Multivariate analysis of the operational parameters and environmental factors of an industrial solar pond

S. Platikanov[a], R. Tauler[a], J.L. Cortina[b,c] and C. Valderrama[b,c,*]

[a]Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona, 18-26, 08026 Barcelona, Spain

[b]Chemical Engineering Department, Universitat Politècnica de Catalunya UPC•BarcelonaTECH,

[c]Barcelona Multi-Scale Science and Engineering Research Center, BarcelonaTECH

*Correspondence should be addressed to: César Valderrama

Universitat Politècnica de Catalunya-Barcelona TECH

C/ Eduard Maristany, 10-14, 08930 Barcelona, Spain

Email: cesar.alberto.valderrama@upc.edu

Abstract

The stability of an industrial Salinity Gradient Solar Pond (SGSP) has been studied by applying classical Principal Component Analysis (PCA) on three datasets with a different number of variables and time evolution. Temperature, density and heat extraction have been measured in the upper convective (UCZ), non-convective (NCZ) and low convective (LCZ) zones in a 500 m^2 solar pond, located near Granada, South of Spain. Local environmental weather conditions have been also considered in the data analysis. Two operational seasons of the solar pond were considered in order to establish: 1) PCA exploratory models for 2014-2015 operational period and 2) to validate the obtained results during the 2015-2016 operational period. PCA results showed that three factors explained the variance in the monitored stability gradients. The first factor was related to the major effect of the seasonal temperature variation on the entire stability gradient. The second factor, related to the diurnal temperature variation, solar irradiance and wind variables, showed a strong impact on the solar pond temperature and salinity gradients and could affect strongly the UCZ and its border with the NCZ. The third factor affected the stability of salinity and temperature gradients at the LCZ and its border with the NCZ. The latter was related to the increase in temperature and salinity at the bottom of the solar pond, which suggests special attention during the initial formation and settlement of the salinity gradient and the subsequent heat extraction activities. This paper shows PCA modelling as a powerful tool for solar pond operation process surveillance and control.

Keywords: Principal Component Analysis; modelling; multiparametric datasets; salinity gradient

1. Introduction

Global warming and, consequently, the climate crisis have motivated the need for a low-carbon economic model and an energy transition that raises the need to increase the global demand for clean and cheap energy. The renewable energy sources as the alternatives to fossil fuels could be utilized given the potential of each geographic region (Kasaeian et al., 2020). Solar energy is widely available and is among the cleanest forms of renewable energy. However, the utilization of solar energy is associated with significant challenges due to its low energy density and intermittency characteristics. The salinity gradient solar pond (SGSP) provides a solution to these problems by employing a large collection area and storage system (Tabor, 1981; Chakrabarty et al., 2020). Some of SGSP's features such as low cost of construction, simplicity in design, and integrated heat storage have promoted the use of solar pond technology as a low cost thermal energy storage system, based on the collection and storage of solar radiation as a heat source for different applications (Valderrama et al., 2016; Kumar et al., 2020). The main advantage of solar ponds is their long-term thermal energy storage capability, which can supply sufficient heat along the entire year (Alcaraz et al., 2018a, Alcaraz et al., 2018b). The classical SGSP is characterized by three different layers, the upper convective zone (UCZ), the middle non-convective zone (NCZ), and the lower convective zone (LCZ) (Zangrando, 1980).

The upper convective zone (UCZ) is the topmost layer of the solar pond. It should be relatively thin and low salinity water. The non-convective zone (NCZ) is the next layer, below the upper convective zone. This layer is the thickest zone and has to be characterized with gradually increasing density along increasing pond’s depth. This layer serves as thermal insulating and plays a significant role in the effectiveness of the capture and storage of solar energy, keeping this layer stable minimizes the heat losses from the bottom LCZ. Finally, the LCZ layer has the highest salinity, near saturation. When the solar pond works properly, convective currents are prevented by the salinity gradient. Thus, the absorbed solar energy should enter from UCZ throughout NCZ and it is stored in the bottom part of the LCZ, where heat can be extracted from. The main purpose of the

NCZ is to act as an insulator to prevent heat from escaping into the UCZ, thus maintaining higher temperatures in the deeper areas (LCZ). The temperature differences between the top and the bottom of the solar ponds can be as high as 50–60 °C (Tundee et al., 2010). There are many examples of practical use of the heat energy (stored in the solar pond), such as the heating of buildings, power production and water desalination purposes (Alcaraz et al., 2018c; Ganguly et al., 2018).

The correct operation of solar ponds is characterized by stable salinity and temperature gradients, which is generally linked to maximizing the thickness of the gradient (NCZ) and avoiding seasonal variability of the three layers. The efficiency of any solar pond in capturing energy depends on the stability of salinity and thermal gradients. Avoiding the appearance of convective forces near the boundary zones (UCZ and NCZ, and NCZ and LCZ), will maintain the stability of the salinity gradient and allow adequate heat transfer to the lower layer.

The deterioration of the solar pond operational conditions is usually related with a reduction of the thickness of the NCZ layer, because the salinity gradient in this zone is destroyed and the heat transfer is altered (Montalà et al., 2019).

The alteration of the boundaries between the different salinity gradient zones of the solar pond has been considered to be the main source of instability (Leblanc et al., 2011). The changes in environmental variables (e.g., temperature, rain, wind) are expected (Alcaraz et al., 2018a) to affect the overall stability of the salinity and temperature gradients of the solar pond, especially in the UCZ layer and in its boundary with the NCZ layer.

The LCZ layer of the solar pond can also be disturbed by convective forces during operational procedures, such as during heat extraction or salt addition along the maintenance stages (Montalà et al., 2019).

Some previous works have reported the use of deterministic differential equations for the calculation of stability indices (Leblanc et al., 2011; Lu et al., 2004; Alenezi 2012). However, this approach is restrictive in terms of the parameters used for analysis and does not take into account all those that potentially affect stability of the solar pond in a wide time scale. In our previous study (Montalà, et al., 2019) the salinity gradient stability of the Granada solar pond (500 m^2) were reported. The analysis was based on the salinity/temperature stratification in water, which occurs when masses of water at different properties, such as salinity, density or temperature, form different layers without mixing. Results reported provided insights on the sources of instability and provided a tool to control of the salinity gradient stability.

Principal Component Analysis (PCA) is the most popular method in multivariate statistical analysis of environmental data, which is based on the assumption that in the original data sets, a small number of dominant factors (components) with significant influence exist, describing the main sources of variation in the studied system. It is used to explain the complex relationships and/or interactions existing among multiple variables and samples (observations), like in the analysis of environmental monitoring data sets. Usually, the application of PCA allows the investigation of the temporal (seasonal) variations, environmental weather impacts and the monitoring of the patterns/trends in the recorded data sets (Platikanov et al., 2019).

In this work, a new approach for the analysis of the stability of an industrial scale solar pond (Granada, Spain) using PCA, (Jolliffe 2002) is presented. In the case of SGSP systems, measurements of salinity (density) and temperature gradients along the solar pond depth produce two data vectors (profiles): i) one vector for the temperatures and ii) one for the density. When many consecutive observations are recorded for different days/seasons/years, these two vectors can be stored in two data matrices. In these data matrices, the columns represent the measurements of the temperature and density at the different depths of the solar pond, and the rows will represent the different monitoring times (time-stamp). Additional information can be added when environmental conditions, heat extraction and maintenance processes are also monitored for the same period of time and this information is concatenated to the salinity and thermal gradients in the matrices. In this way, PCA as a bilinear decomposition method is very useful simultaneous analysis of the measured data and will allow the extraction of: (i) hidden information about the correlations between both temperature and salinity gradient on one side and the environmental factors and operational variables on the other side; and (ii) information on the most important temporal variations and patterns at different levels of detail, diurnal to seasonal.

2. Material and Methods

2.1 Description of the Granada Salinity Gradient Solar Pond

The solar pond was constructed in the Solvay Minerales facilities in Granada (South Spain) in 2014. Details on the design, construction and operation were reported by Alcaraz et al., (2018a). The solar pond was constructed to deliver the heat needed to preheat the water (> 60 °C) used in the mineral flotation unit. Some features of this solar pond are: the total area of the pond is 500 m^2 (20 × 25 m) with a depth of 2.2 m. The thickness of the LCZ, NCZ and UCZ was 0.6 m, 1.4 m and 0.2 m, respectively. In the LCZ, the density was kept almost constant for 10 months with an average value of 1203 kg/m^3. The heat extraction was carried out through a heat exchanger (PE pipe with an internal diameter of 28 m) located at the LCZ with a total length of 1200 m, which was divided into six independent spirals of 200 m. The solar pond was installed in a mine facility devoted to produce celestine ($SrSO_{4(s)}$). The processed rock, with a celestine content of 30-50%, is milled and then concentrated up to a content of 90% by using a flotation stage. The aqueous solution containing the flotation reagents should be heated to 60-65ºC. Before the installation of the solar pond, water was heated using a boiler fed with gasoil. The solar pond was integrated with the flotation unit by connecting a pipe from the freshwater tank that travels through the LCZ of the solar pond and joins the existing pipe line. A view of the experimental solar pond in Granada is shown in Figure 1.

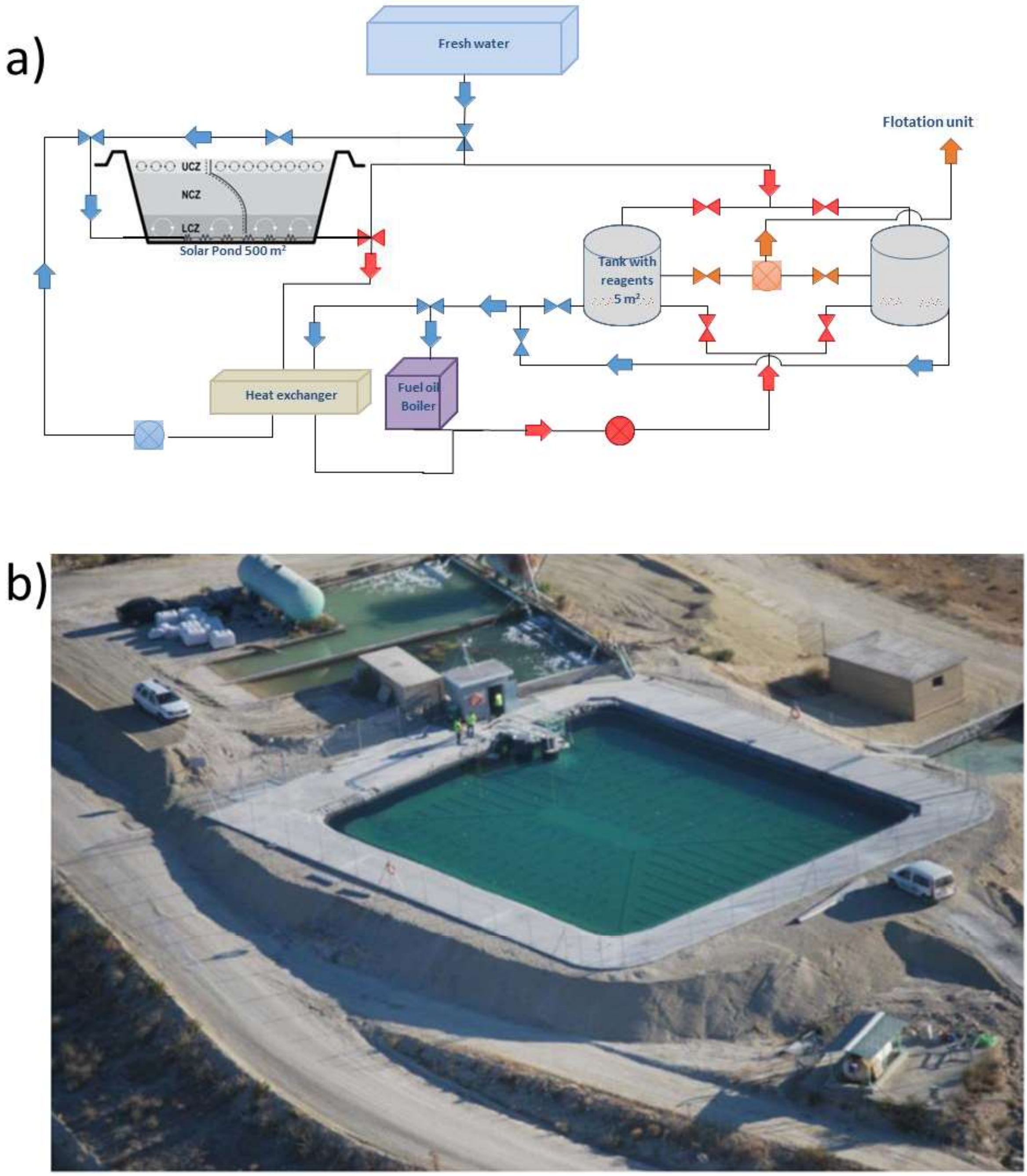

Figure 1. Schematic view showing: a) the integration of the solar pond with the mineral flotation process in Solvay Minerals facilities and b) view of the 500 m^2 solar pond at Solvay Minerales facilities (Granada, Spain)

2.2 Principal Component Analysis

The solar pond data, i.e. salinity and temperature gradients; environmental variables together with the heat extraction data, were arranged in the previously described augmented (concatenated) data matrices and subsequently were analysed by PCA methodology (Jollife, 2002). In this work, PCA results will show the hidden, underlying processes governing the stability of salinity gradient of the solar pond.

According to the PCA model, the original experimental data matrix **D**, is decomposed using a bilinear model, giving two orthogonal matrices, **T** scores (mapping the samples on the principal components), $\mathbf{P}^{\mathbf{T}}$ loadings (mapping the measured variables on the principal components) and **E** is the matrix of residuals (unexplained variance) as in Equation (1)

$$\mathbf{D} = \mathbf{T}\mathbf{P}^{T} + \mathbf{E} \quad (1)$$

The number of principal components in the PCA model (rows of **T** and columns of $\mathbf{P}^{\mathbf{T}}$) in this study was selected on the basis of the two following criteria: i) the sizes of the eigenvalues associated with the principal components and ii) meaningful process explanation of score and loading profiles.

$\mathbf{P}^{\mathbf{T}}$ *loading* profiles show the possible correlations among the environmental conditions, operational heat extraction data, salinity and temperature gradients of the solar pond. **T** scores give the mapping (projection) of the samples (time-stamp measurements) on the principal components. In this study, the scores will give important information about the temporal changes and distribution (measurements over time, time distribution) of the samples during the different campaigns.

PCA also provides various diagnostic tools (Bakeev, 2010) to monitor solar pond stability by very effective and convenient analytic and graphic possibilities for detecting abnormalities that may occur during the solar pond gradient evolution over time and to follow the impact of the environment variables and heat extraction.

Amongst the most popular tools for are: T^2 hotelling values (leverage), which present the sum of the normalized squared scores calculated with PCA, and the Q residuals, which are the measure of the difference, or residual, between a measurement and its projection on the k principal components retained in the model. Both, when plotted, provide very useful charts for the detection of unusual events (Wise and Gallagher, 1996).

Leverages can be used to find very important observations as well as detecting potential out-of-control measurements by calculation of the statistical confidence limits for the values of T^2 setting the threshold line separating in-control from out-of control measurements.

The Q statistics can be used to indicate how well a particular measurement conforms to the model. It gives a measure of the difference, or residual, between a measurement and its projection on the *k* principal components retained in the model. Measurements with very high residuals are not well explained by the model. Confidence limits can be calculated for the model residuals and can serve as threshold giving a limit for the in-control state.

In summary, the application of PCA to the different datasets collected in this work will give a comprehensive overview about the possible correlation among the different variables and measured variables, the discovery of unknown meaningful time trends, the detection of unexpected events and the recovery of valuable information about the stability of salinity gradient of the solar pond technology.

2.3 Dataset organization

Historical temperature (ºC) data were collected from the solar pond and arranged in the data matrix, **Dt** (Ntimes, 40), with 40 measurements (in the columns) every 5 cm from the bottom (t1) to the surface (t40) of the solar pond. These 40 measurements were measured in three different periods on time scale and defined as Ntimes. Density concentrations (g/cm^3) values were collected in the **Ds** (Ntimes, 22) data matrix**,** which has 22 measurements every 10 cm from bottom (s0.1) to 2.2m on the surface (s2.2) throughout the water body. A third data matrix has the heat extraction data (**Dx**(Ntimes, 3), with 3 variables) and a forth data matrix has the environmental variables (weather station) (**Dw** (Ntimes,11), with 11 variables) during 2014-2016 in two operational campaigns. The heat extraction variables include, the time of extraction (*x1*) measured in seconds, the water inflow measured (*x2*) in kg/min, and heat transfer Q (*x3)* measured in MJ. The environmental variables include: air temperature (*w1*) in ºC; relative humidity (*w2*) in %; solar irradiance (*w3*) in W/m^2; accumulated solar irradiance (*w4*) in MJ/m^2; accumulated solar irradiance (*w5*) in kWh/m^2; average wind speed (*w6*) in m/s; maximal wind speed (*w9*) in m/s; average wind direction (*w7*) in degrees as low values mean winds coming usually from North-North East directions and high degree values mean winds coming usually from South-South West directions; standard deviation of wind direction (*w8*) in degrees; wind direction SMM (*w10*) in degrees; and accumulated daily rainfall (*w11*) in mm.

The most important aspect to be considered before their joint analysis of these data sets is the alignment of the different measured variables in accordance to the time frequency of their measurement, i.e. the alignment of the Ntimes observations.

The different datasets described above were arranged in two-dimensional tables or data matrices, where observations/measurements (ordered by specific time-stamp, Ntimes) are in the rows and measured variables/variables are in the columns of a data table (data matrix). Thus, these two-dimensional data tables can be then analysed using existing multivariate statistical and chemometrics bilinear methods (Massart et al., 1998). Collected data was arranged, as shown in Figure 2, in different data sets called respectively:

a) Dataset 1 or **DtDw** (Ntimes,51): Row-wise augmented (concatenated) data matrices containing data about temperature measurements in depth (40 variables from 0 bottom to 2.2 m at the surface) and environmental variables (11). The time-stamp alignment was adjusted for every 30 minutes during the two monitoring campaigns: Campaign 1 took place from 15/07/2014 – 30/04/15 underlining the first solar pond operational period and the concatenated matrix was with dimensions 9349 Ntimes (in the rows) x 51 variables (in the columns). The Campaign 2 or the second solar pond operational period took place from 12/09/2015 till 25/04/2016 and the second concatenated matrix dimension was 5413 Ntimes (in the rows) x 51variables (in the columns).

b) Dataset 2 or **DxDtDw**(Ntimes,54): Row-wise augmented (concatenated) data matrices containing data about heat extraction data (3 variables), temperature measurements in depth (40) and the environmental variables (11). The time-stamp of alignment for these matrices was on daily basis during the same two solar pond operational periods. The first concatenated matrix was with dimensions 290 Ntimes (rows) x 54 variables (columns) and the second concatenated matrix was 227 Ntimes (rows) x 54 variables (columns) corresponding to the two campaigns above.

c) Dataset 3 or **DxDtDwDs**(Ntimes,76): Row-wise augmented (concatenated) data matrices containing data about heat extraction (with 3 variables), temperature measurements in depth (with 40 values), the environmental variables (with 11 variables) and salt concentration measurements (with 22 measurements) in depth. The time-stamp alignment for these datasets was once per week over the same two monitoring campaigns. Thus, the new concatenated matrices were with dimensions 33 Ntimes (rows) x 76 variables (columns) for the first solar pond operational period, and 29 Ntimes (rows) x 76 variables (columns) for the second solar pond operational period.

d) Reduced **DxDtDs** taken from Dataset 3. This dataset included a reduced number of variables – 65, measured one time per week for during the first operational period. The included variables were 3 heat extraction variables, temperature measurements in depth (40 variables) and salinity gradient in depth (22 variables). This new dataset was divided into a calibration dataset including the same 33 observations, and the external validation dataset included 29 observations for the same number of variable during the second operational period.

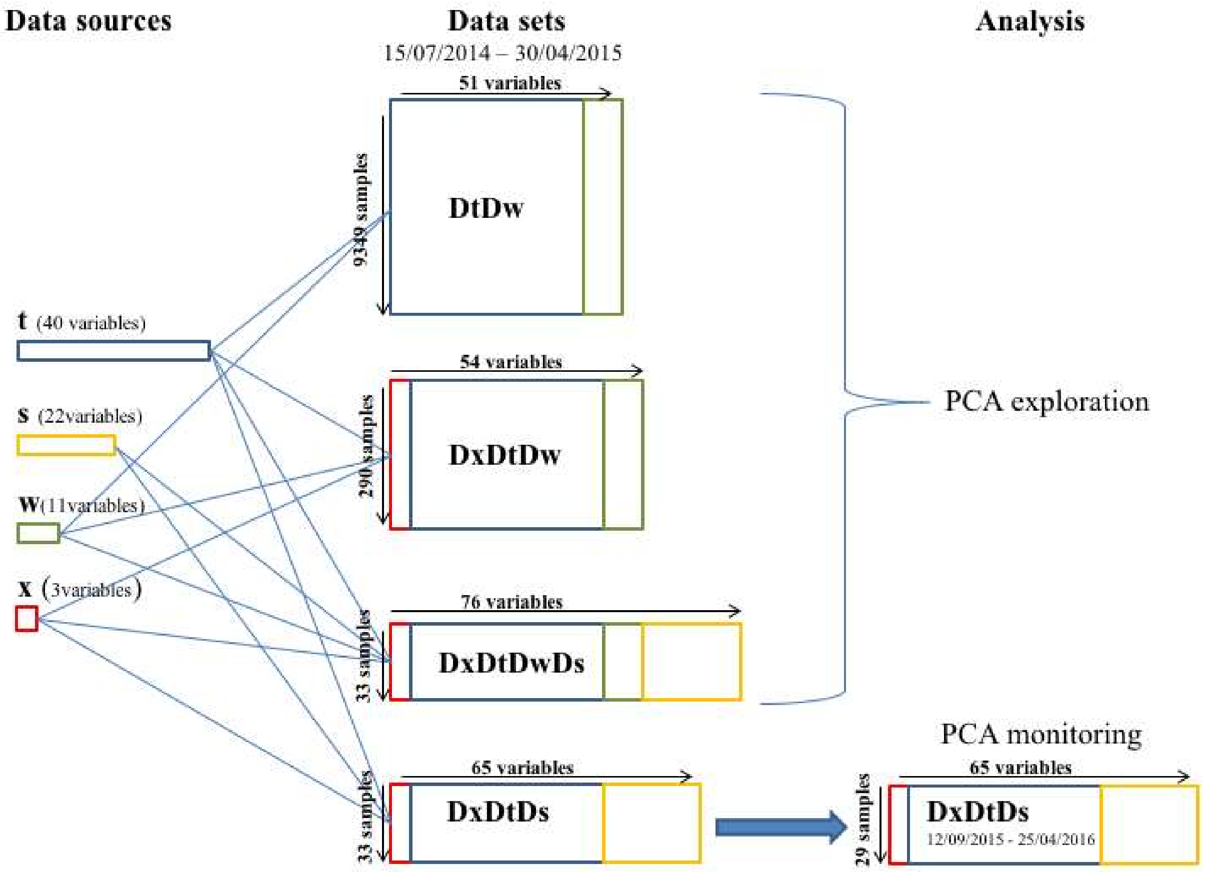

Figure 2. Data arrangement and application of PCA on the different data sets. t temperatures (ºC) collected in the solar pond every 5 cm from the bottom (t1) to the surface (t40) of the solar pond and arranged in the data matrix, Dt (Ntimes, 40), with 40 variables in the columns; s salinity density concentrations (kg/m^3) arranged in the Ds (Ntimes, 22) data matrix, with 22 variables every 10 cm from bottom (s0.1m) to s2.2m on the surface throughout the water body; w environmental variables from a weather station) arranged in Dw (Ntimes,11) matrix, with 11 variables; x heat extraction data (Dx(Ntimes, 3), with 3 variables).

3. Results and Discussion

The first three datasets arranged in this way provide some advantages in their consequent multivariate analysis. For example, PCA of Dataset 1 would provide very detailed information about the time resolution in their PC scores, since data records were made at every 30 minutes. In contrast, PCA of Dataset 2 and Dataset 3, would provide better information about possible variable interactions in the PC loadings, although the span of measurements in time (one per day for Dataset 2 and one per week for Dataset 3) in this case would miss important temporal information present in Dataset 1. For this reason, a global study of the three datasets is attempted in this work.

The PCA monitoring strategy proposed in this work for stability analysis of the salinity and temperature gradients, relies on a model, which is built with control data (when the solar pond is considered to behave with stable temperature and salinity gradients), during the first period of the solar pond operation in 2014-15 years. Then, the established model is validated on data from the second operation period during 2015-16 allowing to inspect the consequent PCA graphics for abnormal measurements and unexplained variance of investigated variables included in datasets. Since, the environmental variables were not controllable, the model validation approach, developed in this work, included only observations of heat extraction, temperature measurements in depth and salinity gradient in depth organized in the reduced **DxDtDs** taken from Dataset 3.

3.1 Seasonal variability and factors affecting the stability of the solar pond during the first operational period. Datasets 1-3 PCA results

Three component PCA models have been obtained in the analysis of Datasets 1, 2 and 3. All three models explain more than 80% of the total variance in the corresponding datasets. In particular, for the Dataset 1, **DtDw** row-wise augmented data matrix, (with solar pond's temperatures in depth, **Dt** data matrix, and environmental variables, **Dw** data matrix) the explained variance was close to 87%; for Dataset 2, **DxDtDw** row-wise augmented data matrix (with heat extraction variables, **Dx** data matrix, in addition to those in **DtDw** data matrix) the explained variance was 84%; for Dataset 3, **DxDtDwDs** data matrix (with the salinity concentrations in depth **Ds** in addition to those in **DxDtDw** data matrix**)** the explained variance was 82%. The % of explained information decreases in the order **DtDw** > **DxDtDw** > **DxDtDwDs**.

Figure 3 shows the representation of the PC1 loadings (Fig. 3a, 3c and 3e) and scores (Fig. 3b, 3d and 3f) obtained in the PCA of each of the three datasets.

This PC1 can be related to the effect of the changes in the seasonal environmental temperature on the solar pond temperature gradient in the three datasets. In the plot of the three corresponding loadings plots (Fig. 3 a, c and e), a strong positive correlation is observed between the environmental temperature (*w1*), the solar irradiance (*w3*) and the solar pond temperature variables (*t1-t40*). Moreover, the same plots reveal the presence of an inverse correlation between the solar pond temperatures and the environmental humidity. As expected, the humidity increases outside the solar pond when average air temperatures drop.

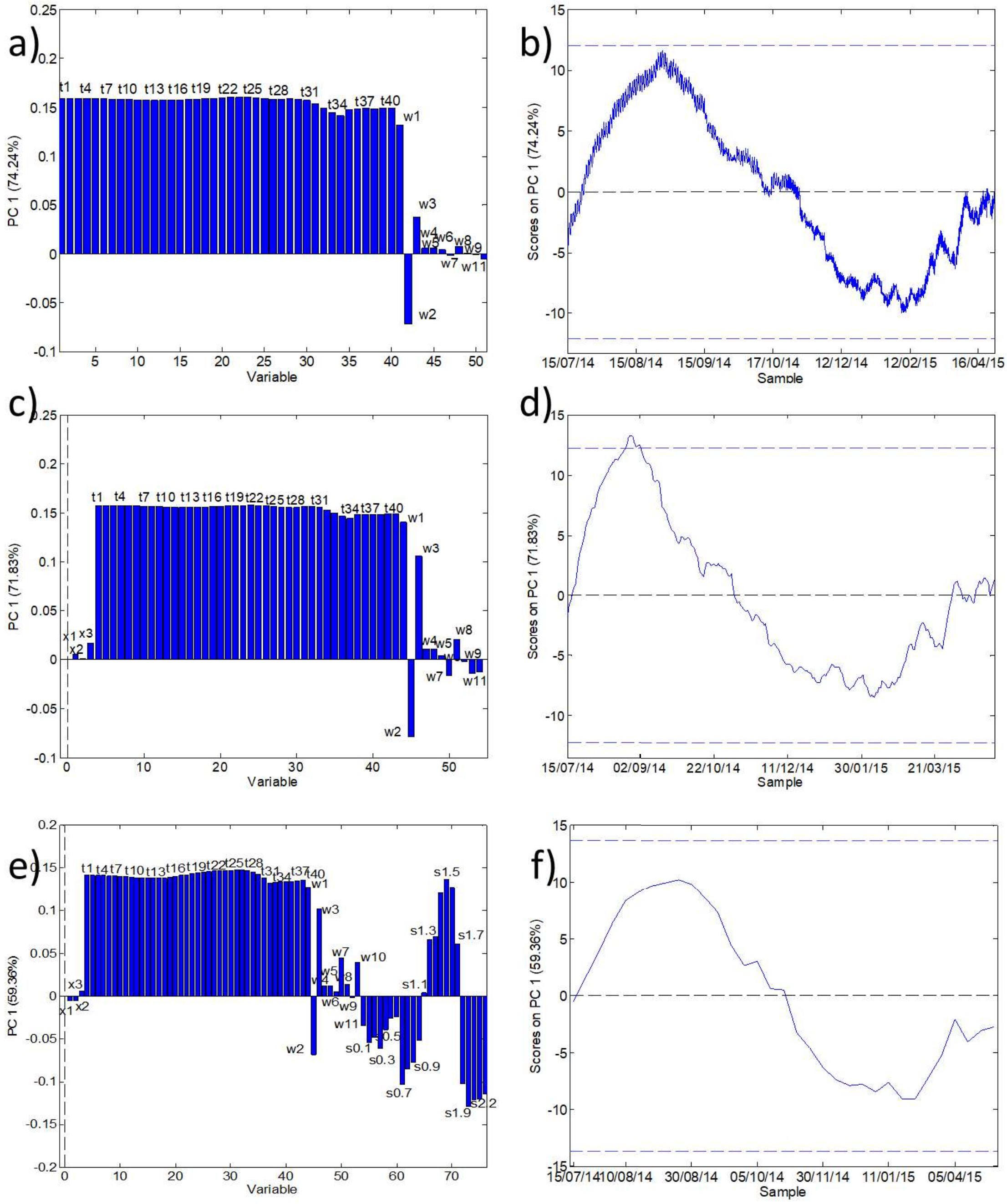

Figure 3. PCA results of the three data sets with the different variables measured at the Granada solar pond during the first operation season 15/07/2014 – 30/04/2015. PC1 loadings (a) and scores (b) in the analysis of **DtDw** data matrix; PC1 loadings (c) and

scores (d) in the analysis of **DxDtDw** data matrix; PC1 loadings (e) and scores (f) in the analysis of **DxDtDwDs** data matrix. Solar pond temperature variables: t1 (bottom) –t40 (surface) in ºC; Salinity density variables: s01 (bottom) - s2.2 (surface) in g/cm^3; Heat extraction variables: time of extraction (*x1*) measured in seconds, water inflow (*x2*) measured in kg/min and heat transfer Q (*x3)* measured in MJ.

Environmental variables: air temperature (*w1*) in Cº; relative humidity (*w2*) in %; solar irradiance (*w3*) in W/m^2; accumulated solar irradiance (*w4*) in MJ/m^2; accumulated solar irradiance (*w5*) in kWh/m^2; average wind speed (*w6*) in m/s; maximal wind speed (*w9*) in m/s; average wind direction (*w7*) in degrees; standard deviation of wind direction (*w8*) in degrees; wind direction SMM (*w10*)in degrees; and accumulated daily rainfall (*w11*) in mm.

The scores of PC1 in the three plots (Fig. 3 b, d and f) are very similar on shape among them, while explaining 74%, 72 and 60% of the variance and reflect the presence of a seasonal time series pattern with a smooth sigmoidal trend over the entire period of investigation. Positive scores are observed in summer, then gradually decrease to negative scores in winter and return back to positive signs in spring. At the beginning of this monitoring period, in August-September, the solar pond temperatures rose from the bottom to the top levels of the pond due to the effect of stronger solar irradiance (*w3*-*w5* have positive loadings) and higher average air temperatures (*w1*positive loadings). In contrast, starting from November, coinciding with the decrease of temperatures, daylight time and sunlight irradiances (inversely correlated to humidity, w2), the trend was the opposite, going down, with the solar pond temperature gradients at the lowest temperatures in February. From March, the temperatures outside and inside the solar pond rose again with the longer daytime and stronger solar irradiance. The loadings of the heat extraction variables (*x1*-*x3*) in Fig. 3c are featureless, showing that heat extraction had insignificant contribution to stability of the temperature gradient for PC1. In Fig. 3e the impact of the changes in the average air temperatures on the salinity gradient can be analysed in detail. Three solar pond salinity zones can be distinguished in this figure. The LCZ is between 0.1 and 0.6 m from the bottom; the NCZ is between 0.6 and 1.7 m from the bottom; and the UCZ is between 1.7 and 2.2m from the bottom. The salinity concentrations at 1.5-1.6 m depth from bottom (with positive loadings in Fig. 3e) are inversely correlated with the salinity concentrations at the surface of the solar pond between 1.8-2.1m (with negative loadings). This inverse correlation is produced in the boundary between UCZ and NCZ zones. Salinity concentrations around 1.5 m from the bottom shows positive correlation with average air temperatures (*w1*) and the solar pond temperature measurements, meaning that the salinity concentrations increase at this zone with the increase of temperature. On the contrary, in the UCZ, salinity concentrations are inversely correlated to the average air temperatures. All this suggests that water evaporation was probably taking place at the pond surface and fresh water was added to compensate the evaporation losses and as a result, a salt increase occurred at 1.5-1.6 m depth (in the upper part of the NCZ). Looking at the corresponding scores plot on Fig. 3f, it can be observed that this fact happened in the summer (positive scores), while in the winter, the results suggest the salt concentrations in the UCZ increased with lower average air temperatures, due to the less evaporation losses. The lower average air temperatures could also impact the bottom part of the solar pond. The salt concentrations between 0 and 0.6/0.7 m from the bottom reported negative loadings. Therefore, they are inversely correlated with average air temperatures, suggesting that in the winter the LCZ increased its extension, while the NCZ reduces its thickness. It can be related to the less diffusion of salt in winter time and also due to the maintenance of the salinity gradient, that in this case, by adding salt into the LCZ could affected the stability at the boundary with the NCZ.

The more interesting changes in terms of temperature and density gradients are located in the critical boundaries between the pond regions LCZ-NCZ and NCZ-UCZ as has been previously reported (Montalà et al., 2019). These changes are not desired for the proper operation of the solar pond and they should be avoided. The continuous monitoring of the average air temperature impact on the salinity gradient is required. However, our findings suggest that this variation is in a large extension a natural process, and therefore it is difficult to control. Moreover, the maintenance of the solar pond, adding fresh water, on the solar pond surface, to compensate evaporation losses at the UCZ and adding salt to the LCZ to compensate the salt diffusion to the pond surface are also critical processes that should be monitored and controlled according to the variation of the temperature and salinity gradients in order to avoid disturbances as has been demonstrated.

The PC2 loadings and scores obtained in the analysis of the three datasets are shown in Figure 4. This PC2 explains 8.5% of the variance of **DtDw**; 7.1% of the variance of **DxDtDw** and 13% of the variance of **DxDtDwDs.**

The loadings for the solar irradiance (*w3-w5)*, average air temperature (*w1*) and wind direction (*w7* and *w8*) variables show the highest positive loadings on the three plots of Fig. 4a, c and e. All they are positively correlated to UCZ and upper NCZ temperatures measurements (*t34-t40*). In contrast, the humidity shows large negative loadings in the three plots, meaning that it is inversely correlated to the UCZ temperatures (Fig. 4 a, c and e). Consequently, the daylight solar, temperature fluctuations and wind direction could affect significantly the temperatures from UCZ layer and affect the boundary with NCZ. The two wind velocity variables (*w6* and *w9*) show negative correlation with the temperature measurements at UCZ and upper NCZ and the variables for wind direction (*w7* and *w8*) in Fig. 4c and 4e. The logical explanation is that in sunny days with winds coming from directions with high degree values (South-South West) the UCZ layer and upper NCZ of the pond can be heated strongly. On the contrary, at nights and when strong winds occur in the area, these layers are cooling very fast. These results coincide with the permanent daily-night temperature strong fluctuations, seen on the scores plot of Fig. 4b. The trend lines of the scores in Fig. 4d and 4f also show the process of surface cooling due to strong winds coming from direction with low degree values (North-North East direction) during the autumn and winter months.

It is worth noting, that the UCZ is affected by the maintenance actions necessary to compensate for evaporation losses and increased salinity in this layer. These trends, identified by the PCA methodology, confirm the behaviour observed in the operation of the solar pond (Valderrama et al, 2011; Bernad et al, 2013, Alcaraz et al, 2018a).

Loadings plot in Fig. 4c, suggests that heat extraction in the solar pond was done only during the day, and not at night, although this process does not have a significant effect on the temperature gradient near the boundary of the LCZ and NCZ, with loadings close to 0.

Loadings plot in Fig. 4e shows that salinity gradient was relatively stable in LCZ, which spreads from bottom to the 0.6 m depth; NCZ area is located between 0.6-1.8 m depth and UCZ, between 1.9-2.1 m. A strong inverse correlation in salinity is observed between the sublayers at 1.7-1.8 m and those at 1.9-2.2 m, pointing out the strong effect of solar irradiation and winds effect (with positive loadings), on the boundary between NCZ and UCZ. The relative humidity (*w2* with negative loading) is again an inverse correlation to the solar irradiance and wind effects, confirming the observation that on sunny and windy days its values decreases.

The scores plot of Fig. 4d and 4f provide some additional information about the salinity gradient operational conditions in UCZ during the first operation season. The changes from positive to negative scores in the initial stage of the season show the occurrence of higher density values in the upper part of the NCZ (1.7-1.8 m) in comparison to the UCZ (1.9-2.2 m). Coinciding with the higher daily environmental temperatures for July and August, it is logical to suggest a water evaporation took place at UCZ.

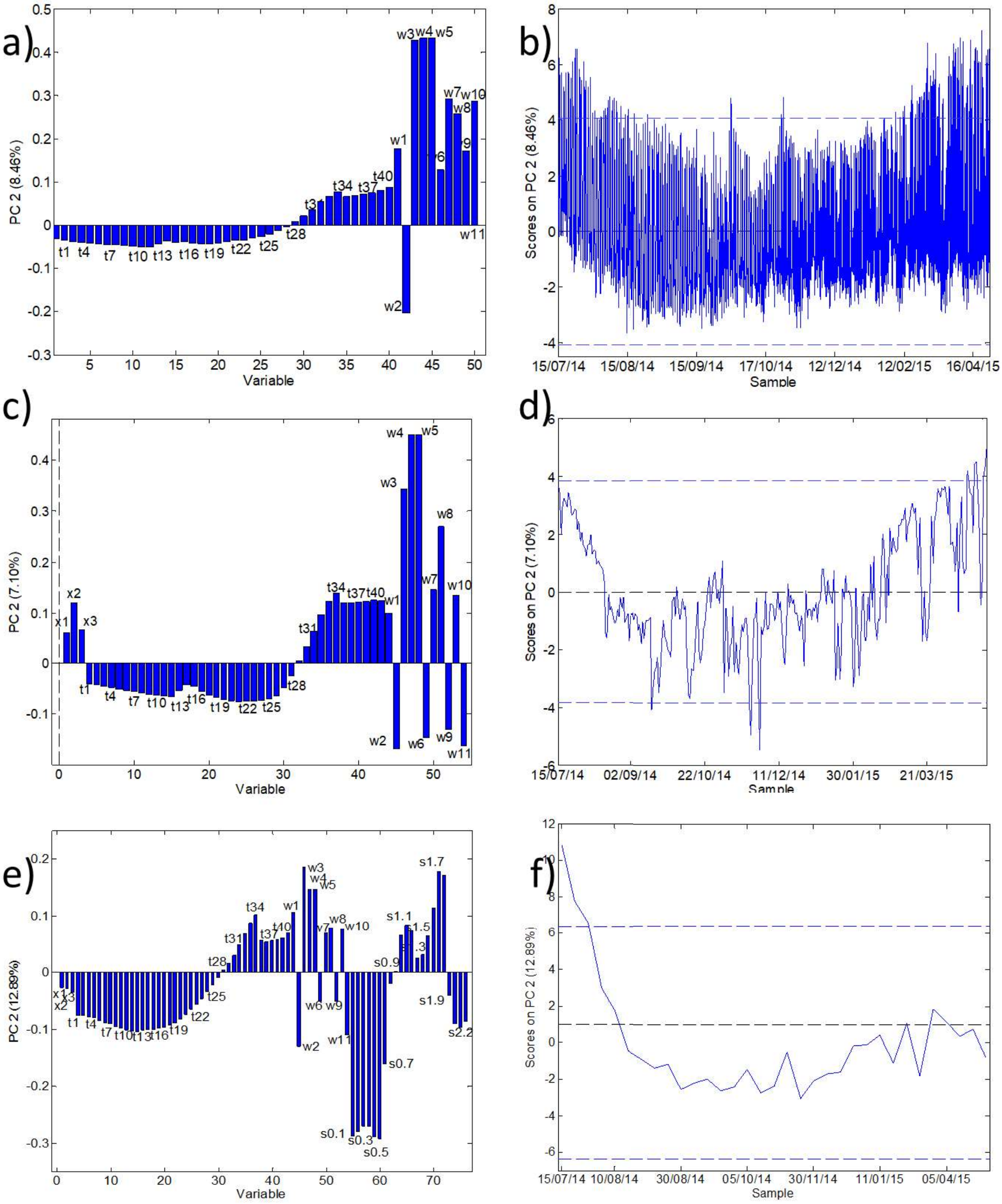

Figure 4. PCA results of the three data sets with the different variables measured at the Granada solar pond during the first operation season 15/07/2014 – 30/04/2015. PC2 loadings (a) and scores (b) in the analysis of **DtDw** data matrix; PC2 loadings (c) and

scores (d) in the analysis of **DxDtDw** data matrix; PC2 loadings (e) and scores (f) in the analysis of **DxDtDwDs** data matrix.

Solar pond temperature variables: t1 (bottom) –t40 (surface) in ºC; Salinity density variables: s01 (bottom) - s2.2 (surface) in g/cm^3; Heat extraction variables: time of extraction (*x1*) measured in seconds, water inflow (*x2*) measured in kg/min and heat transfer Q (*x3)* measured in MJ.

Environmental variables: air temperature (*w1*) in Cº; relative humidity (*w2*) in %; solar irradiance (*w3*) in W/m^2; accumulated solar irradiance (*w4*) in MJ/m^2; accumulated solar irradiance (*w5*) in kWh/m^2; average wind speed (*w6*) in m/s; maximal wind speed (*w9*) in m/s; average wind direction (*w7*) in degrees; standard deviation of wind direction (*w8*) in degrees; wind direction SMM (*w10*) in degrees; and accumulated daily rainfall (*w11*) in mm.

The plots of the PC3 loadings and scores obtained in the analysis of the three datasets are given in Figure 5. This PC3 explains 4.5% of the information in **DtDw**; 5.3% of the information in **DxDtDw** and 9.6% of the information in **DxDtDwDs.** These smaller amounts of data variance explained by this PC are not so significant as for PC1 and PC2. However, some interpretation of them can still be given.

On the loadings plot of Fig. 5a, two large regions of the temperature gradient show inverse correlation between them defining the temperature gradient in the NCZ of the pond. The temperatures measured from the UCZ to the middle of the NCZ in the solar pond show positive loadings, while the temperature measurements from the middle of the NCZ to the bottom of the LCZ have negative loadings. The boundary between LCZ and NCZ is not clearly defined on this loadings plot. Changes in the temperature values of the NCZ are in range between t10-t34. This suggests that temperatures increased at the bottom of the NCZ. Temperatures in the UCZ varied from t37 to t40 (at the pond surface) as has been discussed for PC2. The humidity variable (*w2*), associated with cloudy days or in the nights, is in positive correlation with temperatures measurements in LCZ and the bottom half of NCZ layers of the pond and in inverse correlation with the temperatures measurements in UCZ, suggesting heat accumulation in LCZ and bottom half of NCZ, while UCZ is cooling.

The analysis of the PC3 scores plot of Fig. 5b reveals the strong day-night fluctuations (relevance of the impact of humidity variable). Moreover, the temperatures increased in LCZ and bottom part of NCZ (heat transfer from the bottom to the middle part of the NCZ) rapidly from July to August and then followed a steadier trend until the end of the monitored period. These results suggest that heat extractions activities during night hours instead only during daily hours, especially in the summer time, can prevent excess heat accumulation at the boundary of the NCZ and LCZ, thus reducing the capacity of solar pond storage.

Fig. 5c shows that the heat extraction activities had little or no effect on the temperature gradient, as there were no positive nor negative correlations with the temperature measurements (they show loadings close to 0) of the solar pond. The corresponding scores plot on Fig. 5d shows clearly the days (the observations with the highest score values) when larger amounts of heat were extracted.

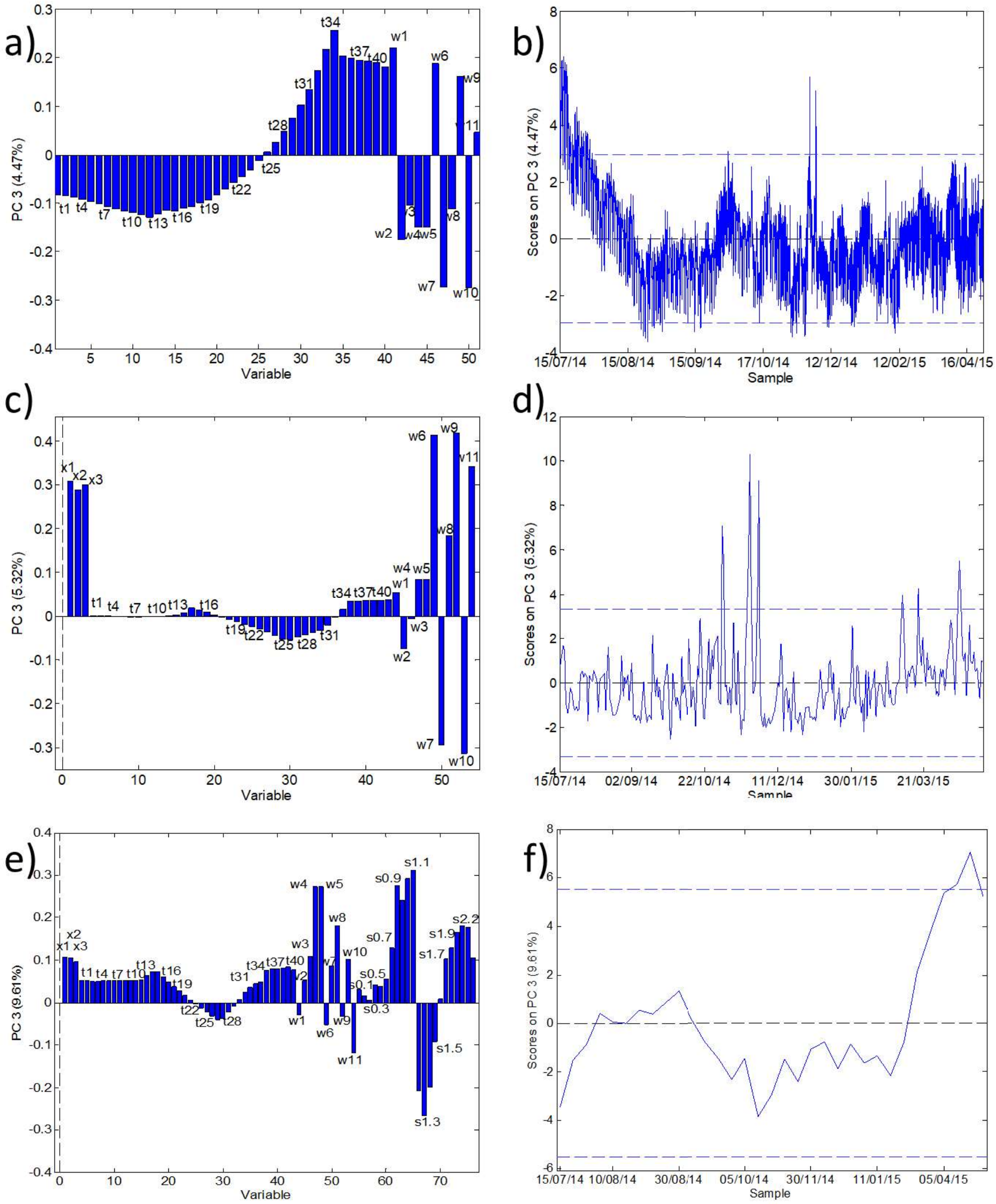

Figure. 5. PCA results of the three data sets with the different variables measured at the Granada solar pond during the first operation season 15/07/2014 – 30/04/15. PC3 loadings (a) and scores (b) in the analysis of **DtDw** data matrix; PC3 loadings (c) and scores (d) in the analysis of **DxDtDw** data matrix; PC3 loadings (e) and scores (f) in the analysis of **DxDtDwDs** data matrix.

Solar pond temperature variables: t1 (bottom) –t40 (surface) in ºC; Salinity density variables: s01 (bottom) - s2.2 (surface) in g/cm^3; Heat extraction variables: time of extraction (*x1*) measured in seconds, water inflow (*x2*) measured in kg/min and heat transfer Q (*x3)* measured in MJ.
Environmental variables: air temperature (*w1*) in Cº; relative humidity (*w2*) in %; solar irradiance (*w3*) in W/m^2; accumulated solar irradiance (*w4*) in MJ/m^2; accumulated solar irradiance (*w5*) in kWh/m^2; average wind speed (*w6*) in m/s; maximal wind speed (*w9*) in m/s; average wind direction (*w7*) in degrees; standard deviation of wind direction (*w8*) in degrees; wind direction SMM (*w10*) in degrees; and accumulated daily rainfall (*w11*) in mm.

Finally, Fig. 5e shows that not only energy was accumulated in the lower NCZ (explained above in Fig. 5c), but also the density values were consequently increasing (propagation) up to 1.1 m from the bottom. These loadings plot in Fig. 5e also shows that the density values at 1.3- 1.5 m were in inverse correlation with the density values, from the bottom of the solar pond, showing the reduction of the NCZ thickness between 1.1 and 1.5 m, and also the possible formation of two sublayers within the NCZ with different density. Looking at the scores plot of Fig. 5f, this evolution of forming two sublayers of different density values at NCZ began immediately after the start of solar pond its first operation season.

3.2 Evaluation of PCA for monitoring of solar pond operational procedures using the heat extraction variables, temperature and salinity gradients

PCA model statistics examination is very common issue for the detection of unusual events (Bakeev, 2010). In the case of SGSP process monitoring, these unusual events can be produced by possible strong influence of environmental parameters, inadequate heat extraction and unexplained changes in thermal or salinity concentration gradients. The PCA model statistics (highlights these unusual events with significant differences between measurements and their projections on the k principal components retained in the PCA model along time (see Section 2.2). Samples found with large leverages and high values of Q residuals can indicate influential observations as well as detecting potential out-of-control samples. The observed samples with large T^2 hoteling values and Q residuals values are not well explained by the established PCA model and they can be considered as unusual events. If a sample is flagged as an event, the next step includes examination of the so-called Q residuals contribution plot (Wise and Gallagher, 1996) to detect which variable/parameter contributes strongly to the overall Q value for the considered sample. Confidence limits can be also calculated for further examination to provide threshold values to define regular or outlying conditions. When the sample measurements are under control, Q residuals should display small values within these confidence limits. Unusual (outlying) events are displayed outside these confidence limits.

Once, a PCA model is built, it provides effective analytic and graphic options (T^2 hoteling and Q residuals) to detect anomalies. In this study, this approach is extended to the data from the second operational period using a PCA model built with data from the first operational period. It is worth mentioning, that in order to proceed correctly, data from the first operational period should include only values that were considered to be in control, while any unusual measurements required to be cleaned before. In this study, it was verified that data from the first operational period represented a steady state operation process without containing significant anomalies. A period of time, between the first and the second operational periods, was omitted since the solar pond did not show operational stability.

The new PCA model was built using the auto scaled, reduced variables ***DxDtDs*** dataset, with three principal components explaining more than 89 % of the data variance with the same three dominant factors explained in the previous sections. The order of the captured variance was PC1 67 % > PC2 14 % > PC3 8.6%. PC1 on Fig. 6a describes the seasonal temperature changes; PC2 on Fig. 6c described the changes in the salinity gradient and the heat accumulation and distribution from the LCZ to NCZ. PC3 on Fig. 6e describes variance as the heat accumulation at UCZ close to the boundary with NCZ layer during days with a strong solar irradiance and wind (explained also in the previous section).

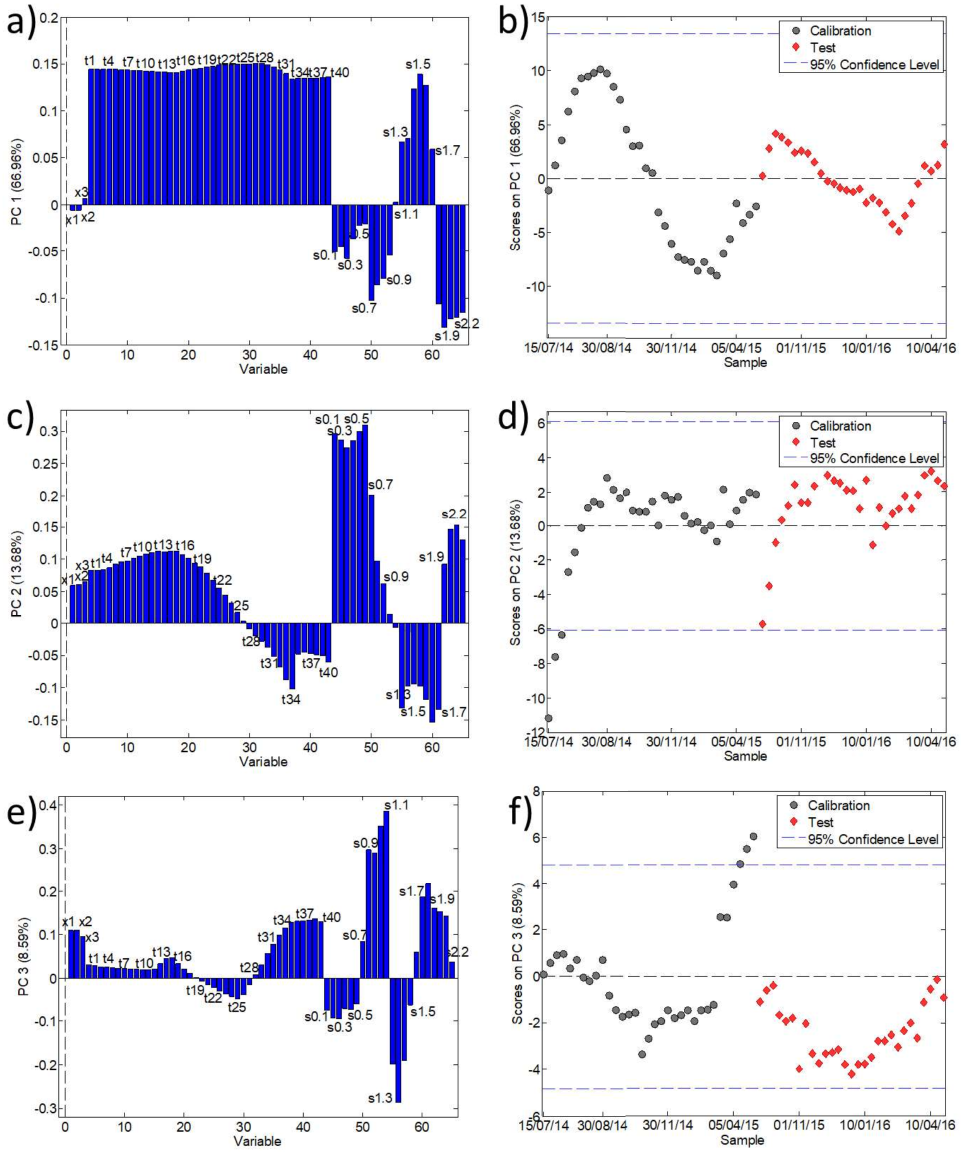

Figure 6. PCA model validation results of the reduced data sets **DxDtDs** measured at the Granada solar pond during 2014-15 (calibration) and 2015-2016 (external validation, test). PC1 loadings (a) and scores (b); PC2 loadings (c) and scores (d); PC3 loadings (e) and scores (f) in the analysis.

Solar pond temperature variables: t1 (bottom) –t40 (surface) in ºC; Salinity density variables: s01 (bottom) - s2.2 (surface) in g/cm^3; Heat extraction variables: time of extraction (*x1*) measured in seconds, water inflow (*x2*) measured in kg/min and heat transfer Q (*x3)* measured in MJ.

While only the first PC1 (Fig. 6a) shows normal functioning of the solar pond, the second (Fig. 6d) and the third factors (Fig. 6f) explain unwanted sources of variation for the solar pond operation processes and consequently its stability. In Fig. 6d, the first three samples from the calibration set (black dots) are observed outside the threshold boundaries (blues dotted lines) suggesting that still the solar pond was not stabilized. In a similar way, the scores plot of Fig. 6f shows that the last samples are observed as abnormal, suggesting that the solar pond already was not operating properly. However, the three PCs have been maintained in the PCA model for its consequent validation on the new data from the second operation season, because PC2 and PC3 can be useful to evaluate the extent of influence of these unwanted factors on the solar pond stability during the second period of operation.

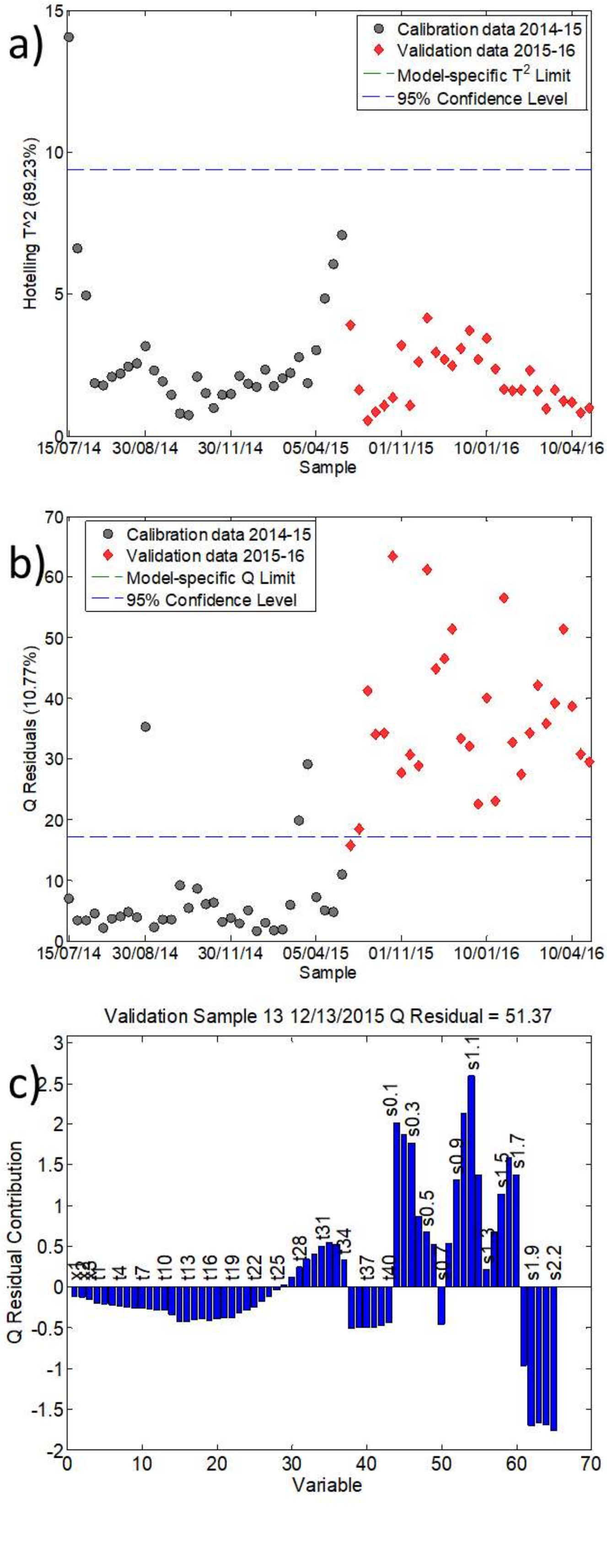
a)
Calibration data 2014-15
Validation data 2015-16
Model-specific T^2 Limit
95% Confidence Level
Hotelling T^2 (89.23%)
15/07/14 30/08/14 30/11/14 05/04/15 01/11/15 10/01/16 10/04/16
Sample
b)
Calibration data 2014-15
Validation data 2015-16
Model-specific Q Limit
95% Confidence Level
Q Residuals (10.77%)
15/07/14 30/08/14 30/11/14 05/04/15 01/11/15 10/01/16 10/04/16
Sample
Validation Sample 13 12/13/2015 Q Residual = 51.37
c)
Q Residual Contribution
Variable

Figure 7. Plots of: a) Hotelling's T^2; b) Q residuals; and c) Q variable contributions for observations from the first operation season (calibration dataset) and from the second operation season (external validation dataset).

Solar pond temperature variables: t1 (bottom) – t40 (surface) in ºC; Salinity density variables: s01 (bottom) - s2.2 (surface) in g/cm^3; Heat extraction variables: time of extraction (*x1*) measured in seconds, water inflow (*x2*) measured in kg/min and heat transfer Q (*x3)* measured in MJ.

The scores plots (Fig. 6b, 6d and 6f) for these three PCs, together with the T^2 hotelling values are plot in Fig. 7a and 7b show the samples for calibration (data from the 1st operation season) in black dots and the samples for the external validation (2nd operation season) in red diamonds. All samples from the second period (external validation data) resulted to be under control with values below the threshold at 95% confidence level, shown with the upper blue dotted line on the three PCs scores plots and the T^2 Hotelling values plot. This fact means that the process during the second operation season of solar pond followed similar trends as during the first operation season. There was not great difference in the impact of the average air temperature on the stability of the two operational periods. The impact of the increase of temperature and density at LCZ and NCZ and of the strong solar irradiance and winds were also similar during both operation seasons.

The Q-statistics results depicted in Fig. 7b, shows all samples from the second operation season above the 95% confidence interval threshold due to unexplained by the PCA model variance in some variables. The Q contributions of the all monitored variables for validation sample number 13 (Fig. 7c) from 13/12/2015 reveal why this sample was situated well above the established Q statistics threshold. For this sample, the density values from LCZ and NCZ 0.1-1.7 m presented high positive scores. This fact can be interpreted with the higher density at these zones, LCZ and NCZ, during the second operation season in respect to that observed during the first operation season. On the contrary, the density values from the UCZ, present negative scores on this plot. Hence, the salinity concentrations at UCZ were lower during the second period than during the first period.

Both, the operation seasons can be assumed as two separate batch experiments and PCA model validation could show clearly the different starting operational condition in the second period.

PCA served for statistical process control can serve as a tool for maintaining the operation of solar ponds. The classical univariate statistical approach (the actual control at many facilities nowadays) has focused on the control of one variable at a time (density or temperature). Thus, the obtained results are not very informative for analysis of the maintenance of the solar pond gradient as a process. The PCA results of this study pointed out that this process depends on multiple operational and environmental variables exhibiting various interactions with the solar pond gradient.

The proper strategy (before to establish any PCA model for process monitoring) would consider similar starting operational conditions for every consequent new batch experiment. These operational conditions should include an establishment of an initial density gradient with similar concentrations and profile in depth. The initial stage of each new operational period is critical for the solar pond settlement and for the further PCA process control monitoring. Once, the solar pond is stabilized and PCA results, visible on the scores plot, suggest that the process is in control, the main operational efforts should be focused on the proper maintaining of the solar pond and a mitigation of possible undesired effects with environmental parameters. At this moment, the correct strategy implies a monitoring for observations with unexplained variance in some of their variable measurements. The monitoring program is also an important aspect for the proper control of the solar pond operational process and it should be focused on the local environmental conditions. The PCA results of this study revealed the day-night solar temperature dynamics, wind velocity and wind direction to have influence on the solar pond gradient in Granada. Continuous measurements, several times per day, should offer a proper timeframe to detect the undesired impact on the gradient of any sudden changes in the weather conditions. From the results obtained in this study, the importance of environmental factors, the extraction of heat and the maintenance of the gradient is observed. In this sense, some guidelines can be followed to track the stability of the salinity gradient solar pond: i) the boundaries regions UCZ-NCZ and NCZ-LCZ are the points where instability can be identified, a frequent control of the depth of each zone, especially the NCZ, and also the variability in these boundaries regions is recommended every month; ii) the impact of heat extraction on the LCZ should be controlled during cold months, temperature differences between NCZ and LCZ can be a source of instability that can be avoided, then heat extraction can be used to regulate the temperature in the LCZ; iii) salinity gradient maintenance operations are key to controlling the depth of NCZ, but can also provide sources of instability, the addition of salt in LCZ and fresh water in UCZ should be planned depending on the season and also the current level of stability.; iv) among the different environmental factors, the wind is by far the one that most impacts stability, it is recommended that the choice of sites for the construction of the solar pond be in sites not affected by strong winds.

4. Conclusions

PCA has been used to analyse simultaneously as advantage the changes in stability gradient of an industrial solar pond in terms of the salt concentrations (density) and temperature changes produced during two operations seasons in correlation to the influence of environmental factors and/or heat extraction procedures. PCA as an efficient method can highlighted the existence of strong correlations between the temperature and salinity gradients, in relation to the changing environmental variables, and with the heat extraction when all monitored parameters are aligned in various datasets in matrix form. During the two operations seasons the Granada solar pond supplied 79 and 94 MJ, with an efficient of 10 and 12%, respectively, the salinity gradient was established with a density in the NCZ that ranged from 1014 to 1204 kg/m^3, and decreased from 1020 to 1185 kg/m^3 after the gradient was considered destroyed. Three similar sources of operational process variation were identified that impact the stability of the SGSP. As expected, the major factor (variance source) was the changes of seasonal temperatures, defining the lowest (in winter) and highest (in summer) temperatures inside the solar pond. The second factor was the effect from daily-night solar irradiations, i.e. the diurnal temperature fluctuations in combination with the wind currents, which affected strongly the UCZ and as a consequence, the boundary between UCZ and NCZ. As a result, water evaporation takes place at UCZ, which produces a salt accumulation in the upper parts of the NCZ layer. Then, the necessary fresh water added to the UCZ also can potentially contribute to the instability of this boundary, thus making necessary to develop appropriate methods seasonally dependent to accurately ensure that this operation would not affect the stability of the salinity gradient. A third resolved factor explained the temperature and density increase between LCZ and the lower half of NCZ layer due to the insufficient heat extraction and/or to the improper conditioning of the initial state of the LCZ salinity layer settlement. It is recommended to include the monitoring of this third factor, since its trend with the time may alert the irreversible deterioration of the heat extraction potential from the solar pond. This study proposes a new strategy as a monitoring tool for the simultaneous analysis of multiple variables and operational procedures of a solar pond through the application of PCA. This strategy offers a promising approach to improve your control and management of the different gradient maintenance and heat extraction processes.

Acknowledgements

SP and RT would like to gratitude the Spanish Ministry of Science and Innovation for the PID2019-105732GB-C21 and CTQ2017-82598-P projects financial support. CV and JLC would like to gratitude the Spanish Ministry of Economy and Competitiveness (MINECO) for the Waste2Product project (CTM2014-57302-R) and by the R2MIT projects (CTM2017-85346-R) and Catalan Government (2017-SGR-312), Finally, the authors gratefully acknowledge personnel from Solvay Minerales and Solvay Martorell facilities for practical assistance, especially to M. Gonzalez, C. Gonzalez, C. Aladjem, and M. Giménez for their valuable cooperation.

References

Alcaraz, A., Montala, M., Valderrama, C., Cortina, J.L., Akbarzadeh, A., Farran, A. 2018a. Thermal performance of 500 m^2 salinity gradient solar pond in Granada, Spain under strong weather conditions. Solar Energy 171, 223-228.

Alcaraz, A., Montalà, M., Valderrama, C., Cortina, J.L., Akbarzadeh, C., Farran, A., 2018b. Increasing the storage capacity of a solar pond by using solar thermal collectors: heat extraction and heat supply processes using in-pond heat exchangers. Sol. Energy 171, 112–121.

Alcaraz, A., Montala, M., Cortina, J.L., Akbarzadeh, Aladjem, C., Farran, A., Valderrama, C.2018c. Design, construction, and operation of the first industrial salinity-gradient solar pond in Europe: An efficiency analysis perspective. Solar Energy 164, 316-326.

Alenezi, I., 2012. Salinity Gradient Solar Ponds: Theoretical Modelling and Integration with Desalination. PQDT – UK Irel. https://doi-org.recursos.biblioteca.upc.edu/May, 2012.

Bakeev, K. 2010.Process Analytical Technology: Spectroscopic tools and implementation strategies for the chemical and pharmaceutical industries 2nd Ed. Wiley

Bernad, F., Casas, S., Gibert, O., Akbarzadeh, A., Cortina, J.L., Valderrama, C., 2013. Salinity gradient solar pond: Validation and simulation model. Solar Energy 98, 366-374.

Chakrabartya. S.G., Wankhede, U.S., Shelke, R.S., Gohil, T.B. Investigation of temperature development in salinity gradient solar pond using a transient model of heat transfer. Solar Energy 202, (2020), 32-44

Ganguly, S., Date, A., Akbarzadeh, A., 2018a. Investigation of thermal performance of a solar pond with external heat addition. J. Sol. Energy Eng. 140, 024501–024506.

Jolliffe, I.T., 2002. Principal Component Analysis 2nd Ed, Springer Verlag, Berlin, Germany.

Kasaeian, A., Sharifi.S., Yan, W-H. Novel achievements in the development of solar ponds: A review. Solar Energy 174, (2018), 189-206.

Kumar, A., Verma, S., Das, R. Eigenfunctions and genetic algorithm based improved strategies for performance analysis and geometric optimization of a two-zone solar pond. Solar Energy 211, (2020), 949-961.

Leblanc, J., Akbarzadeh, A., Andrews, J., Lu, H., Golding, P., 2011. Heat extraction methods from salinity-gradient solar ponds and introduction of a novel system of heat extraction for improved efficiency. Sol. Energy 85, 3103-3142.

Lu, H., Swift, A.H.P., Hein, H.D., Walton, J.C. 2004. Advancements in salinity gradient solar pond technology based on sixteen years of operational experience J. Sol. Energy Eng., 126, 759-767.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., de Jong, S., Lewi, P.J., Smeyers-Verbeke, J., 1998. Handbook of Chemometrics and Qualimetrics. Elsevier Science, Amsterdam.

Montalà. M., Cortina, J.L., Akbarzadeh, A., Valderrama, C. 2019. Stability analysis of an industrial salinity gradient solar pond. Solar Energy 180, 216-225.

Platikanov, S., Baquero, D., González, S., Martín-Alonso, J., Paraira, M., Cortina, J.L., Tauler, R. 2019. Chemometric analysis for river water quality assessment at the intake of drinking water treatment plants. Sci. Total Environ., 667, 552-562.

Tabor, H. Solar ponds. Sol. Energy, 27 (1981), pp. 181-194

Tundee, S. Terdtoon, P. Sakulchangsatjatai, P. Singh, R. Akbarzadeh, A. 2010 Heat extraction from salinity-gradient solar ponds using heat pipe heat exchangers Sol. Energy, 84, 1706-1716.

Valderrama, C., Gibert, O., Arcal, J., Solano, P., Akbarzadeh, A., Larrotcha, E., Cortina, J.L. 2011. Solar energy storage by salinity gradient solar pond: Pilot plant construction and gradien control. Desalination 279, 445-450.

Valderrama, C., Luis Cortina, J., Akbarzadeh, A., 2016. Solar Ponds. In: Storing Energy: With Special Reference to Renewable Energy Sources, pp. 273–289. https://doi.org/10.1016/B978-0-12-803440-8.00014-2.

Wise, B., Gallagher, N.1996. The Process Chemometrics Approach to Process Monitoring and Fault Detection, J. Process Contr. 6, 329 – 348.

Zangrando, F., 1980. A simple method to establish salt gradient solar ponds. Sol. Energy 25, 467–470. http://dx.doi.org/10.1016/0038-092X(80)90456-9.