed a degradation of state of health fastest than the comercial battery controller. A change in the charging modes lead to a battery recovery that was also monitored by the proposed algortihm, and control actions are proposed that lead the BESS to work in normal conditions.

Keywords: Diagnosis, battery energy storage system, principal component analysis, state of health.

1. Introduction

Electricity service is of great importance for the development of a nation. However, more than 1.1 billion of the world's population is without access to electricity [1]. 85% of this population live in rural areas where carrying the interconnection of the electricity grid can be very expensive [1]. Some of these populations use diesel generators to face that problem. Nevertheless, market advances in renewable energies is leading to the replacement of these generatos with stand-alone photovoltaic systems or hybrid renewable generators [2]. In these cases, the most used technology to store energy is battery energy storage systems (BESS). In off-grid applications, energy storage is a crucial component for ensuring reliability of these stand-alone renewable energy systems [3]. These systems primarily use lead acid batteries as a way to store electricity [4]. Besides,

this technology is low cost, mature and has efficient recycling processes [5]. Advances in lithium-ion technology are allowing this type of batteries to be used in renewable energy systems [6-8]. However, in 2015 the lead acid battery market grew to $37 billion [9, 10], so this technology is still being used as shown in the research developed in [11], and it will be working for long time.

During their operation, batteries are subjected to different charging and discharging processes, stress, temperature changes that affect their state of health (SOH). SOH reduction involves loss of capacity and an increase of internal resistance. SOH is used to know the conditions in which a battery has evolved from its start-up [12, 13], and it is generally defined in function of the current maximum capacity and its nominal capacity [14, 15]. Making a SOH diagnosis of the battery allows the operator to detect possible battery failures and to perform timely maintenance or replacement of the battery [16]. The study of the battery internal parameters opens a research field in order to make a better battery management system. Previous researches show different methods to identify the internal parameters of a battery [17-21]. These parameters are associated with the capacity, internal resistance and battery voltage. However, there are no previous research reports that analyze the evolution over time of these parameters and their relationship with battery SOH. This paper proposes the study of the internal parameters of a BESS and its evolution over time in order to generate useful information to perform SOH monitoring and diagnosis.

Because number of data to be managed is relatively high, it is very common in engineering to use a strategy to find correlation structures between the variables. One of these strategies is principal component analysis (PCA). According to [22, 23], PCA is a multivariate statistical process control technique that reduces the dimensionality of an original space that is established from the historical data of the measured variables to a space of lower dimensionality, searching linear combinations between the variables that best describe the process trend. PCA is based on the decomposition of the covariance matrix of the process variables along the directions that best explain the main causes of variability of the information analyzed [24-26].

In this paper a SOH diagnosis of a BESS is made by applying PCA to the internal parameters of a BESS. The methodology used to identify and extract the internal parameters of a BESS is explained in [17], and control actions are proposed that lead the BESS to work in normal conditions. The BESS has been working for three years during which data have been colleted with a sampling time of 15 minutes, and is located in a renewable energy laboratory at Chocó – Colombia [27]. The BESS is composed of 24 Techno Sun 2V-OPzS-TCH2765 (C_{10} = 2160 Ah; C_{120} = 2765 Ah) electrochemical accumulators connected in series. Regarding to Chocó department, it is considered as one of the rainiest regions in the world, with rainfall exceeding 9000 mm of annual precipitation and an average temperature of 27 ^{0}C. It is located in the western part of the country, in the region of the Pacific plain at 04^0 00’ 50’’ and 08^0 41’ 3’’ north latitude and 76^0 02’ 5’’ and 77^0 53’ 3’’ west longitude. Chocó is a Colombian department with the largest number of municipalities with non-interconnected areas, where three-quarters of the department belong to rural areas with a rurality index between 43.65 and 73.67%. In the non-interconnected areas the electric power service is done through Diesel generators installed by the Institute for Planning and Promoting Energy Solutions for Non-Interconnected Zones (IPSE), with an operating time between 6 and 12 hours. In the rest of the Chocó, the electric power supply is made through the national interconnected system.This paper is organized as follows: Section 2 describes the battery model used in this research. In Section 3 the PCA model is described and the mathematical foundations are indicated.The control actions carried out by the commercial controller are shown in Section 4. In Section 5 the PCA model is constructed and a diagnosis is made to new observations. Finally, the conclusions are presented in Section 6.

2. Methodology

2.1. Overview

To carry out this diagnosis, we first used the battery model developed by Copetti [18] to represent the voltage and the state of charge of the BESS. Then we identify the BESS internal parameters by using particle swarm optimization (PSO) algorithm and are explained in [17].

2.2. PSO algorithm description

PSO is an evolutionary computation technique that emulates animal behavior through particles with certain position and velocity in a search space, where the population is called swarm, and each member of the swarm is called a particle. In PSO, each of the particles retains the record of the location of the best particle that provides the position/location value and is denoted *pbest* [28-29]. After collecting the pbest information for each particle the algorithm uses these values to find the optimal global solution that is denoted by *gbest* [28-29]. However, despite the similarity with other heuristic techniques such as genetic algorithms (GA), PSO does not imply mutation and crossover function. PSO finds the best global solution by adjusting the movement vector of each particle according to its personal best position (cognitive aspect, second term) and the best global position (social aspect, third term) of the particles in the whole swarm in each iteration.

The first term is only the current velocity of the particle and can be considered as an impulse term. The second term that is associated with a local search is proportional to the vector $\left(pbest_i^k - x_i^k\right)$ and points from the current position of the particle to its best personal position. The third term that is associated with a global search is proportional to $\left(gbest^k - x_i^k\right)$ and points to the position of the best global particle [29-31].

$$v_i^{k+1} = \omega \cdot v_i^k + c_1 r_1 \left(pbest_i^k - x_i^k \right) + c_2 r_2 \left(gbest^k - x_i^k \right) \quad (1)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (2)$$

Where, r_1 and r_2 are random numbers in the range [0, 1]. k is the current number of iterations. ω is known as the inertial constant or inertial weight and its recommended value is slightly less than 1 (a non-negative number) which regulates the search range of the solution space. c_1 and c_2 are the acceleration constants that determine the amount of particles directed towards an ideal position. One of them is called the cognitive component and the other is designated as a social component. The significance of these two constants is that they determine to what extent pbest and gbest affect the movement of the particles. The recommended value for these two constants is approximately 2 [30-32].

To the identified parameters a PCA model is applied and the causes that are affecting the model are identified. Finally, control actions are proposed that lead the BESS to work in normal conditions.

2.3. The model

The model devolped by Copetti [18] can be used to describe the dynamics of a BESS. The state of charge (SOC) and the battery capacity can be computed the next three equations:

$$\mathrm{SOC}(t) = \mathrm{SOC}_0 - \frac{1}{C(t)} \int_0^t \eta_c I(t) \partial t \quad (3)$$

$$C(t) = \frac{C_N \times C_{tcoef} \times k_{c10}}{1 + A_{cap}\left(\frac{|I(t)|}{I_N}\right)^{B_{cap}}}\left(1 + \alpha_c \Delta T + \beta_c \Delta T^2\right) \quad (4)$$

$$I_N = \frac{C_N}{m} \quad (5)$$

Where βc = 0 ^{0}C; $\Delta T = T - Tref$ is the battery temperature variation regarding to a reference temperature (T_{ref} = 25 ^{0}C); I_N is the battery nominal current; C_N is the battery nominal capacity (at m hours); $0 \le SOC(t) \le 1$, C_{tcoef}, A_{cap} and B_{cap} are empirical parameters, and k_{c10} is the battery capacity reduction coefficient.

The battery internal resistance is expressed by equations (6) and (7)

$$R_{\text{int}}(T) = \frac{n_s}{C_N}(R_1 + R_2 + R_3)\left(1 - \alpha_T \Delta T(t)\right) \quad (6)$$

$$R_{\text{int}}(T) = \frac{n_s}{C_N}\left(\frac{P_1}{1 + |I_{bat}(t)|^{P_2}} + \frac{P_3}{(1 - SOC(t))^{P_4}} + P_5\right)\left(1 - \alpha_T \Delta T(t)\right) \quad (7)$$

Where, R_1 and R_2 represent the influence of the charge/discharge current I_{bat}, R_3 represents the Ohmic effect [33], and n_s is the number of connected batteries in series. P_{1dc}, P_{2dc}, P_{3dc}, P_{4dc}, and P_{5dc} are parameters related to resistive losses [19]. Discharging and charging voltages can be described by equations (8) and (9). These equations were taken from [17].

$$V_{dc}(t) = n_s\left[V_{bodc} - K_{bodc}\left(1 - SOC(t)\right)\right] - n_s \frac{|I_{bat}(t)|}{C_{10}} \times \left(\frac{P_{1dc}}{1 + |I_{bat}(t)|^{P_{2dc}}} + \frac{P_{3dc}}{SOC(t)^{P_{4dc}}} + P_{5dc}\right) \times \left(1 - \alpha_{rdc} \Delta T(t)\right) \quad (8)$$

$$V_c(t) = n_s\left[V_{boc} + K_{boc}\left(SOC(t)\right)\right] + n_s \frac{|I_{bat}(t)|}{C_{10}} \times \left(\frac{P_{1c}}{1 + |I_{bat}(t)|^{P_{2c}}} + \frac{P_{3c}}{\left(1 - SOC(t)\right)^{P_{4c}}} + P_{5c}\right) \times \left(1 - \alpha_{rc} \Delta T(t)\right) \quad (9)$$

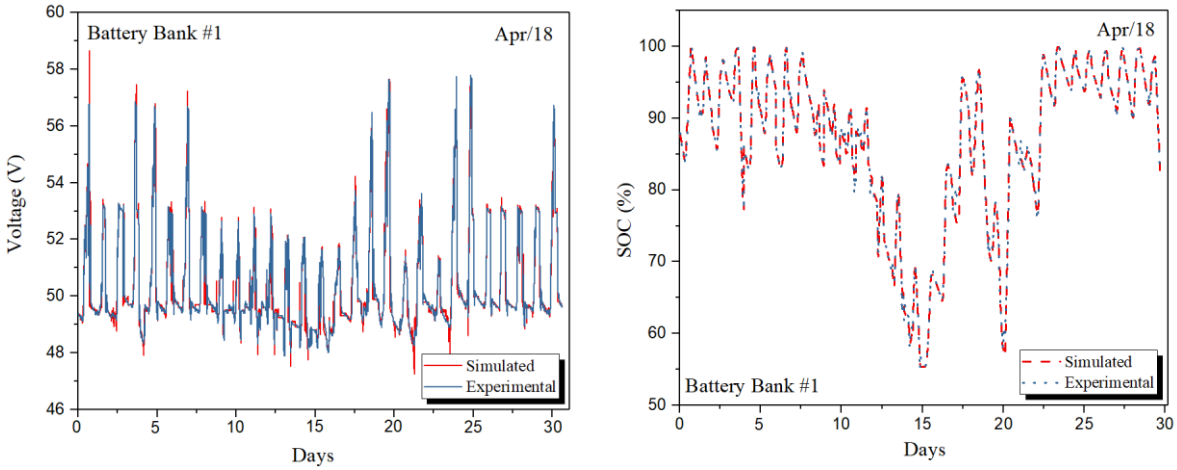

Figure 1. BESS voltage, SOC measurements, and simulation results. Case study: BESS located in the renewable energy laboratory at Chocó – Colombia. BESS voltage signals (left), and BESS SOC signals (right).

In order to use the model, each parameter can be fitted to represent the real behavior of the system. To identify the internal parameters of the BESS we use the methodology presented in [17]. In [17], an improvement of the PSO algorithm is used to fit the data and the results show an accurate response of the model as it is shown in Figure 1. As [17] have already proved the effectiveness of the method, this paper does not give more details about parametter identification.

The identified parameter set for each day is the intial data to perform the SOH diagnosis of the lead acid BESS by PCA.

3. Principal component analysis (PCA)

PCA is a technique used to reduce the dimensionality of a data set that tries to lose as little information as possible [23, 34]. PCA allows determining differences between samples, which are the variables that contribute most to differentiate the samples and detect correlations between variables. When a PCA is applied, a new system variable called principal component (PC) is generated. The first PC contains most of the information contained in the original data matrix; the second PC provides a maximum of the resulting residual variance, and so on [35, 36]. PCA results can be represented in "scores" and "loadings" charts [37]. The scores chart represents the differences and the similarities between samples and it is used to find clusters between them. The loadings chart determines the correlation between variables. The univariate representation of the loading p_a contributes to identify the variables of greatest influence on each PC and to associate a practical interpretation for the components. Through a bivariate loadings chart (p_a vs $p_{a'}$) a partial representation of the usual correlation structure between the variables of the original matrix X is obtained. In this representation, the location of the variables in the same direction and away from the origin is an indicator of positive correlation. Conversely, its location in the opposite direction indicates negative correlation. Variable groups with low correlation will be located in perpendicular directions, distant from the origin. The similarity between the X records can be evaluated through bivariate charts of the score (t_a *vs* $t_{a'}$). Average individuals will be located in the chart center, while the eccentric individuals are the most contributing individuals for each PC.

3.1. Practical application of PCA

To monitor a multivariate system such as the internal parameter set of a BESS, it is essential to know widely its habitual behavior and the correlation structure between its variables. These elements can be characterized using data collected during a reference period (Phase I) in which the system has worked under the regular or desired condition. Data colleted on Phase I, are used to adjust a statistical model that reproduces the habitual behavior of the BESS. Later, in Phase II, the Phase I model is used to evaluate the concordance of future observations. In PCA, the observation matrix X (Eq. (10)) is decomposed into a set of A (A < rank(X)) matrices of rank 1 [38].

$$X = \sum_{a=1}^{A} t_a * p_a^T + \sum_{a=A+1}^{rank(X)} t_a * p_a^T = T * P^T + E \qquad (10)$$

$P \epsilon R^{K \times A}$ is the loading matrix composed of the eigenvectors (p_a) associated with the highest A values (λ_a) of the covariance matrix of X. K indicates the number of variables; A is the number of principal components, and N indicates the number of samples. Vectors p_a (a = 1,2,…A) contain the coefficients of the original variables of X in the principal components (PCs) and define the A directions of highest variability. $T \epsilon R^{N \times A}$ is the score matrix, which contains the location of the projections of the rows of X over the A-dimensional subspace (T = XP). Each column of T is a vector t_a (a= 1, 2,…A), which represents the ath PC. The matrix $E \epsilon R^{N \times K}$ stores the residuals obtained by predicting X from $T*P^T$ ($E = X - T * P^T$). So, E (residual matrix) contains the information of X that is not explained by the PCA model.

PCA is a scale-dependent technique therefore the results of the PCA can vary significantly depending on the scale of measurement of the variables. In multivariate monitoring it is very frequent that the scales of the variables are notoriously different, for which it is necessary to perform a previous work of scaling the columns of the matrix X. A very common procedure is to standardize (center with respect to your average and scale by the standard deviation) each of the columns of the matrix X. With this preprocessing, all variables are expected to have the same weight in the construction of the PCA model. It is also common to find that the data matrix X is made up of small subsets or blocks of variables, each of which contains information about a particular feature of the system. In these cases, the contribution of each block of variables can be leveled out through a scaling procedure per block. In this preprocessing each variable ($X_{.j}$) is centered with respect to its average value ($\bar{X}_{.j}$), and scaled by the product between its standard deviation ($S_{.j}$) and the square root of the number of parameters ($\sqrt{K_h}$) that make up the h-th block (to which the variable belongs). This scaling is represented in equation (11). A matrix of transformed data (X^*) is generated with this procedure in which each block of parameters provides the same level of variability.

$$X_{.j}^* = \frac{X_{.j} - \bar{X}_{.j}}{\sqrt{K_h} * S_{.j}} \qquad (11)$$

In the construction of the PCA model, the user must decide the number of PCs (A) that there will be used for the analysis. In order to determine this number, a quality criterion for representing the data in the new subspace is generally used. One of the most used criteria is the cumulative percentage of variance (CPV) [39] (Eq. (12)) explained by the model.

$$CPV\ (A) = \frac{\sum_{a=1}^{A} \lambda_a}{\sum_{a=1}^{rank(X)} \lambda_a} 100\% \qquad (12)$$

Once the *A* components are selected, the elements of the PCA model (*T, P, E*) are used to understand the PCs system, to study the correlation between the variables of *X*, to evaluate the similarity between the records and to identify extreme observations in the reference data set. The extreme observations identification is made using the information contained in the score matrix (*T*) and in the residuals matrix (*E*). The PCA model transforms each observation (row) of the matrix X^*, $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})$ into a projections vector in PCs $t_i = (t_{i1}, t_{i2}, \ldots, t_{iA})$, and a residue vector $e_i = (e_{i1}, e_{i2}, \ldots, e_{ik})$. This pair of vectors are summarized in two independent statistics: Hotelling's T^2 (Eq.(13)) and sum of square prediction error (SPE) (Eq. (14)).

$$T_i^2 = t_i^T \Theta^{-1} t_i = \sum_{a=1}^{A} \frac{t_{ia}^2}{\lambda_a} \qquad (13)$$

$$SPE_i = e_i^T * e_i = \sum_{j=1}^{K} e_{ik}^2 \qquad (14)$$

Where $\Theta = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_A)$ is the diagonal matrix of the *A* eigenvalues associated with the PC retained in the PCA model. Under the assumption that the score vector follows a multivariate normal distribution, a upper control limit (UCL) greater than (1-α)% for the T^2 statistic can be set at (Eq. (15)):

$$UCL_{T_\alpha^2} = \frac{\left(N^2-1\right)A}{\left(N-A\right)N} F_{\alpha,A,N-A} \qquad (15)$$

Where $F_{\alpha,A,N-A}$ is the critical value of a random variable of the *F* (Fisher-Snedecor) distribution, with *A* degrees of freedom in the numerator and *N* - *A* degrees of freedom in the denominator, and α is the level of significance [27, 39, 40]. Zhao et al. [41] suggests setting the control limit for the SPE statistic on the critical value of a weighted chi-squared distribution($g\chi_h^2$). Where g is the weight and *h* is the degrees of freedom. Nomikos et al. [42] suggest to approximate the value of the degrees of freedom h and of the weighting g, equaling the mean and the variance of the $g\chi_h^2$ distribution with the sampling average (*m*) and the sampling variance (*v*) of the SPE statistic in the observations of Phase I. Therefore the control limit to (1 - α) % for SPE can be set as (Eq. (16)):

$$UCL(SPE) = \frac{v}{2m} \chi^2_{\frac{2m^2}{v},\alpha} \qquad (16)$$

Where, $g = v/2m$ and $h = 2m^2/v$

T^2 and SPE are calculated for each row of the matrix X^* and they are plotted together with the control limits. These two statistics contain complementary information to identify extreme observations in the data set. However, their interpretation is notably different. A point that is signaled by the statistic T^2 corresponds to an observation whose projections in the space of the components are very extreme. This means that this observation has an abnormal value in one or more of the variables, but maintains the direction of the usual correlation structure. In the PCA model this point is labeled as an extreme observation. Besides, a point that is indicated by the SPE statistic corresponds to an observation whose prognosis through the PCA model is significantly different from its observed value. In this case, one or more of the variables have an abnormal

value that does not correspond to the usual correlation structure. This point is labeled as extreme observation outside the PCA model. The presence of these extreme points can significantly condition the results obtained in the PCA model. Therefore, when they are identified in phase I it is convenient to remove them from the observation matrix and adjust the model again. In phase II, the final PCA model of phase I is used to monitor new observations of the process. In this phase, each new multivariate observation x_{new} is projected on the subspace of the PC. Then, the score vector ($t_{new} = x_{new}P$) and the forecast error vector $e_{new} = x_{new} - t_{new}P^T$ are calculated. With that information, both SPE and Hotelling's T^2 statistics are calculated and the observation is represented in the two control charts. If some of the two statistics exceed the control limit, a change signal is generated and the variable involved in the change must be identified. In the monitoring of multivariate processes, contribution plots are one of the most used tools to diagnose the variables associated with a signal of change. When the observation is indicated by the SPE chart, the contribution of the variable $j = 1, 2, \ldots K$, on the SPE statistic, corresponds to the square of its forecast error (the error sign is retained to facilitate the interpretation), as it's shown in Eq. (17):

$$C_j^{SPE} = sign\left(e_{new,j}\right) * e_{new,j}^2 \tag{17}$$

If the observation is indicated by the T^2 statistic, then the contribution of variable j to the T^2 statistic is calculated as (Eq. (18)):

$$C_j^{T^2} = \sum_{a=1}^{A} \frac{\mathrm{t}_{new,a} * _{new,j} * p_{j,a}}{\lambda_a} \tag{18}$$

Therefore, those variables with the greatest contribution to the statistic can be identified as potential cause of the signal disturbance. In addition to the Hotelling's T^2, SPE and contribution plots, the graphical representation of the projection of the new observations in the space of each PC, together with the bivariate loadings plots are very useful to identify the extreme observations and the possible associated variables that cause the problem.

4. Controller actions

A commercial controller carries out the battery management: Sunny Island™ inverter. The controller applies a charging strategy (boost, full, float or equalization) that will depend on the SOC and the configuration of the battery. Figure 2, shows the different Sunny Island™ charging phases. In boost charge and full charge mode a charging voltage of 2.4 V is applied to the BESS cells, meanwhile in float charge and equalization charge mode a voltage of 2.25 V and 2.5 V respectively is applied to the BESS cells.

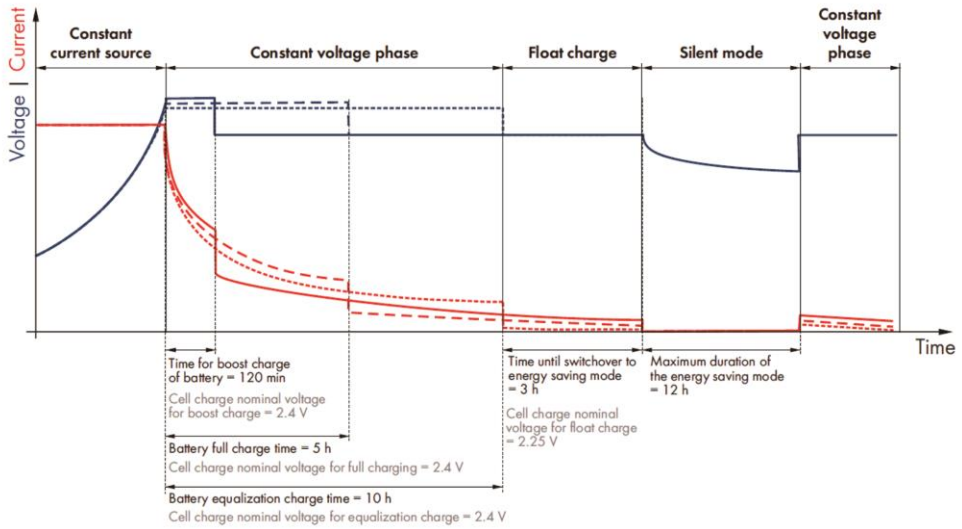

Figure 2. Sunny Island™ inverter charging phases for an AGM battery [43].

When the BESS operates under normal conditions, the controller maintains a float charge. However, due to different operating conditions, the controller can use other charging modes more agresive: boost charge, full charge, or equalization charge. The different charging modes that the BESS experiences can lead to an increase in the current (see Figure 3), and therefore an increase in battery temperature. This increase in current and temperature can lead to a premature deterioration of the battery.

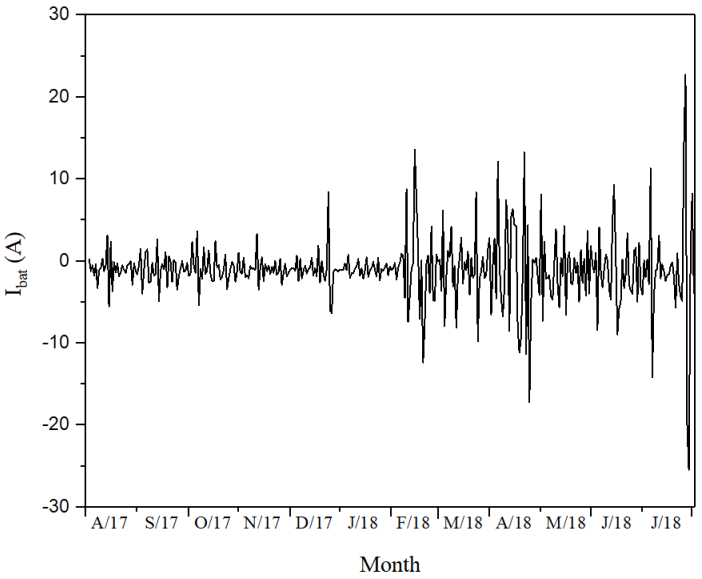

Figure 3. One-year measurements of BESS current taken from the renewable energy laboratory at Chocó – Colombia.

Taking into account this abnormal behavior of the BESS current, we proceed to make a diagnosis by PCA, find the possible causes and propose solutions that lead the BESS to work under normal conditions.

5. Application of PCA to BESS.

5.1. Phase I: Construction of the PCA model

The characterization of the normal behavior of the internal parameters of the BESS is carried out through a PCA model. As reference data, the daily identifications of the model parametters corresponding to the month of October 2017 (31 records) are used. Each record contains the identification of 13 parameters for the charging equations and other 13 for the discharging equations of the BESS. These parameters are grouped into three blocks, depending on their relationship with key features of BESS. Table 1 presents the definition of the parameters considered and the related feature[1].

Table 1. Parameters considered for the diagnosis of the BESS.

Parameters	Feature/Block
P_1, P_2, P_3, P_4, P_5, α_r	Internal Resistance
C_{tcoef}, A_{cap}, B_{cap}, α_c, K_{c10}	Capacity
V_{boc} , K_{boc}	Open circuit voltage

So, the set of reference data (X) corresponds to a matrix of 31 registers and 26 parameters that have to relate to the BESS SOH. In this initial matrix each parameter has its own scale of measurement. Additionally, as Table 1 shows, the number of parameters that make up each block is not homogeneous. Therefore, there is an imbalance in the contribution of each block of parameters on the total variability of the reference data matrix. To level the contribution of each block on the total variability, a scaling procedure (as the one described in section 2) is performed per block. With the preprocessed data matrix (X^*), a first PCA model is built, whose goal is to verify the stability of the reference data and to identify atypical or influential points in the PCA model.

Hotelling's T^2 and SPE control charts, with control limits established under a 1% level of significance are used for the diagnosis of observations. As shown in Figure 4. a), the first 5 PCs retain 80.25% of the total variability. Atypical or influential observations are not presented in the reference data set (see Figure 4b)).

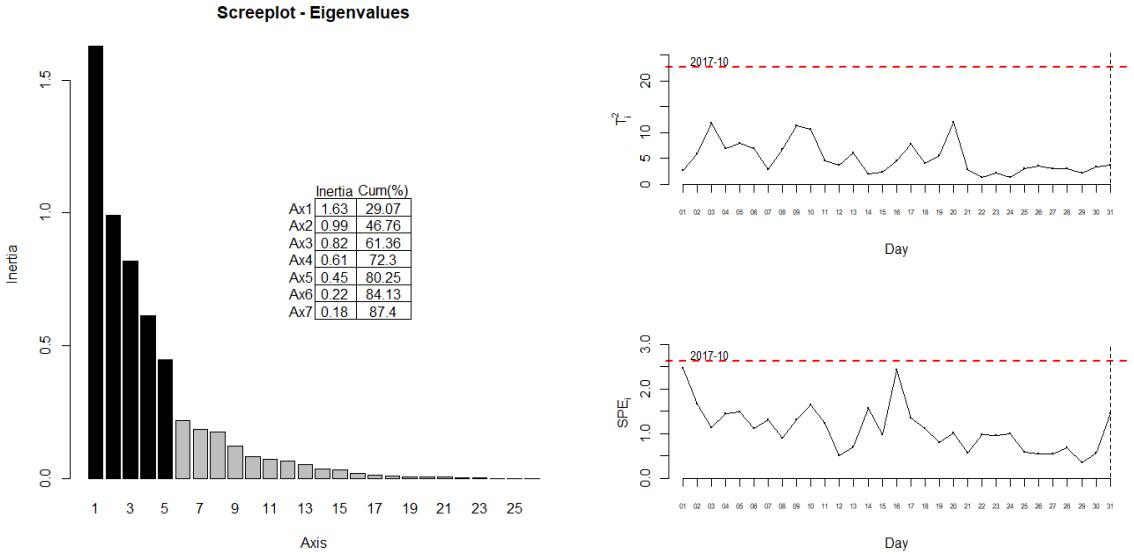

Figure 4. a) Plot of the final PCA model. b) Hotelling's T^2 and SPE control charts for the filtered reference data.

[1] Note that each parametter has a value for the charging process and other for the discharging process.

The set of PC makes up a set of indicators of the BESS SOH. They are computed from the information contained in the parameters weighted according to the coefficients illustrated in Figure 5. In each state (charging and discharging) of the BESS the parameters that have more weight are: K_{boc}, K_{c10}, C_{tcoef}, A_{cap}, B_{cap}. When working with 5 PCs, each pametter evolves over a 5-dimensional space and it is usual to use the projections over each 2-dimensional subspace to show the evolution. Threfore, 5 PC, 10 PC combinations are generated (1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5) with the same number of charts.

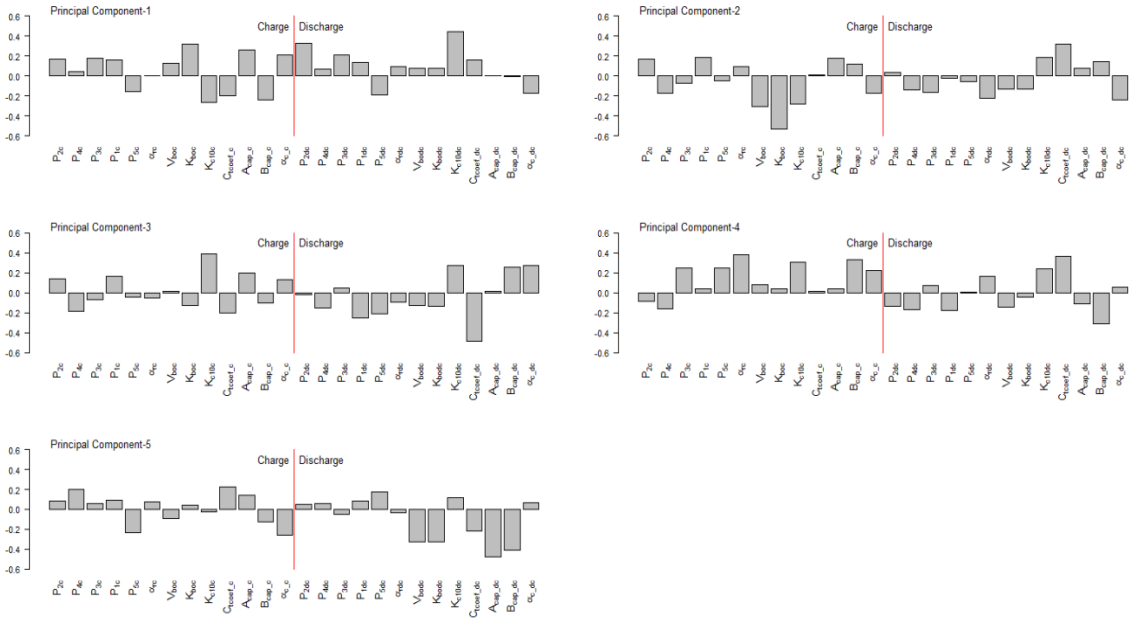

Figure 5. Coefficients (loadings) of the parametters in each PC.

5.2. Phase II - Diagnosis of new observations.

The PCA model identified at phase I is used to diagnose new observations of the internal parameters of the BESS. Figure 6 illustrates the diagnosis algorithm flow. Each day, a set of new observations are organized in a *Xnew* data matrix, whose columns receive the same transformation as in Eq. (10). Then, Hotelling´s T_i^2 (where *i* stand for the day) is computed for this data set and it is monitored with a control chart.

Note that Y axis of the Hotelling's T^2 control chart determines the similarity between two multidimensional random variables, i.e. is the distance (Mahalanobis distance) of the new data with respect to the original model. The difference with the Euclidean distance is that Mahalanobis distance takes into account the correlation between the random variables. The Hotelling's T^2 control chart with control limits is presented in Figure 7. In order to monitor the alarm degree, this paper proposes splitting Hotelling's T^2 control chart into two regions: Level 0 (no alarm), and alarm level 1. Below the established control limit (1% false alarm rate), the BESS is assumed to operate under normal conditions. At alarm level 1, the system still shows good performance but its performance does not match the training conditions. However, when the difference greater than *6* T^2 then, the system working outside its normal conditions. Therefore, if new data in Hotelling's T^2 are placed on level 0, the system is working normally. When there are 3 or more consecutive points and uprising trends on alarm level 1, it means that the identified parametters are different enough from the original ones. Therefore, the diagnosis procedure starts. The complete diagnosis of the situation requires the analysis of the parametters causing the alarm. Therefore, after the diagnosis of the alarm zone, the contribution of each parameter to Hotelling's T^2 has to be studied.

If the points are located in an alarm zone, then inputs have to be checked. If there are abnormal weather conditions or an abnormal demand, then the training conditions are extremely different to the current ones so it is not possible to conclude that there is or not a failure. So, no action is required.

i. If external conditions are similar from the PCA training month, and the alarm signal is generated by C_{tcoef}, A_{cap}, B_{cap}, α_c, K_{c10} parameters then the battery is entering a process of loss of capacity that may be due to: low electrolyte level, electrolyte stratification, low levels of SOC, shedding, sulfation, or degradation of the active matter, all associated with the degradation of battery.

Check the electrolyte level: If the level is low, then fills with distilled water to the level recommended by the manufacturer and verify that the specific density of each cell is that specified by the manufacturer (between 1.265 ± 0.50 Kg/l to 25 ^{0}C). Apply a full charge (full charge time = 5 h). The goal is to charge the battery until it reaches a SOC > 90%. When the battery is charged, apply an equalization charge (equalization charge time = 10 h) to compensate for the differences that may occur in the SOC of the battery cells and finally change float charge mode. If after applying this procedure the battery capacity is recovered, and $T^2 < 3.2$, then the system works in normal conditions, therefore change to float charge mode. If the electrolyte level is not low, then check the SOC. If the SOC is between 70% and 80% then apply a boost charge until the battery reaches 85% and 90% of its current capacity (time for boost charge of battery = 2 h). Then perform a full charge (full charge time = 5 h) until the SOC > 90%. If the capacity is recovered, then the system works under normal conditions, therefore change to float charge mode. If the SOC > 80% then, apply an equalization charge (equalization charge time = 10 h). If the battery capacity is recovered, then the system works under normal conditions, therefore change to float charge mode.
If after applying all these procedures the battery capacity is not recovered, then the battery has a failure that can be due to detachment of active material from the electrodes (shedding), sulphation or degradation of the active matter, among others, that they can not be recovered with controller actions. Therefore the battery must be brought to maintenance or in its effect; it must be replaced by another of the same feature.

ii. If the alarm signal is generated by P_1, P_2, P_3, P_4, P_5, α_r parameters then the battery is presenting resistive losses that are associated with prolonged operating conditions at very low SOC that lead to degradation processes such as corrosion, generating an increase in internal resistance and therefore an increased risk of failure.

Check the SOC: If the SOC is very low, i.e., SOC < 60%, then apply a boost charge until the battery is between 85% and 90% of its current capacity (time for boost charge of battery = 3 h). Then change from boost charge to full charge (full charge time = 5 h). Once the battery is charged, perform an equalization charge (charge equalization time = 10 h). If after applying this procedure the battery internal resistance decreases, then the system works in normal conditions, therefore change to float charge mode. If SOC > 60%, apply a boost charge until the battery is between 85% and 90% of its current capacity (time for boost charge of battery = 2 h). Then change from boost charge mode to full charge (full charge time = 5 h). If the battery internal resistance decreases, then the system works under normal conditions, therefore change to float charge mode. If after applying these procedures the internal resistance does not decrease, then the battery has a fault that it can not recover with controller actions. Therefore, the battery must be replaced by another with the same feature.

iii. If the alarm signal is generated by the V_{boc}, K_{boc} parameters, then is a cause of battery failure and can not be recovered with controller actions. Therefore the battery must be brought to maintenance or in its effect; it must be replaced by another of the same feature. If the open circuit voltage parameters are inside range, then the system works under normal conditions, therefore change to float charge mode.

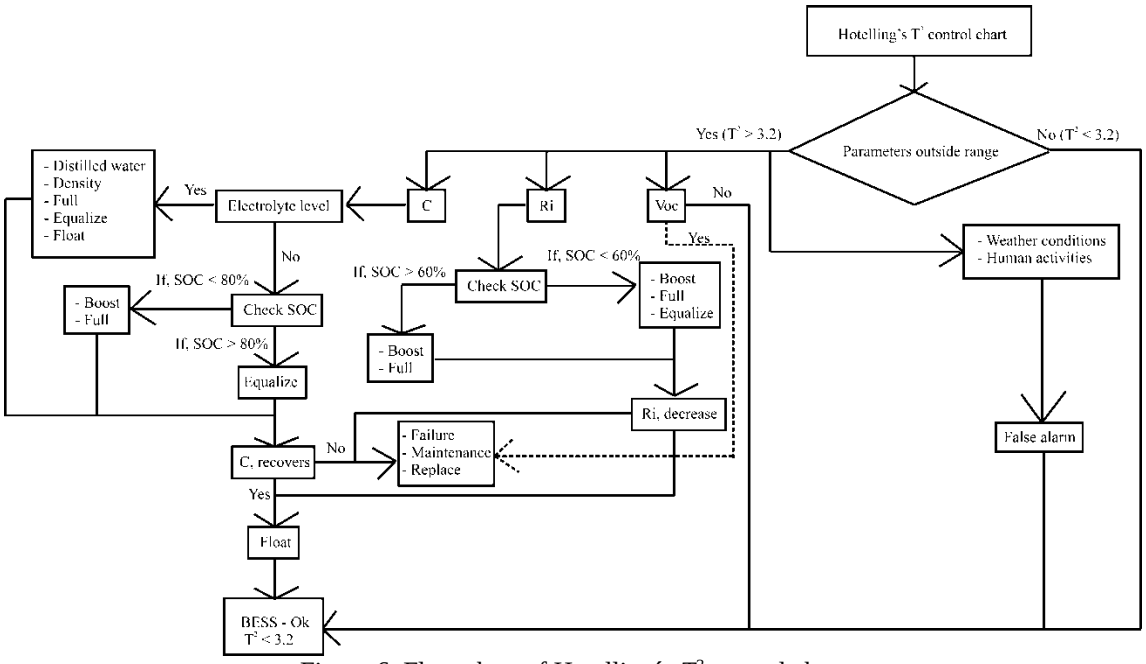

Figure 6. Flow chart of Hotelling´s T^2 control chart

To illustrate the procedure, monitoring is initially performed for the month of November 2017 (see Figure 7). A logarithmic scale is used for a better visualization of T^2 progress. The higher the value of T^2, the greater the distance between the observation and the target value.

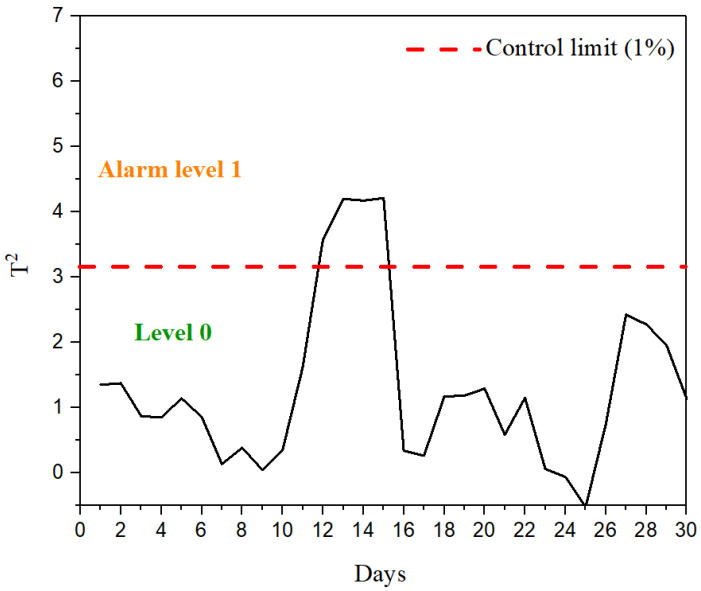

Figure 7. Hotelling's T^2 for November 2017 monitoring.

According to the diagnosis, the system exhibits unusual behavior on days 12 to 15 and inside the PCA model for the rest of days. Therefore, the next step should be the identification of the variables that have caused the unsual behavior and, then, try to explain this abnormal behavior. This analysis can be done by observing the contribution plots or by analysing the projections of the observations in the space of the principal components.

Figure 8 shows the contributions plots to the Hotelling's T^2 obtained for the alarms generated on November 12 to 15, 2017.

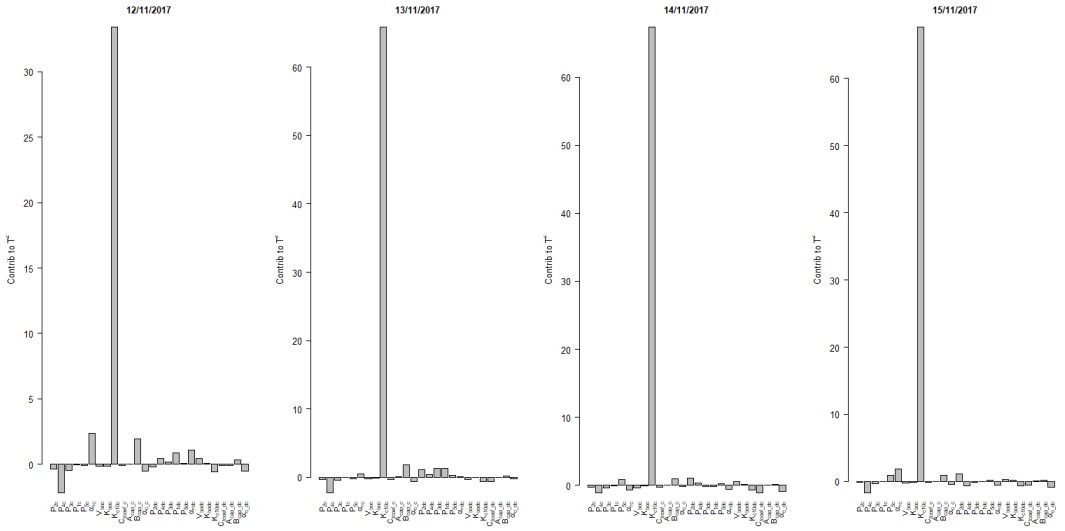

Figure 8. Contribution plots to the Hotelling's T^2 of the extreme observations within the PCA model.

According to this chart, there was an unusual increase in the K_{c10} parameter in the charge phase of BESS during four days. This parameter was included in the model for the first time in [17] and it allows to adjust the battery model proposed in [18] reaching a lower measurement error between the experimental and simulated voltage of the BESS. The other way of interpreting an alarm in the Hotelling's T^2 control chart is to project the observations on the space of the PCs; e.g: November (11/2017) projections chart has been superimposed with the month of training (October/2017) as shown in Figure 9.

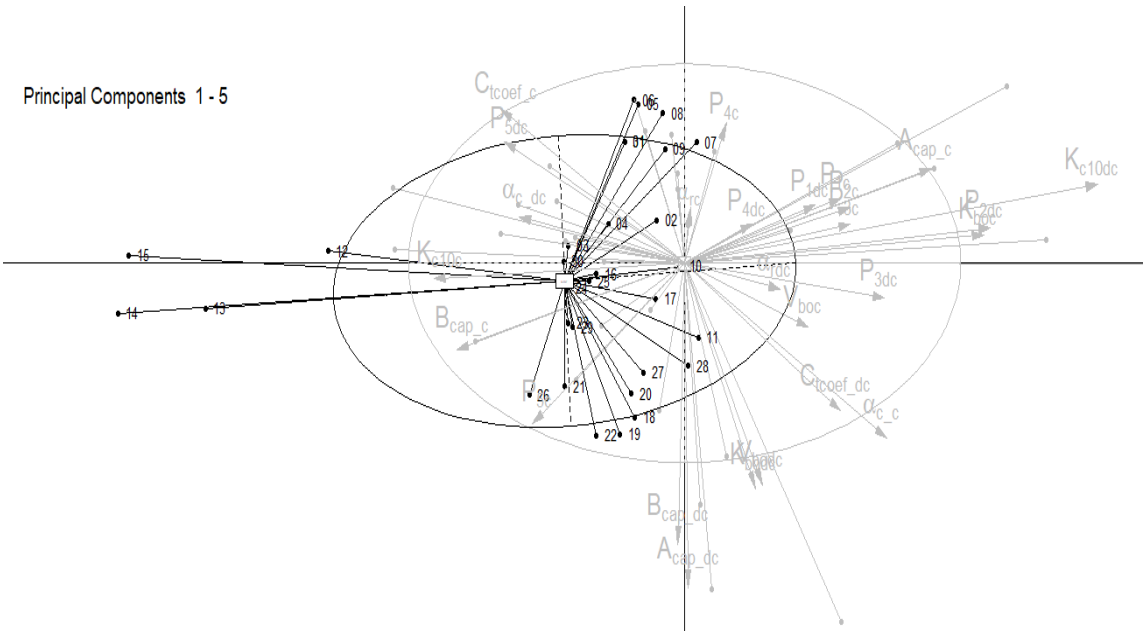

Figure 9. Projections of November observations on the space of the PCs.

In this month the centroid of the observations moves in the direction of the parameter K_{c10c}, specifically on days 12-15 November, which were marked as alarm by the T^2 control chart (Figure 7), and coincides with the

diagnosis obtained in the contributions plots to T^2 (Figure 8). However, the alarms presented during these days they are associated with an increase in the K_{c10} parameter due to a SOC below 85% which caused the controller to apply a boost charge.

5.3. Application to real data observation

At this point, the PCA model is used to monitor the SOH of the system during a longer observation period, from November 2017 to July 2018. Figure 10 shows the diagnosis of the Hotelling's T^2 for the new observations set.

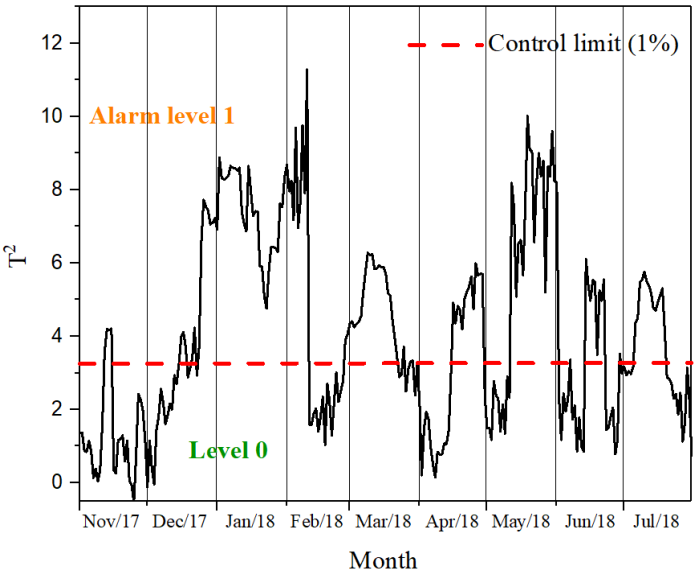

Figure 10. Hotelling's T_i^2 monitoring from November 2017 to July 2018.

Figure 10 shows that on the third week of the month of December, an alarm signal was generated because the BESS had a SOC of less than 80%, which caused an increase of the parameter K_{c10} in the charge mode. The increase of this parameter shows that the BESS is entering a process of loss of capacity. The controller takes a week to change boost charge to float charge. However, it does not manage to lower the alarm level. According to shown in Figure 6, the following actions should have been applied: a boost charge until the battery reaches 85% and 90% of its current capacity (time for boost charge of battery = 2 h). Then perform a full charge (full charge time = 5 h) until the SOC > 90%. The same procedure should have been applied in the month of January.

February is the month with the greatest statistical distance between the new data and the original model (reaching up to 11 distance units), which generates a higher alarm rate. This significant increase was due to work carried out with welding equipment inside of renewable energy laboratory that led to the BESS reaching a SOC below 70%, increase in BESS current greater than 2A, and an increase of the P_1 parameter in the discharge mode. During this period, the controller applied a boost charge, then changing to float charge until the BESS to working under normal conditions. Because there was an abnormal demand, then the training conditions are extremely different to the current ones so it is not possible to conclude that there is or not a failure. So, no action is required. However, subjecting the BESS to very high discharge rates over a long period is leading to a deterioration of health of the battery.

In March the alarms continue at a lower level. However between days 5 and 10 there is an increase in the K_{c10} parameter in charge mode. The controller applies a boost charge for 5 days until it reaches a SOC equal to 97% and then change to float charge until the alarm level falls to level 0. During the first 15 days of April, BESS worked under normal conditions. However, the decrease in SOC (< 70%) caused an increase in the α_c parameter in the discharge mode. α_c is a parameter that is associated to the temperature variation and battery capacity. During this month the controller begins with a float charge, then goes to a boost charge (from 10th to 18th) and finally performs a full charge (from 19th to 25th). Again in the month of May work was done with welding equipment inside the renewable energy laboratory. Therefore, the same decisions must be applied as in the month of February declaring this level of alarms false. Between the days 14th and 22nd of the month of June there are alarms caused by the increase of α_c parameter in the discharge mode. The rest of the days BESS works under normal conditions. The controller applies float charge (from 1 to 13 days), then boost charge (14 to 22 days) and ends with float charge. However, due to the depth of discharge uncontrolled to which the BESS was submitted during the month of May, an equalization charge should have been applied.

Finally in the month of July alarms (from 6 to 19 days) were by the increase of P_2 parameter in charge mode. During this month there was a greater increase in the BESS current (I_{bat} > 15 A). The controller applies float charge (from 1 to 9 days), then boost charge (10 to 19 days) and ends with an equalization charge.

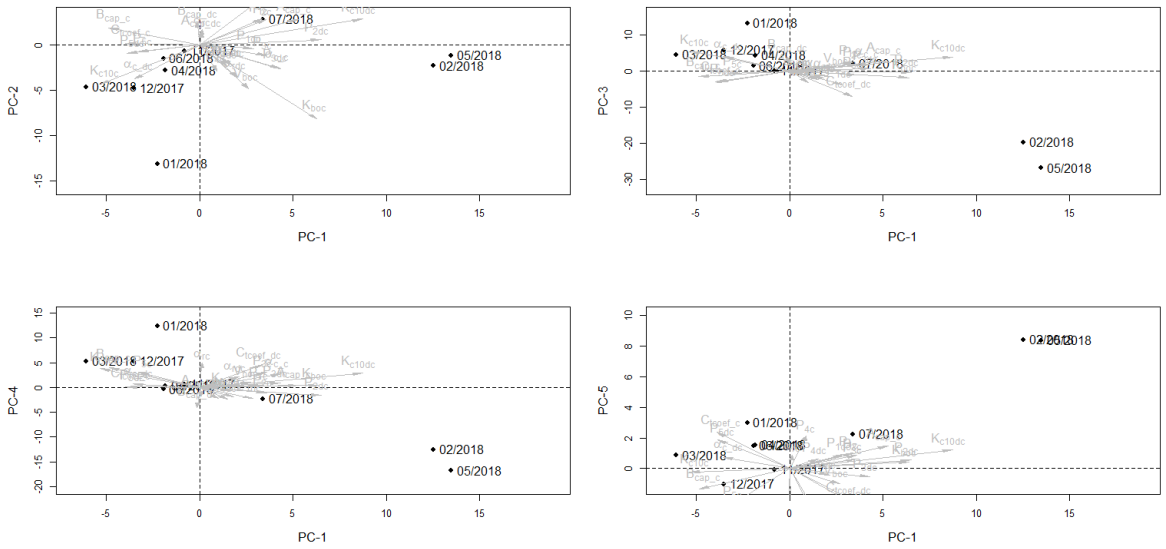

Figure 11. Projections of the centroids of the observations from November 2017 to July 2018 on the space of the PCs.

The unusual increase of these parameters caused a movement of the centroid of the observations (see Figure 11) and a deterioration of the BESS SOH during the first half of the year 2018 (see Figure 12). The deterioration of health is related to a progressive loss of the BESS operational capacity. However, since the month of May the BESS shows recovery in its SOH because the controller applied an equalization charge to the BESS to neutralize the differences that occur in SOC of individual battery cells due to the different behavior of the battery cells. The controller performs an equalization charge when the sum of all discharges since the last equalization charge is 30 times the nominal capacity of the battery [43].

The SOH is estimated by Eq. (19)

$$SOH = \frac{C_{current\max}}{C_N} \times 100\% \qquad (19)$$

Where, $C_{current\ max}$ is current maximum battery capacity. At the time of manufacture is considered *SOH = 100%*, and is considered *SOH = 0%* when the battery capacity to store and supply energy decreases below predefined threshold [44].

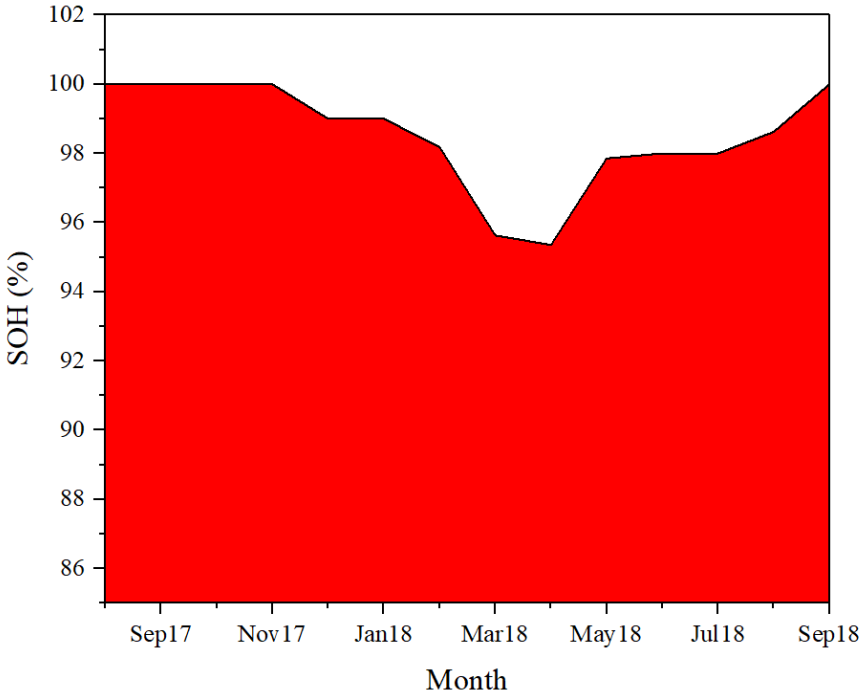

Figure 12. BESS state of health (SOH) estimation of the renewable energy laboratory at Chocó – Colombia.

6. Conclusions

PCA is a versatile method capable of providing a general diagnosis of the BESS. The diagnosis is based on the study of a parameter characterizing the battery. These parameters can be identified from experimental data collected each day with an evolutionary algorithm devoped in previous research. Daily, the new parameters are projected over the PCA model, which was trained on a normal working month. The analysis reveals that in the last semester of the period under study, the BESS presented alarms associated with deterioration in SOH as a result of the different operating modes of the controller. The effect over PCA projections is a movement of the observations centroid with respect to the training period. The goal of the controller is the recovery of the SOH. However, the use of aggressive charging modes increases the current, and therefore increases the BESS temperature. The effect of these increases, if they are manteined over time, is a loo on battery health.

Moreover, the diagnosis proposed in this paper has found that once the alarm is generated the controller takes several days to perform the charge change. Therefore, the use of the PCA diagnosis algorithm will lead to a faster change in the operation mode to recover the battery quicker, so the BESS will suffer lower currents and lower temperature increases.

The methodology presented also allows the detection of false alarms in the BESS, such as those registered in the months of February and May. Threefore, no corrective should be taken if the algorithm rises a false alarm diagnosis.

Acknowledgments: The authors would like to acknowledge the research project "Implementación de un programa de desarrollo e investigación de energías renovables en el departamento del Chocó, BPIN 2013000100285 (in Spanish)" and the Universidad Tecnológica del Chocó (in Spanish). The authors would like to thank the anonymous reviewers as well as the editor for their valuable comments that have greatly improved the final version of the paper.

Author Contributions: All authors have worked on this manuscript together and all authors have read and approved the manuscript.

References

[1] World Energy Outlook (WEO-2017) Special Report: Energy Access Outlook-International Energy Agency (IEA).

https://www.iea.org/publications/freepublications/publication/WEO2017SpecialReport_EnergyAccessOutlook.pdf [Accessed on 03.05.2018]

[2] Perera, A. T. D., Attalage, R. A., Perera, K. K. C. K., & Dassanayake, V. P. C. (2013). Designing standalone hybrid energy systems minimizing initial investment, life cycle cost and pollutant emission. Energy, 54, 220-230.

[3] Krieger, E. M., Cannarella, J., & Arnold, C. B. (2013). A comparison of lead-acid and lithium-based battery behavior and capacity fade in off-grid renewable charging applications. Energy, 60, 492-500.

[4] Dali, M., Belhadj, J., & Roboam, X. (2010). Hybrid solar–wind system with battery storage operating in grid-connected and standalone mode: control and energy management–experimental investigation. Energy, 35(6), 2587-2595.

[5] Aksakal, Can; Sisman, Altug. On the Compatibility of Electric Equivalent Circuit Models for Enhanced Flooded Lead Acid Batteries Based on Electrochemical Impedance Spectroscopy. Energies, 2018, vol. 11, no 1, p. 118.

[6]Astaneh, Majid, et al. A novel framework for optimization of size and control strategy of lithium-ion battery based off-grid renewable energy systems. Energy Conversion and Management, 2018, vol. 175, p. 99-111.

[7] Nge, C. L., Ranaweera, I. U., Midtgård, O. M., & Norum, L. (2019). A real-time energy management system for smart grid integrated photovoltaic generation with battery storage. Renewable energy, 130, 774-785.

[8] Dhundhara, Sandeep; Verma, Yajvender Pal; Williams, Arthur. Techno-economic analysis of the lithium-ion and lead-acid battery in microgrid systems. Energy Conversion and Management, 2018, vol. 177, p. 122-142.

[9] Pillot, C. (2015, March). The rechargeable battery market and main trends 2014–2025. In 31st International Battery Seminar & Exhibit.

(Available from: http://www.avicenne.com/pdf/Fort_Lauderdale_Tutorial_C_Pillot_March2015.pdf). [Accessed on 03.07.2018].

[10] Yang, J., Hu, C., Wang, H., Yang, K., Liu, J. B., & Yan, H. (2017). Review on the research of failure modes and mechanism for lead–acid batteries. International Journal of Energy Research, 41(3), 336-352.

[11] J. Li, Optimal sizing of grid-connected photovoltaic battery systems for residential houses in Australia, Renewable Energy (2018), https://doi.org/10.1016/j.renene.2018.09.099

[12] Tao, L., Ma, J., Cheng, Y., Noktehdan, A., Chong, J., & Lu, C. A review of stochastic battery models and health management. Renewable and Sustainable Energy Reviews, 2017, vol. 80, p. 716-732.

[13] Talha, Muhammad; Asghar, Furqan; Kim, Sung Ho. A Neural Network-Based Robust Online SOC and SOH Estimation for Sealed Lead–Acid Batteries in Renewable Systems. Arabian Journal for Science and Engineering, 2018, p. 1-13.

[14] Marchildon, Jacques; Doumbia, Mamadou Lamine; Agbossou, Kodjo. SOC and SOH characterisation of lead acid batteries. En Industrial Electronics Society, IECON 2015-41st Annual Conference of the IEEE. IEEE, 2015. p. 001442-001446.

[15] Li, X., Shu, X., Shen, J., Xiao, R., Yan, W., & Chen, Z. An on-board remaining useful life estimation algorithm for lithium-ion batteries of electric vehicles. Energies, 2017, vol. 10, no 5, p. 691.

[16] Nguyen, Thanh-Tuan; Tran, Van-Long; Choi, Woojin. Development of the intelligent charger with battery State-Of-Health estimation using online impedance spectroscopy. Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium on. IEEE, 2014. p. 454-458.

[17] Ariza, H. C., Banguero, E., Correcher, A., Pérez-Navarro, Á., & Morant, F. (2018). Modelling, Parameters Identification and Experimental Validation of a Lead Acid Battery Bank Using Genetic Algorithms. Energies, 11(9), 2361; https://doi.org/10.3390/en11092361 (registering DOI).

[18] Copetti, J. B., Lorenzo, E., & Chenlo, F. (1993). A general battery model for PV system simulation. Progress in Photovoltaics: Research and applications, 1(4), 283-292.

[19] Guasch, D., & Silvestre, S. (2003). Dynamic battery model for photovoltaic applications. Progress in Photovoltaics: Research and applications, 11(3), 193-206.

[20] Blaifi, S., Moulahoum, S., Colak, I., & Merrouche, W. (2016). An enhanced dynamic model of battery using genetic algorithm suitable for photovoltaic applications. Applied Energy, 169, 888-898.

[21] Blaifi, S., Moulahoum, S., Colak, I., & Merrouche, W. (2018). Monitoring and enhanced dynamic modeling of battery by genetic algorithm using LabVIEW applied in photovoltaic system. Electrical Engineering, 100(2), 1021-1038.

[22] Gao, Z., Cecati, C., & Ding, S. X. (2015). A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. IEEE Transactions on Industrial Electronics, 62(6), 3757-3767.

[23] Ferrer, A. (2007). Multivariate statistical process control based on principal component analysis (MSPC-PCA): Some reflections and a case study in an autobody assembly process. Quality Engineering, 19(4), 311-325.

[24] Jiang, Q., Yan, X., & Zhao, W. (2013). Fault detection and diagnosis in chemical processes using sensitive principal component analysis. Industrial & Engineering Chemistry Research, 52(4), 1635-1644.

[25] Fan, J., & Wang, Y. (2014). Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis. Information Sciences, 259, 369-379.

[26] Garcia-Alvarez, D., Fuente, M. J., & Sainz, G. I. (2012). Fault detection and isolation in transient states using principal component analysis. Journal of Process Control, 22(3), 551-563.

[27] Banguero, E., Aristizábal, A. J., & Murillo, W. (2017). A Verification Study for Grid-Connected 20 kW Solar PV System Operating in Chocó, Colombia. Energy Procedia, 141, 96-101.

[28] Shi, Yuhui, et al. Particle swarm optimization: developments, applications and resources. In Proceedings of the 2001 congress on evolutionary computation (IEEE Cat. No. 01TH8546). IEEE, 2001. p. 81-86.

[29] Rahman, Md Ashiqur; Anwar, Sohel; Izadian, Afshin. Electrochemical model parameter identification of a lithium-ion battery using particle swarm optimization method. Journal of Power Sources, 2016, vol. 307, p. 86-97.

[30] Yang, X., Chen, L., Xu, X., Wang, W., Xu, Q., Lin, Y., & Zhou, Z.. Parameter identification of electrochemical model for vehicular lithium-ion battery based on particle swarm optimization. Energies, 2017, vol. 10, no 11, p. 1811.

[31] Kai, H., Yong-Fang, G., Zhi-Gang, L., Hsiung-Cheng, L., & Ling-Ling, L. Development of Accurate Lithium-Ion Battery Model Based on Adaptive Random Disturbance PSO Algorithm. Mathematical Problems in Engineering, 2018, vol. 2018.

[32] Venter, Gerhard; Sobieszczanski-Sobieski, Jaroslaw. Particle swarm optimization. AIAA journal, 2003, vol. 41, no 8, p. 1583-1589.

[33] Layadi, T. M., Champenois, G., Mostefai, M., & Abbes, D. (2015). Lifetime estimation tool of lead–acid batteries for hybrid power sources design. Simulation Modelling Practice and Theory, 54, 36-48.

[34] Rahmani, M., & Atia, G. K. (2017). Coherence pursuit: Fast, simple, and robust principal component analysis. IEEE Transactions on Signal Processing, 65(23), 6260-6275.

[35] Bro, R., & Smilde, A. K. (2014). Principal component analysis. Analytical Methods, 6(9), 2812-2831.

[36] Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2017). Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. Trends in Food Science & Technology.

[37] Soh, W., Kim, H., & Yum, B. J. (2018). Application of kernel principal component analysis to multi-characteristic parameter design problems. Annals of Operations Research, 263(1-2), 69-91.

[38] Deng, X., Tian, X., Chen, S., & Harris, C. J. (2018). Nonlinear process fault diagnosis based on serial principal component analysis. IEEE transactions on neural networks and learning systems, 29(3), 560-572.

[39] De Ketelaere, B., Hubert, M., & Schmitt, E. (2015). Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data. Journal of Quality Technology, 47(4), 318-335.

[40] Vanhatalo, E., Kulahci, M., & Bergquist, B. (2017). On the structure of dynamic principal component analysis used in statistical process monitoring. Chemometrics and Intelligent Laboratory Systems, 167, 1-11.

[41] Zhao, C., Wang, F., Gao, F., Lu, N., & Jia, M. (2007). Adaptive monitoring method for batch processes based on phase dissimilarity updating with limited modeling data. Industrial & engineering chemistry research, 46(14), 4943-4953.

[42] Nomikos, P., & MacGregor, J. F. (1995). Multivariate SPC charts for monitoring batch processes. Technometrics, 37(1), 41-59.

[43] Design of Off-Grid Systems with Sunny Island 4.4M / 6.0H / 8.0H Devices. Version 2.3. SMA. 1-42. http://files.sma.de/dl/1353/Designing-OffGridSystem-PL-en-23.pdf [Accessed on 15.07.2018]

[44] Ungurean, L., Cârstoiu, G., Micea, M. V., & Groza, V. (2017). Battery state of health estimation: a structured review of models, methods and commercial devices. International Journal of Energy Research, 41(2), 151-181.