INDEPENDENT JOURNAL OF MANAGEMENT & PRODUCTION (IJM&P)
http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

SHORT-TERM SOLAR RADIATION FORECASTING BY USING AN ITERATIVE COMBINATION OF WAVELET ARTIFICIAL NEURAL NETWORKS

Julio Cesar Royer
Federal Institute of Parana (IFPR), Brazil
E-mail: julioroyer@gmail.com

Volmir Eugênio Wilhelm
Federal University of Paraná (UFPR), Brazil
E-mail: volmirw@gmail.com

Luiz Albino Teixeira Junior
Latin American Integration Federal University (UNILA), Brazil
E-mail: luiz.a.t.junior@gmail.com

Edgar Manuel Carreño Franco
Western Parana State University (Unioeste), Brazil
E-mail: emfra.unioeste@gmail.com

Submission: 23/10/2015
Accept: 06/11/2015

ABSTRACT

The information provided by accurate forecasts of solar energy time series are considered essential for performing an appropriate prediction of the electrical power that will be available in an electric system, as pointed out in Zhou et al. (2011). However, since the underlying data are highly non-stationary, it follows that to produce their accurate predictions is a very difficult assignment. In order to accomplish it, this paper proposes an iterative Combination of Wavelet Artificial Neural Networks (CWANN) which is aimed to produce short-term solar radiation time series forecasting. Basically, the CWANN method can be split into three stages: at first one, a decomposition of level p, defined in terms of a wavelet basis, of a given solar radiation time series is performed, generating $r+1$ Wavelet Components (WC); at second one, these $r+1$ WCs are individually modeled by the k different ANNs, where $k>5$, and the 5 best forecasts of each WC are combined by means of another ANN, *producing the combined forecasts of WC; and, at third one, the combined forecasts WC are simply added, generating the forecasts of the underlying solar radiation data. An iterative algorithm is proposed for iteratively searching for the optimal values for the CWANN parameters, as we will see. In order to evaluate it, ten real solar radiation time series of Brazilian system were modeled here. In all statistical results, the CWANN method has achieved remarkable greater forecasting performances when compared with a traditional ANN (described in Section 2.1).*

Keywords: solar radiation time series, wavelet decomposition, artificial neural networks, forecasts.

1. INTRODUCTION

The conversion of solar energy into electrical energy is one of most promising alternatives to generate electricity from clean and renewable way. It can be done through large generating plants connected to a transmission system or by means of small generation units for the isolated systems. The Sun provides annually to the Earth's atmosphere, approximately, $1.5\text{x}10^{18}$ kWh of energy, but only a fraction of this energy reaches the Earth's surface, due to the reflection and absorption of sunlight by the Earth's atmosphere (SINGH; CHAUDHARY; THAKUR, 2011).

One problem of renewable energy, for instance, wind and solar energies is the fact that the production of these sources dependents on meteorological factors. Particularly, in the case of solar energy, the alternation of day and night, the seasons, the passage of clouds and rainy periods cause great variability and discontinuities in the production of electricity. In addition, it is needed to have capable devices of storing energy during the day in order to make it available during the night such as battery banks or salt tanks, as pointed out by Wittmann *et al.* (2008).

Thus, the safe economic integration of alternative sources in the operation of the electric system depends on accurate predictions of energy production so that operators may make decisions about the maintenance and dispatch of generating units that feed the system as a whole.

Among the techniques employed in solar radiation forecasting, it can be pointed out that the *Auto-Regressive Integrated with Moving Average* (ARIMA) (PERDOMO *et al.,* 2010), the *Artificial Neural Networks* (ANN) (DENG *et al*., 2010; YANLING *et al*., 2012; YONA; SENJYU, 2009; ZERVAS, *et al*., 2008; ZHANG;

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

BEHERA, 2012), the Kalman Filter (CHAABENE; AMMAR, 2008) and the different ways of combining wavelet orthonormal basis and ANN (CAO *et al*., 2009; ZHOU *et al.*, 2011; TEIXEIRA JR., *et al.*, 2015).

The wavelet methods combined with several types of predictive models (as the ANNs) have been proposed, achieving remarkable accuracy gains. Basically, the wavelet methods consist of auxiliary pre-processing procedures of the data in question, which can be accomplished generally in two ways: by decomposition (as in TEIXEIRA JR., *et al.*, 2013) or by noise shrinkage (as in MALLAT, 2009) of the time series to be forecasted.

Particularly, several studies show the predictive gains achieved by combining wavelet decomposition and ANN approaches, as in: Krishna *et al*. (2011), who applied to forecast river flows; Liu *et al*. (2010), Catalão *et al*. (2011), who modeled wind time series; Teixeira Junior *et al*. (2015), who worked with time series of solar radiation; and Minu *et al*. (2010), who studied time series of number of terrorist attacks in the world.

Due to the complexity to predict the solar radiation time series, two aspects should be accounted for. Firstly, although it is well-known that ANNs integrated with wavelet decompositions, referred to here as wavelet ANN, commonly lead to remarkable predictive gains, their best configuration are obtained in a manual way - and not iteratively by means of a computational algorithm.

Secondly, the adoption of individual forecasting methods (as the ANNs and wavelet ANNs) in forecasting processes underestimates the structural risk exhibited mainly by non-stationary time series (as is the case of solar energy time series). However, as pointed out by Firmino *et al.* (2014), this trouble is accomplished when is performed a combination of unbiased forecasts provided by individual forecasting methods.

This paper put forwards an interactive *Combination of Wavelet Artificial Neural Networks* (CWANN) which is aimed to produce short-term solar radiation forecasting. Summarily, the CWANN method can be split into three steps: in Step 1, a decomposition of level *r*, defined in terms of a wavelet basis, of a given solar radiation time series is performed, generating $r + 1$ Wavelet Components (WC); secondly, these $r + 1$ WCs are individually modeled by the k different ANNs, where

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

$k > 5$, the *5* best forecasts are combined by means of other ANNs, and the best combined one is selected; and, thirdly, all the combined forecasts of the WCs are added, producing the solar radiation time series forecast.

For this, an iterative algorithm is used to iteratively find the optimal CWANN parameters, as it will be seen. For illustrating it in a real case, ten time series of global horizontal solar radiation of Brazilian system were modeled here. All statistical results have shown that CWANN method has achieved better accuracy performances when compared with the ANN and wavelet ANN described, respectively, in Sections 2.1 and 2.2, and used in Teixeira Jr. et al (2015).

So, this paper this is divided into five sections. In Sections 2, there are introduced theoretical aspects about the wavelet decomposition and ANNs. Section 3 describes the CWANN method. The main statistical results are exposed and commented upon in Section 4. In Section 5, the paper is closed.

2. REVIEW OF LITERATURE

The purpose of this section is to present a brief review of some basic concepts which are needed for defining the CWANN method, described in Section 3. It starts, in Section 2.1, by describing the wavelet decomposition of level *r*, which is the algorithm adopted in initial step of the CWANN method. This is followed by the description of the (feed-forward multi-layer perceptron) Artificial Neural Networks (ANNs), in Section 2.2, that are used to model individually each WC (produced by a wavelet decomposition of level *r*) and to perform the proposed linear combination of forecasts, as we will see in Section 3.

2.1. Wavelet Decomposition of Level *r*

Let l^2 denote a collection of all scalar-valued complex infinite sequences with finite energy (that is, $l^2 := \{y_t\ (t \in \mathbb{Z}):\ \sum_{t \in \mathbb{Z}} |y_t|^2 < \infty\}$), equipped with an usual inner product $\langle\ ;\ \rangle$ (as in Kubrusly, 2012). According to Kubrusly and Levan (2006), any subset $h_n\ (n \in \mathbb{Z})$ in l^2 is an orthonormal basis of l^2 if holds the following axioms: (i) orthogonality: $\langle h_n, h_m \rangle = 0$, whenever $n \neq m$, where $n, m \subset \mathbb{Z}$; (ii) normality: $\|h_n\| = 1$, where $n \in \mathbb{Z}$; (iii) completeness: $\langle x, h_{n} \rangle = 0$, if $x = 0$. Thus, based on Kubrusly (2012), if $h_n\ (n \in \mathbb{Z})$ is an orthonormal basis l^2, then any y_t of the sequence $y_t (t \in \mathbb{Z})$ in l^2 can orthogonally expanded in terms of $h_n\ (n \in \mathbb{Z})$, as follows:

ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

$$y_t = \sum_{n \in \mathbb{Z}} \langle y_t, h_n \rangle h_n \tag{1}$$

where $\langle y_t, h_n \rangle$ denotes the usual inner product of $y_t (t \in \mathbb{Z})$ and h_n. An element $\omega(.) \in l^2$ is called an orthonormal wavelet function if, and only if, the functions $\omega_{m,n}(.)$, where $(n, m) \subset \mathbb{Z}x\mathbb{Z}$, defined by

$$\omega_{m,n}(t) := 2^{\frac{m}{2}} \omega(2^m(t) - n) \tag{2}$$

- called an orthonormal wavelet functions, form an orthonormal basis for l^2. Accordingly, any y_t of the sequence $y_t (t \in \mathbb{Z})$ in l^2 admits the expansion in (3).

$$y_t = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle y_t, \omega_{m,n}(t) \rangle \omega_{m,n}(t) \tag{3}$$

Now, for each $n, m \in \mathbb{Z}$, the projection of y_t onto $\omega_{m,n}(.)$ is given by $\langle y_t, \omega_{m,n}(t) \rangle \omega_{m,n}(t)$. According to Levan and Kubrusly (2003), this can be considered as a *detail variation* of y_t at scale 2^m and time-shift n. In effect, for each $m \in \mathbb{Z}$, the projection $y_{D_m,t}$ of y_t at scale 2^m is defined by the following partial sum

$$y_{D_m,t} = \sum_{n \in \mathbb{Z}} \langle y_t, \omega_{m,n}(t) \rangle \omega_{m,n}(t). \tag{4}$$

This, in turn, can be regarded as a *WC of detail* of y_t at scale at scale 2^m (as in Mallat, 2009). Therefore, any $y_t (t \in \mathbb{Z}) \in l^2$ is the sum of all its WC of detail over all scales.

On the other hand, an element $\phi(.) \subset l^2$ is called scaling function if, and only if, the functions $\phi_{m,n}(t)$, where $(n, m) \in \mathbb{Z}x\mathbb{Z}$, defined by

$$\phi_{m,n}(t) = 2^{\frac{m}{2}} \phi(2^m(t) - n) \tag{5}$$

are such that $\langle \phi_{m',n'}(t), \phi_{j,k}(t) \rangle = 0$, whenever $m' = j$ e $n' \neq k$, and $\langle \phi_{m',n'}(t), \phi_{j,k}(t) \rangle \neq 0$, if otherwise. For each $m \in \mathbb{Z}$, the projection $y_{A_m,t}$ of y_t is given by the following partial sum

$$y_{A_m,t} = \sum_{n \in \mathbb{Z}} \langle y_t, \phi_{m,n}(t) \rangle \phi_{m,n}(t) \tag{6}$$

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

This can be referred to as the WC of approximation of y_t at scale at scale 2^m (MALLAT, 2009). Based on Levan and Kubrusly (2006), an y_t in the sequence $y_t(t \in \mathbb{Z}) \in l^2$ can be orthogonally expanded as in (7).

$$y_t = y_{A_{m_0},t} + \sum_{m=m_0}^{+\infty} y_{D_m,t} \tag{7}$$

The expansion in (7) in usually referred to as "wavelet decomposition" of y_t. According to Teixeira Jr *et al.* (2015), any finite time series, denoted by $y_t(t = 1, \ldots, T)$, can be decomposed as in (8).

$$y_t = y_{A_{m_0},t} + \sum_{m=m_0}^{m_0+(r-1)} y_{D_m,t} \tag{8}$$

The expansion in (8) is usually regarded as a *wavelet decomposition of level* r of the state y_t.

2.2. Artificial Neural Networks

The *Artificial Neural Networks* (or simply ANNs) are very flexible computing frameworks for modeling and forecasting a broad range of stochastic time series, because they just requires they exhibit either linear and non-linear auto-dependence structures. As is the case of most statistical linear models, the stationarity property are not required by ANN approaches (as in HAMILTON, 1994).

Another important aspect is that the ANNs are universal approximators of compact (i.e., closed and bounded) support functions, as pointed out by Cybenko (1989). Thus, since a time series y_t $(t = 1, \ldots, T)$ that depends on its own past may be seen as points a compact support, it follows that the ANNs are capable to approximate (for modeling or forecasting) it with a high degree of accuracy. According to Zhang (2003), their predictive power comes from the parallel processing of the information from the data. In addition, the ANN models are largely determined by the stochastic characteristics inherent in the time series.

In perspective, the *feed-forward multi-layer perceptron ANNs* (referred to from now on as ANNs) are the most widely used prediction model for time series forecasting. Particularly, a single hidden layer ANN is characterized by a network composed by three layers of simple processing units numerically connected by acyclic links. The relationship between the output at instant *t*, denoted by y_t, and the

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

p-lagged inputs, represented by the sequence $y_{t-k}\ (k = 1, \dots, p)$, has got the following mathematical representation:

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j\, g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij}\, y_{t-i}\right) + \varepsilon_t \qquad (9)$$

where $\alpha_j\ (j = 0, 1, \dots, q)$ and $\beta_{ij}\ (i = 0, 1, \dots, p;\ j = 0, 1, \dots, q)$ are the (single hidden layer) ANN parameters, which are often called the connection weights; p is the number of input nodes; q is the number of hidden nodes; ε_t is the approximation error at time *t*; and $g(.)$ is here a logistic function, although it would be possible to adopt another transfer function (please, see Haykin (2001) for more details). The logistic function is widely used as the hidden layer transfer function in forecasting processes and its mathematical representation is given by

$$g(x_t) = \frac{1}{1+\exp(-x_t)} \qquad (10)$$

where $x_t := \beta_{0j} + \sum_{i=1}^{p} \beta_{ij}\, y_{t-i}$ and $exp(.)$ is the exponential function with Euler's basis (as in HAYKIN, 2001). Due to $g(.)$ is a non-linear transfer function, the ANN model, in (9), in fact performs a non-linear mapping from the past observations $y_{t-k}\ (k = 1, \dots, p)$ to the future state y_t. Equivalently, the model in (9) can be rewritten, as follows:

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, w) + \varepsilon_t \qquad (11)$$

where w denotes a vector of all ANN parameters and $f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, w)$ is the model determined by the network structure and connection weights. Indeed, the neural network is equivalent to a non-linear auto-regressive model.

In practice, w is an unknown vector of ANN parameters and hence needs to be adjusted. So, in order to find the optimal solution $\hat{w}$, accounting for some criteria, for the vector of ANN parameter w, some optimization algorithm must be employed. Although there are several methodologies in specialized literature, maybe the Levenberg-Marquardt's algorithm (as in ADAMOWSKI; KARAPATAKI, 2010) might be considered most used for this assignment. The minimization in-sample squared error mean (i.e., $min_w \sum_{t=1}^{T} \varepsilon_t^2$) is usually used as numerical criteria. Thus, it is

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

desired that the solution $\hat{w}$ of this optimization problem is the argument that minimizes the $\sum_{t=1}^{T} \varepsilon_t^2$. Once obtained $\hat{w}$, it has

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \hat{w}) + \hat{\varepsilon}_t \tag{12}$$

where $f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \hat{w})$ is the final ANN output at instant *t* which consists of forecast, denoted by $\hat{y}_t$, of the state y_t, and $\hat{\varepsilon}_t$ is its forecasting error.

3. PROPOSED METHODOLOGY

Let y_t $(t = 1, \dots, T)$ denotes a solar radiation time series that exhibits auto-dependence structure, and assume it is required to produce its out-of-sample forecasts. With this purpose, the proposed iterative methodology, referred to as a *Combination of Wavelet Artificial Neural Networks* (or simply CWANN), can be carried out according to the following four steps.

- **Step 1**: a wavelet decomposition of level r (as in Section 2.1), defined in terms of an orthogonal wavelet basis (i.e., the Haar, Daubechies, Minimum-Bandwidth, Fejér-Korovkin, Battle-Lemarie, and Symlet families (as in MALLAT, 2009; DAUBECHIES, 1992; MORRIS; PERAVALI, 1999), of y_t $(t = 1, \dots, T)$ is performed, producing 1 WC of approximation at level m', denoted by $y_{A_{m'},t}$ $(t = 1, \dots, T)$, where $m' \in \mathbb{Z}$, and r WCs of detail at levels m', $m' + 1$, …, $m' + (r - 1)$, denoted by $y_{D_m,t}$ $(t = 1, \dots, T)$, respectively, where $r \in \mathbb{Z}$.

- **Step 2**: Each WC from Step 1 is individually modeled by *k* different ANNs (as in Section 2.2), where $k > 5$. So, here is generated the following sequences of forecasts: $\hat{y}_{A_{m'},t,i}$, $\hat{y}_{D_{m'},t,i}$, …,$\hat{y}_{D_{m'+(r-1)},t,i}$ $(t = t' + 1, \dots, T + h)$, where $i = 1, \dots, k$ and t' represents the degree of freedoms lost till this step. Note that $\hat{y}_{A_{m'},t,i}$ and $y_{D_m,t,i}$, where $m = m', \dots, m' + (r - 1)$, are, respectively, the forecasts of the state $y_{A_{m'},t}$, produced by the *ith* ANN, and of the state $y_{D_m,t}$, generated by the *ith* ANN. In this way, if it is accounted for a wavelet de decomposition of level r of y_t $(t = 1, \dots, T)$, then $(r + 1) \times k$ distinct ANNs are needed for separately modeling them.

INDEPENDENT JOURNAL OF MANAGEMENT & PRODUCTION (IJM&P)
http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

- **Step 3**: The *5* best forecasts of each WC in Step 2 are combined by means of an ANN, generically denoted by $f(.)$, in order to produce its combined forecasts. Mathematically talking, it has:

  $f_{A_{m'}}\left(\hat{y}_{A_{m'},t,1}, \hat{y}_{A_{m'},t,2}, \ldots, \hat{y}_{A_{m'},t,5}, \omega\right) = \hat{y}_{A_{m'},t,C}$, and

  $f_{D_m}\left(\hat{y}_{D_m,t,1}, \hat{y}_{D_m,t,2}, \ldots, \hat{y}_{D_m,t,5}, \omega\right) = \hat{y}_{D_m,t,C}$

  where $\hat{y}_{A_{m'},t,C}$ and $\hat{y}_{D_m,t,C}$ are, respectively, the combined forecasts of $y_{A_{m'}t}$ and $y_{D_{m'}t}$. Indeed, the ANNs $f_{A_{m'}}(.)$ and $f_{D_m}(.)$ provide the predictions $\hat{y}_{A_{m'},t,C}$ and $\hat{y}_{D_m,t,C}$, respectively. Importantly, the ANNs used in Steps 2 and 3 are the same ones described in Section 2.2.

- **Step 4:** The combined forecasts of each WC in Step 3 are simply added, generating the out-of-sample combined forecasts of solar radiation time series, denoted by $\hat{y}_t\ (t = t'' + 1, \ldots, T + h)$, where t'' is the degrees of freedom lost until Step 3. That is:

  $\hat{y}_t = \hat{y}_{A_{m'},t,C} + \sum_{m=m'}^{m'+(r-1)} \hat{y}_{D_m,t,C}$.

  The steps 1, 2, 3 and 4 are repeated for all wavelet orthogonal basis, and for decomposition level r from 1 to 3. The best forecast is selected.

In this paper, the four steps above are together carried out in an iterative way by means of a computational algorithm (schematized in Figure 1) that tests several values for the CWANN parameters exhibit in Steps 1, 2 and 3. As objective function, the minimization of the in-sample (or training) MSE of a given solar radiation time series (i.e., $min \sum_{t''+1}^{T}(y_t - \hat{y}_t)^2$) is adopted. More specifically, by CWANN parameters here, it means: the level of decomposition wavelet r and the wavelet orthogonal basis, in Step 1; and the window length p and the number q of artificial neurons in hidden layer of each ANN, in Steps 2 and 3.

Notice that, in Steps 2 and 3, other ANN parameters (as transfer functions and training algorithm) could also be used as "variables" to be optimized. However, in order to avoid increasingly large periods of training, only the window length and the number of nodes in the hidden layer were chosen for working out as variables to be numerically adjusted.

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

Some aspects were accounted for computational algorithm and deserve then to be commented upon. Firstly, since it can be performed an enormous number of different setups between a wavelet level of decomposition r and a wavelet orthonormal basis, and there is no major analytic role in determining what is the best, it follows that an enormous number of empirical tests are needed for obtaining so.

Secondly, concerning the ANN parameter q, there exists no systematic rule in deciding its optimal value such that all values belonging to a determined range (defined by the decision maker) are numerically tested. Thirdly, to choosing an appropriate number of hidden nodes, another important task of ANN modeling of a time series is the selection of the number of lagged observations, p, the dimension of the input vector.

This is perhaps the most important parameter to be estimated in an ANN model because it plays a major role in determining the auto-dependence structure of the time series. However, there is also no theory that can be used to guide analytically the selection of an optimal value of p. Fourthly, due to the overfitting effect typically found in neural network modeling, a validation sample was used for choosing the best ANNs. By an overfitted ANN, it means an ANN that has a good fit to the in-sample data, but has poor generalization ability in the out-of-sample period.

Figure 1(a) shows the wavelet decomposition flowchart, including the loops to search for the best decomposition parameters (filter and level of decomposition) and the call to the ANN routine which makes the CWs forecast.

The variables used are *SR* (Solar Radiation time series); *ws* and *we* (window start and window end, limits for the window size range); *ns* and *ne* (neurons start and neurons end, the range for the number of neurons in the hidden layer); *is* and *ie* (maximum number of iterations start and end); *numFilters* is the total number of filters (orthogonal wavelet basis) tested; *SRForecast[f, r]* is a matrix with all tested forecasts, to choose the best forecast; *f* and *r* are indexes to indicate filter and the level of decomposition.

Apr is an approximation CW and *Det* is a detail CW, and ANN is a calling to the ANN forecasting procedure described in Figure 3 (b). To this figure, the variables *w*, *n* and *i* are also used, as indices of windows, neurons and iterations ranges; *series*, as the CW to be modeled; *fcTr*, *fcVal* and fc*Tst* as the Training, Validation and

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

Test sample forecasts; *cbTr*, *cbVal*, cbTst as the combined Training, Validation and Test sample forecasts; *mlp* is the Multilayer Perceptron neural network calling procedure and *RMSE* is the routine that calculates the root mean square error, used as the evaluation parameter for the forecast accuracy.

Basically, the procedure (a) makes wavelet decompositions for all possible combinations of the available filters and decomposition levels, using two nested loops. For each CW it calls the ANN procedure (b). This ANN procedure test all possible combinations of window size, neurons on the hidden layer and number of iterations, making ANN forecasts. Then it selects the 5 best ones and combines them using other two nested loops, and selects the best combination. This combination is then returned as the CW forecast. All CW forecasts are added to form the Solar Radiation forecast for one filter and one decomposition level. At the end, the best Solar Radiation forecast is selected as the final forecast.

4. NUMERICAL EXPERIMENTS

In order to compare and numerically illustrate the CWANN, the same numerical experiments carried out in Teixeira Jr *et al.* (2015) were performed here. For this, ten (real) solar radiation time series sampled in ten cities of Brazil and different years were modelled; all of them with hourly data in a period of one year (i.e., 1^{th} January to 31^{th} December), resulting 8760 observations. Their graphical representations of the daily profiles can be seen in Figure 2, and the mean and standard deviation statistics, in Table 1. The solid black line in Figure 2 represents the hourly mean of solar radiation.

INDEPENDENT JOURNAL OF MANAGEMENT & PRODUCTION (IJM&P)
http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

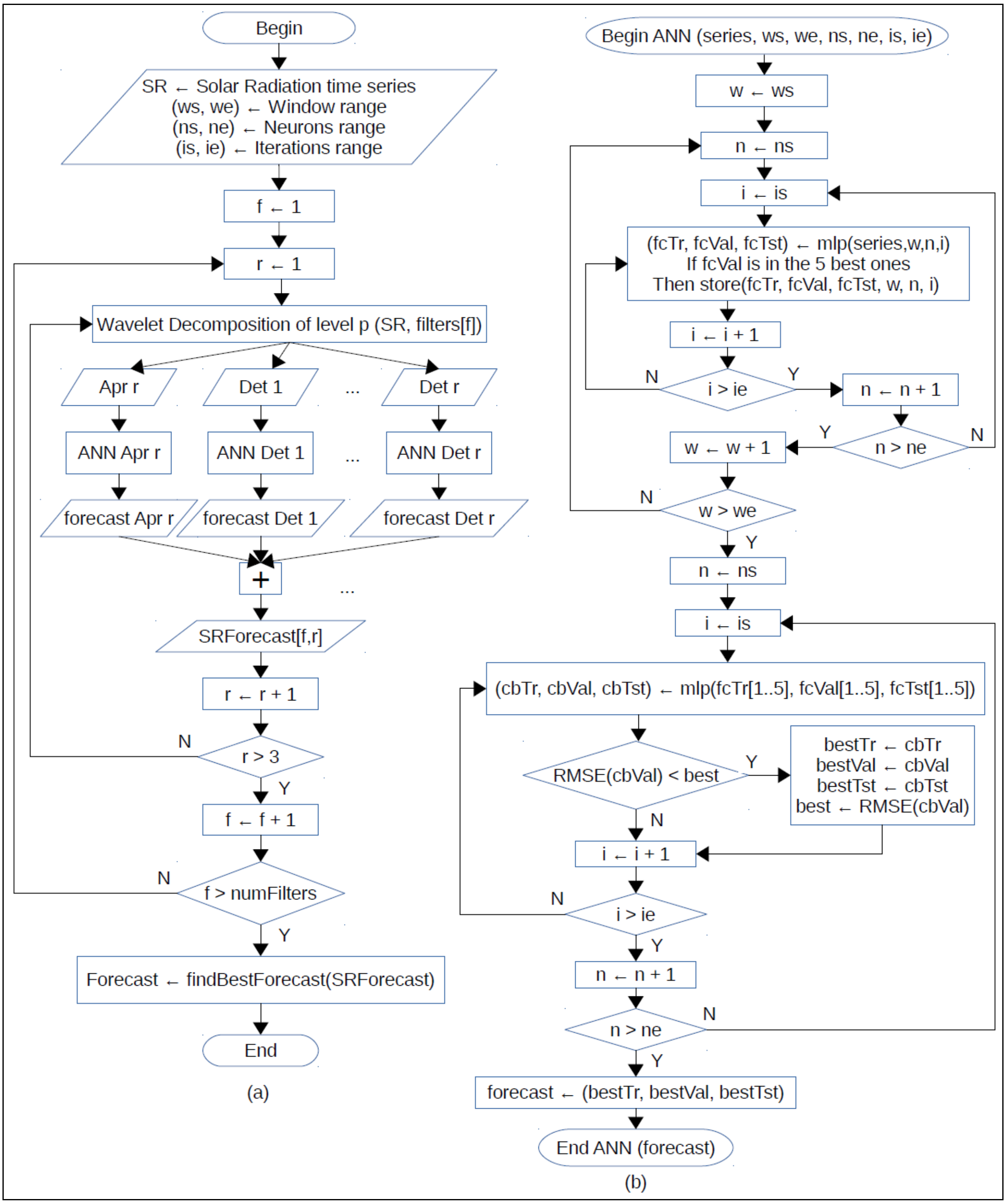

Figure 1: Flowchart of the CWANN method.

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

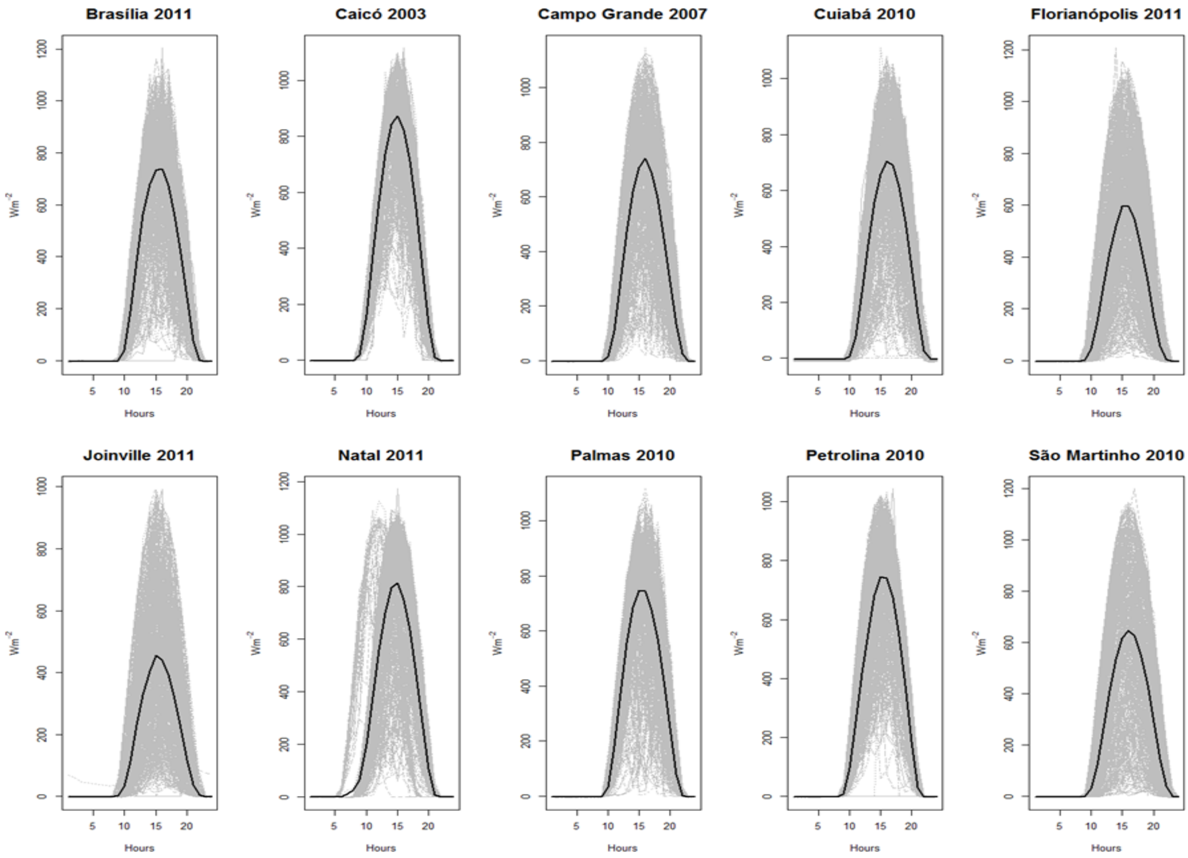

Figure 2: Daily profiles of solar radiation time series.

Table 1: Mean and standard deviation of the global horizontal solar radiation

Meteorological station	Mean W/m^2	Standard Deviation W/m^2
Brasília 2011	219,20	304,15
Caicó 2003	253,82	339,08
Campo Grande 2007	213,34	299,37
Cuiabá 2010	204,51	290,98
Florianópolis 2011	171,77	270,09
Joinville 2011	128,72	210,37
Natal 2011	241,13	334,09
Palmas 2010	220,46	304,90
Petrolina 2010	220,41	302,75
São Martinho 2010	196,04	296,06

4.1. Proposed Computational Algorithm

The computational algorithm described in Section 3 was implemented in R software (R CORE TEAM, 2015) and accounted for the following packages: Waveslim package (WHITCHER, 2015), used to perform the wavelet decomposition method (described in Step 1, in Section 3, and schematized on left side in Figure 1); and the RSNNS package (BERGMEIR; BENITEZ, 2012), employed to make the modelling by ANN models (required in Steps 2 and 3, in Section 3, and schematized on right side in Figure 1).

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

In one hand, the orthonormal basis used in the wavelet decompositions were those available in Waveslim package, i.e.: Haar (db1); Daubechies (db2, db4 and db8); Minimum Bandwidth (mb4, mb8, mb16 and mb24); Fejér-Korovkin (fk4, fk6, fk8, fk14 and fk22); Least Asymmetric or Symlet (la8, la16 and la20); and Battle-Lemarie (bl14 and bl20). The wavelet decomposition levels tested were 1, 2 and 3 (i.e., $r = 1, 2, 3$).

On the other hand, all the ANNs used in Steps 2 and 3 were composed with one hidden layer, hyperbolic tangent activation function in the hidden layer, and linear activation function in the output layer. Regarding the learning algorithm, the scaled gradient conjugate (SGC), which simulates the Levenberg-Marquardt algorithm, was used.

In all simulations, the input patterns of all ANNs were transformed into a point belonging to the bounded and closed interval [-1,1]. The implemented computational algorithm selects the optimum values to the following parameters: wavelet orthonormal basis, wavelet decomposition level, RNA window size, number of neurons in the hidden layer and maximum number of iterations for each CW and for the combined forecast.

Importantly, the time series modelling followed the same approach in Teixeira *et al.* (2015) in order to guarantee a proper comparison of results. Thereby, the training sample was composed of 7008 observations, the next 876 observations were the validation sample, and the 876 remaining data were the test sample.

4.2. Modeling For The 10 Time Series

Table 2 exhibits the *Root Mean Square Deviation* (RMSE) and the coefficient of determination R^2 (as in HAMILTON, 1994) in the test sampling, as well as the best configuration, of the naive predictor (as in HAMILTON, 1994) , a conventional ANN (as in Section 2.1), the Wavelet ANN proposed by Teixeira Jr et al. (2015) and the CWANN method (proposed method).

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

Table 2: Wavelet basis, types of ANN, RMSE and R^2 on the test sample for each time series´ modeling, by Teixeira Junior *et al.* (2015) (*) and by the CWANN method (**)

Local	Result	Wavelet basis	Window length	Neurons in the hidden layer	Maximum iterations number	RMSE Wm^{-2}	R^2
Brasília	*	Naive predictor				143.34	0.7848
	*	without	12	8		107.88	0.8707
	*	db32	15	8		91.29	0.9074
	**	db8	10 to 15	18 to 25	27 to 30	17.05	0.9968
Caicó	*	Naive predictor				134.68	0.8611
	*	without	15	19		66.58	0.9648
	*	db20	15	8		37.61	0.9888
	**	db8	10 to 15	18 to 25	27 to 30	11.79	0.9989
Campo Grande	*	Naive predictor				154.21	0.7898
	*	without	15	10		120.07	0.865
	*	db20	12	8		76.06	0.9458
	**	fk8	10 to 16	18 to 25	27 to 30	19.89	0.9964
Cuiabá	*	Naive predictor				144.84	0.8076
	*	without	10	19		106.28	0.8908
	*	db38	10	12		26.14	0.9934
	**	la20	10 to 25	15 to 25	25 to 30	16.49	0.9974
Florianópolis	*	Naive predictor				143.29	0.8302
	*	without	10	10		109.72	0.8958
	*	db40	8	15		50.12	0.9783
	**	la16	10 to 15	18 to 25	27 to 30	19.22	0.9969
Joinville	*	Naive predictor				111.58	0.8089
	*	without	11	5		92.24	0.8625
	*	db32	12	10		84.34	0.885
	**	la16	10 to 16	18 to 25	27 to 30	15.44	0.9963
Natal	*	Naive predictor				133.36	0.8738
	*	without	15	5		57.32	0.9759
	*	db20	15	13		75.96	0.9577
	**	bl20	10 to 16	18 to 25	27 to 30	8,68	0.9995
Palmas	*	Naive predictor				138.54	0.7780
	*	without	15	10		101.77	0.8727
	*	db40	10	13		60.3	0.9553
	**	db8	10 to 15	18 to 25	27 to 30	16.72	0.9966
Petrolina	*	Naive predictor				121.63	0.8543
	*	without	15	9		75.11	0.9423
	*	db15	9	20		82.51	0.9303
	**	bl14	10 to 16	18 to 25	27 to 30	11.87	0.9986
São Martinho	*	Naive predictor				140.37	0.8694
	*	without	15	20		98.87	0.9329
	*	db13	20	14		19.68	0.9973
	**	la20	10 to 15	18 to 25	27 to 30	17.11	0.9982

5. CONCLUSIONS

In this paper, an iterative *Combination of Wavelet Artificial Neural Networks* (CWANN), detailed described in Section 3, has been put forward with the aim of producing out-of-sample one-step forecasts of solar radiation. An iterative algorithm also is proposed for iteratively searching for the optimal values for the CWANN

ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

parameters, as it was shown. For evaluating it as well as compare it with other methods, 10 real solar radiation time series of Brazilian system were modeled here.

Regarding the adherence statistics RMSE and R^2, Table 1 shows clearly that the CWANN predictions have achieved greater accuracy power than Naive predictor, ANN and the Wavelet ANN proposed by Teixeira Jr *et al.* (2015). Indeed, these results are an important empirical evidence that the optimal numerical adjustment of the non-linear combination of different ANN forecasts by using an ANN integrated with the wavelet decomposition may provide accuracy gains. In addition, the automatic search, not only avoid operational effort, but also contributes to accomplish best forecasts of solar radiation.

Finally, it points out that, even though the mathematical theory associated with the methods (i.e., wavelet decomposition and ANN) that compose the CWANN is relatively complex, they can be straightforwardly implemented in an operational way by using the software R as well as its packages mentioned in the text.

ACKNOWLEDGEMENTS

The authors thank the Araucaria Foundation for partial financial support, project n. 44721 and the IFPR for the possibility of fulltime dedication to the research.

REFERENCES

ADAMOWSKI, J.; KARAPATAKI, C. (2010) Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. **Journal of Hydrological Engineering, ASCE**, v. 15, p. 729-743.

CAO, S.; WENG, W.; CHEN, J.; LIU, W.; YU, G.; CAO, J. (2009) Forecast of Solar Irradiance Using Chaos Optimization Neural Networks. POWER AND ENERGY ENGINEERING CONFERENCE. ASIA-PACIFIC. Shanghai. **Proceedings.** Shanghai: IEEE, 2009.

CATALÃO, J. P. S.; POUSINHO, H. M. I.; MENDES, V. M. F. (2011) Short-Term Wind Power Forecasting in Portugal by Neural Networks and Wavelet Transform. **Renewable Energy**, v. 36, p. 1245-1251.

CHAABENE, M.; AMMAR, B. M. (2008) Neuro-Fuzzy Dynamic Model with Kalman Filter to Forecast Irradiance and Temperature for Solar Energy Systems. **Renewable Energy**. v. 33, n. 7, p. 1435-1443.

CYBENKO, G. (1989) Approximation by Superpositions of a Sigmoidal Function. **Mathematics of Control, Signals, and Systems**. v. 2, p. 303-314.

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

DAUBECHIES, I. (1988) Orthonormal Bases of Compactly Supported Wavelet. **Communications Pure and Applied Math**. v. 41, n. 7, p. 909-996.

DAUBECHIES, I. (1992) **Ten Lectures on Wavelets**. CBMS-NSF Regional Conference Series In Applied Mathematics: SIAM.

DENG, F.; SU, G.; LIU, C.; WANG, Z. (2010) Global Solar Radiation Modeling Using The Artificial Neural Network Technique. POWER AND ENERGY ENGINEERING CONFERENCE. ASIA-PACIFIC. Chengdu. **Proceedings.** Beijing: IEEE, 2010.

HAMILTON, J. D. (1994) **Time Series Analysis**. Princeton University Press.

HAYKIN, S. S. (2001) **Redes Neurais Princípios e Aplicações**. 2ª Ed. Porto Alegre: Bookman.

KRISHNA, B.; SATYAJI RAO Y. R.; NAYAK P. C. (2011) Time Series Modeling of River Flow Using Wavelet Neural Networks. **Journal of Water Resource and Protection**. v. 3, p.50-59.

KUBRUSLY, C. S. (2012) **Spectral Theory of Operators on Hilbert Spaces**. New York: Birkhäuser.

KUBRUSLY, C. S.; LEVAN, N. (2006) Abstratc Wavelets Generated By Hilbert Space Shift Operators. **Advances in Mathematical Sciences and Applications**. v. 16, p. 643-660.

LEVAN, N.; KUBRUSLY, C. S. (2003) A *Wavelet* “Time-Shift-Detail” Decomposition. ***Mathematics and Computers in Simulation***. v. 63, n. 2, p. 73-78.

LIU, H.; TIAN, H. Q.; CHEN, C.; LI, Y. (2010) A Hybrid Statistical Method to Predict Wind Speed and Wind Power. **Renewable Energy**. v. 35, p. 1857-1861.

MALLAT, S. (2009) **A *Wavelet* Tour of Signal Processing**. 3rd ed. Burlington: Academic Press.

MINU, K.K.; LINEESH, M. C.; JESSY, J. C. (2010) Wavelet Neural Networks for Nonlinear Time Series Analysis. **Applied Mathematical Sciences**. v. 4, n. 50, p. 2485-2495.

MORRIS, J. M.; PERAVALI, R. (1999) Minimum-bandwidth discrete-time wavelets. **Signal Processing**. v. 76, n. 2, p. 181-193.

PERDOMO, R.; BANGUERO, E.; GORDILLO, G. (2010) Statistical Modeling for Global Solar Radiation Forecasting in Bogotá. PHOTOVOLTAIC SPECIALISTS CONFERENCE (PVSC). Honolulu. **Proceedings**. Honolulu: IEEE, 2010.

SINGH, C.; CHAUDHARY, R.; THAKUR, R. S. (2011) Performance of advanced photocatalytic detoxification of municipal wastewater under solar radiation – A mini review. **International Journal of Energy and Environment**. v. 2, n. 2, p. 337-350.

TEIXEIRA JUNIOR, L. A.; MENEZES, M. L.; CASSIANO, K. M.; PESSANHA, J. F. M.; SOUZA, R. C. (2013) Residential Electricity Consumption Forecasting Using a Geometric Combination Approach. **International Journal of Energy and Statistics**. v. 1, n. 2, p. 113-125.

http://www.ijmp.jor.br v. 7, n. 1, January - March 2016
ISSN: 2236-269X
DOI: 10.14807/ijmp.v7i1.393

TEIXEIRA JUNIOR, L. A.; SOUZA, R. M.; MENEZES, M. L.; CASSIANO, K. M.; PESSANHA, F. M. P.; SOUZA, R. C. (2015) Artificial Neural Network and Wavelet Decomposition in the Forecast of Global Horizontal Solar Radiation. **Brazilian Operations Research Society**. v. 35, n. 1, p. 1-16.

WITTMANN, M.; BREITKREUZ, H.; SCHROEDTER-HOMSCHEIDT, S.; ECK, M. (2008) Case Studies on the Use of Solar Irradiance Forecast for Optimized Operation Strategies of Solar Thermal Power Plants. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**. v.1, n. 1, p. 18-27.

YANLING, G.; CHANGZHENG, C.; BO, Z. (2012) Blind Source Separation for Forecast of Solar Irradiance. *INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM DESIGN AND ENGINEERING APPLICATION*. Sanya, Hainan. **Proceedings**. Sanya, Hainan: IEEE, 2012.

YONA, A.; SENJYU, T. (2009) One-Day-Ahead 24-Hours Thermal Energy Collection Forecasting Based on Time Series Analysis Technique for Solar Heat Energy Utilization System. TRANSMISSION AND DISTRIBUTION CONFERENCE AND EXPOSITION: ASIA AND PACIFIC. Seoul. **Proceedings**. Seoul: IEEE, 2009.

ZERVAS, P. L.; SARIMVEIS, H.; PALYVOS, J. A.; MARKATOS, N. C. G. (2008) Prediction of Daily Global Solar Irradiance on Horizontal Surfaces Based on Neural-Network Techniques. **Renewable Energy**. v. 33, n. 8, p. 1796-1803.

ZHANG, N.; BEHERA, P. K. (2012) Solar Radiation Prediction Based on Recurrent Neural Networks Trained by Levenberg-Marquardt Backpropagation Learning Algorithm. *INNOVATIVE SMART GRID TECHNOLOGIES (ISGT)*. Washington, DC. **Proceedings**. Washington, DC: IEEE, 2012.

ZHANG, P. G. (2003) Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. **Neurocomputing**. v. 50, p. 159-175.

ZHOU, H.; SUN, W.; LIU, D.; ZHAO, J.; YANG, N. (2011) The Research of Daily Total Solar-Radiation and Prediction Method of Photovoltaic Generation Based on *Wavelet*-Neural Network. POWER AND ENERGY ENGINEERING CONFERENCE (APPEEC). Wuhan. **Proceedings**. Wuhan: IEEE, 2011.