



On meteorological forecasts for energy management and large historical data: A first look

Cyril VOYANT^{1, 2}, Cédric JOIN^{3, 5, 6}, Michel FLIESS^{4, 5}, Marie-Laure NIVET¹, Marc MUSELLI¹
and Christophe PAOLI⁷

1- SPE (CNRS, UMR 6134), Università di Corsica Pasquale Paoli
20250 Corte, France

2- Hôpital de Castelluccio, Unité de Radiothérapie
BP 85, 20177 Ajaccio, France

3- CRAN (CNRS, UMR 7039), Université de Lorraine
BP 239, 54506 Vandœuvre-lès-Nancy, France

4- LIX (CNRS, UMR 7161), École polytechnique
91128 Palaiseau, France

5- AL.I.E.N. (ALgèbre pour Identification & Estimation Numériques)
24-30 rue Lionnois, BP 60120, 54003 Nancy, France

6- Projet NON-A, INRIA Lille – Nord-Europe, France

7- Département de Génie Informatique, Université Galatasaray
No: 36 34357 Ortakoy, Istanbul, Turkey

{voyant, marc.muselli, marie-laure.nivet}@univ-corse.fr,
cedric.join@univ-lorraine.fr, Michel.Fliess@polytechnique.edu,
{cedric.join, michel.fliess}@alien-sas.com, cpaoli@gsu.edu.tr

Abstract. This communication is devoted to a comparison between various meteorological forecasts, for the purpose of energy management, via different time series techniques. The first group of methods necessitates a large number of historical data. The second one does not and is much easier to implement, although its performances are today only slightly inferior. Theoretical justifications are related to methods stemming from a new approach to time series, artificial neural networks, computational intelligence and machine learning. Several numerical simulations are provided and discussed.

Keywords. Meteorology, forecasting, time series, trend, quick fluctuations, multilayer perceptron, persistence, scaled persistence, big data.

1. Introduction and overview of the methods

Numerical weather prediction is a most important topic of investigation which utilizes many advanced and competing tools (see, *e.g.*, [7, 17, 34], and the references therein). Our purpose here, for which rather short-term forecasts are needed, is to improve energy management, in Corsica, a well known French island in the Mediterranean sea.

Remark. Weather forecast for energy management is a rather new field of studies. See, *e.g.*, [1, 3, 17, 19, 24, 30, 35], and the references therein.

We will concentrate on the connections between *time series*, or *TS*, which play a key role not only in meteorology, but also in many other domains, like, for instance, econometrics and medicine, and various topics from computer science, like artificial neural networks, computational intelligence and machine learning (see, *e.g.*, [3, 5, 6, 19, 23, 28]). This communication, which should be read in conjunction with [18], is perhaps among the first ones which take into account the implementation issues by evaluating quite closely their need of *big data*, *i.e.*, of a large number of historical data.

Our multivariable meteorological time series stem from hourly measurements during 11 years in the city of Ajaccio in Corsica (41°55'N and 8°44'E, elev. 0-787 m). They contain four components, *i.e.*, the solar radiation (Sun in Wh/m), the humidity (Hum in %), the temperature, in Celsius degrees, and the wind, in Beaufort scale. Let us emphasize that the solar radiation and the temperature exhibit very strong periodicities, whereas those features are much milder with the two other quantities. All the data are from the French national meteorological

service (*Météo-France*). There are of course a few missing values. They are replaced by appropriate averages. The 9 first years are dedicated to the training if large historical data are needed. The 2 last years are devoted to the cross comparison of predictors. The prediction horizon is 1 hour which is also the time step of our time series. Those quite repetitive behaviors (yearly and daily periodicities) do necessitate revisiting techniques which work otherwise in a more or less satisfactory manner.

Remark. Those repetitive behaviors are often called *seasonalities* in the classic literature on time series, which is mainly devoted to econometrics (see, *e.g.*, [16]). See, also, [25] for a nice historical overview of this literature.

A. Overview of the various techniques

We compare two groups of methods for the TS analysis, with or without a need of a large number of historical data.

1) Settings *without* large historical data

Two types of approaches are considered:

- The first one *without model* (*WM*). This *model-free* setting (see [9, 11, 12], and the references therein) employs new online estimation techniques (see, *e.g.*, [8, 14, 15, 22]).
- The *persistence* (*P*).

Remark. See [10] for an introduction to *model-free control* and to the corresponding *intelligent PID* controllers.

A theorem due to Cartier and Perrin [4] is also utilized in order to show that any TS under a quite weak integrability assumption is a sum of a *trend* and of *quick fluctuations*.¹ Efficient forecasts in financial engineering have been already deduced [9, 12]. The connection with computational intelligence and machine learning, which is developed in [8, 13, 14, 15, 22, 29], shows an easy implementation, and an excellent robustness with respect to noise corruption.

Remark. See [18] for a slightly more detailed summary of this setting, which is also written for meteorological purposes.

The persistence (*P*) method [28], which is a trivial machine learning view point, assumes no change between the forecast and the last measurement.

2) Settings *with* large historical data

A rather precise modeling is needed for the second group, which is therefore quite greedy with data. We are considering four cases:

- The *MultiLayer Perceptron* (*MLP*) is the most popular artificial neural network concerning meteorological TS modeling (see, *e.g.*, [3, 5, 31, 33]). Several studies (see, *e.g.*, [23, 31]) have validated MLP predictions for nonlinear time series case. In order to build this predictor, a fixed number of past measurements are set as inputs of the MLP, *i.e.*, as

juxtaposition in its input layer. The output of the MLP, often called the *output layer*, allows to generate the prediction for the horizon *h* of the future measurement. It reads

$$x(t+h) = f(x(t), x(t-1), \dots, x(t-p))$$

where the function *f*, which is in general nonlinear, depends on the MLP characteristics. The MLP has been computed here with the Neural Network toolbox of Matlab[®]. The characteristics, which are related to previous works, are the following ones:

- one *hidden layer*,
- the activation functions are a continuously and differentiable hyperbolic tangent concerning the hidden layer, and a linear trend concerning the *output layer*,
- the Levenberg-Marquardt learning algorithm, which is a Newton's type method of approximation, with a max fail parameter before stopping training equal to 3.²

The optimization of the number of input nodes is done with the partial autocorrelation factor. The first minimum of the function defines the deepest lag to consider, *i.e.*, if the minimum concerns the lag 3: only $x(t)$, $x(t-1)$ and $x(t-2)$ are set in the input layer of the MLP. Concerning the four situations for our TS, the size of the input layer varies between 4 and 9 nodes: 6 for Sol, 4 for Temp, 9 for Hum and 5 for Wind. The number of hidden neurons is taken equal to the number of input nodes. The results of Section 3. are related to the best networks among six different trainings coupled with a random weights initialization:

- If we take into account a probabilistic description of the TS, a stationary hypothesis is often necessary. It yields a MA-MLP, *i.e.*, a moving average in connection with the MLP in order to deal with a stationary TS [32].
- The average (*A*) method [28], which may be viewed as a technique of computational intelligence, is based on a typical year of moving averages computed during the 9 first years.
- The scaled persistence (*SP*) [32] adds to the above *P* a moving average procedure like the above one. It allows to take into account the fact that the TS is periodical. This methodology takes into account, in the global radiation case, the apparent position of the sun for the next hour.

B. Organization of the paper

After this rather lengthy Section 1., the remaining of our paper is organized as follows. Implementation issues and numerical results obtained in various situations are respectively reported in Section 2. and 3.. Some concluding remarks may be found in Section 4..

1. See [21] for a simpler explanation of this deep mathematical result.

2. Such an early stopping concept is useful in order to prevent overfitting.

2. Some implementation issues

Figure 1 shows for the MLP and for the four TS available the impact of the learning data size. With SP and MA-MLP the average curves are respectively fitted via the functions

$$f_1(x) = 1.743 \exp(-1.611x) + 1.021 \exp(-0.002271x)$$

and

$$f_2(x) = 206.5 \exp(-3.907x) + 1.082 \exp(-0.009358x)$$

In this Figure the the $nRMSE_{min}$ metric, *i.e.*, the minimum of the normalized root mean square error, is related to the maximum data set size, *i.e.*, 9 years. For the global radiation, 4 years are useful. With 6 years the loss is close to 1 %. For the wind speed and the humidity the impact of the learning data size is not important (maybe linked to the less significant periodicities) and for temperature at least 1 year is necessary. In average, if only one half year is used the nRMSE is increased by 25%. Concerning the WM procedure, in order to extend [9, 12] to meteorological predictions, only the fluctuation corresponding to the previous day has to be taken into account. This is achieved via a polynomial $a_0 + a_1t + a_2t^2$ of degree 2. The elementary algebraic manipulations giving a_0, a_1, a_2 are deduced from [14, 15, 22]. The forecast is deduced from the error corresponding to the last measurement.

3. Prediction results

A. Presentation

The normalized root mean square error (nRMSE) and the normalized mean absolute error (nMAE) are presented in Tables 1 and 2. The probabilistic models are derived from the overall training data set. Despite the small improvement due to the extension of the size of the training set we chose to remain in the most favorable situation, *i.e.*, 9 years. Note that the similarity between P and SP in the case of Humidity and Wind is explained by low fluctuations.

B. A preliminary ranking

The interpretation of Tables 1 and 2 allows to define a mono-criterion analysis for the predictor ranking. It is based on the interpretation of the prediction error, *i.e.*, nRMSE and nMAE, concerning our four TS. For each TS, the best predictor, which is related to the lowest error metrics, scores 5 points, the second 4 points, . . . , the worst 0 point. Applying this methodology, the predictors ranking is almost equivalent concerning nRMSE and nMAE and becomes:

$$\text{MA-MLP}(17,17) > \text{SP}(16,17) > \text{MLP}(15,14) > \text{WM}(6,7) > \text{P}(6,6) > \text{A}(0,0)$$

where the two numbers between the parentheses are related to nRMSE and nMAE. The following facts should be stressed:

- MA-MLP is generally better predictor than the single MLP,
- WM is slightly better than P,

- SP yields accurate predictions while its methodology is very simple.

Table 3 shows that the difference between WM and the best predictor is respectively less than 6 and 5 %. WM and P seems moreover equivalent. For a strongly seasonal behavior (Sol and Temp), the gain obtained by employing WM is important. In order to compare the six predictors, it is also possible to construct a statistical test based on the percentage of time where predictors give signal estimations better than 5% and 10%. The results, obtained thanks to the passing rate test, are displayed in Table 4.

This process allows us to determine the methodologies which exhibit a lot of outliers. They are of course most detrimental in practice. If we apply the same ranking methodology as above but now concerning the test passing rate with respect to the percentage of good estimations, the ranking is left unchanged. The prioritization is respected and the best predictor is MA-MLP with a score equal to 17 and the worst is the average with a score equal to 1. Predictions using modeling and requiring large learning historical data, *i.e.*, SP, MLP and MA-MLP, are often better than others. However, WM and P are in some cases most interesting. The two best predictors of each groups, *i.e.*, large and sparse historical data, are MA-MLP and WM, especially for the wind speed. Table 5 represents the seasonal nRMSE for those two predictors.

During summer and autumn, MA-MLP is much better than WM. Nevertheless, WM is the best predictor during the two other seasons. In 37.5% (resp. 62.5%) of the cases WM (resp. MA-MLP) provides the best results. Prediction profiles concerning 100 hours in autumn are shown in Figure 2. All the predictors seem to be equivalent and approximate with a good accuracy the measured data, except for the wind where periodicities are less significant. Let us emphasize that the wind speed is certainly the most difficult meteorological data to forecast among the four situations (see, *e.g.*, [1, 2, 26], and the references therein).

4. Conclusion

MA-MLP, *i.e.*, multilayer perceptron with moving average, and SP, *i.e.*, scaled persistence, are the best predictors concerning the four seasonal TS studied. The WM approach, *i.e.*, without model, although it may be today slightly less efficient, needs much less measurements and is quite easy to implement. Those encouraging results will yield further investigations on settings with less data and their comparisons to more classic ones with a large number of historical data. Note moreover that for an efficient climate control of a building (see, *e.g.*, [27]), where some weather forecasts are also needed, model-free control [10], which was already successfully applied to agricultural greenhouses [20], might be the appropriate tool.

References

- [1] R. Baïle, “Analyse et modélisation multifractales de vitesses de vent. Application à la prévision de la ressource éolienne”, Thèse, Université de Corse, 2010.
- [2] R. Baïle, J.-F. Muzy, P. Poggi, “Short-term forecasting of surface layer wind speed using a continuous random cascade model”, *Wind Energy*, 14: 719–734, 2011.
- [3] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, A. R. Noghrehabadi, “The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data”, *Solar Energy*, 84:1468–1480, 2010.
- [4] P. Cartier, Y. Perrin, “Integration over finite sets”, in F. & M. Diener, Eds, *Nonstandard Analysis in Practice*, Springer, 1995, pp. 195–204.
- [5] S.F. Crone, “Stepwise selection of artificial neural networks models for time series prediction”, *J. Intelligent Syst.*, 14: 99–122, 2005.
- [6] J. G. De Gooijer, R. J. Hyndman, “25 years of time series forecasting”, *Int. J. Forecasting*, 22: 443–473, 2006.
- [7] C. Duchon, R. Hale, *Time Series Analysis in Meteorology and Climatology: An Introduction*, Wiley-Blackwell, 2012.
- [8] M. Fliess, “Analyse non standard du bruit”, *C.R. Acad. Sci. Paris Ser. I*, 342: 797–802, 2006.
- [9] M. Fliess, C. Join, “A mathematical proof of the existence of trends in financial time series”, in A. El Jai, L. Afifi and E. Zerrik, Eds, *Systems Theory: Modeling, Analysis and Control*, Presses Universitaires de Perpignan, 2009, pp. 43–62. Available at <http://hal.archives-ouvertes.fr/inria-00352834/en/>
- [10] M. Fliess, C. Join, “Model-free control”, *Int. J. Control*, 86: 2228–2252, 2013.
- [11] M. Fliess, C. Join, “Towards a new viewpoint on causality for time series”, *Coll. Modél. Contr. Anal. Syst. Prof. A. El Jai, Ifrane*, 2014. To appear in *ESAIM ProcS*. Available at <http://hal.archives-ouvertes.fr/hal-00991942/en/>
- [12] M. Fliess, C. Join, F. Hatt, “A-t-on vraiment besoin d’une approche probabiliste en ingénierie financière?”, in *Proc. Conf. Médit. Ingénierie Sûre Systèmes Complexes*, Agadir, 2011. Available at <http://hal.archives-ouvertes.fr/hal-00585152/en/>
- [13] M. Fliess, C. Join, M. Mboup, “Algebraic change-point detection”, *Appl. Algebra Engin. Communic. Comput.*, 21: 131–143, 2010.
- [14] M. Fliess, C. Join, H. Sira-Ramírez, “Non-linear estimation is easy”, *Int. J. Modelling Identif. Control*, 4:12–27, 2008. Available at <http://hal.archives-ouvertes.fr/inria-00158855/en/>
- [15] M. Fliess, H. Sira-Ramírez, “Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques”, in H. Garnier and L. Wang, editors, *Closed-Loop Parametric Identification for Continuous-Time Linear Systems via New Algebraic Techniques*, Springer, 2008, pp. 362–391.
- [16] C. Gourieroux, A. Monfort, *Séries temporelles et modèles dynamiques (2^e éd.)*, Economica, 1995. English translation: *Time Series and Dynamic Models*, Cambridge University Press, 1997.
- [17] F. O. Hocaoglu, “Stochastic approach for daily solar radiation modeling”, *Solar Energy*, 85: 278–287, 2011.
- [18] C. Join, C. Voyant, M. Fliess, M. Muselli, M.-L. Nivet, C. Paoli, F. Chaxel, “Short-term solar irradiance and irradiation forecasts via different time series techniques: A preliminary study”, in *Proc. 3rd Int. Symp. Environ.-Friendly Energy Appl.*, Paris, 2014. Available at <http://hal.archives-ouvertes.fr/hal-01068569/en/>
- [19] S. Kalogirou, “Artificial neural networks in renewable energy systems applications: A review”, *Renew. Sustain. Energy Rev.*, 5: 373–401, 2001.
- [20] F. Lafont, J.-F. Belmat, N. Pessel, M. Fliess, “A model-free control strategy for an experimental greenhouse with an application to fault accommodation”, *Comput. Electron. Agriculture*, 110: 139–149, 2015.
- [21] C. Lobry, T. Sari, “Nonstandard analysis and representation of reality”, *Int. J. Control*, 81: 519–536, 2008.
- [22] M. Mboup, C. Join, M. Fliess, “Numerical differentiation with annihilators in noisy environment”, *Num. Algo.*, 50: 439–467, 2009.
- [23] A. Mellit, S. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review”, *Prog. Energy Combustion Sci.*, 34: 574–632, 2008.
- [24] Methaprayoon K., W.-J. Lee ; Rasmiddatta S., Liao J.R., Ross, R.J., “Multistage artificial neural network short-term load forecasting engine with front-end weather forecast”, *IEEE Trans. Industry Appl.*, 43: 1410–416, 2007.
- [25] V. Meuriot, *Une histoire des concepts des séries temporelles*, Harmattan–Academia, 2012.
- [26] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, “Wind power forecasting: state-of-the-art 2009”, *Tech. Rep.*, Argonne National Laboratory, 2009. Available at <http://www.dis.anl.gov/pubs/65613.pdf>
- [27] F. Oldewurtel, A. Parisio, C.N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, M. Morari, “Use of model predictive control and weather forecasts for energy efficient building climate control”, *Energy Buildings*, 45: 15–27, 2012.
- [28] C. Paoli, C. Voyant, M. Muselli, M.-L. Nivet, “Forecasting of preprocessed daily solar radiation time series using neural networks”, *Solar Energy*, 84: 2146–2160, 2010.

- [29] H. Sira-Ramírez, C. García-Rodríguez, J. Cortés-Romero, A. Luviano-Juárez, Algebraic Identification and Estimation Methods in Feedback Control Systems, Wiley, 2014.
- [30] C. Voyant, “Prédiction de séries temporelles de rayonnement solaire global et de production d’énergie photovoltaïque à partir de réseaux de neurones artificiels”, Thèse, Université de Corse, 2011.
- [31] C. Voyant, M. Muselli, C. Paoli, M.-L. Nivet, “Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation”, Energy, 36: 348–359, 2011.
- [32] C. Voyant, C. Paoli, M. Muselli, M.-L. Nivet, “Multi-horizon solar radiation forecasting for Mediterranean locations using time series models”, Renew. Sustain. Energy Rev., 28: 44–52, 2013.
- [33] C. Voyant, W. Tamas, C. Paoli, A. Balu, M. Muselli, M.-L. Nivet, G. Notton, “Time series modeling with pruned multi-layer perceptron and 2-stage damped least-squares method”, J. Phys. Conf. Ser., 490: 012040, 2014
- [34] T.T. Warner, Numerical Weather and Climate Prediction, Cambridge University Press, 2010.
- [35] V.M. Zavala, E.M. Constantinescu, T. Krause, M. Anitescu, “On-line economic optimization of energy systems using weather forecast information”, J. Process Contr., 19: 1725–1736, 2009.

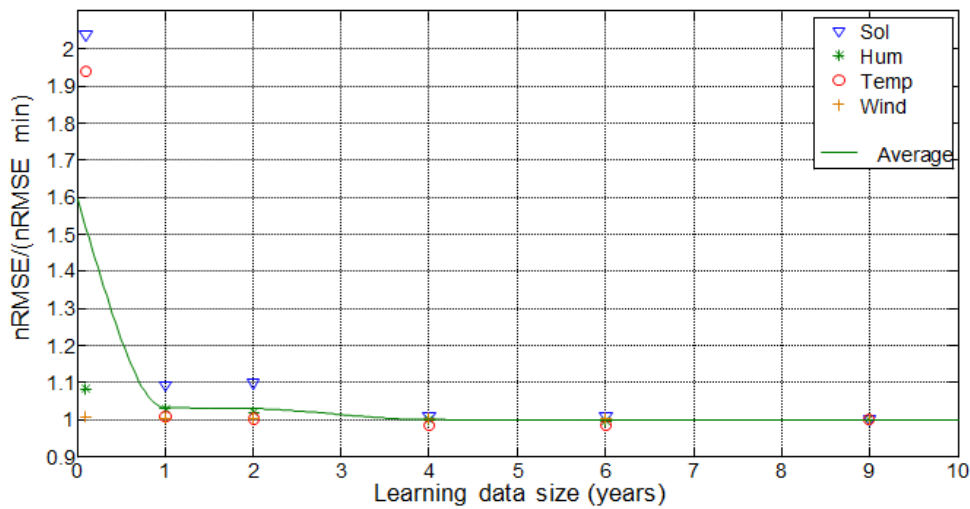


FIGURE 1: Impact of the learning size on the MLP performance

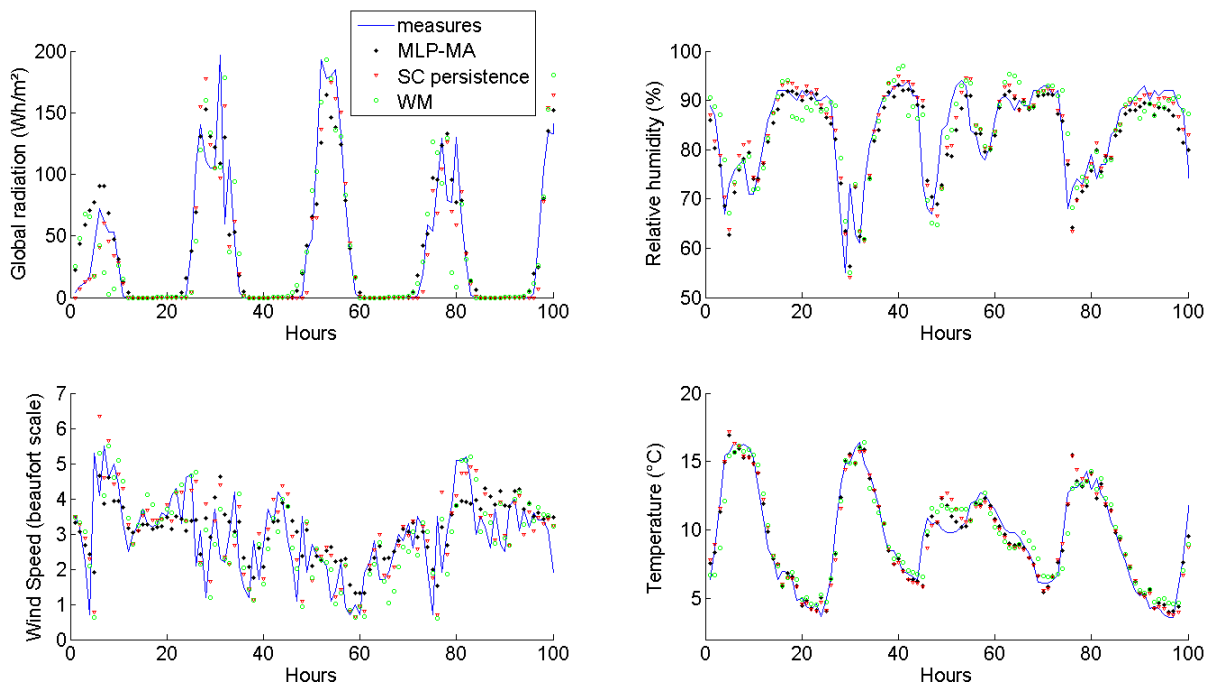


FIGURE 2: Measurement versus prediction

	P	WM	A	SP	MLP	MA-MLP
Sol	0.5073	0.3591	0.5743	0.3186	0.3119	0.3341
Hum	0.0767	0.0814	0.1571	0.0698	0.0733	0.0700
Temp	0.0816	0.0671	0.1881	0.0584	0.0679	0.0583
Wind	0.3943	0.4029	0.501	0.3935	0.3624	0.3517

TABLE 1: nRMSE of the six predictors concerning two years of predictions. In bold the best results

	P	WM	A	SP	MLP	MA-MLP
Sol	0.3022	0.1929	0.3106	0.1552	0.1629	0.1933
Hum	0.0522	0.0602	0.1209	0.0480	0.0502	0.0480
Temp	0.0575	0.0506	0.1478	0.0415	0.0462	0.0402
Wind	0.2819	0.2941	0.3545	0.2808	0.2618	0.2518

TABLE 2: nMAE of the six predictors concerning two years of predictions. In bold the best results

	P	WM	A	SP	MLP	MA-MLP
Sol	19.54/14.7	4.72/3.77	26.24/15.54	0.67/0	0/0.77	2.22/3.81
Hum	0.69/0.42	1.16/1.22	8.73/7.29	0/0	0.35/0.22	0.02/0
Temp	2.33/1.73	0.88/1.04	12.98/10.76	0.01/0.13	0.96/0.6	0/0
Wind	4.26/3.01	5.12/4.23	14.93/10.27	4.18/2.90	1.07/1	0/0

TABLE 3: Difference between all the predictors and the best predictor of each TS in percentage point. In each cells, the first result concerns the nRMSE (%) results and the second the nMAE (%)

	P	WM	A	SP	MLP	MA-MLP
Sol	<i>4%-8%</i>	<i>9%-17%</i>	<i>6%-11%</i>	<i>14%-20%</i>	15%-23%	<i>8%-16%</i>
Hum	<i>63%-82%</i>	<i>55%-81%</i>	<i>27%-53%</i>	<i>66%-85%</i>	<i>64%-84%</i>	66%-86%
Temp	<i>57%-79%</i>	<i>58%-82%</i>	<i>23%-43%</i>	<i>67%-87%</i>	<i>64%-84%</i>	68%-88%
Wind	16%-27%	<i>13%-24%</i>	<i>11%-22%</i>	<i>14%-26%</i>	<i>14%-28%</i>	15%-28%

TABLE 4: Percentages of good estimations (accuracy better than 5% and then 10%). In bold the best results and in italic the worst results

	predictor	Winter	Spring	Summer	Autumn
Sol	MA-MLP	56.37	35.83	25.71	29.28
	WM	41.10	28.34	32.24	47.23
Hum	MA-MLP	13.69	9.45	6.09	6.07
	WM	8.32	7.73	7.41	8.34
Temp	MA-MLP	6.58	9.73	5.27	5.54
	WM	6.59	6.49	6.82	6.61
Wind	MA-MLP	41.05	33.95	30.29	36.84
	WM	40.99	36.19	40.15	39.75

TABLE 5: Seasonal nRMSE for MA-MLP and WM. In bold the lowest nRMSE values.