



**HAL**  
open science

## Solar irradiation nowcasting by stochastic persistence: a new parsimonious, simple and efficient forecasting tool

Cyril Voyant, Gilles Notton

### ► To cite this version:

Cyril Voyant, Gilles Notton. Solar irradiation nowcasting by stochastic persistence: a new parsimonious, simple and efficient forecasting tool. *Renewable and Sustainable Energy Reviews*, 2018, 92, pp.343-352. hal-01783937

**HAL Id: hal-01783937**

**<https://hal.science/hal-01783937>**

Submitted on 2 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Solar irradiation nowcasting by stochastic persistence: a new parsimonious, simple and efficient**  
2 **forecasting tool**  
3

4 Cyril Voyant<sup>1,2</sup>, Gilles Notton<sup>1</sup>

5 1- University of Corsica/CNRS UMR SPE 6134, Campus Grimaldi, 20250 Corte – France

6 2- CHD Castelluccio, radiophysics unit, B.P85 20177 Ajaccio- France  
7

8 \*corresponding author; tel +33495293666, fax +33495293797, [voyant@univ-corse.fr](mailto:voyant@univ-corse.fr)  
9

10 **Abstract**  
11

12 Simple, naïve, smart or clearness persistences are tools largely used as naïve predictors for the global  
13 solar irradiation forecasting. It is essential to compare the performances of sophisticated prediction  
14 approaches with that of a reference approach generally a naïve methods. In this paper, a new kind of  
15 naïve “nowcaster” is developed, a persistence model based on the stochastic aspect of measured solar  
16 energy signal denoted stochastic persistence and constructed without needing a large collection of  
17 historical data. Two versions are proposed: one based on an additive and one on a multiplicative  
18 scheme; a theoretical description and an experimental validation based on measurements realized in  
19 Ajaccio (France) and Tilos (Greece) are exposed. The results show that this approach is efficient, easy  
20 to implement and does not need historical data as the machine learning methods usually employed.  
21 This new solar irradiation predictor could become an interesting tool and become a new member of the  
22 solar forecasting family.

23  
24 **Keywords:** prediction, machine learning, forecasting, persistence, bias-variance

## 25 1. Introduction

26

### 27 1.1. Interest of solar irradiation forecasting

28

29 Over the last ten years, energy market was boosted with the advent of renewable energies and in  
30 particular thanks to solar energy. The main interest of this kind of primary energy is to be easily and  
31 cleanly transformed into electricity particularly via photovoltaic conversion [1], which is the most  
32 flexible form of energy [2]. The main problem concerning the use of solar energy is its continuous  
33 variability relating both to time and space [3,4]. The variability can be divided into two components,  
34 the first one denoted deterministic part and the second one stochastic or random part. If the  
35 deterministic component is generated by the movements of rotation and revolution of the Earth [5], the  
36 stochastic component is generated by weather and cloud occurrences [6]. Solar energy intermittency  
37 has a great influence on the output power of photovoltaic (*PV*) plants, which can fluctuate significantly  
38 in short intervals (related to the random part) and in long intervals (related to daily and yearly seasonal  
39 effects) [7]. This no-controllable intermittence has negative consequences on the management of the  
40 electrical distribution and stability (forcing to limit the penetration rate of such intermittent energy  
41 systems) and on the kWh production costs [8]. One way to solve or to reduce this problem is to  
42 forecast this *PV* output power [9]. A good forecast helps the grid manager to plan the other energy  
43 capabilities to compensate for the *PV* plants power variations [10]. The forecasting quality of the  
44 output *PV* plant is strongly linked to the global horizontal irradiation (*GHI*) forecasting accuracy [11].  
45 Some authors go even further and consider the problem of *PV* output power forecasting and the solar  
46 irradiance forecasting problem as equal [12]. In this paper, a new forecasting tool is developed and  
47 tested in view to assist the electrical grid manager by predicting easily *GHI*.

### 48 1.2. Prediction and Parsimony

49

50 Time series forecasting [13] consists to estimate possible events or their evolutions by using as tools  
51 the past and the present. Before exposing the deferent tools available in order to nowcast *GHI*, it is  
52 important to define the “time series” term and the word “prediction” related to this kind of  
53 mathematical tools [14,15].

54 **Definition 1.1.** Time series: A univariate time series is a sequence of measurements of the same  
55 variable collected over time. Most often, the measurements are made at regular time intervals. The  
56 common notation concerning a time series of *GHI* measurement is  $GHI = \{GHI(t): t \in T\}$  where  $T$  is  
57 the index set.

58 **Remark 1.1.** *GHI* (nondeterministic) time series may be analyzed by assuming they are partly the  
59 manifestations of stochastic (random) processes [16–18] which is a statistical phenomenon consisting  
60 of a collection of random variables ordered in time and evolving according to a priori unknown  
61 probabilistic laws.

62 To succeed a time series prediction, only four conditions must be fulfilled:

- 63 - a certain regularity in the functioning of the studied process,
- 64 - this regularity must provide information on the future,
- 65 - the method chosen to establish the prediction captures a part of this regularity,
- 66 - the prediction will be efficient if and only if the "**noise**" or past irregularities are excluded as  
67 far as possible.

68 Forecasting the solar irradiation from 1 hour to 6 hours (defining the nowcasting [12]) is currently  
69 done using statistical or machine learning methods coupled to time series analysis. Many papers show  
70 that these methods yield similar results [19–21], none appears to outperform other and sometimes  
71 simple methods propose very similar results. According to a review analysis [22], it seems that it is not  
72 interesting to predict with very complex methods because a gain of tenths of a percent on the  
73 forecasting performances has only a small (but not negligible) impact on the grid management.  
74 Moreover, in [8], authors model a fictive solar plant with a nominal capacity of 1000 kW and show  
75 that a large nRMSE reduction from 32% to 28% (-4 percentage points) allows a financial saving close  
76 to 9%, so 70€ per day for the considered installation. In fact, the electricity grid operator needs a  
77 reliable tool which is adaptable for all horizons (between 5 minutes and 6 hours). The ideal case is to  
78 elaborate a tool which does not require a large learning history [15] in order to be quickly deployed on  
79 any site. In this paper we propose a new very simple and parsimonious tool based on the persistence of  
80 stochastic signal. Note that if in the operational case, the prediction with persistence does not need a  
81 large historical data (only a few hours), the present study is a retrospective comparison and is operated  
82 with historical data. The idea behind parsimonious models stems from the 14th century and the  
83 formulation of the Occam's razor [23] stating that "we should use no more parameters than necessary  
84 to explain the model well. There is generally a tradeoff between goodness of fit and parsimony.  
85 Models with many parameters (as machine learning tools [24,25]) tend to have a better fit than high  
86 parsimony models (as persistence), however this is not usually a good thing. Indeed, adding more  
87 parameters usually results in a good model fit for the data at hand, but that same model will likely be  
88 useless for predicting other data sets. In [26] (pp. 103-104), sentences summarize the interpretation  
89 related to simple models results: «Sometimes a simple model will outperform a more complex model .  
90 . . . Nevertheless, I believe that deliberately limiting the complexity of the model is not fruitful when the  
91 problem is evidently complex. Instead, if a simple model is found that outperforms some particular  
92 complex model, the appropriate response is to define a different complex model that captures whatever

93 aspect of the problem led to the simple model performing well». It is essential to correctly study the  
94 simple models before to elaborate more sophisticated approaches. Reference models should be well  
95 chosen to truly and objectively decide on the quality of the forecast.

96

## 97 **2. Machine learning or simple models of persistence**

98

99 Machine learning [27] is a branch of artificial intelligence [28]. It concerns the construction and the  
100 study of systems that can learn from data sets, giving to computers the ability to learn without being  
101 explicitly programmed.

102

### 103 **2.1. Models definitions**

104 With the machine learning tools based predictions, the system is built from a random output (denoted  
105 variable  $y$ ) and a set of random input (denoted variables  $x = \{x_1 \dots, x_n\}$ ). Using a learning sample  
106  $\{y_i, x_i\}_1^N$  of known values of pairs  $(y, x)$ , the aim is to obtain and estimate a model function  $f^*(x)$ ,  
107 among all the functions  $f(x)$  available and which allows to map (as well as possible!)  $x$  to  $y$ . The  
108 objective is reached after an optimization of the expected value ( $\mathbb{E}$ ) of some specified loss functions  
109  $L(y, f(x))$  over the joint distribution of all  $(y, x)$  pairs:

$$110 \quad f^*(x) = \operatorname{argmin}_f(\mathbb{E}(L(y, f(x)))) \quad \text{Equation 1}$$

111 In a regression problem, the loss function  $L(y, f(x))$  includes usually 2-norm or 1-norm distances  
112 respectively computed from the squared-error  $(y - f(x))^2$  (Euclidean norm giving more importance  
113 to large deviations or outliers) and the absolute error  $|y - f(x)|$  (absolute-value norm giving  
114 importance to the trend gap). Typically in the supervised cases, the machine learning methods are  
115 confronted to bias-variance tradeoff and are very user dependent and difficult to make a good use [29].  
116 Is machine learning is overhyped? This question was recently asked in [30], it may be time to consider  
117 other methods of modeling. The simplest method of forecasting the weather, persistence, relies upon  
118 today's conditions to forecast the conditions tomorrow. This can be a valid way of forecasting the  
119 weather when it is in a steady state, such as during the summer season when clouds are rare. This  
120 method of forecasting strongly depends upon the presence of a stagnant weather pattern. Therefore,  
121 with a fluctuating weather pattern, this method of forecasting becomes inaccurate. It can be useful in  
122 both short range forecasts and long range forecasts. The time series of global horizontal irradiation  
123 ( $GHI$ ) is composed by a stochastic part (Cf previous section); often when a machine learning method  
124 is used, a strong condition is necessary: the stationarity of the input data [31]. That means that the joint

125 distribution of  $GHI(t)$  and  $GHI(t+h)$  does not depend on  $t$  but only on  $h$  ( $t, h \in \mathbb{N}^*$ ). To our  
 126 knowledge, it is not proved that the tools used to make the  $GHI$  time series stationary allows to  
 127 correctly respect this condition [32]. It is legitimate to ask: can we really use these methods even if the  
 128 results are consistent ? we have of course not the answer and we would be very embarrassed to answer  
 129 "no" to this question given that we ourselves abundantly study the forecast of  $GHI$  via the data driven,  
 130 machine learning, artificial intelligence and others statistical methods. What is sure is that with the  
 131 persistence there are both advantages: directly usable (without learning and without need of historical  
 132 data) and any hypotheses or conditions concerning the model building. The "classical" persistence is  
 133 not really adapted to the forecast [29] while the smart persistence (integrating a knowledge-based  
 134 model using a clear sky model taking into account the sun position and the average conditions of sky  
 135 state) allows to greatly improve the prediction [19].

136 **Definition 2.1.** Simple persistence: the term persistence (or simple persistence) in time series context  
 137 is related to the notion of memory properties of time series, the model is built for the horizon (look-  
 138 ahead time)  $h$  as  $GHI(t+h) = GHI(t) + \varepsilon_{t+h}$ , where  $t$  is a time index and  $\varepsilon$  denotes the residual. The  
 139 forecast  $\widehat{GHI}$  obtained with this model is  $\widehat{GHI}(t+h) = GHI(t)$ , which states that the expected value  
 140 at horizon  $h$  is equal to the most recent measured value.

141 **Definition 2.2.** Smart persistence: This model is based on the same assumption than persistence model  
 142 but is corrected for the deterministic diurnal variation in solar irradiance, using a knowledge-based  
 143 model  $K(t, h)$ :  $\widehat{GHI}(t+h) = GHI(t).K(t, h)$

## 144 2.2. A short literature review on persistence

145 Numerous studies show the efficiency of these naïve predictions: the persistence. In [33] the  
 146 persistence is extremely detailed and authors wrote «It has been found that for short time horizons,  
 147 beating persistence models is a difficult task » and demonstrated that, often, the persistence is the best  
 148 method to use for the short-casting (<1h) and the now-casting (1h-6h). In several studies, the simple  
 149 persistence allows obtaining very good results [34- for which the difference in term of prediction error,  
 150 compared with machine learning method is lower than 2.5% [35] and in [36] lower than 5%.  
 151 Concerning the comparison between machine learning and smart (or clearness) persistence, this  
 152 difference is even lower, [19] and reach 2% and the authors wrote "for hour ahead solar forecasting,  
 153 the picture is less clear and seems to depend on the sky conditions". For stable clear sky conditions  
 154 (clear skies for instance), the nonlinear methods slightly improve the scaled-persistence. For unstable  
 155 sky conditions, the discrepancy between the machine learning methods and the simple models is more  
 156 pronounced with a 2%  $nRMSE$  difference in average. In [37] and [38] authors showed that the smart  
 157 persistence is a good predictor compared to more complicated methods with an increasing of  $nRMSE$   
 158 of 1%. In [39] and [40] the persistence is sometime as efficient as sophisticated models while in [35] it

159 is better than support vector machine. Note that the smart persistence use depends on the clear sky  
 160 model use as described in [41]. In conclusion, it appears that the persistence should be an interesting  
 161 forecasting. However, keep in mind that the atmospheric dynamics has major importance, and cannot  
 162 be dismissed from the predictors without affecting their performance, especially when the prediction  
 163 time horizon is larger than 1 hour. So, in theory, this kind of prediction based on the persistence of the  
 164 phenomenon is dedicated to the very short horizons and will never be as powerful as models based on  
 165 atmospheric dynamics.

166

### 167 3. Stochastic persistence formalisms

168

169 As all techniques for estimating derivatives of a noisy signal, persistence suffers from a high  
 170 sensitivity to noise (or quick fluctuations). To quantify the noise related to a time series, it is common  
 171 to estimate the Signal Noise Ratio (*SNR*) defined by the ratio between the average of the signal and the  
 172 noise (standard deviation of the time series). It is a multiplicative inverse of the variation coefficient  
 173 [42]. On one year and for an hourly time granularity in Tilos (Greece, 1 hour horizon), *SNR* varies  
 174 between 0.8 and 0.6 respectively in summer and winter. When this parameter is high the persistence or  
 175 smart persistence gives very good results (in summer  $nRMSE=8.7\%$  for smart persistence) but  
 176 becomes less interesting and efficient when *SNR* decreases (in winter  $nRMSE =17.4\%$  for smart  
 177 persistence). In this paper, we propose to modify the persistence estimation considering the fact that  
 178 the studied series are noisy time series and thus the stochastic aspect of the measured signal will be  
 179 taken into account. Note that, as *SNR* varies, the variability varies also, so it is very complicated for a  
 180 machine learning method to take into account all these characteristics without considered additional  
 181 informations such as exogenous data or dummy temporal variables. The simple persistence is  
 182 described in the definition 2.1, for the definition of the smart persistence, the function  $K(t, h)$  (in the  
 183 definition 2.2) is usually defined as the ratio  $CS(t)$  of the solar radiation at the ground level on the  
 184 estimated clear sky solar radiation ( $CS(t)$  is computed using the well-known Solis model [43,44]) as  
 185 described by :

$$186 \quad \widehat{GHI}(t + h) = GHI(t) \cdot \frac{CS(t+h)}{CS(t)} \quad \text{Equation 2}$$

187 In fact, this reference predictor is built from the persistence of the clear sky index ( $CSI(t) =$   
 188  $GHI(t)/CS(t)$ ) and thus:

$$189 \quad \widehat{CSI}(t + h) = CSI(t) \quad \text{Equation 3}$$

190 We notice that one naturally transform the ratio to trend to an additive model by passing to the log (Cf  
 191 Box-Cox transformation, logarithmic transformation is often necessary to stabilize the variance) with  
 192  $\log(\widehat{CSI}(t+h)) = \log(CSI(t))$ , hence although it is never used in solar irradiation forecasting,  
 193 another definition of the smart persistence could be:

$$194 \log(\widehat{GHI}(t+h)) = \log(GHI(t)) + \log(CS(t+h)) - \log(CS(t)) \quad \text{Equation 4}$$

195 As we will see later, the inclusion of atmospheric variables ( $CS$ ) into the prediction process for solar  
 196 radiation will improve its performance. Two definitions (arithmetic and geometric means) are  
 197 necessary to understand the next sections, either a  $GHI$  time series defined by  $\{GHI(t): t \in T\}$  with  
 198  $\max(T) > N$  and  $N \in \mathbb{N}^*$ :

199 **Definition 3.1.** Arithmetic mean at time  $t$  for the series  $x$  denoted  $\overline{x}_N(t) = \frac{1}{N} \sum_{i=0}^{N-1} x(t-i)$

200 **Definition 3.2.** Geometric mean at time  $t$  for the series  $x$  denoted  $\overline{x}_N^G(t) = \left(\prod_{i=0}^{N-1} x(t-i)\right)^{1/N}$

### 201 3.1. Additive scheme of the stochastic persistence

202 It is possible to define a time series as a sum of 2 other series referring to Wold's theorem (or also to  
 203 Cartier Perrin theorem) [45,46]. These theorems say that every covariance-stationary time series can  
 204 be written as the sum of two time series, one deterministic and one stochastic; in our case, we can  
 205 write :  $GHI(t) = T(t) + \sum_{j=0}^{\infty} b_j E(t-j)$  with  $E(t)$  is an uncorrelated sequence which is the  
 206 innovation process (or white noise) that is the input to the linear filter  $\{b_j\}$ .  $b$  is the possibly infinite  
 207 vector of moving average weights and  $T(t)$  is a deterministic time series. Note that the stochastic part  
 208 defines a new time series  $\varepsilon(t) = \sum_{j=0}^{\infty} b_j E(t-j)$  thus in the following, we consider the decomposition  
 209 form  $GHI(t) = T(t) + \varepsilon(t)$ . Theoretically, the  $\varepsilon(t)$  part is not a predictable quantity, all the available  
 210 prediction tools focus on the  $T(t)$  estimation; so concerning the persistence, it seems illogical to apply  
 211 and propagate a random term to the future. In this paper, we propose new definitions of the  
 212 persistence; the first one (Fig 1) is based on a model of knowledge (clear sky model;  $CS$ ) and on the  
 213 arithmetic mean of the difference between  $CS$  and the past  $GHI$  measurements  $\Delta(t) (=CS(t) -$   
 214  $GHI(t))$ .

$$215 \widehat{GHI}(t+h) = CS(t+h) - \overline{\Delta}_N(t) \quad \text{with} \quad \overline{\Delta}_N(t) = \frac{1}{N} \sum_{i=0}^{N-1} (CS(t-i) - GHI(t-i)) \quad \text{Equation 5}$$

216 This averaging allows minimizing the stochastic part of the measurement, it is a low pass filter  
 217 operated by moving average. From Eq (6) it comes  $(\overline{GHI}_N(t))$  and  $(\overline{CS}_N(t))$  arithmetic means of the  
 218  $GHI$ , and  $CS$ ):

$$219 \widehat{GHI}(t+h) = \overline{GHI}_N(t) + CS(t+h) - \overline{CS}_N(t) \quad \text{Equation 6}$$



220 This form of the persistence is equivalent to the reduced definition:

$$221 \hat{\Delta}(t+h) = \bar{\Delta}_N(t) \quad \text{Equation 7}$$

222

223

224

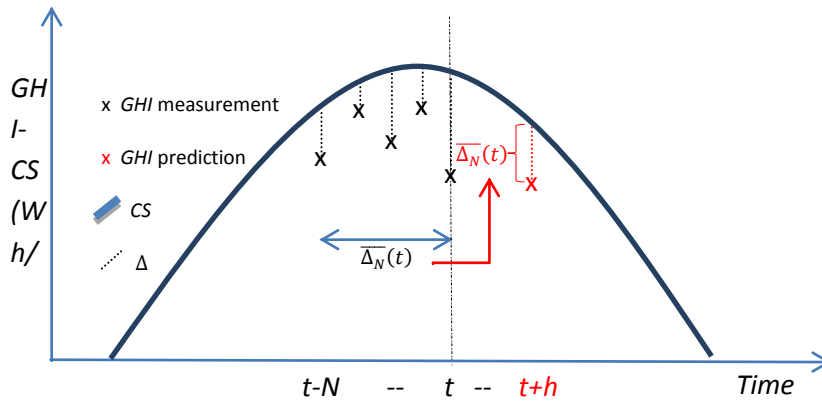
225

226

227

228

229



230

231

232

233

234

235 Figure 1. Principle of the stochastic persistence based on an additive scheme

236

### 237 3.2. Multiplicative scheme of the stochastic persistence

238 Based on the previous subsection, we define now the persistence as a multiplicative scheme (Fig 2)

239 where:

$$240 \widehat{GHI}(t+h) = CS(t+h) \cdot \overline{CSI}_N^G(t) \quad \text{Equation 8}$$

241  $\overline{CSI}_N^G(t)$  is the geometric mean of the ratio to trend denoted clear sky index  $CSI(t) (= \frac{GHI(t)}{CS(t)})$  and

242 computed with:

$$243 \overline{CSI}_N^G(t) = \left( \prod_{i=0}^{N-1} \frac{GHI(t-i)}{CS(t-i)} \right)^{1/N} \quad \text{Equation 9}$$

244

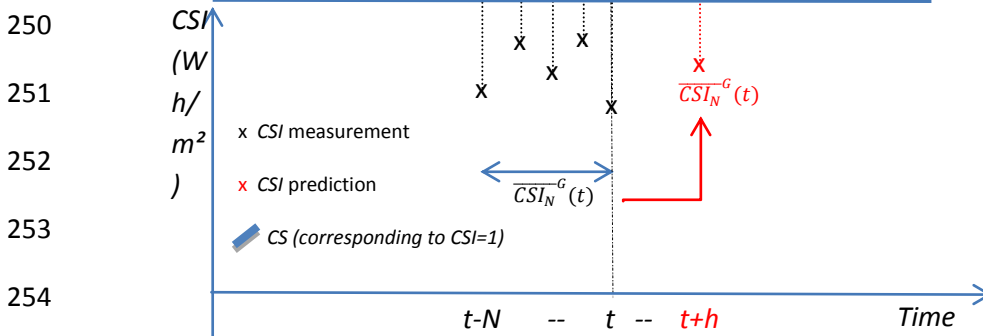
245 Thus, it comes:

246  $\widehat{GHI}(t+h) = \overline{GHI}_N^G(t) \cdot \frac{CS(t+h)}{\overline{CS}_N^G(t)}$  Equation 10

247

248

249



255 Figure 2. Principle of the stochastic persistence based on a multiplicative scheme

256

257 A similar equation (equation 11) to the smart persistence case is obtained. It is a generalized formula  
 258 of the CSI persistence minimizing the random part of the measurement. If  $N=1$  the equation is  
 259 equivalent to the classical smart persistence equation. The reduced form is:

260  $\widehat{CSI}(t+h) = \overline{CSI}_N^G(t)$  Equation 11

261

### 262 3.3. Stochastic persistence optimization

263 In Table 1 are summarized all the models defined in the previous sections.

<b>Definitions</b>	
<b>Simple persistence (P)</b>	$\widehat{GHI}(t+h) = GHI(t)$
<b>Smart persistence (SP)</b>	$\widehat{CSI}(t+h) = CSI(t)$ or $\widehat{GHI}(t+h) = GHI(t) \cdot \frac{CS(t+h)}{CS(t)}$
<b>Stochastic persistence : additive mode (StP<sup>+</sup>)</b>	$\widehat{\Delta}(t+h) = \bar{\Delta}_N(t)$ or $\widehat{GHI}(t+h) = \overline{GHI}_N(t) + CS(t+h) - \overline{CS}_N(t)$
<b>Stochastic persistence : multiplicative mode (StP<sup>x</sup>)</b>	$\widehat{CSI}(t+h) = \overline{CSI}_N^G(t)$ or $\widehat{GHI}(t+h) = \overline{GHI}_N^G(t) \cdot \frac{CS(t+h)}{\overline{CS}_N^G(t)}$

264 Table 1. Short summary concerning the persistence models.

265 In the additive and multiplicative cases, the optimal solution consists in quantifying  $N$  which  
 266 minimizes the prediction error (mean square error;  $MSE$ ), then:

$$267 \operatorname{argmin}\{MSE(\widehat{GHI}(t+h) - GHI(t+h) | N \in \mathbb{N}^*)\} \quad \text{Equation 12}$$

268 This is a simple optimization problem that does not require the use of optimization algorithm; the  
 269 exhaustive (or brute-force) search is easily and quickly achievable (in practice  $N \in [1:100]$ )

### 270 3.4. Theoretical validation of the stochastic persistence

271 In order to understand the stochastic persistence interest, it is necessary to remind the  $MSE$   
 272 decomposition (variance and bias [47]) as described below:

$$273 MSE = Var(\widehat{GHI}(t+h)) + \left( Bias(\widehat{GHI}(t+h), GHI(t+h)) \right)^2 \quad \text{Equation 13}$$

$$274 \text{With } Var(\widehat{GHI}(t+h)) = \mathbb{E} \left[ (\widehat{GHI}(t+h) - \mathbb{E}[\widehat{GHI}(t+h)])^2 \right] \quad \text{Equation 14}$$

$$275 \text{And } \left( Bias(\widehat{GHI}(t+h), GHI(t+h)) \right)^2 = \mathbb{E} \left[ (\mathbb{E}[\widehat{GHI}(t+h)] - GHI(t+h))^2 \right] \quad \text{Equation 15}$$

276

277 Whether one uses the smart persistence or one of the two stochastic persistences described previously,  
 278 the biases are similar; indeed,  $\mathbb{E}[\widehat{GHI}(t+h)]$  is identical in the three cases, only the variance changes  
 279 and is directly impacted by the mean of  $CSI$  or  $\Delta$ . In conclusion, the variance part and so  $MSE$   
 280 decreases with the stochastic persistence use. A validation can be operated considering two  
 281 components of the measured  $GHI$  signal (mean and noise). Note that here the determinist part ( $T$ ) is  
 282 not the  $CS$  described previously but only an average value (or trend) and the noise a random variable  
 283 ( $\varepsilon$ ). So with the condition  $T(t) + \varepsilon(t) < CS(t)$ , we obtain:

$$284 GHI(t) = T(t) + \varepsilon(t) \text{ with } t \in \mathbb{N}^*, T \in \mathcal{R}^+ \text{ and } \varepsilon \in \mathcal{R} \quad \text{Equation 16}$$

285 In the additive case,  $\mathbb{E}[\varepsilon(t)] = 0 \forall t$  according to the white noise definition around a signal. Moreover  
 286 with the stochastic persistence, there will be a persistence of the trend but not of the noise, so:  
 287  $\lim_{N \rightarrow \infty} \hat{\varepsilon}(t+h) = 0 \forall t, h$ . For the smart persistence ( $SP$ ) and the stochastic persistences ( $StP^+$  and  
 288  $StP^x$ ), we obtain three important results:

$$289 \textbf{Result 1: } Var_{StP}(\widehat{GHI}(t+h)) < Var_{SP}(\widehat{GHI}(t+h)) \quad \text{Equation 17}$$

$$290 \textbf{Result 2 : } Bias_{SP}(\widehat{GHI}(t+h), GHI(t+h)) = Bias_{StP}(\widehat{GHI}(t+h), GHI(t+h)) \quad \text{Equation 18}$$

$$291 \textbf{Result 3 : } nRMSE_{StP}(\widehat{GHI}(t+h)) < nRMSE_{SP}(\widehat{GHI}(t+h)) \quad \text{Equation 20}$$

292 The proofs of these three results are available in the **annex** of this paper. The most important result  
293 (the third) induces that the stochastic persistence improves (at least in theory) the prediction done with  
294 the smart persistence or the simple persistence.

295

## 296 **4. Results**

297

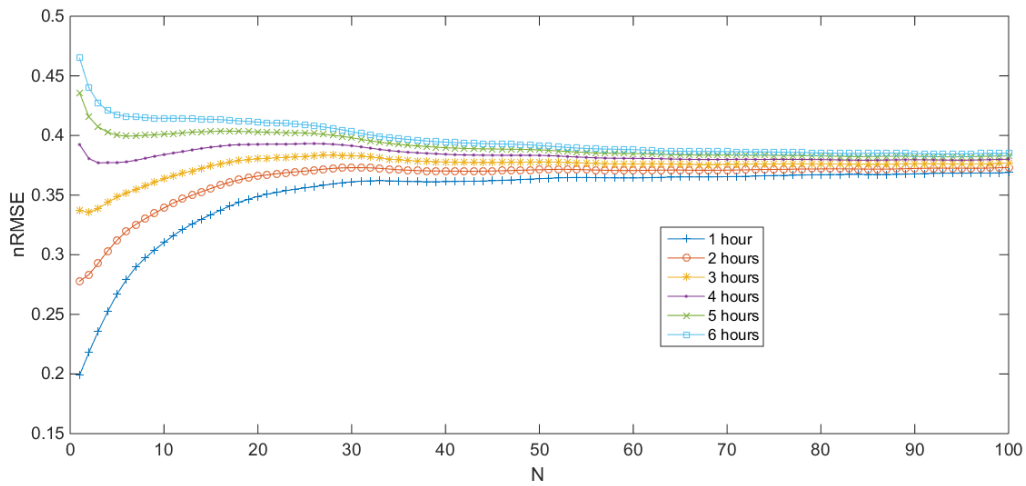
298 The forecasting of *GHI* needs usually a cleaning and a preparation of the dataset. Mistakes often  
299 appear in the temporal series of solar data due to problems with the acquisition system; an automatic  
300 quality check used in the frame of GEOSS project (Group on Earth Observation System of System)  
301 [48] has been applied to the data. The process to estimate the quality of the data [49] and the procedure  
302 applied to flag suspicious or erroneous measurements is described in detail in [50]. Then, we applied a  
303 filter on the datasets which remove all the data that correspond to a solar elevation angle lower than  
304  $10^\circ$ , in order to removing the night hours [19]. In this section, we will compare the prediction results  
305 of stochastic persistence with those related to two well know machine learning tools: *ARMA* (more  
306 precisely *AR* with *MA* part) and *MLP*. The training and optimization phases of these models are classic  
307 and interesting reader could find all the methodologies of prediction in [14,19]. In order to objectively  
308 compare the results, we propose the k-fold sampling use [51], the dataset is divided in ten samples  
309 (each with 80% of the total available data) and every sample is used at least one time for the training  
310 (only for the machine learning tools and not for the different persistences) and one time for the test  
311 (20% of the data and use for all models). This method induces to avoid the problems which can results  
312 of measurements of the dataset. The datasets used in this study are time series of measurements of  
313 global horizontal solar irradiation (*GHI*) in two different sites with different meteorological situations.  
314 The first dataset is provided by the station of Ajaccio from 1998 to 2009 (Corsica, France,  $41^\circ55$  N,  
315  $8^\circ44$  E, 4m asl), it is located near the Mediterranean Sea (100 m) and nearby mountains (1000 m  
316 altitude at 40 km from the site). This specific geographical configuration and the island context make  
317 the nebulosity difficult to forecast. The Mediterranean climate is characterized by hot summers with  
318 abundant sunshine and mild, dry and clear wintersand. The second one is constituted by measurements  
319 in Tilos from 2015 to 2016 (Tilos Island, Greece,  $36^\circ24$  N,  $27^\circ22$  E, 96m asl), which is a small island  
320 in the Dodecanese archipelago, the tallest mountain is about 650 m high and the cloud occurrences are  
321 much less frequent than in Ajaccio. These stations are equipped with pyranometers (CM 11 Kipp &  
322 Zonen) and standard meteorological sensors (pressure, temperature, etc.), the solar data are measured  
323 and stored with a time step equal to 1 min.

### 324 **4.1. Hourly time granularity in Ajaccio**

325

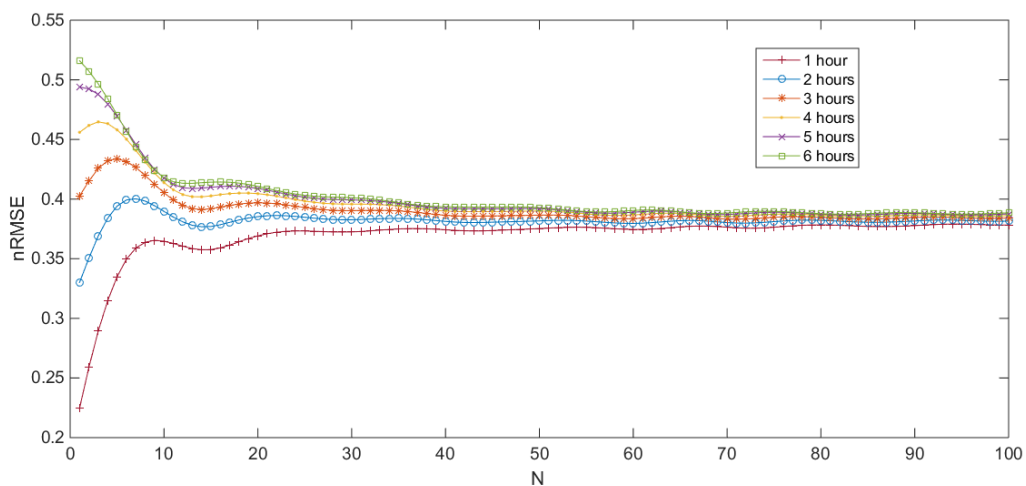
326 In this subsection, the data of Ajaccio are used. The tested horizons are from 1 hour to 6 hours. Before  
 327 to expose the results of prediction, we propose to develop the optimization step of the  $StP^+$  and  $StP^x$   
 328 formalisms. In Figs 3 and 4 are represented the prediction errors in term of size of sliding windows.  
 329 The optimization concerns the  $N$  parameter described in Eq 12. The optimized models are related to  $N$   
 330 giving the lowest value of  $nRMSE$ . For example, in the first figure and concerning the horizon 1 hour,  
 331 the  $StP^x$  constructed with  $N=1$ , gives the best performance. In this case, the  $StP^x$  is equivalent to  
 332 classical  $SP$  estimator (see Eq 11). We remark that the higher the horizon is, the higher the optimized  
 333  $N$  value is in the two cases  $StP^x$  and  $StP^+$ , but also that a known conclusion is verified: the prediction  
 334 error increases with the horizon. The additive mode is less reliable than the multiplicative mode with  
 335 the hourly time granularity. This kind of stochastic time series seems follow a multiplicative scheme.

336



337

338 Figure 3. Prediction error (from horizon 1 hours to 6 hours in Ajaccio) in term of size of the sliding  
 339 window concerning the  $StP^x$  (Eq. 12)



340

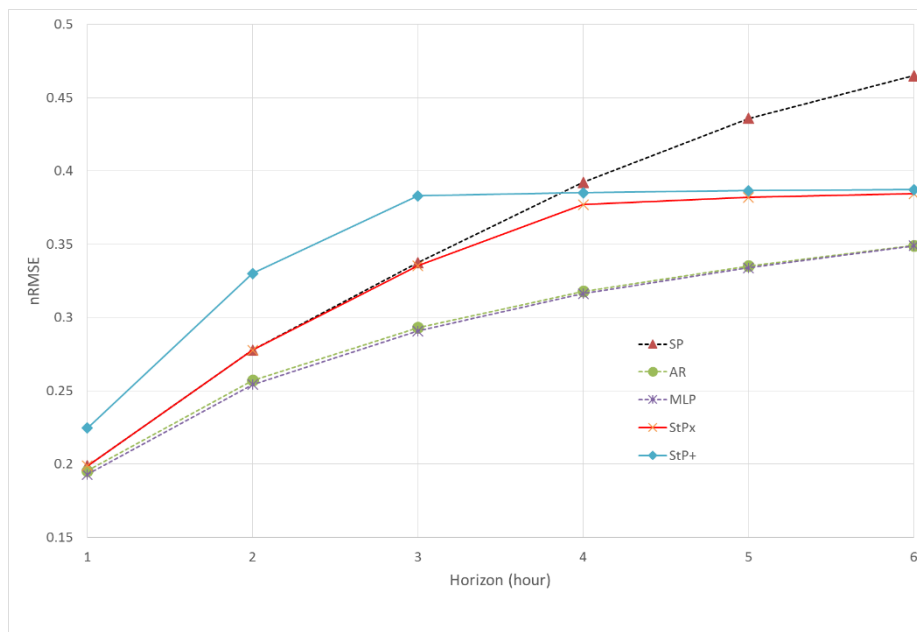
341 Figure 4. Prediction error (from horizon 1 hour to 6 hour in Ajaccio) in term of size of the sliding  
 342 window concerning the  $StP^+$  (Eq. 12)

343 The best configurations for each horizon are given in the Table 2.

	1 hour	2 hours	3 hours	4 hours	5 hours	6 hours
$StP^+$	1	1	71	84	83	82
$StP^x$	1	1	2	3	94	93

344 Table 2. Value of the optimized  $N$  for each horizon (Ajaccio)

345 Now the stochastic persistences are optimized, it is essential to compare the results of prediction with  
 346 classical methodologies ( $SP$ ,  $AR$  and  $MLP$ ). Fig 5 shows the errors related to all these models.



347

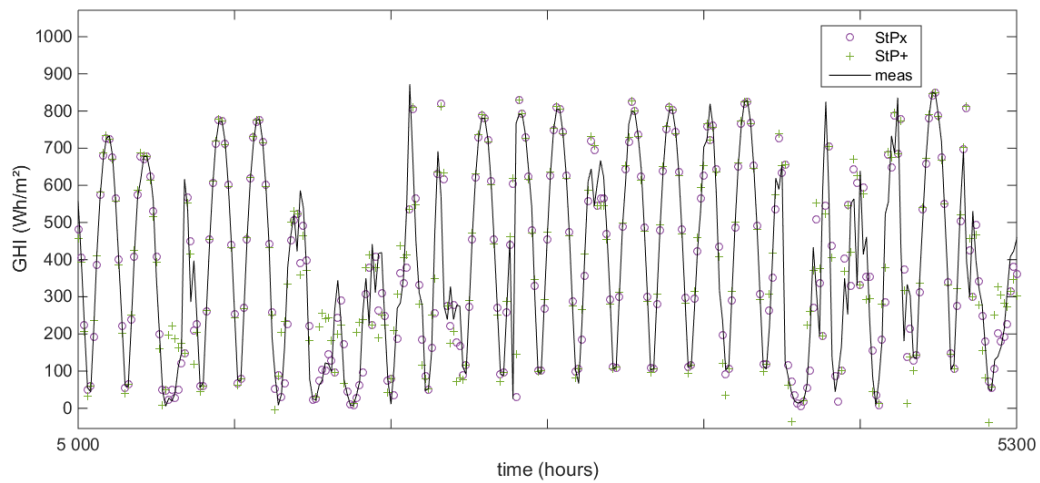
348 Figure 5.  $nRMSE$  Vs horizon for the 5 studied models (Ajaccio)

349 If the  $AR$  and  $MLP$  models give the best results, using  $StP^x$  allows to improve the prediction related to  
 350  $SP$ . The numerical values of  $nRMSE$  are given in Table 3. We note that  $StP^+$  and mainly  $StP^x$  represent  
 351 a high improvements compared to  $SP$  and  $P$  for long horizons (from  $h+4$  to  $h+6$ ) and reach a good  
 352 level of accuracy compared with  $AR$  and  $MLP$ .

horizons	$P$	$SP$	$AR$	$MLP$	$StP^+$	$StP^x$
1 hour	0.3442	0.1988	0.1954	0.1929	0.2246	0.1988
2 hours	0.5981	0.2778	0.2570	0.2543	0.3301	0.2778
3 hours	0.8061	0.3375	0.2931	0.2908	0.3830	0.3353
4 hours	0.9662	0.3923	0.3179	0.3163	0.3852	0.3770
5 hours	1.0741	0.4359	0.3352	0.3340	0.3865	0.3820
6 hours	1.1269	0.4650	0.3492	0.3490	0.3873	0.3844

353 Table 3.  $nRMSE$  for all models (Ajaccio)

354 It is now interesting to show profiles of prediction in order to visually verify the quality of the  
355 stochastic persistence. In Figure 6 is shown the 1 hour horizon case concerning  $StP^+$  and  $StP^x$ .



356

357 Figure 6. Profile of prediction related to  $StP^+$  and  $StP^x$  between the 5000<sup>th</sup> and 5300<sup>th</sup> hour (spring in  
358 Ajaccio)

359 We see a very good accordance between predictions and measures even when the variability is  
360 important. In the next subsection, we will verify the conclusion drawn here with another kind of time  
361 granularity and another location.

362

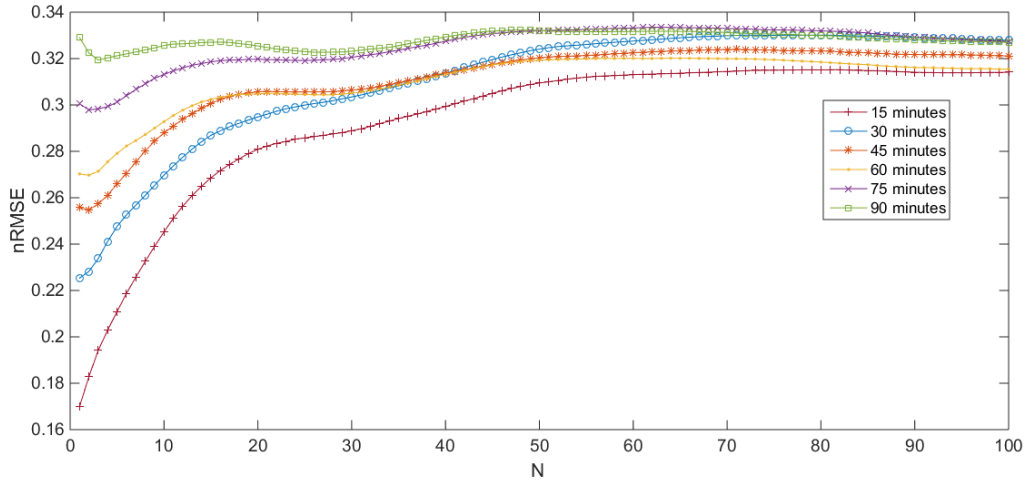
363

364

#### 365 4.2. 15 minutes time granularity in Tilos

366

367 As for Ajaccio using hourly data, for Tilos with 15 min data, the first step is to optimize the stochastic  
368 persistences using Figs 7 and 8.

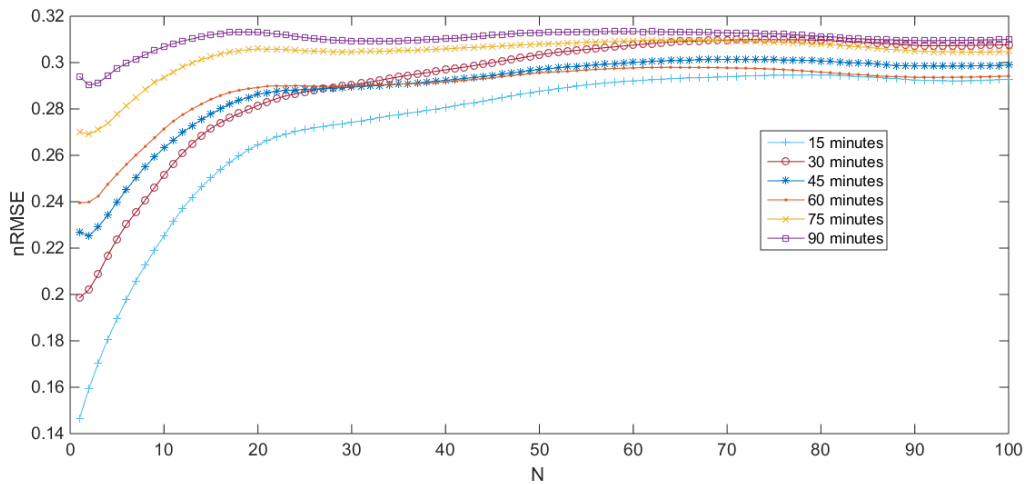


369

370

371

Figure 7. Prediction error (from horizon 15 minutes to 90 minutes in Tilos) in term of size of the sliding window concerning the  $StP^x$  (Eq 12)



372

373

374

Figure 8. Prediction error (from horizon 15 minutes to 90 minutes in Tilos) in term of size of the sliding window concerning the  $StP^+$  (Eq 12)

375

376

377

378

379

380

381

382

383

384

385

The conclusions are similar to the hourly case for Ajaccio, but here the additive mode seems more relevant than the multiplicative mode. In fact, we think this phenomenon is related to the clear sky estimation. In hourly case, it is less important to have a precise  $CS$  function (smoothing related to the hourly aggregation). The hourly sum tends to minimize the impact of the quality of the  $CS$  modeling. For lower time granularity, the consequence of using a good clear sky model becomes very important with the use of multiplicative mode and the division by  $CS$  (ratio to trend). Indeed, introducing briefly the condition number of a problem as tool measuring how the output value of the modeling can change for a small change in the input argument, we can certainly consider that our  $CS$  estimation is not efficient for the concerning problem. It is really difficult to improve the  $CS$  modeling because a lot of parameters (not always available) change during the year, the day and each hour. So as minimal conclusion, we can consider that the additive scheme is the most interesting when the time granularity

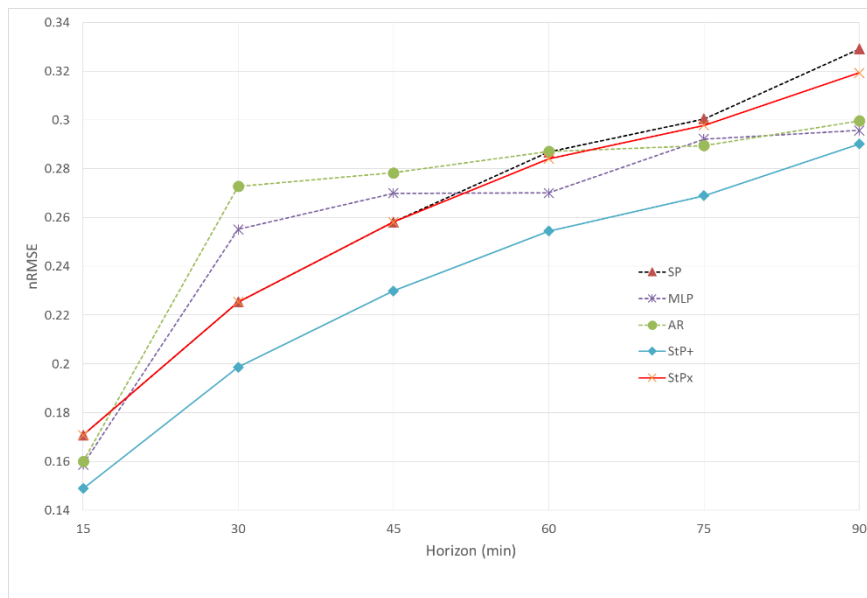


386 decreases. Concerning the multiplicative case, the *CS* estimation induces an ill-conditioned problem  
 387 not really performant. In Table 4 are listed the optimized value of *N* parameters.

	15 min	30 min	45 min	60 min	75 min	90 min
<i>StP</i> <sup>+</sup>	1	1	2	2	2	3
<i>StP</i> <sup>x</sup>	1	1	2	1	2	2

388 Table 4. Value of the optimized *N* for each horizon (Tilos)

389 Now the stochastic persistences are optimized, we can compare the prediction errors with the other  
 390 reference models (*SP*, *AR* and *MLP*). Figure 9 shows this comparison concerning 6 horizons from 15  
 391 minutes to 90 minutes.



392

393 Figure 9. *nRMSE* vs horizon for the 5 studied models (Tilos)

394 Bellow 60 minutes, the three persistences (*SP*, *StP*<sup>x</sup> and *StP*<sup>+</sup>) give better results than sophisticated  
 395 methods. The best model is, for all the horizon, *StP*<sup>+</sup>. The *nRMSE* values are given in Table 5.

396

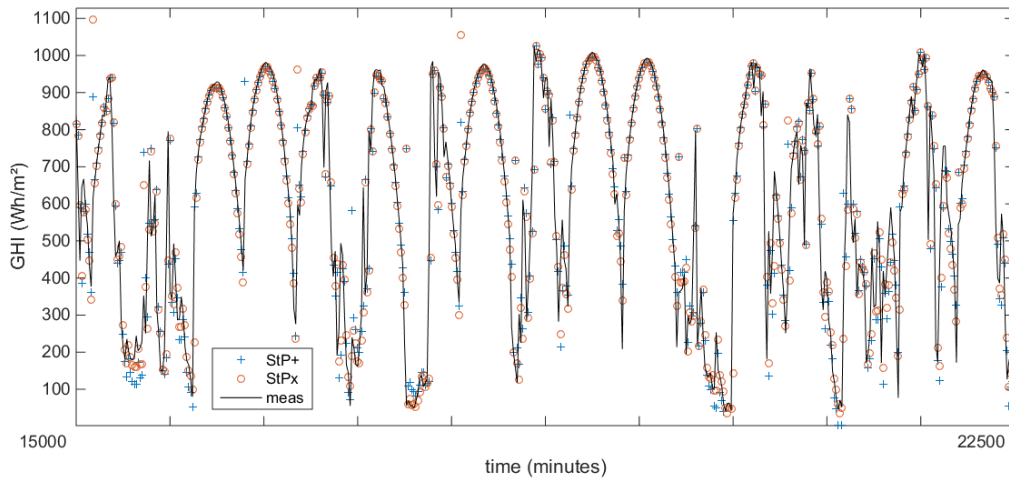
horizons	<i>P</i>	<i>SP</i>	<i>AR</i>	<i>MLP</i>	<i>StP</i> <sup>+</sup>	<i>StP</i> <sup>x</sup>
15 min	0.1929	0.1708	0.1600	0.1587	0.1489	0.1708
30 min	0.2804	0.2254	0.2728	0.2552	0.1986	0.2254
45 min	0.3419	0.2582	0.2784	0.2695	0.2299	0.2582
60 min	0.3988	0.2869	0.2872	0.2701	0.2545	0.2841
75 min	0.4528	0.3005	0.2896	0.2922	0.269	0.2979
90 min	0.5013	0.3291	0.2997	0.2957	0.2902	0.3194

397

Table 5. *nRMSE* for all models (Tilos)

398 We observe that the prediction is more reliable in the 15 minutes case than in the hourly case.  
399 Probably because there are some weather afterglow for very short duration. For this time granularity,  
400 the stochastic persistence model is the best whatever the time horizon is.

401 As for the hourly case in Ajaccio, in Fig 10, is presented a comparison between measured and  
402 forecasted values on a period of several days for  $StP^+$  and  $StP^x$  (15 minutes ahead).



403  
404 Figure 10. Profile of prediction related to  $StP^+$  and  $StP^x$  between the 15000<sup>th</sup> and 22500<sup>th</sup> minutes  
405 (spring in Tilos)

406 We observe a high accordance between measurements and predictions and that predicted values of  
407  $GHI$  by  $StP^+$  are better than ones predicted by  $StP^x$ .

## 408 409 5. Conclusion

410  
411 A new forecasting methodology was presented, it is based on the assumption that  $GHI$  signal has two  
412 components: a stochastic and a deterministic parts. Two stochastic methods were developed an  
413 additive and a multiplicative schemes. The stochastic persistence allows to easily establish  $GHI$   
414 prediction with a good accuracy without the need of large historical data collection.

415 The stochastic persistence was experimentally tested in two sites Ajaccio, Corsica, France and Tilos,  
416 Greece with two time granularities (1 hour and 15 min). It appeared that the results obtained by  
417 stochastic persistence model are systematically better than those obtained with classical or smart  
418 persistences; For 1 hour horizon, they are relatively close to those obtained with some sophisticated  
419 machine learning tools. For other prediction horizons and time granularities (15 minutes), the data  
420 driven methods are less interesting than the stochastic persistence in the additive mode.

421 In the hourly case, the stochastic persistence should be considered as naïve predictor in order to  
422 compare and valid more sophisticated methods of machine learning. Moreover, the reliability of the  
423 multiplicative stochastic persistence method, is not so very far from those obtained by sophisticated  
424 methods.

425 It would probably be interesting to construct error metric related to this tool, especially a new version  
426 of the skill score which is actually the most common parameter in the production of global radiation.

427 For the 15 minutes case, the stochastic persistence gives very good results mainly with the additive  
428 scheme. Some investigation related to clear sky modeling concerning the very short time granularity  
429 should be undertaken with the goal to improve the multiplicative scheme of the stochastic persistence  
430 and to valid the conclusions drawn here.

431 Thus the developed forecasted tool (with its two versions, additive and multiplicative) showed very  
432 good performances for a forecasting method that does not need a long and rare set of historical data  
433 and complicated training phase for nowcasting purpose.

434 **Annex**

435

436 **Result 1:**  $Var_{StP}(\widehat{GHI}(t+h)) < Var_{SP}(\widehat{GHI}(t+h))$

437 **Result 2:**  $Bias_{SP}(\widehat{GHI}(t+h), GHI(t+h)) = Bias_{StP}(\widehat{GHI}(t+h), GHI(t+h))$

438 **Result 3:**  $nRMSE_{StP}(\widehat{GHI}(t+h)) < nRMSE_{SP}(\widehat{GHI}(t+h))$

439 Proof of results 1

440 We have:

441  $Var_{SP}(\widehat{GHI}(t+h)) = Var_{SP}(\widehat{T}(t+h) + \hat{\varepsilon}(t+h))$

442 Related to the variance definition, we can write:

443  $Var_{SP}(\widehat{GHI}(t+h)) = \mathbb{E}[(\widehat{T}(t+h) + \hat{\varepsilon}(t+h) - \mathbb{E}[\widehat{T}(t+h) + \hat{\varepsilon}(t+h)])^2]$

444 If  $\hat{\varepsilon}(t+h)$  is a white noise, so:

445  $Var_{SP}(\widehat{GHI}(t+h)) = \mathbb{E}[(\widehat{T}(t+h) + \hat{\varepsilon}(t+h) - \mathbb{E}[\widehat{T}(t+h)])^2]$

446 Signifying that  $Var_{StP}(\widehat{GHI}(t+h)) = \mathbb{E}[(\widehat{T}(t+h) - \mathbb{E}[\widehat{T}(t+h)])^2] < Var_{SP}(\widehat{GHI}(t+h))$

447

448 Proof of results 2

449 Related to the bias definition, we have:

$$\left( Bias_{SP}(\widehat{GHI}(t+h), GHI(t+h)) \right)^2 = \mathbb{E} \left[ \left( \mathbb{E}[\widehat{T}(t+h) + \hat{\varepsilon}(t+h)] - T(t+h) + \varepsilon(t+h) \right)^2 \right]$$

450 If  $\hat{\varepsilon}(t+h)$  is a white noise, we also can consider:

451  $\left( Bias_{SP}(\widehat{GHI}(t+h), GHI(t+h)) \right)^2 = \mathbb{E} \left[ \left( \mathbb{E}[\widehat{T}(t+h)] - T(t+h) + \varepsilon(t+h) \right)^2 \right]$

452

453 So, we can write that:

454  $\left( Bias_{StP}(\widehat{GHI}(t+h), GHI(t+h)) \right)^2 = \left( Bias_{SP}(\widehat{GHI}(t+h), GHI(t+h)) \right)^2$

455 Proof of results 3

456 The results 1 and 2 lead to  $MSE_{StP}(\widehat{GHI}(t+h)) < MSE_{SP}(\widehat{GHI}(t+h))$ , according to the definition  
457 of the normalized root mean square error ( $nRMSE$ , []), we obtain:

458 
$$nRMSE_{StP}(\widehat{GHI}(t+h)) < nRMSE_{SP}(\widehat{GHI}(t+h))$$

459

460 **Bibliography**

461

462 [1] Paulescu M, Paulescu E, Gravila P, Badescu V. *Weather Modeling and Forecasting of PV*  
463 *Systems Operation*. London: Springer London; 2013.

464 [2] Azofra D, Martínez E, Jiménez E, Blanco J, Azofra F, Saenz-Díez JC. Comparison of the  
465 influence of photovoltaic and wind power on the Spanish electricity prices by means of artificial  
466 intelligence techniques. *Renew Sustain Energy Rev* 2015;42:532–42. doi:10.1016/j.rser.2014.10.048.

467 [3] Polo J, Gastón M, Vindel JM, Pagola I. Spatial variability and clustering of global solar  
468 irradiation in Vietnam from sunshine duration measurements. *Renew Sustain Energy Rev*  
469 2015;42:1326–34. doi:10.1016/j.rser.2014.11.014.

470 [4] Hoff TE, Perez R. Modeling PV fleet output variability. *Sol Energy* 2012;86:2177–89.  
471 doi:10.1016/j.solener.2011.11.005.

472 [5] Badescu V. *Modeling solar radiation at the earth's surface: recent advances*. Springer; 2008.

473 [6] Perez R, Kivalov S, Schlemmer J, Hemker Jr. K, Hoff TE. Short-term irradiance variability:  
474 Preliminary estimation of station pair correlation as a function of distance. *Sol Energy* 2012;86:2170–  
475 6. doi:10.1016/j.solener.2012.02.027.

476 [7] De Felice M, Petitta M, Ruti PM. Short-term predictability of photovoltaic production over  
477 Italy. *Renew Energy* 2015;80:197–204. doi:10.1016/j.renene.2015.02.010.

478 [8] Cros S, Buessler E, Huet L, Sébastien N, Schmutz N. The benefits of intraday solar irradiance  
479 forecasting to adjust the day-ahead scheduled PV power. 2015.

480 [9] Sylvain Cros OL. *Clear sky models assessment for an operational PV production forecasting*  
481 *solution* 2013.

482 [10] Almeida MP, Perpiñán O, Narvarte L. PV power forecast using a nonparametric PV model.  
483 *Sol Energy* 2015;115:354–68. doi:10.1016/j.solener.2015.03.006.

484 [11] Diagne HM, Lauret P, David M. *Solar irradiation forecasting: state-of-the-art and proposition*  
485 *for future developments for small-scale insular grids*, n.d.

486 [12] Paulescu M, Mares O, Eugenia P, Stefu N, Pacurar A, Calinoiu D, et al. Nowcasting solar  
487 irradiance using the sunshine number. *Energy Convers Manag* 2014;79:690–7.  
488 doi:10.1016/j.enconman.2013.12.048.

- 489 [13] De Gooijer JG, Hyndman RJ. 25 years of time series forecasting. *Int J Forecast* 2006;22:443–  
490 73. doi:10.1016/j.ijforecast.2006.01.001.
- 491 [14] Voyant C, Motte F, Fouilloy A, Notton G, Paoli C, Nivet M-L. Forecasting method for global  
492 radiation time series without training phase: comparison with other well-known prediction  
493 methodologies. *Energy* 2017;120:199–208.
- 494 [15] Join C, Voyant C, Fliess M, Muselli M, Nivet ML, Paoli C, et al. Short-term solar irradiance  
495 and irradiation forecasts via different time series techniques: A preliminary study, 2014.
- 496 [16] Hocaoglu FO. Stochastic approach for daily solar radiation modeling. *Sol Energy*  
497 2011;85:278–87. doi:10.1016/j.solener.2010.12.003.
- 498 [17] Kristensen. Parameter estimation in stochastic grey-box models. *Automatica* 2004.
- 499 [18] Mustacchi C, Cena V, Rocchi M. Stochastic simulation of hourly global radiation sequences.  
500 *Sol Energy* 1979;23:47–51. doi:10.1016/0038-092X(79)90042-2.
- 501 [19] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. A benchmarking of machine learning  
502 techniques for solar radiation forecasting in an insular context. *Sol Energy* 2015;112:446–57.  
503 doi:10.1016/j.solener.2014.12.014.
- 504 [20] Pedro HTC, Coimbra CFM. Nearest-neighbor methodology for prediction of intra-hour global  
505 horizontal and direct normal irradiances. *Renew Energy* 2015;80:770–82.  
506 doi:10.1016/j.renene.2015.02.061.
- 507 [21] Kaplanis SN. New methodologies to estimate the hourly global solar radiation; Comparisons  
508 with existing models. *Renew Energy* 2006;31:781–90. doi:10.1016/j.renene.2005.04.011.
- 509 [22] Perez R, Hoff T, Dise J, Chalmers D, Kivalov S. The Cost of Mitigating Short-term PV  
510 Output Variability. *Energy Procedia* 2014;57:755–62. doi:10.1016/j.egypro.2014.10.283.
- 511 [23] Domingos P. The Role of Occam’s Razor in Knowledge Discovery. *Data Min Knowl Discov*  
512 1999;3:409–25. doi:10.1023/A:1009868929893.
- 513 [24] Divya U, Pasupathi C. Survey on Machine Learning Approaches for Solar Irradiation  
514 Prediction. *Int J Eng Sci Res Technol* 2014;3:478–82.
- 515 [25] Perera KS, Aung Z, Woon WL. Machine Learning Techniques for Supporting Renewable  
516 Energy Generation and Integration: A Survey. *Data Anal. Renew. Energy Integr.*, Springer; 2014, p.  
517 81–96.
- 518 [26] Bayesian Learning for Neural Networks | Radford M. Neal | Springer. n.d.

- 519 [27] Voyant C, Notton G, Kalogirou S, Nivet M-L, Paoli C, Motte F, et al. Machine learning  
520 methods for solar radiation forecasting: A review. *Renew Energy* 2017;105:569–82.
- 521 [28] Mellit A, Kalogirou SA, Hontoria L, Shaari S. Artificial intelligence techniques for sizing  
522 photovoltaic systems: A review. *Renew Sustain Energy Rev* 2009;13:406–19.  
523 doi:10.1016/j.rser.2008.01.006.
- 524 [29] Voyant C, Paoli C, Muselli M, Nivet M-L. Multi-horizon solar radiation forecasting for  
525 Mediterranean locations using time series models. *Renew Sustain Energy Rev* 2013;28:44–52.
- 526 [30] Quora. Is Machine Learning Overhyped? Forbes n.d.  
527 <https://www.forbes.com/sites/quora/2016/07/18/is-machine-learning-overhyped/> (accessed September  
528 27, 2017).
- 529 [31] Kim TY, Oh KJ, Kim C, Do JD. Artificial neural networks for non-stationary time series.  
530 *Neurocomputing* 2004;61:439–47. doi:10.1016/j.neucom.2004.04.002.
- 531 [32] Marty C, Philipona R. Clear-sky index to separate clear-sky from cloudy-sky situations in  
532 climate research. *Geophys Res Lett* 2000;27:2649–2652.
- 533 [33] Antonanzas J, Osorio N, Escobar R, Urraca R, Ascacibar FJ, Antonanzas F. Review of  
534 photovoltaic power forecasting. *Sol Energy* 2016;136. doi:10.1016/j.solener.2016.06.069.
- 535 [34] Ahmad A, Anderson TN, Lie TT. Hourly global solar irradiation forecasting for New Zealand.  
536 *Sol Energy* 2015;122:1398–408. doi:10.1016/j.solener.2015.10.055.
- 537 [35] Sanfilippo A, Martín Pomares L, Mohandes N, Perez-Astudillo D, Bachour D. An adaptive  
538 multi-modeling approach to solar nowcasting. *Sol Energy* 2016;125:77–85.  
539 doi:10.1016/j.solener.2015.11.041.
- 540 [36] Martín Pomares L, Zarzalejo L, Polo J, Navarro A, Marchante R, Cony M. Prediction of  
541 global solar irradiance based on time series analysis: Application to solar thermal power plants energy  
542 production planning. *Sol Energy* 2010;84:1772–81. doi:10.1016/j.solener.2010.07.002.
- 543 [37] Wang Y, Wu L. On Practical Challenges of Decomposition-Based Hybrid Forecasting  
544 Algorithms for Wind Speed and Solar Irradiation. *Energy* 2016;112.  
545 doi:10.1016/j.energy.2016.06.075.
- 546 [38] Voyant C, Haurant P, Muselli M, Paoli C, Nivet M-L. Time series modeling and large scale  
547 global solar radiation forecasting from geostationary satellites data. *Sol Energy* 2014;102:131–42.



- 548 [39] Chu Y, Coimbra CFM. Short-term probabilistic forecasts for Direct Normal Irradiance. *Renew*  
549 *Energy* 2017;101:526–36.
- 550 [40] Chu Y, Li M, T.C. Pedro H, F.M. Coimbra C. Real-time prediction intervals for intra-hour  
551 DNI forecasts. *Renew Energy* 2015;83. doi:10.1016/j.renene.2015.04.022.
- 552 [41] Urraca R, Antonanzas J, Alía Martínez M, Ascacibar FJ, Antonanzas F. Smart baseline models  
553 for solar irradiation forecasting. *Energy Convers Manag* 2016;108:539–48.  
554 doi:10.1016/j.enconman.2015.11.033.
- 555 [42] Brown CE. Coefficient of Variation. *Appl. Multivar. Stat. Geohydrology Relat. Sci.*, Springer,  
556 Berlin, Heidelberg; 1998, p. 155–7. doi:10.1007/978-3-642-80328-4\_13.
- 557 [43] Ineichen P. A broadband simplified version of the Solis clear sky model. *Sol Energy*  
558 2008;82:758–62. doi:10.1016/j.solener.2008.02.009.
- 559 [44] Mueller RW, Dagestad KF, Ineichen P, Schroedter-Homscheidt M, Cros S, Dumortier D, et al.  
560 Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module. *Remote Sens*  
561 *Environ* 2004;91:160–74. doi:10.1016/j.rse.2004.02.009.
- 562 [45] Masani P. Shift invariant spaces and prediction theory. *Acta Math* 1962;107:275–90.  
563 doi:10.1007/BF02545791.
- 564 [46] Cartier P, Perrin Y. Integration over finite sets. *Nonstandard Anal. Pract.*, Springer, Berlin,  
565 Heidelberg; 1995, p. 185–204. doi:10.1007/978-3-642-57758-1\_9.
- 566 [47] Neville J, Jensen D. A bias/variance decomposition for models using collective inference.  
567 *Mach Learn* 2008;73:87–106. doi:10.1007/s10994-008-5066-6.
- 568 [48] GEOSS n.d. <https://www.earthobservations.org/geoss.php> (accessed July 26, 2017).
- 569 [49] Korany M, Boraïy M, Eissa Y, Aoun Y, Abdel Wahab MM, Alfaro SC, et al. A database of  
570 multi-year (2004–2010) quality-assured surface solar hourly irradiation measurements for the  
571 Egyptian territory. *Earth Syst Sci Data* 2016;8:105–113.
- 572 [50] David M, Ramahatana F, Trombe P-J, Lauret P. Probabilistic forecasting of the solar  
573 irradiance with recursive ARMA and GARCH models. *Sol Energy* 2016;133:55–72.
- 574 [51] Wiens TS, Dale BC, Boyce MS, Kershaw GP. Three way k-fold cross-validation of resource  
575 selection functions. *Ecol Model* 2008;212:244–55. doi:10.1016/j.ecolmodel.2007.10.005.
- 576

577

578

579

580 **List of Tables**

581

582 Table 1. Short summary concerning the persistence models.

583 Table 2. Value of the optimized  $N$  for each horizon (Ajaccio)

584 Table 3.  $nRMSE$  for all models (Ajaccio)

585 Table 4. Value of the optimized  $N$  for each horizon (Tilos)

586 Table 5.  $nRMSE$  for all models (Tilos)

587

588 **List of figures**

589

590 Figure 1. Principle of the stochastic persistence based on an additive scheme

591 Figure 2. Principle of the stochastic persistence based on a multiplicative scheme

592 Figure 3. Prediction error (from horizon 1 hours to 6 hours in Ajaccio) in term of size of the sliding  
593 window concerning the  $StP^x$  (Eq. 12)

594 Figure 4. Prediction error (from horizon 1 hour to 6 hour in Ajaccio) in term of size of the sliding  
595 window concerning the  $StP^+$  (Eq. 12)

596 Figure 5.  $nRMSE$  Vs horizon for the 5 studied models (Ajaccio)

597 Figure 6. Profile of prediction related to  $StP^+$  and  $StP^x$  between the 5000<sup>th</sup> and 5300<sup>th</sup> hour (spring in  
598 Ajaccio)

599 Figure 7. Prediction error (from horizon 15 minutes to 90 minutes in Tilos) in term of size of the  
600 sliding window concerning the  $StP^x$  (Eq 12)

601 Figure 8. Prediction error (from horizon 15 minutes to 90 minutes in Tilos) in term of size of the  
602 sliding window concerning the  $StP^+$  (Eq 12)

603 Figure 9.  $nRMSE$  vs horizon for the 5 studied models (Tilos)

604 Figure 10. Profile of prediction related to  $StP^+$  and  $StP^x$  between the 15000<sup>th</sup> and 22500<sup>th</sup> minutes  
605 (spring in Tilos)

606

607

608