24-hours ahead global irradiation forecasting using Multi-Layer Perceptron

Cyril Voyant, Prisca Randimbivololona, Marie Laure Nivet, Christophe Paoli, Marc Muselli

24-hours ahead global irradiation forecasting using Multi-Layer Perceptron

Cyril Voyant[1*], Prisca Randimbivololona[3], Marie Laure Nivet[2] Christophe Paoli[2] and Marc Muselli[2]

1- CH Castelluccio, Radiophysics Unit, B.P.85 20177 Ajaccio - France

2- University of Corsica/CNRS UMR SPE 6134, Campus Grimaldi, 20250 Corte – France

3- Observ'ER, 146 rue de l'Université, 75007 Paris - France

*Corresponding author: phone/fax: +33(0)4 952 936 66/937 97, voyant@univ-corse.fr

Abstract. The grid integration of variable renewable energy sources implies that their effective production could be predicted, at different times ahead. In the case of solar plants, the driving factor is the global solar irradiation (sum of direct and diffuse solar radiation projected on a plane (Wh/m²)). This paper focuses on the 24-hours ahead forecast of global solar irradiation (i.e. hourly solar irradiation prediction for the day after). A method based on artificial intelligence using Artificial Neural Network (ANN) is reported. The ANN hereafter considered is a Multi-Layer Perceptron (MLP) applied to a pre-treated time series (TS). Two architectures are tested; it is shown that the most relevant is based on a multi-output MLP using endogenous and exogenous input data. A real case 2-years TS is computed and the MLP results are compared with both a statistical approach (AutoRegressive-Moving Average model; ARMA) and a reference persistent approach. Results show that the prediction

error estimate (nRMSE) can be reduced by 1.3 points with an ANN compared to ARMA and by 7.8 points compared to the naïve persistence.

1. Introduction

The deployment of Renewable Energy Source (RES) responds to strategic objectives, including independent and secured energy sourcing, air quality improvement and greenhouse gas emissions reduction. Therefore RES, and particularly for wind and solar energies, should significantly increase their share in the energy mix since conventional and renewable energies complement each other to a certain extent. The main advantage of PhotoVoltaic (PV) and wind energies is their inexhaustible aspects. Their main drawbacks are their variability (with respect to season, day/night, atmosphere state namely inherent absorption and scattering) and low ratio of energy produced over surface covered. The characterization of a RES (i.e. local wind or solar conditions) helps to design RES power plants in the development phase (Zhou *et al.*, 2008, Mellit *et al.*, 2009, Haurant *et al.*, 2010 and Jenn, 2011). Furthermore, during exploitation phase, power generation prediction at various times ahead is required by both the electrical system operator and the producer, respectively for system management and portfolio optimization (Santarelli *et al.*, 2004, Mwale *et al*, 2004, Ni *et al.*, 2006, Ipsakis *et al.*, 2009 and Chen *et al.*, 2011, Franco and Salza. 2011). In operational conditions, different time-horizons are required, from short term to medium term (see Figure 1):

- Real Time optimization: 15 min to 30 min. Ramp rates vary from one technology to another. Typically base load generation units (coal, nuclear) have a small ramp rate, whereas high peak load units (CCGT, gas-fired turbine) can be fully ramped up or down within an hour;

- Infra Day market: a few hours. Electricity producers consider the latest forecast and finely optimize their portfolio;

- Day ahead market: next day by hourly step. Electricity producers optimize their portfolio.

The energy mix management can be improved and critical events for the day coming can be anticipated if a skillful estimate of RES production 24-hours ahead for each hour (*h+24* prediction) is available. Therefore, transmission system operator can take appropriate measures to overcome any critical events. Figure 1 summarizes the link between energy production and prediction horizon.

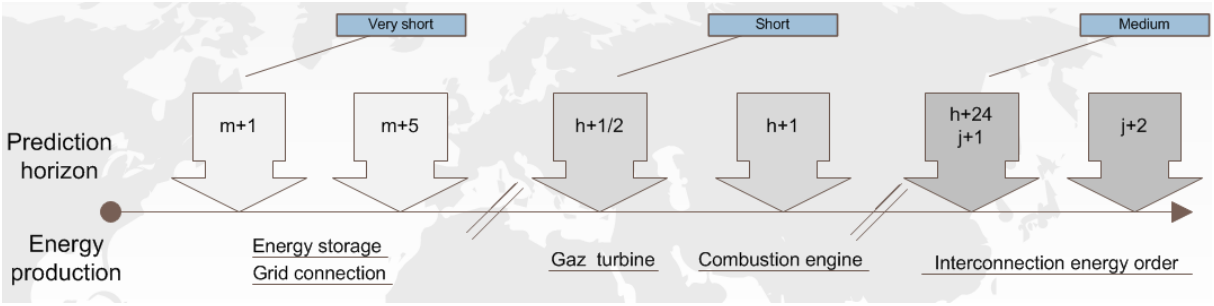


Figure 1. Impact to the intermittent energies prediction horizons on the energy management

In practice, the global irradiation (sum of direct and diffuse solar radiation on a plane ; Wh/m²) forecast is the name given to the process used to predict the amount of solar energy available in the current and near terms. Although many methods have been studied and developed worldwide, they all require knowledge of the previous state of atmospheric variables to forecast the next trend. This paper deals with the 24-hours ahead forecast of global solar irradiation on a horizontal surface i.e. hourly solar irradiation prediction for the day after (Chaouachi *et al.*, 2009). Conventional prediction tools are based on stochastic processes using Time Series (TS). TS is a sequence of data points measured successively in time and at regular time intervals in a consistent manner. According to the literature, Auto Regressive Moving Average (ARMA), Bayesian inferences, Markov chains, k-Nearest-

Neighbors predictors and Artificial Neural Network (ANN) are among the most reliable and robust predictors (Hamilton, 1994, Abrahart and See, 1998, Michaelides *et al.*, 2001, and De Gooijer and Hyndman, 2006). ANN and ARMA methods are often considered to be the most interesting prediction methods (Alados *et al.*, 2007, Altandombayci and Golcu, 2009 and Balestrassi *et al.*, 2009). Focusing on solar global irradiation, an optimized ANN or ARMA model can forecast daily and hourly time series with acceptable errors (Behrang *et al.*, 2010 and Voyant *et al.*, 2012a). For more distant horizons (i.e. more than two hours in the hourly case or two days in the daily case) predictions are less skillful. Indeed, Hawking (1998) states that the overall elements of space-time are not especially related to each other. As an extension to this concept: the degree of correlation between events decreases when spatial or temporal distances increase. A proof is given by the Maximal Lyapunov Exponent (MLE; Cao, 1995) theory, beyond a characteristic time (computed from MLE and called Lyapunov horizon; LH), prediction is almost impossible. In Ajaccio and for an hourly global irradiation time series, the LH of predictability is estimated at about 24-hours (calculation by classical methodology based on auto-mutual information by Cao's method (Cao, 1997) and correlation dimension by Taken estimator (Parlitz, 1995)). Beyond this limit, the stochastic models are valid but not relevant. The intrinsic characteristics of the series are not sufficient to accurately describe the distant phenomena. Therefore, 24-hours ahead global irradiation forecast cannot reach the same accuracy as a prediction over a shorter time horizon such that horizon *h+1* or *h+2*. That is perhaps the reason why very few studies are dealing with this horizon. Nevertheless research recently gained interest in the *h+24* time prediction horizon, especially using MLP approach (MultiLayer Perceptron), a feed forward ANN model that maps a set of input data onto a set of appropriate output data. Mellit and Pavan (2010) recommend using daily global irradiation (endogenous) and temperature (exogenous) as inputs to the model, and

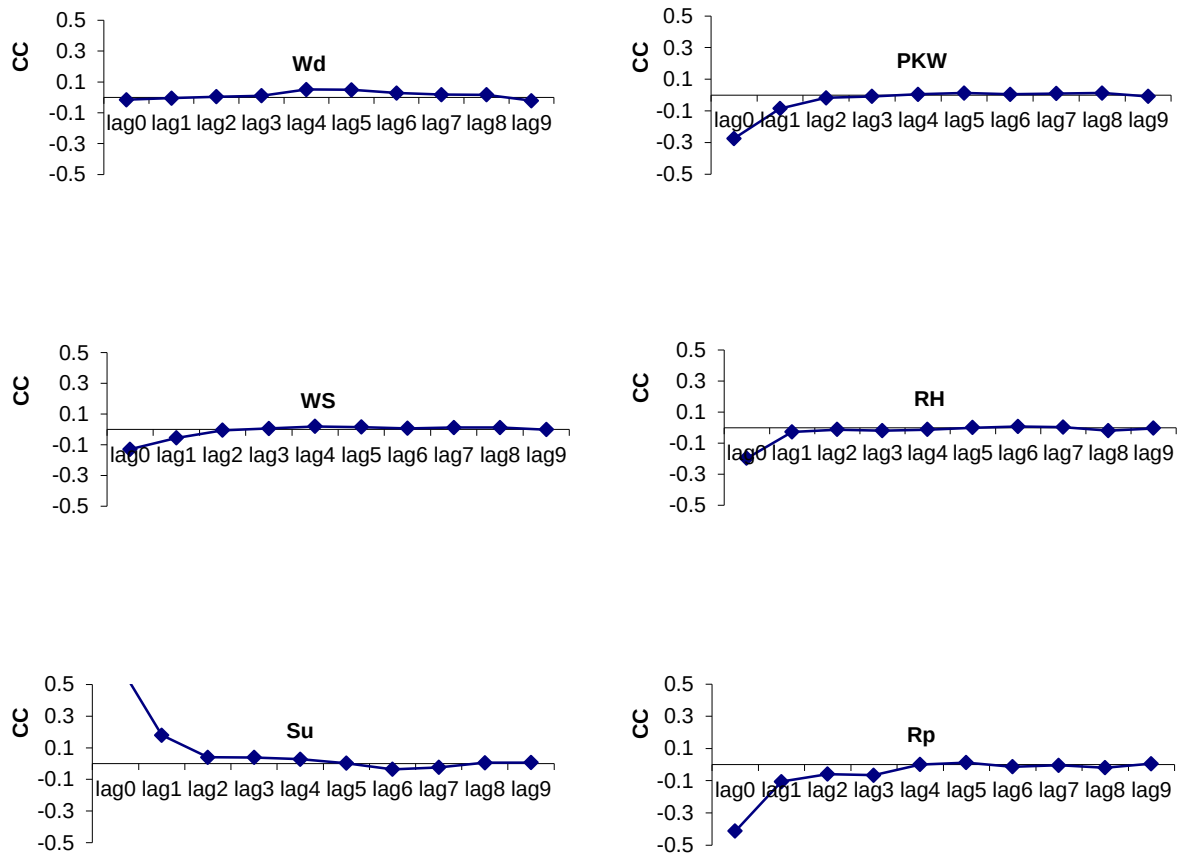
a daily time based increment in the considered month. In the PhotoVoltaic case (PV), Cococcioni *et al.* (2011) have presented that the MLP with endogenous data can be used like tool for technicians of PV installation to correctly configure the forecasting model according to the particular installation characteristic. This present study considers also an MLP-approach. This architecture choice is based on previous study (Paoli *et al.*, 2010 and Voyant *et al.*, 2011) and literature results (Reddy, 2003, Mubiru, 2008 and Mubiru *et al.* 2008).

In next sections, this paper focuses on an *h+24* prediction method that optimizes the trade-off between accuracy and complexity of the model. The following steps will be presented and assessed: architecture, stationarization, uni- and multi-variate modeling of MLP networks as well as ARIMA and persistence models. Results are discussed in the last Section.

2. Available data and analysis

Experimentation data were provided by Ajaccio's meteorological center (Météo France), which coordinates are 41°5'N and 8°5'E (seaside, 4 m asl). It faces the Gulf of Ajaccio with mountain behind. Summer days are hot and dry, and winter days have mild temperature. During autumn and spring, violent storm episodes may occur. Experimentation data cover the period from 1999 to 2008. The data set has been split in two parts: a training data set covering the first eight years of data, and a test data set covering the two remaining years (2007 and 2008). The configuration of the MLP model is finalized through a so-called training phase. The training phase is used to set MLP model parameters, and the validation phase compares MLP model results to real data. Dealing with global solar irradiation, a day has been defined to last 9 hours and to range from 8 am to 4 pm (solar time). Missing data (<

4%) were replaced with the average of their corresponding hourly data at a given day over the period covered by the set of data. The selection of exogenous variables is based on the literature and previous studies (Raschke *et al.*, 1987, Eltahir and Humphries, 1998, Baigorria *et al.*, 2004, Lopez *et al.*, 2005, Badescu, 2008, Crone and Kourentzes 2010, Voyant *et al.*, 2011). Figure 2 shows the Pearson cross-correlation between clear sky index (defined in section 4-3) and exogenous variables for the site of Ajaccio for a 1-day time lag. The cross correlation at lag 1 of *Wd* (wind direction) is the correlation between clear sky index at day *d* and *Wd* at day *d-1*.



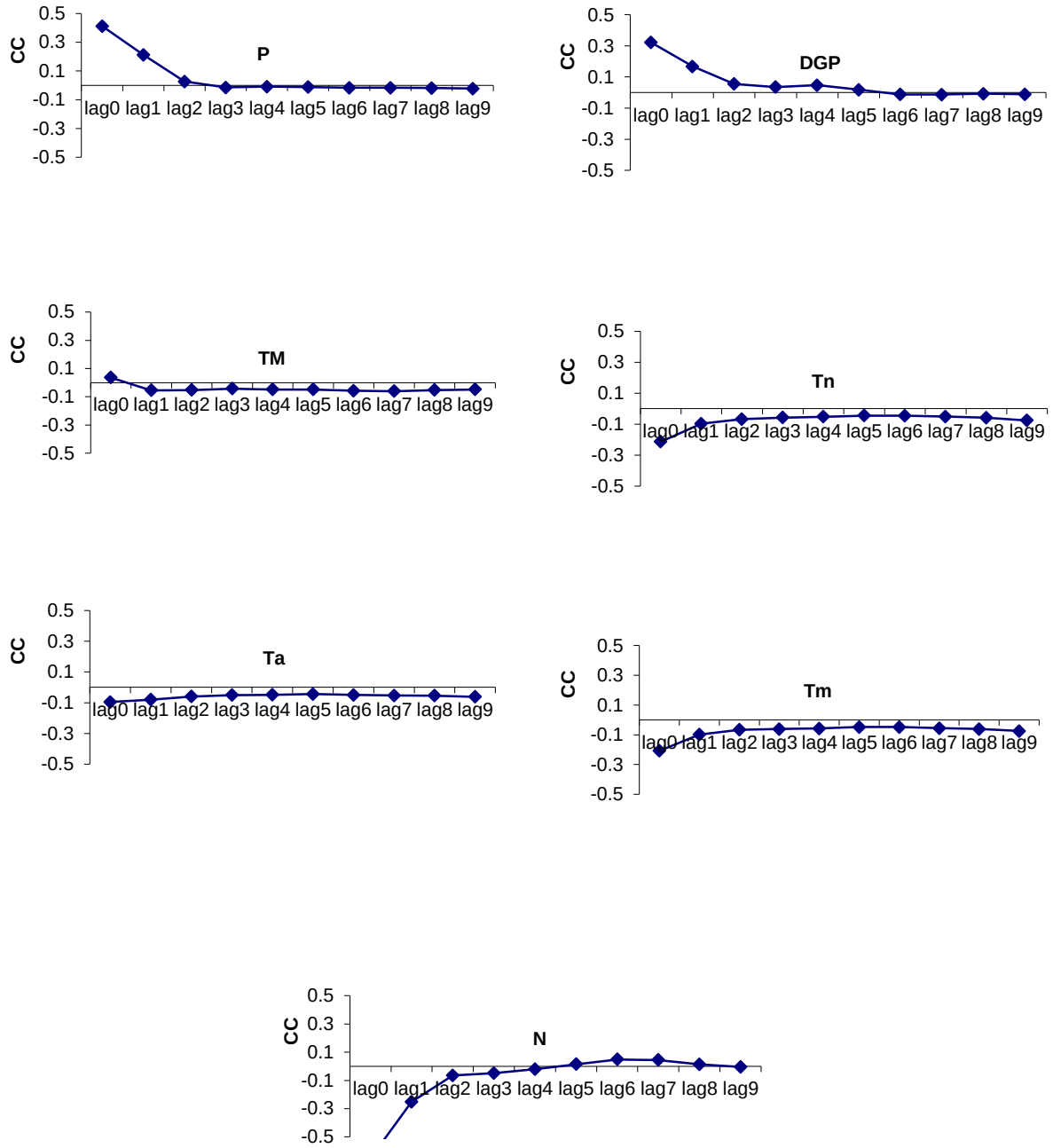


Figure 2. Pearson cross-correlation between the clear sky index, and exogenous variables for Ajaccio stations. (wind direction *Wd*, peak of wind speed *PKW*, wind speed *Ws*, relative humidity *RH*, sunshine duration *Su*, rain precipitation *RP*, pressure *P*, differential pressure *DGP*, ambient temperature average *Ta*, night temperature *Tn* , max *TM* and min *Tn* temperatures and nebulosity *N*)

Only the most relevant parameters related to global irradiation or clear sky index have been considered in this study. The relevant parameters are:

- the pressure (*P*, Pa; average measured by numerical barometer during one hour), the nebulosity (*N*, Octas) and the precipitations (*RP*, mm, five cumulative measures of six minutes during half an hour) for the nine considered hours ;

- the daily mean value of pressure and its gradient (difference between the mean pressure of day *d* and day *d-1*), nebulosity and precipitations ;

- the differential pressure between the first and last hours of sunshine.

Next section gives a detailed description of the three considered approaches: MLP, ARIMA (reference estimator) and persistence (naïve estimator).

3. The prediction methodologies

This section gives an overview of the principles behind MLP, ARMA and persistence prediction methods. The Multi-Layer Perceptron (MLP) is a type of ANN architecture and is the main typology described in this paper. The next section summarizes the main principles of this model and presents the considered test methodologies for the 24h-ahead forecast (optimization, variable selection and stationarization). Forecasting methods based on persistence (naive forecaster) and ARMA models (reference forecasters) are also introduced and described.

3.1 Artificial neural networks

Although many ANN architectures exist (Coulibaly *et al.*, 1999, Hu *et al.*, 2002, Benghanem and Mellit, 2010, Cao and Cao, 2006), MLP remains the most popular (Cybenko, 1989, Hornik *et al.*, 1989, Al-Alawi and Al-Hinai, 1998, Crone, 2005, Bosh, *et al.*, 2008, Behrang *et al.*, 2010). An MLP is made of several layers: one input layer, one or several intermediate layers and one output layer. Hidden and outputs layers are composed of artificial neurons. These functional elements are parameterized functions of variables called input (x_j) multiplied by a weight (regression coefficients called ω_{ij}^1). The transfer function (or activation function; g) is applied to the linear combination $x_j \omega_{ij}^1$, which leads to the following output: $y_i = g(\sum_{j=1}^{n} x_j \, \omega_{ij}^1 + b_i^1)$. The b_i^1 term is a particular weight called bias (it is an offset constant). The output vector of a considered layer becomes the input vector of the following layer until reaching the output layer. Depending on the signal intensity arriving to the activation function, the neuron will be « on » (the output is 1) or « off » (the output is 0 or -1). Note that other intermediate values are authorized outside the Boolean outputs ($\in [-1,1]$ or $[0,1]$). It can be said that the activation function « fuzzifies » the outputs. In most cases, a single neuron is of no interest. However, interconnected together single neurons build a network of neurons which can solve complex problems such that classification, pattern recognition, time series prediction, etc. For TS forecast, the MLP input vector is made of a settled number of past values and the MLP output is the predicted value (Iqdour *et al.*, 2006 and Paoli *et al.*, 2010). Weighted coefficient and bias values (from the linear combinations of input layers) are set through the learning phase. The testing phase is used to test the robustness and the reliability of the model: the validation set is checked against results obtained with the model. The information flows from the input layer to the output layer. The most

commonly used artificial neural network is the feed-forward MLP with one hidden layer and one output layer (Hornik, *et al.*, 1989, Ito, 1991). It is often used for modeling and forecasting TS (Jain *et al.*, 1996, Moreno *et al.*, 2008). Several studies (Kalogirou, 2001, Zhang and Qi., 2005 and Min *et al.*, 2008) have validated this approach based on ANN as a non-linear model to describe TS. The principle of MLP for TS forecasting is detailed in Figure 3. The methodology is based on the sliding window principle.

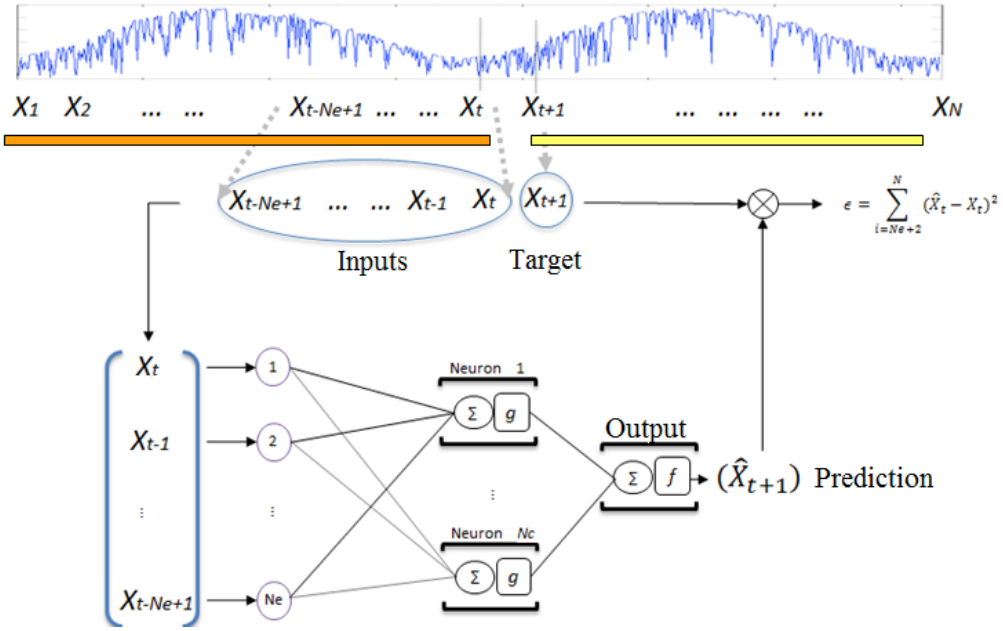


Figure 3. MLP and TS prediction with N_e inputs, N_c hidden nodes and one output

In this case the output of the system is governed by Equation 1 and 2 (N_e inputs, N_c hidden nodes, one output, ω the weights, b the bias, f the activation function of the output layer and g the activation function of the hidden layer).

$$\hat{X}_{t+1} = f\left(\sum_{i=1}^{Nc} y_i\, \omega_i^2 + b^2\right) \qquad \text{Equation 1}$$

$$y_i = g\left(\sum_{j=1}^{Ne} X_{t-j+1}\, \omega_{ij}^1 + b_i^1\right) \qquad \text{Equation 2}$$

The MLP methodology is a non-linear regression model based on the interpretation of the past global irradiation values. The MLP used with one output node associated with the identity activation function: $y = x$ (architecture frequently used) allows establishing a regression as described in the Equation 3.

$$\hat{X}_{t+1} = \sum_{i=1}^{Nc}(g(\sum_{j=1}^{Ne} X_{t-j+1}\,\omega_{ij}^{1} + b_i^1))\,\omega_i^2 + b^2 \qquad \text{Equation 3}$$

In this study, MLP has been compiled with the Matlab© software and its Neural Network toolbox. The main characteristics chosen and related to previous works (Voyant *et al.*, 2011) are: the hyperbolic tangent (hidden) and identity (output) activation functions, the Levenberg-Marquardt (Dreyfus, 2004) learning algorithm with a max fail parameter before stopping training set to 3 (early stopping method limiting the overtraining). In this study, training, validation and testing data sets (Matlab© parameters) were respectively set to 80%, 20% and 0% (Matlab© parameters) and inputs are normalized between – 0.9 and + 0.9. These three phases concern the eight first years of the global solar irradiation values covered in the set of data. The testing phase for prediction results uses the remaining last two years of the data set. It is critical to identify best-suited input variables to optimize the MLP. Therefore a systematic approach has been adopted by browsing all realistic configurations (the maximum number of parameters is limited). The final choice is based on the cross-comparison of considered configurations prediction error (mono-criterion analyze). The architecture with the lowest prediction error defines the final MLP topology. The considered error estimation is the nRMSE (normal Root Mean Square Error) defined by $\sqrt{< (x - \hat{x})^2 > / < x^2 >}$. In order to apply the MLP approach to the h+24 time-horizon, the h+1 horizon methodology has been transposed (optimization, learning and stationarization). To predict the nine hours of the next days, three alternatives were considered:

- MLP committee : nine MLP models each of them dedicated to one hour;

- Multi-output MLP : a nine-output MLP model; and

- A one-output MLP model used nine times (see Figure 4).

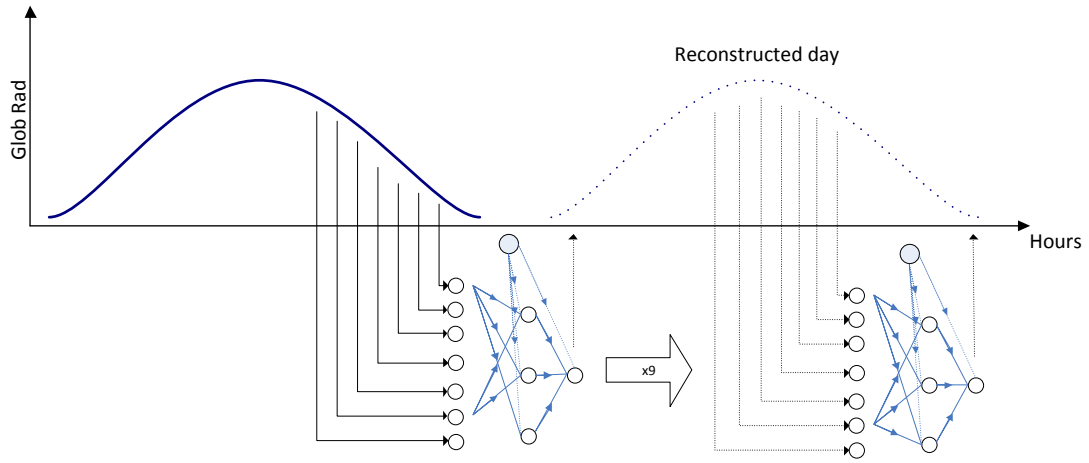


Figure 4. Method based on the nine consecutive MLP forecasters at h+1 horizon

The first method is derived from the observation that cloud occurrences and mist appear often at the same hours. The second approach is more conventional and assumes that global solar irradiation for each hour is related to previous measurements. The last method is intellectually simple, each prediction is considered chronologically like inputs of the network. Considering the prediction error increases exponentially when predictions are set as network input as it is explained in APPENDIX 1, only the two first alternatives have been studied. The details of the two chosen methods are:

- **MLP committee:** one predictor is dedicated to one hour only. Nine time series are extracted from the original TS (one time series per considered hour). The input vector is a set of chronologically ordered measurements related to the same hour at previous days.

Figure 5 gives a good illustration of the nine time series (focus on 14:00 time series). Its prediction is then equivalent to a daily horizon forecast, and the input vector to predict time 14:00 on day *d+1* is made up with observations at time 14:00 on days *d*, *d-1*, …, *d-*N_e. The predicted day can be reconstructed by concatenating the nine sunshine hours predictions (Figure 6.a).

-**Multi-output MLP:** This approach is a conventional methodology based on the sliding window principle. Measures of global irradiation are chronologically ordered to build up the MLP input vector (Figure 6.b).

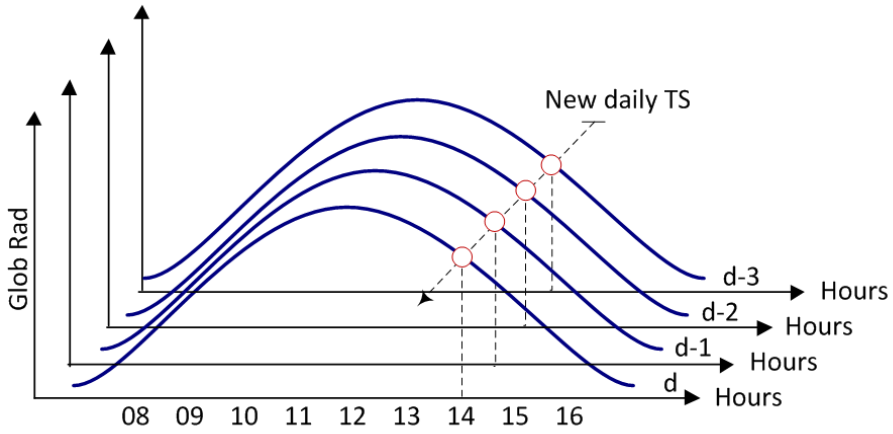


Figure 5. Daily decomposition and reduction of the hourly TS by nine daily TS. The considered hours are 8:00 to 16:00, the yellow circles represent the 14:00 TS for the days *d*,*d-1*, *d-2* and *d-3*

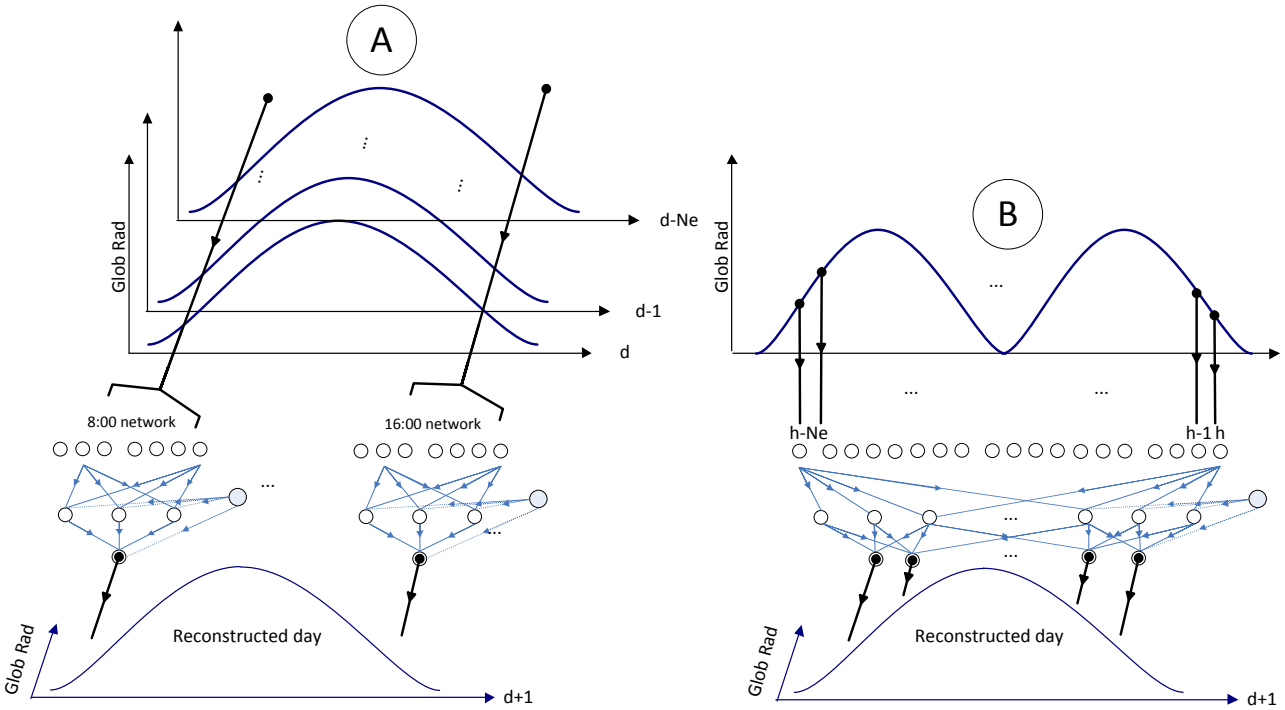


Figure 6. Reconstituted daily global radiation with nine independent predictors (a) and one MLP with nine outputs (b)

3.2 Other forecasters

In this paper, the MLP methodologies defined in section 4.1 (Figure 6) are compared to the persistence approach, which uses the last measure as prediction output, and conventional stochastic ARMA approach, which is the most commonly studied predictor for TS modeling (Brockwell and Davis, 1991 and Bourbonnais and Terraza, 2008). Both are based on an approach similar to that of Figure 6.a and daily decomposition. Nine concatenated predictions form the global solar irradiation for the day after. Note that these methodologies are compatible with the clearness or clear sky index (see the section 4.3). Equation 4 presents the

persistence principle where a second order correction is operated to take into account the theoretical global radiation difference between time *t* and *t*+1.

$$\hat{X}_{t+1} = X_t . \frac{H_{gh}(t+1)}{H_{gh}(t)} \quad \text{Equation 4}$$

The ARMA model is based on two elementary models: the MA model (moving average model, Equation 5) and the AR model (autoregressive model, Equation 6). They are defined like a regression on the last residues and the last measures.

$$MA(q) \rightarrow x_t = \sum_{i=0}^{q} \theta_i . \epsilon_{t-i} = \theta(L)\epsilon(t),\ \forall t \in \mathbb{Z} \quad \text{Equation 5}$$

$$AR(p) \rightarrow x_t = \sum_{i=1}^{p} \varphi_i . x_{t-i} + \epsilon_t = [\varphi(L)]^{-1}\epsilon_t,\ \forall t \in \mathbb{Z} \quad \text{Equation 6}$$

A more general so-called SARIMA model is issued from AR and MA models. It is defined in Equation 7. The ARMA model (the most popular for TS prediction) is a particular case of SARIMA with seasonal parameters *P*, *D* and *Q* set to zero.

$$SARIMA\ (p,d,q)(P,D,Q) \rightarrow \varphi_p(L)\ \varphi_P(L)^s\ (1-L^s)^D (1-L)^d x_t = \theta_q(L)\theta_Q(L)^s\ \epsilon_t \quad \text{Eq 7}$$

3.3 The need to make the TS stationary

The TS transformation to make the series stationary (a transformation hereafter called pre-treatment) has been discussed for instance in the literature (Brokwell *et al.*, 1991). This methodology must be applied to the ARMA process (Bourbonnais and Terraza, 2008) and to MLP TS prediction (Zhang and Qi., 2005). According to Hornick *et al.* (1989) and Cybenko (1989), any network can be considered as a compact universal approximator. Furthermore, they are asymptotically stationary and are not characterized

by a divergent behavior or a variance increase over time (Hornik *et al.* 1989). In practice, to use an MLP, the input data must be stationary (Ito, 1991), or slowly varying. However, a MLP network can also simulate a non-stationary process, if and only if, it operates on a finite time interval. In this case, the system cannot learn all the characteristics of the studied process, and the predictions made outside the training sample and associated performances will be poor. One of the solutions to overcome the problem of non-stationary data consists to transform the initial TS to make it stationary (Crone, 2005). The stationarization method used in this paper is based on the ratio to trend (Bourbonnais and Terraza, 2008). When the top-of-atmosphere irradiation is used, the clearness index (k) is generated (Mellit *et al.*, 2009) and when a ground clear sky model is used, the clear sky index (CSI) is generated (Ineichen, 2006 and 2008, Voyant *et al.*, 2012a). These methods based on the coupling between MLP and knowledge models permit to switch from a black box approach to a so-called "grey box" approach (Dreyfus, 2004). The global methodology to make TS stationary is defined in the Figure 7.

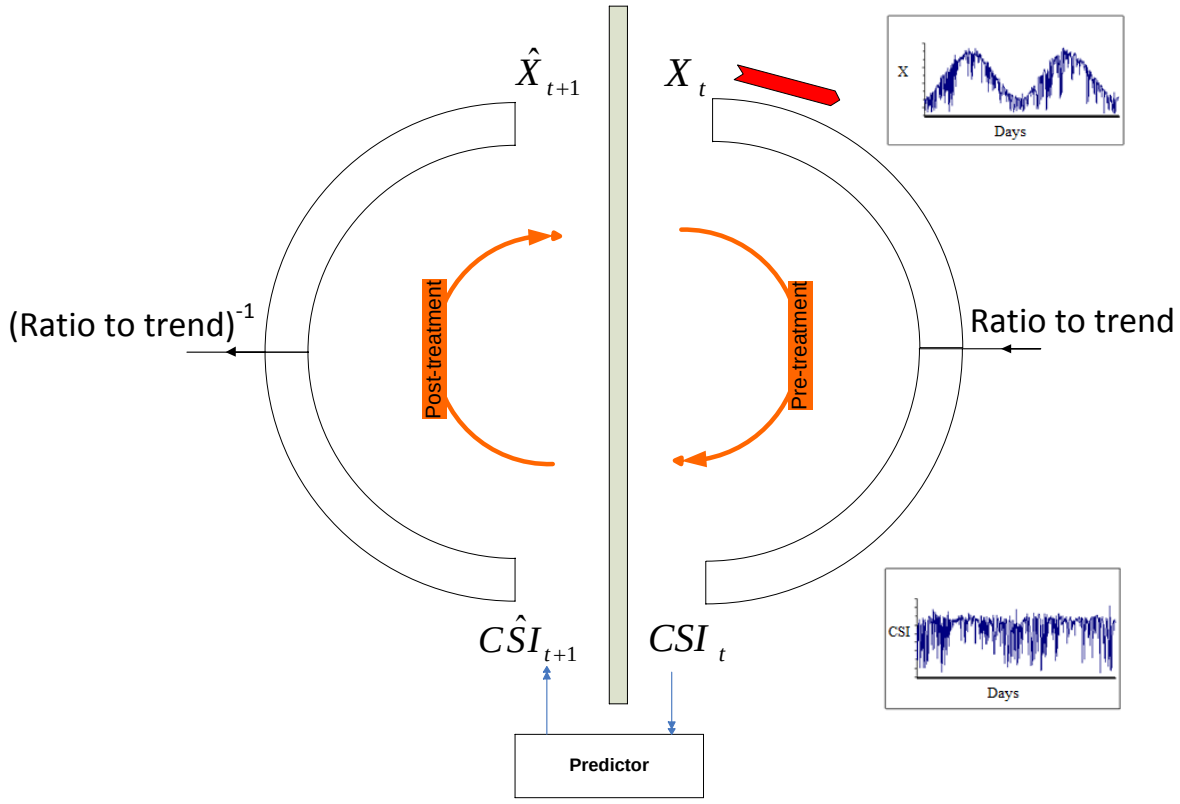


Figure 7. Ad-hoc methodology and the procedure impact on the TS trend

This section described the global methodology of a prediction based on an MLP approach. Next paragraph gives a methodology to increase the MLP performance with a meteorological data set as input.

3.4 The MLP and the multivariate analyze

Section 3 has shown that some exogenous data are correlated with the CSI. These data can then be used to optimize the MLP architecture and improve prediction results. The ability to consider exogenous parameters as inputs is specific to MLP, and is not possible with standard ARMA (or persistence). The next section presents results obtained for models parameters optimization and their prediction results. The goal is to build an

h+24 global solar irradiation predictor from the naive model (i.e. persistence), the reference model (ARMA) or the MLP models.

4. Results

Considering that the aim of this study is to find the best compromise between accuracy and complexity of the models, results are exposed step by step. The first results presented in this section are the optimal MLP and ARMA models for the *h+24* time horizon after the optimization process. In a second time, in order to avoid heavy computation time for MLP and ARMA architectures optimization and final results, only endogenous data are considered. Finally exogenous inputs are taken into account and their benefits are evaluated.

4.1 MLP and ARMA optimization

The optimization mainly consists in determining input variables and hidden nodes. Data used for models optimization cover the period from 1999 to 2006. Years 2007 and 2008 are used to compare results of MLP, ARMA and persistence models. Set parameters are: one hidden layer using the hyperbolic tangent as activation function and the Levenberg-Marquardt optimization algorithm (Suratgar, 2005). Selection of endogenous variables for the input layer is based on a systematic approach, which consists in testing all possible combinations between 1 to 20 input nodes and 1 to 20 hidden neurons. This selection process has been applied to three TS: the original TS (without stationarization process; *X*), the clearness index TS (*k*) and the clear sky index TS (*CSI*). No other stationarity rules, e.g. seasonal adjustments (Bourbonnais, 1998), have been tested due to the high level of complexity compared to the *h+1* or *d+1* cases [Voyant *et al.*, 2012]. Table 1 shows results of the optimization process for the nine MLP described in Figure 6.a. For N_e inputs ($\in$[1,20]), N_c hidden nodes ($\in$[1,20]) and

one output, the nomenclature used is Endo$^{1\text{-}Ne}$xN_cx1. In order to prevent the risk to optimize the MLP based on a local minimum during the training step, four simulations are carried on for each N_c and N_e values (only average is represented).

	networks	*X*	*k*	*CSI*
MLP committee nine networks	8:00	*Endo*$^{1\text{-}3}$x4x1	*Endo*$^{1\text{-}3}$x4x1	*Endo*$^{1\text{-}11}$x8x1
	9:00	*Endo*$^{1\text{-}10}$x4x1	*Endo*$^{1\text{-}8}$x3x1	*Endo*$^{1\text{-}8}$x3x1
	10: 00	*Endo*$^{1\text{-}11}$x4x1	*Endo*$^{1\text{-}20}$x2x1	*Endo*$^{1\text{-}20}$x2x1
	11:00	*Endo*$^{1\text{-}3}$x8x1	*Endo*$^{1\text{-}2}$x4x1	*Endo*$^{1\text{-}3}$x3x1
	12:00	*Endo*$^{1\text{-}20}$x1x1	*Endo*$^{1\text{-}20}$x1x1	*Endo*$^{1\text{-}20}$x1x1
	13:00	*Endo*$^{1\text{-}12}$x7x1	*Endo*$^{1\text{-}19}$x1x1	*Endo*$^{1\text{-}1}$x14x1
	14:00	*Endo*$^{1\text{-}18}$x2x1	*Endo*$^{1\text{-}20}$x1x1	*Endo*$^{1\text{-}3}$x19x1
	15:00	*Endo*$^{1\text{-}11}$x5x1	*Endo*$^{1\text{-}12}$x1x1	*Endo*$^{1\text{-}4}$x10x1
	16:00	*Endo*$^{1\text{-}13}$x3x1	*Endo*$^{1\text{-}1}$x1x1	*Endo*$^{1\text{-}6}$x5x1
Multi-outputs MLP	Only one	*Endo*$^{1\text{-}27}$x1x9	*Endo*$^{1\text{-}27}$x1x9	*Endo*$^{1\text{-}27}$x2x9

Table 1. Optimization of the nine independent MLP and the multi-output MLP

For the single multi-outputs MLP (Figure 6.b), optimization results are also given in Table 1 ($N_e \in [1,27]$, $N_c \in [1,27]$)). The 27 inputs visible for the three modes (*X*, *k* et *CSI*) correspond to three days of nine hours each (=3x9). Considering this table, no rule seems to exist. The endogenous inputs and the hidden nodes numbers present large variations. Note that the use of the *CSI* seems to increase the number of hidden neurons. Concerning the ARMA models, according to tests results and literature (De Gooijer, 2006), only the AutoRegressive part seems to be relevant. It seems that the Moving Average part has no positive influence on the prediction error. Models optimization is based on the analysis of autocorrelograms (correlation dimension (Hamilton, 1994)). A student test (Bourbonnais and Terraza, 2008) on the index p of the AR(p) regression coefficients called φ_i (for $i \in [1,p]$) shows that only the first lag is significantly different of zero (Figure 8). Although there are

some particular points e.g. lag 3 of the 8:00 TS, by desire to homogenize, make efficient and simple, a lag one configuration for the nine models (AR(1)) is adopted.

Figure 8. Autoregression coefficients of the nine AR(*p*) with the *CSI* forecasting mode (the red line is the significance limit)

4.2 Univariate predictions

Prediction models defined in section 5.1 are tested using Ajaccio meteorological center 2007 and 2008 data. Table 2 shows relative errors for the nine daily series approach using ARMA and MLP models and for the multi-output model using the MLP model only. The nRMSE shown are related to the daily mean of global solar irradiation.

Models		Annual	Winter	Spring	Summer	Autumn
Persistence		35.1	54.8	35.2	28.0	40.4
ARMA	*k*	29.1	44.6	29.2	24.0	33.2
	CSI	28.6	44.2	28.6	23.1	32.8
MLP committee	*k*	28.5	44.6	28.8	22.9	32.8
	CSI	28.2	44.1	28.6	22.4	33.2
Multi-outputs MLP	*k*	27.9	44.2	**27.9**	22.2	32.7
	CSI	**27.8**	**42.8**	28.4	**22.0**	**31.3**

Table 2. nRMSE (%) of predictions with the nine independent networks, with the mupti-output MLP and AR model (bold characters are related to the best results)

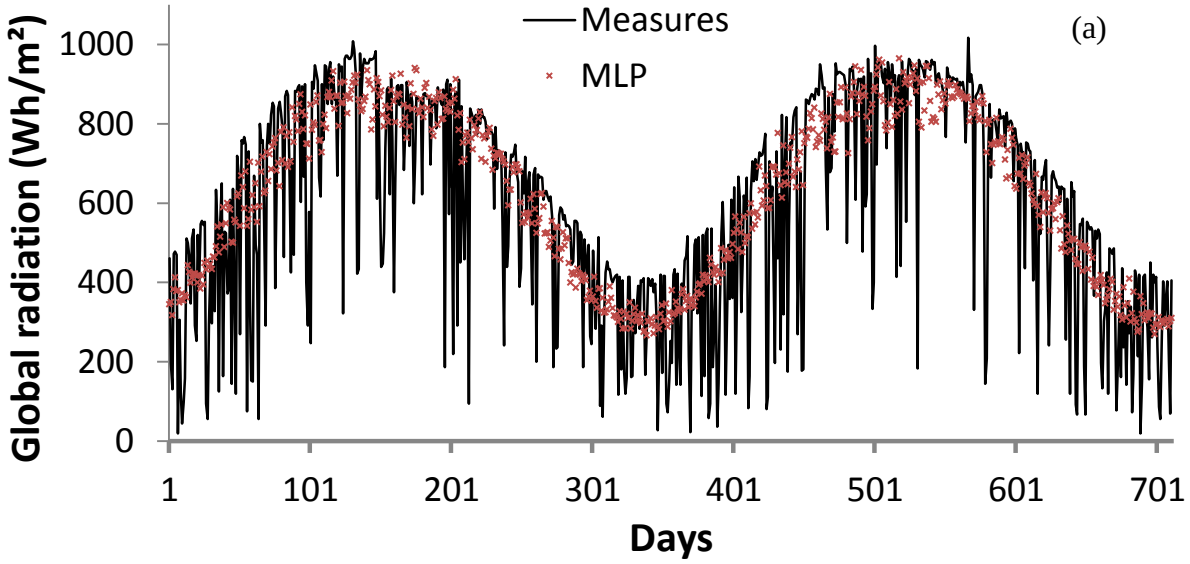
Sophisticated approaches such as ARMA or MLP give far better results than the naive persistence predictor, especially at wintertime. The unskilled performance of the persistence model is due to the choice of the prediction horizon. Indeed, the persistence model is tailored for short-term time horizons. If we consider sub-hourly or one-minute time horizons, the probability of climate conditions at time *t* to be very close to climate conditions at time *t-1* is high. In most cases, the MLP predictors have better results than the ARMA models. But the difference is small, except for summertime where the MLP presents a +0.7 point improvement compared to the ARMA model. In most cases, the best predictions results are obtained when the *CSI* has been applied to the TS. As a complement, Table 2 also presents results from the multi-output MLP approach.

The use of a stationary process introduces a significant difference in the prediction results. Indeed, except for spring, the *CSI* approach gives better results than the original approach which does not apply any stationarization process to the TS. Moreover, using one MLP only (nine outputs) improves the accuracy of the prediction. In the *CSI* case, results are improved by +0.4 point over the two-year period; during winter time, the gain is lower than 1.5 points. Using measures covering the entire sunshine duration of the day, rather than one considered time, improves the MLP results since extra information is brought into the model. For example, if yesterday at 08:00 the sky was cloudy but cleared away afterwards, then:

- in the case of multi-outputs MLP, this occurrence will be smoothed by the information brought by the other eight hours of the day. The cloudiness of the sky will have a minor impact on the prediction for time 08:00. The noise is averaged on the nine sampled hours;

- in the case of the nine MLP, the information of cloudy skies at 08:00 yesterday will have a major influence on the prediction of time 08:00.

Figures 9 and 10 illustrate the comparison between the two best predictors of each type: nine *ARMA* with one output and one MLP with nine outputs and the use of the *CSI* mode. These Figures only show the predictions related to the 12th hour (one prediction per day).



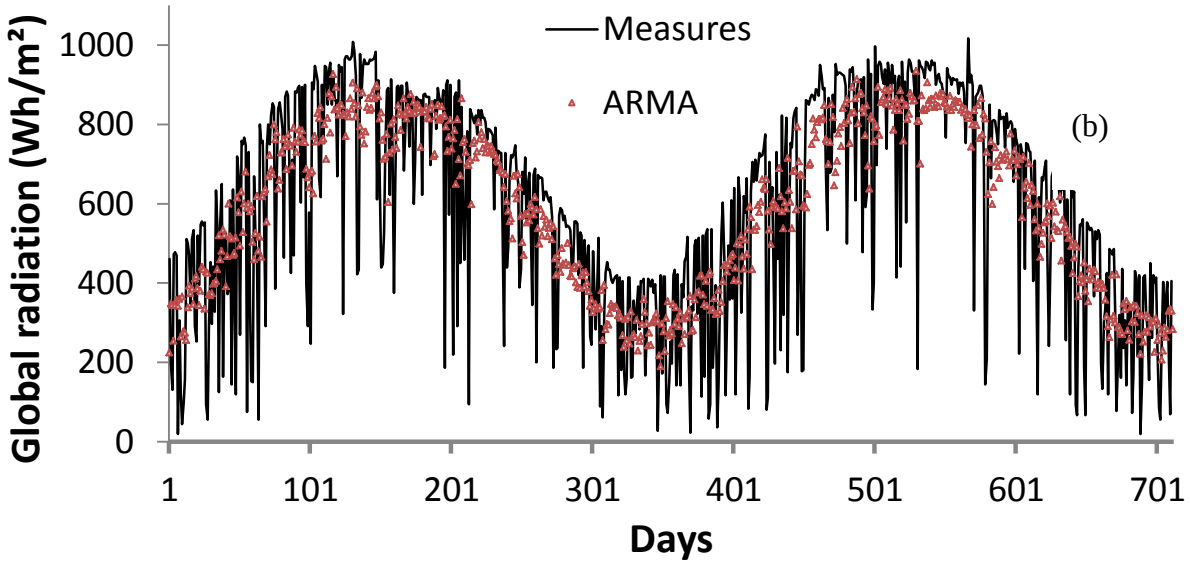


Figure 9. Best predictors comparison concerning the 12:00 TS. (a) the daily profile of the MLP predictions and measured TS, (b) the daily profile of the ARMA predictions and measured TS

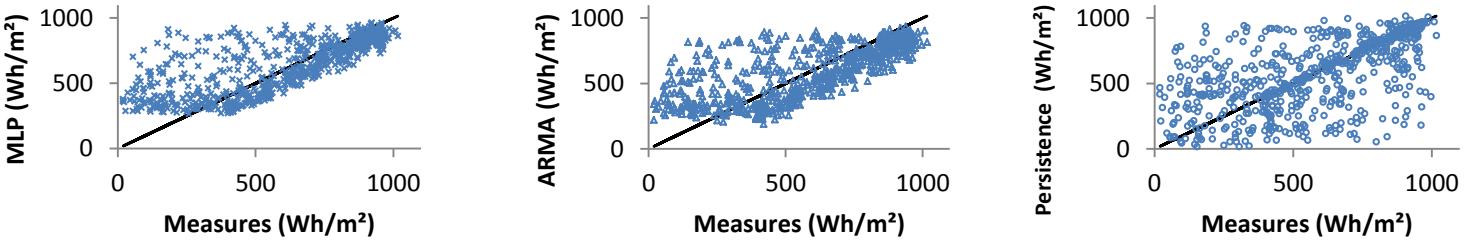


Figure 10. Best predictors comparison concerning the 12:00 TS. y=x graphical comparison of MLP model, AR model and persistence model

The squared nature of the prediction error (nRMSE) taken during the optimization phase allows to decrease the large differences but to increase the number of smaller error (Ahlburg, 1992, Voyant, 2012). It is certainly for this reason that for ARMA or for MLP, the predictions seems to describe a smoothed curve, “no risks taken” with small deviation from a central average value. During winter, the persistence model gives a prediction with a large variability and becomes unusable, whereas ARMA and MLP behave similarly. During summer, the three predictors present similar behaviors, but these results are local and cannot be generalized. Ajaccio is a seaside location with low nebulosity in summer. To improve the results for other seasons, the use of exogenous data could be useful (Paoli *et al.*, 2010 and Sfetsos and Coonick, 2000).

In the next part, multivariate regressions are proposed and exogenous data are introduced in the models to study their influence on the quality of the prediction.

4.3 Multivariate prediction

Unlike ARMA and persistence models, which are restricted to endogenous input data, MLP allows the use of exogenous input data (see section 3.4). These new models are based on the networks optimized for the endogenous case (see section 4.2). The exogenous variables are selected using the Pearson correlation coefficient (Hamilton, 1994). Among all the data presented in Section 2, one can observe that pressure gradient, daily difference pressure and rain precipitation at step *d* are not correlated enough to the global radiation at step *d+1*. Thus, these data are not included in the input nodes of the MLP. Therefore, in addition to the endogenous data as stated in section 3, the input vector may include the pressure (at lag one) between 8:00 and 16:00 (nine components), the nebulosity for the same hours for the lag one and two (eighteen components) and the daily mean nebulosity at lag two (two components). The impact of exogenous data on prediction quality is indicated in Table 3 (here only *CSI* mode is used considering it was the best pre-treatment) in terms of nRMSE related to the daily mean of global radiation.

Models		Annual	Winter	Spring	Summer	Autumn
Persistence		35.1	54.8	35.2	28.0	40.4
ARMA		28.6	44.2	28.6	23.1	32.8
MLP endo	MLP committee	28.2	44.1	28.6	22.4	33.2
	Multi-outputs MLP	27.8	42.8	**27.4**	22.0	**31.3**
MLP endo-exo	MLP committee	28.0	**42.4**	28.8	22.4	31.9
	Multi-outputs MLP	**27.3**	**42.4**	27.8	**21.7**	**31.3**

Table 3. Impact of the exogenous data on the prediction quality for the nine MLP, and the multi-output MLP (bold characters are the best results) in the case of the *CSI* pre-treatment

Exogenous variables contribute positively to multi-outputs MLP, the largest improvement being reached in the case of the 9-output MLP. During winter the gain is 1.7

points, and during autumn results are improved by 1.3 points. However during spring and summer, the error has barely decreased of 0.2 points. Figure 11 presents the detail of hourly errors for the best-case MLP (one network with nine outputs, *CSI* and exogenous variables), ARMA (nine AR with one output and the *CSI*) and the persistence models.

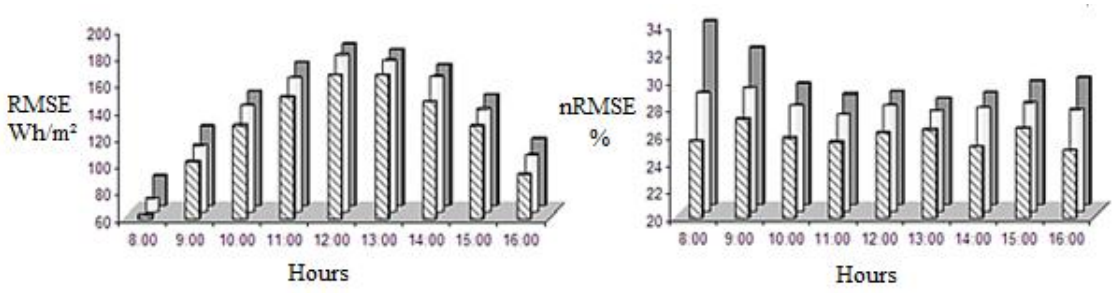


Figure 11. Hourly error (RMSE and nRMSE) of prediction for the MLP (foreground and hatched), ARMA (second ground and white) and the persistence (background and grey)

The trends of the RMSE and nRMSE curves are completely different. In the case of nRMSE, the normalization is carried out using the quadratic mean of the TS (8:00…16:00). The largest RMSE is reached at 12:00 for the three tested predictors (between 160-200 Wh/m²). The nRMSE obtained with MLP and ARMA is 'quasi-uniformly' distributed between the nine hours studied, while for the persistence, the nRMSE is more important during the low global irradiation hours. The RMSE of the three predictors don't exceed 200 Wh/m². Only in the MLP case, the nRMSE is lower than 28%.

5. Conclusion

In this paper, we have shown that MLP modeling of the global solar irradiation TS can be applied to the *h+24* time horizon prediction. The results shown demonstrate a higher accuracy with MLP models than with the persistence method. They are also consistent with

Mellit and Pavan (2010) conclusions, and allow a generalization of their approach. With this work, the ANN is now justified for a new location, with different architecture, an ad-hoc stationary process and results confronted to naïve and reference estimators (respectively persistence and ARMA models). Indeed, the use of MLP to predict the *h+24* global radiation horizon is interesting but the chosen architectures and stationarization modes, and how to use the multivariate analyze, modify greatly the results. Different levels of complexity can be implemented, namely: multi-output MLP (with or without exogenous data), MLP committee (with or without exogenous data) or ARMA model. Even if the choice between the three methodologies is scientifically difficult; an industrial decision maker can probably find a positive financial impact for a gain of 1.3 points in case of MW installations. Note that for the three methods, the complexity of the methodology is not really a barrier to its use, one time the learning phase done the model is applicable during a lot of years. The ANN and ARMA can be seen like a simple non-linear regression very easy to implement, even with a simple spreadsheet. This paper has been restricted to two architectures. As a perspective to this study, the implementation of hybrid architectures may lead to more accurate results. It would probably be interesting to combine the two approaches, by integrating data related to other hours or predictions related to other hours within the input layer of the MLP. Similarly, the use of other exogenous input data (including numerical weather forecast or other types of physical measurements) should be further investigated.

6. Acknowledgements

We thank the French National Meteorological Organization (Météo-France) for providing the data used in this study.

7. APPENDIX: The error decomposition

The purpose of this appendix is to demonstrate that the use of a one-output MLP nine times (Figure 4) is theoretically unjustified. In a simple case such that the *h+2* time horizon prediction, this methodology provides under certain assumptions, worse results than a methodology based on the MLP committee (Figure 6.a). Considering a single MLP trained over a given data set, y_i being the measures available and $\hat{y}_i$ the prediction of this value simulated with the MLP, the MSE, performance of an estimator ANN is given by (*E* is the expected value):

$$MSE = E\{(y_i - \hat{y}_i)^2\} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad \text{Equation 8}$$

In the following, the error decomposition will be applied to the simple case of the error generated for the 2-hours ahead prediction.

At first, instead of considering the nine MLP with one output (Figure 6.a), let's consider the system with only two MLP, each forecasting a series (called y^A and y^B). The series are supposed to be stationary. The decomposition scheme of the initial series y_i ($i \in [1, N]$) may be written from the two series y_j^A and y_j^B ($j \in [1, N/2]$):

$$y_i = sin^2\left(i\frac{\pi}{2}\right).y^A_{\frac{i+1}{2}} + cos^2\left(i\frac{\pi}{2}\right).y^B_{\frac{i}{2}}$$ Equation 9

The decomposition is also valid during the prediction, as shown in Equation 10.

$$\hat{y}_i = sin^2\left(i\frac{\pi}{2}\right).\hat{y}^A_{\frac{i+1}{2}} + cos^2\left(i\frac{\pi}{2}\right).\hat{y}^B_{\frac{i}{2}}$$ Equation 10

The MSE definition implies that:

$$MSE^{comm} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 = \frac{1}{N}\left(\sum_{j=1}^{\frac{N}{2}}\left(y_j^A - \hat{y}_j^A\right)^2 + \sum_{j=1}^{\frac{N}{2}}\left(y_j^B - \hat{y}_j^B\right)^2\right)$$

$$= \frac{MSE^A + MSE^B}{2}$$ Equation 11

If the series are stationary and if the two prediction methodologies are equivalent the two predictions errors are equivalent. In this case, the global error is calculated considering that $MSE^A = MSE^B = MSE^{h+1}$.

$$MSE^{comm} = MSE^{h+1}$$ Equation 12

The result is equivalent for one MLP with two outputs building two series of prediction, the first for only the *h+1* horizon and the second for only the *h+2* horizon.

If now, one MLP with one output is considered (like in Figure 4, but with only *h+2* horizon), the *h+1* prediction is taken as input to forecast the *h+2* prediction. The prediction becomes the input and the approach is then similar to a loop. The scheme decomposition of

the prediction may be written from the two series $\hat{y}_j^{h+1}$ and $\hat{y}_j^{h+2}$ ($j \in [1, N/2]$) as shown in Equation 13.

$$\hat{y}_i = sin^2\left(i\frac{\pi}{2}\right).\hat{y}_{\frac{i+1}{2}}^{h+1} + cos^2\left(i\frac{\pi}{2}\right).\hat{y}_{\frac{i}{2}}^{h+2} \quad \text{Equation 13}$$

For this prediction the MSE is described by:

$$MSE^{loop} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - sin^2\left(i\frac{\pi}{2}\right).\hat{y}_{\frac{i+1}{2}}^{h+1} - cos^2\left(i\frac{\pi}{2}\right).\hat{y}_{\frac{i}{2}}^{h+2}\right)^2 \quad \text{Equation 14}$$

After some modifications on the decomposed expression, it follows:

$$MSE^{loop} = \frac{1}{N}\sum_{j=1}^{N/2}\left(y_{2j-1} - \hat{y}_j^{h+1}\right)^2 + \frac{1}{N}\sum_{j=1}^{N/2}\left(y_{2j} - \hat{y}_j^{h+2}\right)^2 \quad \text{Equation 15}$$

Considering the MSE^A expression and the precedent equation, MSE^{loop} becomes:

$$MSE^{loop} = \frac{1}{2}(MSE^{h+1} + E\{\left(y_{2j} - \hat{y}_j^{h+2}\right)^2\}) \quad \text{Equation 16}$$

This equation is equivalent to:

$$MSE^{loop} = \frac{1}{2}(MSE^{h+1} + MSE^{h+2}) \quad \text{Equation 17}$$

To compute the prediction $\hat{y}_j^{h+2}$ it is necessary to use the prediction $\hat{y}_j^{h+1}$, the MSE^{h+2} is greater than MSE^{h+1}. Considering that $MSE^{h+2} = (1+\varepsilon).MSE^{h+1}$ with $\varepsilon > 0$. The total MSE is equal to:

$$MSE^{loop} = \frac{1}{2}(MSE^{h+1} + (1+\varepsilon).MSE^{h+1}) = \frac{2+\varepsilon}{2}.MSE^{h+1} \quad \text{Equation 18}$$

In the *h+2* case: $MSE^{loop} > MSE^{comm}$. Therefore, there is no interest in using the loop approach. Next part of the appendix focuses on the *h+9* horizon ranking and on the prediction error behavior (divergence) for the loop methodology based on one MLP with one output.

As seen previously, the MSE for the MLP system and the multi-outputs MLP are decomposed by the same scheme. For these two predictors and for the *h+2* time horizon, the total MSE is lower than the error obtained with the MLP built with one output (loop methodology; the *h+1* prediction is taken like input to forecast the *h+2* prediction, etc.). This result is valid if an hourly prediction is equivalent to a daily prediction (same noise and same stationarity) and if all the hourly TS used are equivalent for the prediction process. With all these hypotheses the result ($MSE^{loop} > MSE^{comm}$) can be generalized to the *h+9* time horizon.

8. References

Abrahart RJ, See L. 1998. Neural Network vs. ARMA Modelling: constructing benchmark case studies of river flow prediction.

Ahlburg D. 1992. Error measures and the choice of a forecast method. International Journal of Forecasting. 8(1):99–100.

Alados I, Gomera MA, Foyo-Moreno I, Alados-Arboledas L. 2007. Neural network for the estimation of UV erythemal irradiance using solar broadband irradiance. International Journal of Climatology. 5;27(13):1791–9.

Al-Alawi SM, Al-Hinai HA. 1998. An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renewable Energy. 14(1-4):199–204.

Altandombayci O, Golcu M. 2009. Daily means ambient temperature prediction using artificial neural network method: A case study of Turkey. Renewable Energy. 34(4):1158–61.

Badescu V. 2008. Modeling solar radiation at the earth's surface: recent advances. Springer.

Baigorria GA, Villegas EB, Trebejo I, Carlos JF, Quiroz R. 2004. Atmospheric transmissivity: distribution and empirical estimation around the central Andes. International Journal of Climatology. 1;24(9):1121–36.

Balestrassi P, Popova E, Paiva A, Marangonlima J. 2009. Design of experiments on neural network's training for nonlinear time series forecasting. Neurocomputing. 72(4-6):1160–78.

Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR. 2010. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Solar Energy. 84(8):1468–80.

Benghanem M, Mellit A. 2010. Radial Basis Function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia. Energy. 35(9):3751–62.

Bosch J, Lopez G, Batlles F. 2008. Daily solar irradiation estimation over a mountainous area using artificial neural networks. Renewable Energy. 33(7):1622–8.

Bourbonnais R, Terraza M. 2008. Analyse des séries temporelles : application à l'économie et à la gestion. 2e éd. Paris: Dunod.

Brockwell PJ, Davis RA. 1991. Time series: theory and methods. 2nd ed. New York: Springer-Verlag.

Cao L. 1997. Practical method for determining the minimum embedding dimension of a scalar time series. Physica D: Nonlinear Phenomena 110, no. 1-2: 43-50.

Cao JC, Cao SH. 2006. Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis. Energy. 31(15):3435–45.

Chaouachi A, Kamel RM, Ichikawa R, Hayashi H, Nagasaka K. 2009 . Neural Network Ensemble-Based Solar Power Generation Short-Term Forecasting. World Academy of Science, Engineering and Technology. 54.

Chen Y-H, Chen C-Y, Lee S-C. 2011. Technology forecasting and patent strategy of hydrogen energy and fuel cell technologies. International Journal of Hydrogen Energy. 36(12):6957–69.

Cococcioni M, D'Andrea E, Lazzerini B. 2011. 24-hour-ahead forecasting of energy production in solar PV systems. 11th International Conference on Intelligent Systems Design and Applications (ISDA), 1276 -1281,

Coulibaly P, Anctil F, Bobée B. 1999. Prévision hydrologique par réseaux de neurones artificiels : état de l'art. Can. J. Civ. Eng. 26(3):293–304.

Crone SF. 2005. Stepwise Selection of Artificial Neural Networks Models for Time Series Prediction.

Crone SF, Kourentzes N. 2010. Feature selection for time series prediction – A combined filter and wrapper approach for neural networks. Neurocomputing. 73(10-12):1923–36.

Cybenko G. 1989. Approximation by superpositions of a sigmoidal function. Math. Control Signal Systems. 2(4):303–14.

De Gooijer JG, Hyndman RJ. 2006. 25 years of time series forecasting. International Journal of Forecasting.;22(3):443–73.

Dreyfus G. 2004. Réseaux de neurones: Méthodologie et applications. 2sd éd. Eyrolles;.

Eltahir EAB, Humphries Jr. EJ. 2004. The role of clouds in the surface energy balance over the Amazon forest. International Journal of Climatology. 30;18(14):1575–91.

Franco A and Salza P. 2011. Strategies for optimal penetration of intermittent renewables in complex energy systems based on techno-operational objectives. Renewable Energy. 36, pp.743-753

Hamilton J. 1994. Time series analysis. Princeton N.J.: Princeton University Press.

Haurant P, Muselli M, Oberti P, Pillot B, Thibault C. 2010. Multicriteria Decision Aiding for Selection of Photovoltaic Plants on Farming Fields in Corsica. EU PVSEC Proceedings. Valencia, Spain. p. 5267 - 5270.

Hawking S. 1998. A Brief History of Time. Updated and Expanded Tenth Anniversary ed. Bantam Doubleday Dell Publishing Group.

Hornik K, Stinchcombe M, White H. 1989. Multilayer feedforward networks are universal approximators. Neural Networks. 2(5):359–66.

Hu Y, Hwang J-N. 2002. Handbook of neural network signal processing. Boca Raton: CRC Press.

Ineichen P. 2006. Comparison of eight clear sky broadband models against 16 independent data banks. Solar Energy. 80(4):468–78.

Ineichen P. 2008. A broadband simplified version of the Solis clear sky model. Solar Energy. 82(8):758–62.

Ipsakis D, Voutetakis S, Seferlis P, Stergiopoulos F, Elmasides C. 2009. Power management strategies for a stand-alone power system using renewable energy sources and hydrogen storage. International Journal of Hydrogen Energy. 34(16):7081–95.

Iqdour R, Zeroual A. 2006. The MLP Neural Networks for Predicting Wind Speed. Marrakech, Morocco.

Ito Y. 1991. Representation of functions by superpositions of a step or sigmoid function and their applications to neural network theory. Neural Networks. 4(3):385–94.

Jain AK, Jianchang Mao, Mohiuddin KM. 1996. Artificial neural networks: a tutorial. Computer. 29(3):31–44.

Jenn Jiang H. Policy review of greenhouse gas emission reduction in Taiwan. Renewable and Sustainable Energy Reviews. 15(2):1392–402.

Kalogirou S. 2011. Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews. 2001;5(4):373–401.

López G, Batlles FJ, Tovar-Pescador J. 2005. Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy. 30(9):1675–84.

Mellit A, Kalogirou SA, Hontoria L, Shaari S. 2009. Artificial intelligence techniques for sizing photovoltaic systems: A review. Renewable and Sustainable Energy Reviews. 13(2):406–19.

Mellit A, Pavan AM. 2010. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. Solar Energy. 84(5):807–21.

Michaelides SC, Pattichis CS, Kleovoulou G. 2001. Classification of rainfall variability by using artificial neural networks. International Journal of Climatology. 1;21(11):1401–14.

Min Qi, Zhang GP. 2008. Trend Time Series Modeling and Forecasting With Neural Networks. IEEE Trans. Neural Netw. 19(5):808–16.

Moreno-Munoz A, de la Rosa JJG, Posadillo R, Bellido F. 2008. Very Short Term Forecasting of Solar Radiation. PVSC: 2008 33RD IEEE PHOTOVOLTAIC SPECIALISTS CONFERENCE, VOLS 1-4. p. 949–53.

Mubiru J. 2008. Predicting total solar irradiation values using artificial neural networks. Renewable Energy. 33(10):2329–32.

Mubiru JMwale D, Yew Gan T, Shen SSP. 2004. A new analysis of variability and predictability of seasonal rainfall of central southern Africa for 1950–94. International Journal of Climatology. 1;24(12):1509–30.

Ni M, Leung MKH, Sumathy K, Leung DYC. 2006. Potential of renewable hydrogen production for energy supply in Hong Kong. International Journal of Hydrogen Energy. 31(10):1401–12.

Paoli C, Voyant C, Muselli M, Nivet M-L. 2010. Forecasting of preprocessed daily solar radiation time series using neural networks. Solar Energy. 84(12):2146–60.

Parlitz U. 1995. Nonlinear Time Series Analysis, Proceedings of the 3rd international specialist workshop on Nonlinear Dynamics of Electronic Systems, NDES'95, University College Dublin, Ireland, 28-29. pp. 179-192

Raschke E, Gratzki A, Rieland M. 1987. Estimates of global radiation at the ground from the reduced data sets of the international satellite cloud climatology project. Journal of Climatology. 1;7(3):205–13.

Reddy K. 2003. Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy Conversion and Management. 44(15):2519–30.

Santarelli M, Calì M, Macagno S. 2004. Design and analysis of stand-alone hydrogen energy systems with different renewable sources. International Journal of Hydrogen Energy. 29(15):1571–86.

Sfetsos A, Coonick AH. 2000. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Solar Energy. 68(2):169–78.

Suratgar, A. A., M. B. Tavakoli, et A. Hoseinabadi. 2005. Modified Levenberg-Marquardt method for neural networks training. World academy of science, Engineering and Technology 6: 46–48.

Voyant C, Muselli M, Paoli C, Nivet M-L. 2011. Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation. Energy. 36(1):348–59.

Voyant C, Muselli M, Paoli C, Nivet M-L. 2012. Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. Energy.

Voyant C. Prédiction de séries temporelles de rayonnement solaire global et de production d'énergie photovoltaïque à partir de réseaux de neurones artificiels. Thèse de Doctorat, université de Corse 2012.

Zhang GP, Qi M. 2005. Neural network forecasting for seasonal and trend time series. European Journal of Operational Research. 160(2):501–14.

Zhou K, Ferreira JA, de Haan SWH. 2008. Optimal energy management strategy and system sizing method for stand-alone photovoltaic-hydrogen systems. International Journal of Hydrogen Energy. 33(2):477–89.

9. List of Figures and tables

Figure 8. Autoregression coefficients of the nine AR(*p*) with the *CSI* forecasting mode (the red line is the significance limit)

Figure 9. Best predictors comparison concerning the 12:00 TS. (a) the daily profile of the MLP predictions and measured TS, (b) the daily profile of the ARMA predictions and measured TS

Figure 10. Best predictors comparison concerning the 12:00 TS. y=x graphical comparison of MLP model, AR model and persistence model

Figure 11. Hourly error (RMSE and nRMSE) of prediction for the MLP (foreground and hatched), ARMA (second ground and white) and the persistence (background and grey)

Table 1. Optimization of the nine independent MLP and the multi-output MLP

Table 2. nRMSE (%) of predictions with the nine independent networks, with the mupti-output MLP and AR model (bold characters are related to the best results)

Table 3. Impact of the exogenous data on the prediction quality for the nine MLP, and the multi-output MLP (bold characters are the best results) in the case of the *CSI* pre-treatment

	networks	*X*	*k*	*CSI*
MLP committee nine networks	8:00	$Endo^{1\text{-}3}$x4x1	$Endo^{1\text{-}3}$x4x1	$Endo^{1\text{-}11}$x8x1
	9:00	$Endo^{1\text{-}10}$x4x1	$Endo^{1\text{-}8}$x3x1	$Endo^{1\text{-}8}$x3x1
	10: 00	$Endo^{1\text{-}11}$x4x1	$Endo^{1\text{-}20}$x2x1	$Endo^{1\text{-}20}$x2x1
	11:00	$Endo^{1\text{-}3}$x8x1	$Endo^{1\text{-}2}$x4x1	$Endo^{1\text{-}3}$x3x1
	12:00	$Endo^{1\text{-}20}$x1x1	$Endo^{1\text{-}20}$x1x1	$Endo^{1\text{-}20}$x1x1
	13:00	$Endo^{1\text{-}12}$x7x1	$Endo^{1\text{-}19}$x1x1	$Endo^{1\text{-}1}$x14x1
	14:00	$Endo^{1\text{-}18}$x2x1	$Endo^{1\text{-}20}$x1x1	$Endo^{1\text{-}3}$x19x1
	15:00	$Endo^{1\text{-}11}$x5x1	$Endo^{1\text{-}12}$x1x1	$Endo^{1\text{-}4}$x10x1
	16:00	$Endo^{1\text{-}13}$x3x1	$Endo^{1\text{-}1}$x1x1	$Endo^{1\text{-}6}$x5x1
Multi-outputs MLP	Only one	$Endo^{1\text{-}27}$x1x9	$Endo^{1\text{-}27}$x1x9	$Endo^{1\text{-}27}$x2x9

Table 1. Optimization of the nine independent MLP and the multi-output MLP

Models		Annual	Winter	Spring	Summer	Autumn
Persistence		35.1	54.8	35.2	28.0	40.4
ARMA	*k*	29.1	44.6	29.2	24.0	33.2
	CSI	28.6	44.2	28.6	23.1	32.8
MLP committee	*k*	28.5	44.6	28.8	22.9	32.8
	CSI	28.2	44.1	28.6	22.4	33.2
Multi-outputs MLP	*k*	27.9	44.2	**27.9**	22.2	32.7
	CSI	**27.8**	**42.8**	28.4	**22.0**	**31.3**

Table 2. nRMSE (%) of predictions with the nine independent networks, with the mupti-output MLP and AR model (bold characters are related to the best results)

Models		Annual	Winter	Spring	Summer	Autumn
Persistence		35.1	54.8	35.2	28.0	40.4
ARMA		28.6	44.2	28.6	23.1	32.8
MLP endo	MLP committee	28.2	44.1	28.6	22.4	33.2
	Multi-outputs MLP	27.8	42.8	**27.4**	22.0	**31.3**
MLP endo-exo	MLP committee	28.0	**42.4**	28.8	22.4	31.9
	Multi-outputs MLP	**27.3**	**42.4**	27.8	**21.7**	**31.3**

Table 3. Impact of the exogenous data on the prediction quality for the nine MLP, and the multi-output MLP (bold characters are the best results) in the case of the CSI pre-treatment