Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 3216–3222

# Spatial Estimation of Solar Radiation Using Geostatistics and Machine Learning Techniques

**A.Núñez-Reyes** * **S.Ruiz-Moreno** *

* *Escuela Técnica Superior de Ingeniería. Departamento de Ingeniería de Sistemas y Automática. Camino de los Descubrimientos, s/n 41092 Sevilla, Universidad de Sevilla, Spain (e-mail: anreyes@us.es, sarruimor1@alum.us.es)*

**Abstract:** In large solar fields, where the control system is distributed, it is important to know the values of solar radiation in the complete area. Local solar radiation can be obtained by means of static sensors, using e.g. a wireless sensor network or movable sensors with drones for the general obtainment of variables. In this paper, solar radiation estimation is accomplished using Ordinary Kriging and distance weighting, and an alternative method is presented, which is based on a non-supervised competitive artificial neural network called Self-Organizing Map. This neural network generates a map with the most representative nodes and their weights, which are used to obtain the spatial variability of solar radiation in the area.



*Keywords:* Distributed Control and Estimation, Machine Learning, Sensors Networks.

## 1. INTRODUCTION

Since the Kyoto Protocol in 1997 to the present day, different legal and mandatory strategies have been carried out in order to deal with climate change and minimize its impact in nature and society. The fight against climate change is the greatest impulse for renewable energies and most energy efficiency technologies.

In this context, there is much research using thermic, photovoltaic and thermoelectric solar energy, where the knowledge of true solar radiation is crucial when improving the efficiency of solar plants, thus optimizing control: Camacho et al. (2012), Núñez-Reyes et al. (2005). The studies in Castilla et al. (2011) compare several approaches of predictive control to optimize thermal comfort in a building where the solar radiation considered is measured by a meteorological station placed on the roof, as well as other significant environmental variables. Núñez-Reyes et al. (2017) addresses the problem of integration of $PV$ plants in the electricity market. It proposes an optimal planning strategy using Model Predictive Control ($MPC$) (Camacho and Bordons (2004)) with restrictions that maximize the economic benefit. This paper considers the solar radiation as an input to the control strategy.

Numerous robust control strategies have been designed as well in order to obtain a controller that keeps the temperature of the heat transfer fluid at the desired value addressing the disturbances in solar radiation, reflectiveness of the mirrors and the collectors inlet temperature, given that the solar radiation varies along the day, causing changes in the dynamics of the plant and strong disturbances during the process (Limon et al. (2008)).

Huang et al. (2019) highlights that the output power of a photovoltaic ($PV$) system is related to the external environmental factors of the installation location, among which the solar radiation becomes more relevant. Therefore, the solar radiation conditions constitute significant reference data which allows us to evaluate the initial establishment of a photovoltaic system, as well as long-term performance.

Gallego and Camacho (2012a) and Gallego and Camacho (2012b) suggest the need of knowing global effective radiation. The pyrheliometer offers a measure of the local irradiance, but the extrapolation to the rest of the field is reasonable when the size of the collector field is small. When it is large, assuming that the solar radiation remains the same in the whole field can damage the efficiency and control of the plant. This is because scattered clouds can affect the local sensor whilst the rest of the plant is receiving direct radiation or vice versa.

In this work, we approach the problem described above, using several techniques that allow us to analyze the spatial variability of solar radiation with the aim of estimating its value on a surface. This will be carried out from samples of the area of interest.

This paper uses two well-known methods for solar radiation estimation: the Ordinary Kriging interpolation and the Inverse Distance Weighting. These two methods are applied to two different scenarios for obtaining the spatial variability of solar radiation. One scenario is based on Geostatistics and the other one in Machine Learning techniques. The main contribution of this research consists in reaching the second scenario with a SOM, which obtains the spatial correlation of solar radiation in the particular area.

Geostatistics is a statistical technique used for estimating spatially correlated data (Cressie (1993), Isaaks and Srivastava (1990)), known as "the art of modelling spatial


10.1016/j.ifacol.2020.12.1092

data". Its importance resides in the fact that it allows the description of the "spatial continuity" of variables and estimation of values very close to the real ones in unknown points.

On the other hand, the Self-Organizing Map (SOM) is a non-supervised artificial neural network that supplies an automatic method of data analysis by producing a low dimensional representation (map) of an input space of high dimension with non-external supervision (Kohonen (1990), Kohonen (1995), Vesanto (1999)).

Several versions of SOM models have been applied to numerous disciplines (Kohonen (2013)) such as financial analysis, oceanography, meteorology, bioinformatics and image recovery.

The employment of a SOM in the field of meteorology has been carried out from different perspectives. Chang et al. (2010) applies a SOM for evaluating the daily variability of evaporation based on meteorological variables as solar radiation, wind velocity, sunshine hour and humidity. In Sobri et al. (2018), the SOM was utilized to perform the characterization of cloud cover index that was subsequently used for evaluating solar radiation.

This document is organized as follows. After the introduction, section 2 describes the techniques employed. Section 3 presents the data, the proposed scenarios and the results. Finally, some conclusions drawn from the work and acknowledgments are presented.

## 2. TECHNIQUES APPLIED

### *2.1 Geostatistics*

Geostatistics can be considered as a discipline that addresses the statistical analysis of spatially distributed variables. A regionalized variable ($RV$) is one that is distributed in the space in a way that it presents a spatial structure of correlation (Samper and Carrera (1996), Matheron (1962)). Examples of regionalized variables are climatological variables such as solar radiation, temperature and humidity. From a mathematical point of view, a $RV$ is a function $z(c)$ that adopts one value for each point $x$ of the space.

A geostatistical analysis consists of two phases. In the first one, a study of the spatial dependence of the variable is made by calculation of the experimental and theoretical variogram. During the second phase, Kriging techniques are used (Krige (1951)) to estimate the values of the variables in unsampled spots.

*Variogram*
The variogram shows the spatial variability of the data, considering that the nearest places will have more similar values of the attributes than the ones corresponding to places away from each other. The variogram function $\gamma(h)$ of a random function $Z(x)$ quantifies its spatial correlation structure.

From the data, we can calculate the experimental or sample variogram as:

$$\gamma^*(h) \simeq \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \tag{1}$$

where $Z(x_i)$ corresponds to the experimental values at the points $x_i$ of which data are available from both $x_i$ and $x_i + h$; $N(h)$ is the number of pairs of points separated by a distance $h$. The variogram $\gamma(h)$, is independent of the location $x$ but depends on the module and direction of the vector $|h|$.

With the experimental variogram, we can obtain the value of $\gamma(h)$ for certain values of $h$. However, in most practical applications it is required the knowledge of all values of $\gamma$ for each $h$.

For that reason, an adjustment of a theoretical model variogram is made to the experimental variogram. Some of the most used models are the exponential, the spherical, the monomial and the gaussian model. For the adjustment, it is common to use methods like least squares or maximum likelihood, among others.

*Kriging Interpolation. Ordinary Kriging*
Kriging is a technique of interpolation that provides the best linear unbiased estimator ($BLUE$), as well as an error estimator called the Kriging variance, which depends on the variogram calculated and the locations of the original data (Journel and Huijbregts 1978).

The estimation for every not sampled point ($x_0$) is calculated as a linear combination of the $n$ available data $Z(x_i)$, Cressie (1993):

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{2}$$

with

$$\sum_i \lambda_i = 1 \tag{3}$$

$Z^*(x_0)$ is said to be the best predictor because the weights ($\lambda_i$) are obtained in such a way that they minimize the variance of the estimation error. For the minimization, it is applied the method of Lagrange multipliers and the values of $\lambda_i$ are obtained solving the following system of equations:

$$\begin{pmatrix} 0 & \gamma_{12} & \gamma_{13} & \dots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \gamma_{23} & \dots & \gamma_{2n} & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \gamma_{n3} & \dots & 0 & 1 \\ 1 & 1 & 1 & \dots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ 1 \end{pmatrix} \tag{4}$$

where $\gamma_{ij}$ is the variogram function $\gamma(h)$, calculated for the distance between points $x_i, x_j$: $\gamma_{ij} = \gamma(x_i - x_j)$; $\gamma_{i0}$ is the value of the variogram calculated for the distance between the sample observation $i$-th and the point where the estimation is desired: $\gamma_{i0} = \gamma(x_i - x_0)$; $\mu$ is the Lagrange multiplicator employed for the given restriction in the equation 3.

The variance of the estimation error of the Ordinary Kriging ($OK$) is subsequently obtained taking into account that $E(Z^* - Z) = 0$,

$$\sigma_k^2 = Var(Z^* - Z) = E[(Z^* - Z)^2] \tag{5}$$

Thus obtaining the following:

$$Var(Z^* - Z) = \sum_i \lambda_i \gamma(x_i - x_0) + \mu \tag{6}$$

### 2.2 The Inverse Distance Weighted (IDW) Technique

The $IDW$ method is used to estimate the unsampled values and it is based on the fact that the weight is a function of the inverse distance (Philip and Watson (1982), Watson and Philip (1985)). The general equation of $IDW$ is:

$$Z^*(x_0) = \frac{\sum_{i=1}^{n} Z(x_i)(\frac{1}{d_{i0}^k})}{\sum_{i=1}^{n}(\frac{1}{d_{i0}^k})} \quad (7)$$

where $Z^*(x_0)$ is the estimated value at the point $x_0$, $Z(x_i)$ represents the known values at the point $i$; $d_{i0}$ is the distance between the known point $x_i$ and the unknown $x_0$; $n$ the number of known points applied in the estimation and $k$ is the power function designating the degree of weights, which, according to Dixon et al. (2015), must equal 2.

### 2.3 Self-Organizing Map

The $SOM$ neural network, also known as Kohonen map or Kohonen net after its author (Kohonen (1982)), is a type of artificial neural network which learning is non-supervised and competitive. Regardless of the application granted, the $SOM$ produces a map (usually bidimensional), a discretized representation of the input space, resulting in a dimensionality reduction.

The $SOM$ network must discover correlations and categories in the input space without any concrete objective or tendency requirement at the output. Besides, the competitive learning classifies the input data intending to activate only one of the output neurons (or a group of neighbors) in the case of an input pattern.

The Kohonen self-organizing map is constituted by two levels of neurons: one at the input and one at the output. The input layer consists of $N$ neurons, one corresponding to each input variable, and the output layer is conformed by $M$ neurons, organized as a bidimensional map. The connections between layers are always forward and the connectivity is total. Each input layer $i$ is connected to each one of the output neurons $j$ by a weight $w_{ij}$, so that the output neurons have an associated vector of weights $W_j$ known as reference vector or codebook vector, which is usually randomly initialized and that constitutes the prototype vector of the category represented by the output neuron $j$.

The aim is having all the weights spatially correlated so that the nearest points in the grid will be more similar between each other than to the ones more distant.

The learning process of a $SOM$ consists in two steps:

First, the activation function is applied, calculating the similarity between the input vectors $x$ and the reference vectors $W_j$, using, for example, the euclidean distance, which will be useful for the activation of the output neurons.

$$||x - W_j|| = min||x - W_j|| \quad (8)$$

Once found the best matching unit ($BMU$), the nearest vector, i.e. the greatest activation or the smallest euclidean distance, is selected as the winner vector. In this way, the $BMU$ and its neighbors move close to the vector $x$ in the state space, in topological way and with an attraction index named learning rate $LR(k)$, which is decreasing in time and can be linear, potential or inverse time. This is accomplished with the actualization of the weights of the $BMU$ and the nearest neighbors.

In a similar way to the learning rate, the neighboring function $\epsilon(k)$ also declines with time and space when it distances from the winner. There are various types of neighboring functions and the most common is the gaussian. The learning rule can be denoted as:

$$w_j(k+1) = w_j(k) + LR(k)\epsilon(k)\,(x(k) - w_j(k)) \quad (9)$$

where $k$ denotes the learning iteration at each instant. Note that the number of learning steps must be specified at the beginning of the training. $x(k)$ represents the input pattern presented at the moment. This learning procedure leads to a topologically ordered mapping of the presented input data.

Similar patterns are assigned to the neighboring regions in the map, whereas the least similar patterns are assigned to the farthest regions.

The main contribution of this paper consists in using the output of the SOM network as a new scenario where the methods described above can be applied: Ordinary Kriging and IDW.

The output of the $SOM$ network consists of a map where the neighboring nodes have more relationship than the more distant ones. This work uses this characteristic to obtain the spatial correlation of solar radiation in the given area, relacionando el variograma con la red SOM ($Var\_SOM$).

## 3. EXPERIMENTAL RESULTS

Starting from a net formed by 49 neurons, we are going to estimate the solar radiation in an area of 50 $km^2$, with a total of 2400 points. For this purpose, two methods of spatial interpolation will be applied: $IDW$ and Ordinary Kriging to two different scenarios. The first one of them consists of the original sensor network (Fig.1) and the second one is based on the map generated by an artificial neural network, such as the self-organizing map (Fig.2). Both scenarios will be explained in this section.

### 3.1 Data

The solar radiation data have been obtained through CAMS Radiation Service, on the SoDa portal, a service framed in the Copernicus program, the European Union's Earth Observation Programme, with the aim of obtaining data on the state of our planet and the environment.

As the aim of this work does not focus on knowing the accuracy of the measure, but on having a great amount of localized data of solar radiation, the CAMS service is completely acceptable, given that we can obtain the solar radiation in every geographic spot of the planet with a minimum step time of 1 minute, since 2004 until present day.

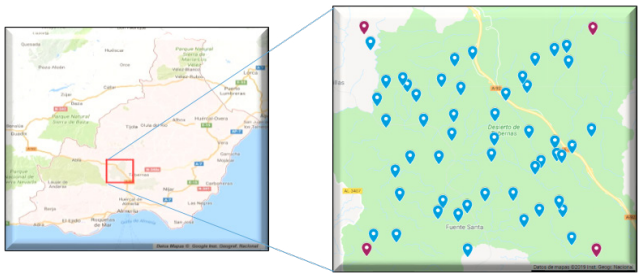

Fig. 1. Scenario 1. Original Sensor Network (*ORG*)

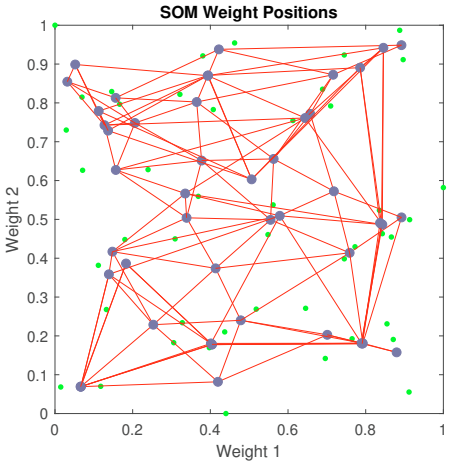


Fig. 2. Scenario 2. Self-Organizing Map *SOM*

### *3.2 Scenario 1: ORG*

The sensors network has been assumed in the desert of Tabernas, in Almeria (Spain), which is placed 30 $km$ north of Almeria, at an altitude of 400 $m$ above the sea level, located near the town of Tabernas. We have selected a 50 $km^2$ rectangular zone inside the limits of the region of Tabernas and we have randomly chosen 49 points (Fig. 1) irregularly distributed inside the area of study.

For this work, we have studied the month of June of 2015, obtaining 141120 data with a sample time of 15 minutes.

*Variogram*
It has been conducted the experimental variogram of the data of June and the theoretical variogram (see Fig. 3). The theoretical variogram that best matches the data is the monomial or potential model (Samper and Carrera (1996)): $\gamma(h) = a + Kh^{\theta}$, where $\theta \in (0, 2)$ to satisfy the conditions of the models of the variograms. The values obtained are:

$$\gamma(h) = 0.018 + 0.77 * h \tag{10}$$

For $\theta = 1$, the behavior at small distances is similar to the spherical and exponential variogram. The monomial variogram has small-scale fluctuations that increase inversely by reducing $\theta$ and large-scale fluctuations so large that they can seem a drift. Actually, they represent correctly the daily solar radiation curve.

### *3.3 Scenario 2: SOM*

The input to the *SOM* are vectors of three dimensions that correspond to longitude, latitude and solar radiation. The

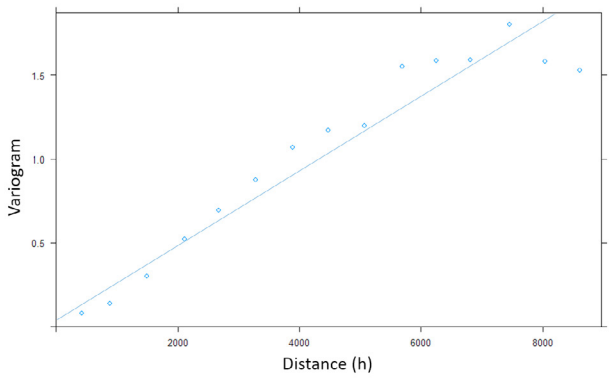


Fig. 3. Variogram month of June

network has been trained with 141120 input vectors and the number of iterations is 1000. For the accomplishment of this work, we have used the neural network toolbox of MATLAB (R2018a) version 11.1. The most important values of the *SOM* are shown in Table 1, the rest of the parameters are assumed by default.

Table 1. Parameters of the *SOM*

| Dimensions | Topology | DistanceFcn | training | Simulation |
|---|---|---|---|---|
| [7 7] | 'hextop' | 'linkdist' | 'trainbu' | 'negdist' |

To measure the quality of the *SOM*, the quantization error (*QE*) and topographic error (*TE*) are considered. The *QE* is the mean value of the distance between each input data and its *BMU*, and the *TE* represents the portion of the data which first and second *BMU* are not adjacent (Kohonen (2001). For the resulting *SOM*, the values of quality are the following: $QE = 0.1032$ y $TE = 0.1542$.

Fig. 2 shows the weights of the first and second attributes of the *SOM* network (black dots). Drawing on the correspondence of these attributes to geographic coordinates in space, it shows the position of the *ORG* net in the first scenario too (green dots).

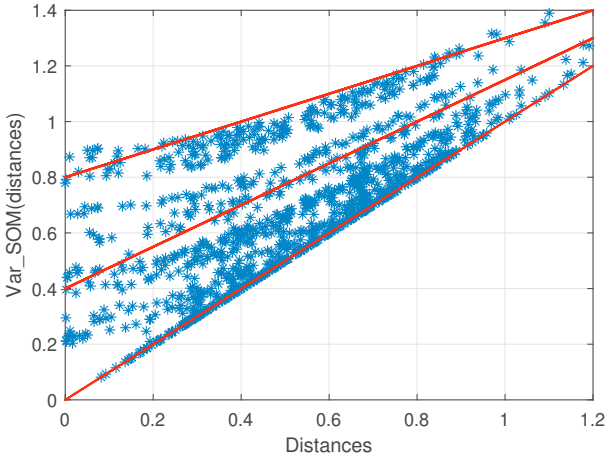


Fig. 4. *SOM* Variogram (*Var_SOM*) month of June

*Variogram of the SOM*
The *Var_SOM* obtained is shown in Fig. 4, where it can

be observed the relation that exists between the distance of the 49 nodes and the weight (Euclidean distance) of solar radiation. As well as with the variogram, increasing the geographical distance increases the weight (distance) of the variables, decreasing the correlation between farther points.

In Bastin et al. (1984) a global variogram, also known as climatological variogram, is described, where invariant terms are distinguished in space and time. Following the aforementioned paper, we can generate a model of theoretical variogram of monomial type with the term that only depends on the distance: $d^\theta$ with $\theta = 1$ and other terms that depend on the intensity of solar radiation $a(c)$ and $b(c)$:

$$VAR\_SOM(d) = a(c) + b(c)d^\theta \tag{11}$$

Table 2 presents the five different classes of the SOM, three of them have been drawn in Fig. 4 in a red line.

Table 2. 9th June at 7:15 a.m.

| Classes | Solar radiation | a(c) | b(c) |
|---|---|---|---|
| 1 | [0-0.2) | 0 | 1 |
| 2 | [0.2-0.4) | 0.2 | 0.85 |
| 3 | [0.4-0.6) | 0.4 | 0.75 |
| 4 | [0.6-0.8) | 0.6 | 0.65 |
| 5 | [0.8-1] | 0.8 | 0.5 |

### *3.4 Results*

As aforementioned, the solar radiation estimation is accomplished by employing Ordinary Kriging and IDW techniques for the two scenarios described. In this way, for each estimated point there are four different estimations:

- Scenario 1 and Ordinary Kriging: $ORG\_OK$
- Scenario 1 and IDW: $ORG\_IDW$
- Scenario 2 and Ordinary Kriging: $SOM\_OK$
- Scenario 2 and IDW: $SOM\_IDW$

Fig. 5 shows the points situated in the space that have been used for the work.

- Scenario 1: $ORG$
- Scenario 2: $SOM$
- Area to estimate: $GRID$
- Validation points of the estimation: $VAL$

Mean absolute error ($MAE$) and root mean square error ($RMSE$) are the evaluated indices in this paper. $MAE$ is the mean absolute value of the deviations:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Z(x_i) - Z^*(x_i)| \tag{12}$$

RMSE is the square root of the mean of the squares of the deviations:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z(x_i) - Z^*(x_i))^2} \tag{13}$$

The estimation of solar radiation in all the area has been made for June 2015. We will analyze the results of 9th June as a representative cloudy day Fig. 6. This graphic represents the true and estimated solar radiation with

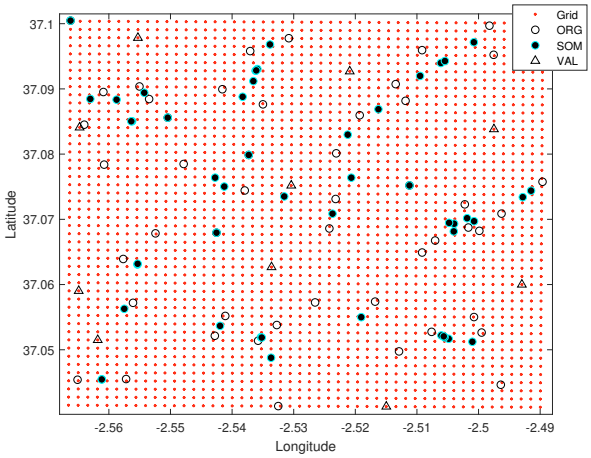


Fig. 5. Geographic points. Scenario 1 ($ORG$), scenario 2 ($SOM$), area to estimate ($GRID$) y validation points ($VAL$)

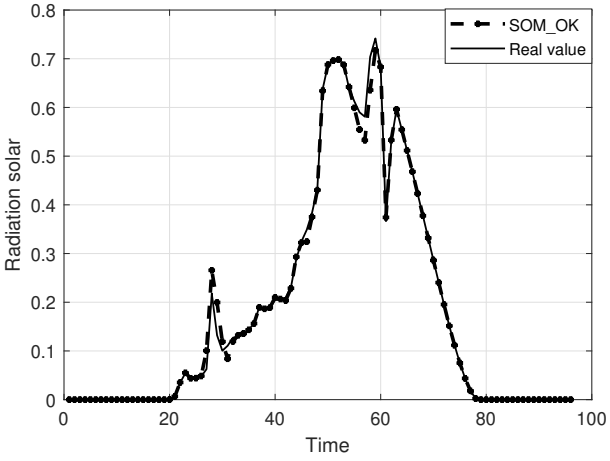


Fig. 6. Solar radiation of 9th June of the point: lng.:−2.5149 y lat.:37.0412

the method Ordinary Kriging and scenario 2 ($SOM\_OK$) throughout the entire day, for one of the 10 validation points ($VAL$).

Henceforth, the results will be shown as a function of the space variables for the day 06/09, at the instant of 07:15 a.m. (Time = 30 in Fig. 6).

Fig. 7 shows the contour plots of the estimated area ($GRID$) for the analyzed moment of the two scenarios ($ORG$ and $SOM$) with the two methodologies used ($OK$ and $IDW$). They show the continuity of solar radiation and similar behavior in both scenarios, where the area of greatest and fewest radiation is similar in all methods. It should be noted that the interpolation of the $SOM\_OK$ turns out to be smoother, the transition areas are less steep than in the other methods.

In Fig. 8, the estimated area ($GRID$) is represented next to the true values (black dots) of the validation network ($VAL$) of the $SOM\_OK$. It can be seen that the surface adapts perfectly to the real values of solar radiation.

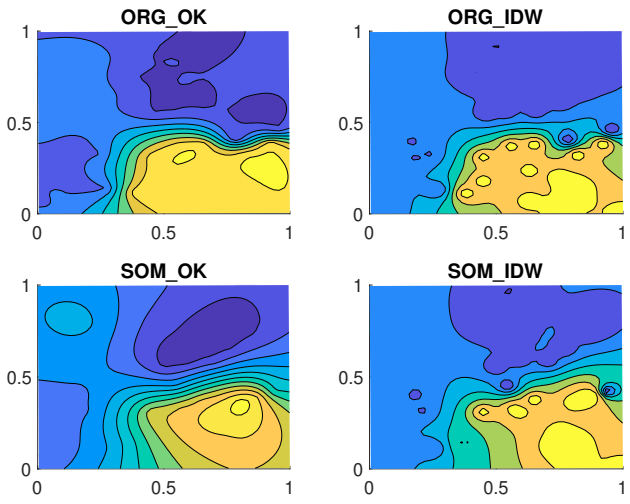


Fig. 7. Solar radiation estimation. Day 9th June at 7 : 15 a.m.

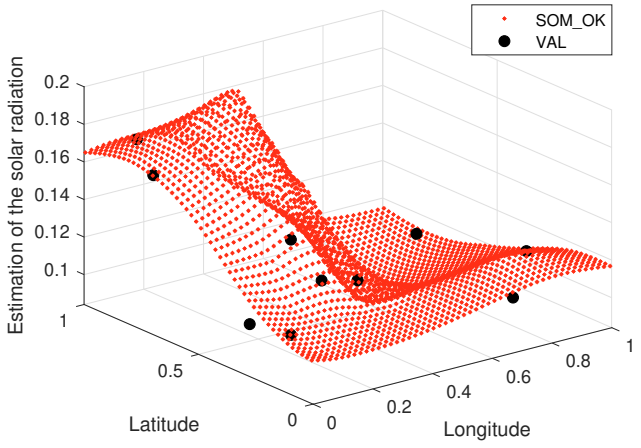


Fig. 8. Solar radiation estimation with $SOM\_OK$. 06/09 at 7 : 15 a.m.

The error obtained in each of the methods is collected in Table 3. It can be observed that the two networks present a very small error, the SOM network being the one that obtains the best results for that particular point. However, the errors obtained in all the points estimated for the two scenarios are quite acceptable and very similar.

Table 3. Day 9th June at 7 : 15 a.m.

| | $ORG\_OK$ | $ORG\_IDW$ | $SOM\_OK$ | $SOM\_IDW$ |
|---|---|---|---|---|
| MAE | 0.0081 | 0.0108 | 0.0076 | 0.0103 |
| RMSE | 0.0129 | 0.0139 | 0.0097 | 0.0121 |

## 4. DISCUSSION

This paper addresses one of the problems that are part of the research, currently under development, of distributed predictive control, based on a space-time model.

It starts from certain measurements of solar radiation taken by static or dynamic sensors to later estimate solar radiation on a large surface, where a hypothetical field of solar collectors is located.

The idea focuses on the fact that each of the controllers, at each sampling time, can take a measure of the solar radiation of the place where it is located in order to obtain the maximum information of the disturbance. For online estimation, both the measurements of the sensors at the instant $t$ and the weights obtained are needed, depending on the selected method.

This paper has provided a new alternative to obtain the spatial variability of a regionalized variable based on a self-organizing map. The $SOM$ network is used to obtain a second scenario where to apply the interpolation of solar radiation.

The results of the applied methods, in the two proposed scenarios, have been very similar, which demonstrates that the use of the $SOM$ is perfectly valid for applying it in the spatial interpolation of regionalized variables.

This work opens up future research lines, including the optimal design of sensor networks or the study of spatial variability for different sampling times, with the aim of using mobile sensors for measuring or extending to other continuous variables in space, such as temperature and humidity.

## ACKNOWLEDGEMENTS

This work was supported by project SI−1838/24/2018 of the Programme: $H2020 - EU.1.1$ EXCELLENT SCIENCE. European Research Council ($ERC$). Topic: $ERC-2017 - ADG - ERC$ Advanced Grant

## REFERENCES

Bastin, G., Lorent, B., Duqué, C., and Gevers, M. (1984). Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. *Water resources research*, 20, 463–470.

Camacho, E.F., Berenguel, M., Rubio, F.R., and Martínez, D. (2012). *Control of solar energy systems*. Springer-Verlag, London.

Camacho, E.F. and Bordons, C. (2004). *Model Predictive Control*. Springer-Verlag, London.

Castilla, M., Álvarez, J.D., Berenguel, M., Rodríguez, F., Guzmán, J.L., and Pérez, M. (2011). A comparison of thermal comfort predictive control strategies. *Energy and Buildings*, 43(10), 2737 – 2746. doi: https://doi.org/10.1016/j.enbuild.2011.06.030.

Chang, F.J., Chang, L.C., Kao, H.S., and Wu, G.R. (2010). Assessing the effort of meteorological variables for evaporation estimation by self-organizing map neural network. *Journal of Hydrology*, 384, 118–129. doi: https://doi.org/10.1016/j.jhydrol.2010.01.016.

Cressie, N. (1993). *Statistics for Spatial Data*. John Wiley & Son, Inc., New York.

Dixon, B., Uddameri, V., and Ray, C. (2015). *GIS and Geocomputation for Water Resource Science and Engineering*. John Wiley Sons.

Gallego, A.J. and Camacho, E.F. (2012a). Adaptative state-space model predictive control of a parabolic-trough field. *Control Engineering Practice*, 20, 904–911. doi:https://doi.org/10.1016/j.conengprac.2012.05.010.

Gallego, A.J. and Camacho, E.F. (2012b). Estimation of effective solar irradiation using an unscented kalman

filter in a parabolic-trough field. *Solar Energy*, 86, 3512–3518. doi:https://doi.org/10.1016/j.solener.2011.11.012.

Huang, C., Chen, Y., Hwang, M., and Kuo, P. (2019). *A simple solar radiation estimation algorithm design and field verification*, 161–163. doi:10.1201/9780429019777-33.

Isaaks, E.H. and Srivastava, R.M. (1990). *An Introduction to Applied Geostatistics*. Oxford University Press, USA.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69. doi:10.1007/bf00337288.

Kohonen, T. (1995). *The Self-Organizing Maps*. Springer-Verlag, Berlin/Heidelberg, Germany.

Kohonen, T. (2001). *The Self-Organizing Maps*. Springer-Verlag.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78, 1464–1480.

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, 37, 52–65. doi: https://doi.org/10.1016/j.neunet.2012.09.018.

Krige, D.G. (1951). *A statistical approach to some basic mine valuation problems on the Witwatersrand*. Thesis, University of the Witwatersrand, Johannesburg.

Limon, D., Alvarado, I., Alamo, T., Ruiz, M., and Camacho, E.F. (2008). Robust control of the distributed solar collector field acurex using mpc for tracking. *Proceedings of the 17th Wolrd Congress. IFAC*.

Matheron, G. (1962). *Traité de geostatistique appliquee. Tome 1 and 2*. Paris.

Núñez-Reyes, A., Marcos, D., Bordons, C., and Ridao, M.A. (2017). Optimal scheduling of grid-connected pv plants with energy storage for integration in the electricity market. *Solar Energy*, 144, 502–516.

Núñez-Reyes, A., Normey, J.E., Bordons, C., and Camacho, E.F. (2005). A smith predictive based mpc in a solar air conditioning plant. *Journal of Process Control*, 15, 1–10.

Philip, G. and Watson, D.F. (1982). A precise method for determining contoured surfaces. *J. Aust. Petrol. Explor. Assoc.*, 22, 205–212.

Samper, F.J. and Carrera, J. (1996). *Geoestadística. Aplicaciones a la hidrología subterránea*. Centro Internacional de Métodos Numéricos en Ingeniería. Universidad Politécnica de Cataluña, Barcelona.

Sobri, S., Koohi-Kamali, S., and Rahim, N.A. (2018). Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, 156, 459–497.

Vesanto, J. (1999). Som-based data visualization methods. *Intell. Data Anal.*, 3, 111–126.

Watson, D.F. and Philip, G. (1985). A refinement of inverse distance weighted interpolation. *Geo-processing*, 2, 315–327.