

This is a post-peer-review, pre-copyedit version of an article published in Neurocomputing. The final authenticated version is available online at: <http://dx.doi.org/10.1016/j.neucom.2015.02.089>

# Prediction of the Italian Electricity Price for Smart Grid Applications

Emanuele Crisostomi<sup>a</sup>, Claudio Gallicchio<sup>b</sup>, Alessio Micheli<sup>b</sup>, Marco Raugi<sup>a</sup>,  
Mauro Tucci<sup>a</sup>

<sup>a</sup>*Department of Energy, Systems, Territory and Constructions Engineering, University of Pisa, Italy.*

<sup>b</sup>*Department of Computer Science, University of Pisa, Italy.*

---

## Abstract

In this paper we address the problem of one day-ahead hourly electricity price forecast for smart grid applications. To this aim, we investigate the application of a number of predictive models for time-series, including methods based on empirical strategies frequently adopted in the smart grid community, Kalman Filters and Echo State Networks (ESNs). The considered methods have been suitably modified to address the electricity price forecast problem. Strategies based on daily re-adaptation of models' parameters are taken into consideration as well. The predictive performance achieved by the considered models is assessed, and the methods are compared among each other on recent real data from the Italian electricity market. As a result of the comparison over three years data, ESN methods appear to provide the most accurate price predictions, which could imply significant economic savings in many smart grid activities, such as switching on power plants to support power generation from renewable sources, electric vehicle recharging or usage of household appliances.

*Keywords:* Electricity Price, Time Series Prediction methods, Smart Grid, Echo State Networks

---

## 1. Introduction

The new “Smart Grid” paradigm is expected to “redefine the concept of what it means to build and operate the grid” [1]. In particular, although human system operators will still have the final say, a plethora of intelligent tools and facilities will be implemented to operate the grid in an automatic

way. The opportunity of exploiting bidirectional power flows, a variety of small-size power plants and time-varying energy prices will be used to provide more refined control tools than simply turning power plants on and off [1].

Under the general umbrella of the Smart Grid, two new concepts of energy clusters are currently emerging: Virtual Power Plants (VPPs) and Microgrids. A VPP is a cluster of dispersed generator units, controllable loads and storage systems, aggregated in order to operate as a unique plant [2], [3]. The microgrid shares similar features with the VPP, but usually the term VPP is used to emphasize the economic aspects, while the microgrid must have the important property of being able to operate in a grid-connected mode, isolated mode, and in a transition between grid-connected and isolated modes [4]. The core of both microgrids and VPPs is the Energy Management System (EMS), that is the intelligent system that has to take the important decisions related to the computation of the optimal power flows (i.e., *power scheduling problem*). Algorithms for computing optimal power flows, according to different multi-objective utility functions can be found in [5], and specific examples of optimal power flows in VPPs are [6] and [7]

The EMS takes the optimal decisions on the basis of three main factors: (i) the energy demand, (ii) the energy availability, and (iii) the electrical energy price. In addition, the optimality of the decisions must consider a long future time horizon and not the instant when the decision is taken (e.g., while it might instantaneously seem convenient to sell energy to the grid, it might be more convenient to store the energy, wait a few hours, and then sell it to the grid at a higher price). However, in order to take optimal decisions considering a future time horizon, it is required to have accurate predictions of the future energy demand, the future energy availability and future electricity prices. Note that these quantities are highly stochastic, and depend on many factors that can not always be predicted with the required accuracy; for instance, the weather conditions affect the energy demand (e.g., air conditioning), the energy availability (due to renewable sources such as solar/wind) and consequently also the electricity price.

### 1.1. Motivations

The electric grid is not the only stakeholder in the smart grid framework who has a concrete interest in developing the ability to accurately predict electric energy prices. The accurate prediction of the electricity price is of great interest also for energy providers. Indeed, on the one side, energy providers entering the one day-ahead electricity market would greatly benefit

of knowing in advance the prices in order to make better bids, and on the other side they would be able to take optimal strategic actions, e.g. switching on a conventional power plant to gain the ability to sell more energy on the next day. Similarly, energy providers have to decide whether surplus power generated from renewable sources should be sold to the grid, or rather stored in appropriate storage systems and rather sold at a later time, upon more favorable conditions.

In this paper we tackle the problem of electrical energy price forecasting one day in advance. In this context, some model-based methods have been frequently used in the past to predict the price of electrical energy. In particular, it has been noticed by many authors that the electrical energy price signal has two evident auto-correlation peaks, i.e. at 24 hours and at seven days, corresponding to the daily trend and to the weekly trend respectively, see for instance [8] for a numerical evaluation of the auto-correlation function. Following the previous result, linear and non-linear regressive functions have been used to predict the electrical energy prices, taking full advantage of the daily and weekly patterns. The objective of this paper is to illustrate and compare different predictive models aimed at forecasting the electricity price, with applications to microgrid control and optimization, with a specific focus on predictive models for time-series processing. On purpose, we compare two predictive methods that do explicitly exploit the aforementioned daily and weekly correlation, an empirical method and a Kalman Filter, with another neural network method that exploits the same correlation pattern, although in an extended temporal context. We believe neural network approaches have some attractive features that can be successfully used to tackle the specific problem of interest here. Specifically, due to the nature of the problem, based on the trend of past price values to estimate the current one, models for time-series prediction tasks seem to find application.

In particular, in this paper we take into consideration one of most prominent approaches for learning in sequential domains, i.e., the class of Recurrent Neural Networks (RNN) [9] models. RNNs implement non-linear dynamical systems and are used to learn variable-term input-output relationships, which are not restricted to any specific fixed-size time window. In this regard, Reservoir Computing (RC) [10, 11] is a paradigm for efficient RNN modeling. Such efficiency mainly stems from the separation between an untrained recurrent component, which provides a dynamical memory of the input history, and a trained recurrent-free output component. Within the RC paradigm we take into specific account the Echo State Network (ESN)

[12, 13], a neural network model which represents a state-of-the-art approach in the context of efficiently learning in sequential domains. Indeed, in recent years ESNs have been successfully applied in a wide range of application areas, including chaotic time series prediction (e.g. [12, 14]), speech recognition (e.g. [15]) autonomous systems modeling (e.g. [16, 17]), control systems (e.g. [18]), Ambient Assisted Living (e.g. [19, 20, 21, 22, 23]), Human Activity Recognition (e.g. [24]), robot localization (e.g. [25]).

The proposed predictive algorithms are tested and compared upon public open-access data, made available by GME<sup>1</sup>. Such data pertains to the Italian market, still not widely considered in literature, and refers to a large and very recent time period, spanning from January 1<sup>st</sup> 2010 to December 31<sup>st</sup> 2013.

### *1.2. State of the art*

Many works in literature have investigated the problem of electricity price forecasting, differing among each other for the different adopted predictive method, for the information used in input, the time horizon of the prediction and the market considered for the application [26]. Among the works on this subject, although applied to different national markets and different from each other according to the aforementioned characteristics, it is worth mentioning applications of generalized auto-regressive conditional heteroscedasticity (GARCH) models [27] and autoregressive integrated moving average (ARIMA) models [28], which are basically linear models, including hybrid variants making use of wavelet transform techniques [29]. Other approaches perform electrical price prediction based on the use of neural networks [30, 31, 32, 33, 34], extreme learning machines [35] and support vector machines (SVMs) [36, 37, 38, 39].

Another recent interesting reference is [40] where, differently from the conventional literature that tries to forecast the normal range electricity prices, a special focus was given on the ability to predict the price spikes. Spikes correspond to the occurrences of prices that are tens or even hundreds of times higher than the normal price. More recently, smart grid applications have raised new interest in the prediction of the electricity prices. For instance, in a forthcoming scenario where the penetration level of Plug-in Electric Vehicles (PEVs) is expected to have a strong impact on the grid, it will be important to develop the ability to recharge the vehicles when the price of the

---

<sup>1</sup><http://www.mercatoelettrico.org/En/Default.aspx>

electricity is small [41]. In this perspective, the work described in [42] aims at designing economic Model Predictive Control (MPC) techniques to minimize the cost of electricity consumption for electric vehicle charge planning. Similarly, reference [8] aims at providing accurate electricity price predictions to support the implementation of optimal residential load control actions. Recently, feed-forward neural networks and RC models have been applied in a different (related) context, focusing on the forecasting of the electricity load [43, 44, 45]. Finally, a first attempt to predict the hourly value of the Italian electrical energy price using, among other methods, SVM techniques, is provided in [46, 39], whose results are outperformed by the models proposed in this paper.

It is important to remark that the electricity price prediction strongly depends on some other variables, e.g., weather forecasts. In particular, the price of electricity depends on the difference between the load (i.e., electricity demand) and the energy that can be produced from renewables sources (i.e., a subset of the electricity offer). When the renewables sources manage to (almost) balance the load, then the price of electricity can be very small and close to zero. On the other hand, when the renewable sources are far from matching the load demand, then it is required to burn fuel (e.g., coal or gas), which is clearly more expensive and the electricity price increases. Therefore, electricity prices could be predicted more accurately if the load demand, the energy production from renewables sources, and the prices of fuels (e.g., gas) were available (in an accurate fashion) as well.

In this work however, we are interested in predicting the electricity price solely from previous values, as a classic time-series prediction task. The obtained results can then be interpreted also as a benchmark result. Therefore, it is worth to note that the results provided by any other prediction strategy that intends to exploit also other information when available (e.g., expected production from renewable sources), should then be compared to the results obtained by our predictive models in order to assess the usefulness of the (possibly highly noisy) extra information used. Moreover, in our opinion, also the comparison among the different approaches proposed, under the aspect of modeling based on the addressed problem of electricity price forecasting, represents an interesting result. It is also worth noticing that although the predictive models proposed in this paper are experimentally evaluated on data from the Italian market, they are not specifically designed for such a market. Thereby, our investigations and results can be exploited also by other experts operating in the sector, under the assumption that the

electrical energy price time-series should have similar patterns in different countries, and that the algorithm most performing in the Italian market is expected to also provide accurate results for other markets.

### 1.3. Outline of the Paper

The rest of this paper is organized as follows. In Section 2 the formulation of the forecasting problem is presented, details on the datasets used are provided, and the algorithms are described both in their basic and improved forms. The results of the experimental analysis are reported and discussed in Section 3, and in Section 4 some final remarks are delineated.

## 2. Problem formulation, Dataset and Algorithms

The problem of electricity price forecasting considered in this paper is formulated in Section 2.1. The dataset used to train, validate and compare the different proposed algorithms is described in Section 2.2. The basic algorithms and the basic versions of the predictive models proposed to predict the electrical energy price are presented in Section 2.3, whereas further advancements proposed to such predictive models in order to tailor them to the characteristics of the temporal dynamics in the application of interest are described in Section 2.4.

### 2.1. Problem formulation

In this work we consider the hourly purchase price for end-customers of electrical energy (euros per *MWh*). Accordingly, we assume that  $\mathbf{p}_d \in \mathbb{R}^{24}$  is a (column) vector of 24 elements that contains the price of electrical energy corresponding to day  $d$ , i.e.  $\mathbf{p}_d = [p_d(1), p_d(2), \dots, p_d(24)]^T$ , where 24 corresponds to midnight. Analogously, the vector  $\hat{\mathbf{p}}_d \in \mathbb{R}^{24}$  contains the *predicted* electrical energy prices.

In this paper, we are primarily interested in investigating the ability of different predictive models to forecast the electrical energy prices one day in advance. We use, as a starting point, some results in literature that establish that hourly prices are strongly correlated at a day scale and at a week scale, while correlations at other time scales can be neglected [8]. We remark here that such a feature is typical of the electrical energy price, but is not in general true for other market purchase prices. Based on such an assumption,

Table 1: Summary of model notations.

$\mathbf{p}_d$	24-component vector of hourly purchase price of electrical energy at day $d$
$p_d(h)$	electrical energy price at day $d$ , hour $h$
$\widehat{\mathbf{p}}_d$	24-component vector of <i>predicted</i> hourly purchase price of electrical energy at day $d$
$\widehat{p}_d(h)$	<i>predicted</i> electrical energy price at day $d$ , hour $h$
$\alpha_{1,d}, \alpha_{2,d}$	scalar parameters used in the empirical approach at day $d$
$\mathbf{y}_d, \mathbf{n}_d, \mathbf{v}_d$	parameters of the dynamical model underlying the electricity price time series used by the KF (explanation in the text)
$\mathbf{P}, \mathbf{Q}, \mathbf{K}, \mathbf{R}$	parameters of the KF (explanation in the text)
$\mathbf{u}_d, \mathbf{x}_d$	input and state vectors, respectively, of the ESN for day $d$

we can say that a basic formulation, representing the starting point for the formulation of the prediction problem is given by the following equation:

$$\widehat{\mathbf{p}}_{d+1} = \mathbf{f}(\mathbf{p}_d, \mathbf{p}_{d-6}). \quad (1)$$

which expresses the underlying assumed relationship (indicated as a generic function  $\mathbf{f}$  in Equation 1) between the electricity price for any given day, and the electricity price one day before and seven days before.

The objective of this paper is to compare different prediction functions  $\mathbf{f}$  on their ability to provide accurate predictions of the real electrical energy price. As it will be discussed in more detail in Section 3, we shall use three different indexes, namely, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Average Percentage Error (MAPE), to compare the different algorithms. For the sake of simplicity and clarity in the description of the different predictive models considered in this paper, we adopt a notation which is made as much uniform as possible among the different approaches. The main notations used throughout the paper are summarized in Table 1.

## 2.2. Dataset

We compare the algorithms according to their ability to predict the electricity price (euros per *MWh*) in the Italian market (Single National Price,

corresponding to purchase price for end customers). Hourly prices and historical data are freely available in the GME website. In particular, we take into consideration the electricity price data from January 1st 2010 to December 31st 2013, for a total number of 1461 days. The resulting dataset consist in a multidimensional time-series  $\mathcal{D} = \{\mathbf{p}_d | d = 1, \dots, 1461\}$ . Based on preliminary investigations on historical data from previous years (2008 to 2010), aimed at optimizing the size of the portion of the dataset used as *training set*, we found out that the best length is exactly 365 days, i.e., one year, in the case that the empirical linear regression algorithms were used (see Section 2.3.2). Therefore, data pertaining to year 2010 is used as training set, i.e.,  $\mathcal{T} = \{\mathbf{p}_d | d = 1, \dots, 365\}$ . All the remaining data pertaining to years 2011, 2012 and 2013 are used as *test set*, i.e., they are used to assess (and compare) the generalization performance of the predictive models.

Figure 1 shows a hourly representation of the electricity cost for the period spanned by the test set data. As can be seen from the figure, predicting the electricity price is a challenging task as there are not clearly visible paths, e.g., seasonal paths, and due to the many spikes corresponding to days, and sometimes even to hours, when the electricity price is very high or very low. In particular, it can be noted that in 2013 there are many spikes of low prices. This is due to the fact that there has been an amazing increased production of electrical energy from renewable sources in the last couple of years in Italy, especially from solar and wind sources. Therefore, in some specific days of the year, e.g., summer Sundays, the quantity of energy produced from solar plants almost matched the overall national load, and thus the price of electrical energy approached zero (as the cost of energy produced from renewable sources is close to zero). On the other hand, note that when the energy produced from renewable sources is a low fraction of the overall national load, then it is required to generate electrical energy from conventional sources (e.g., burning gas, oil, coal), and thus there are both fuel/carbon and operation costs that increase the price of the electrical energy. While an accurate prediction of the electrical load and of the power that can be generated from renewable sources is expected to be a valuable information for the prediction of electricity price, in this paper we are interested in evaluating how accurate can the prediction be if no other information apart from the historical series is available.



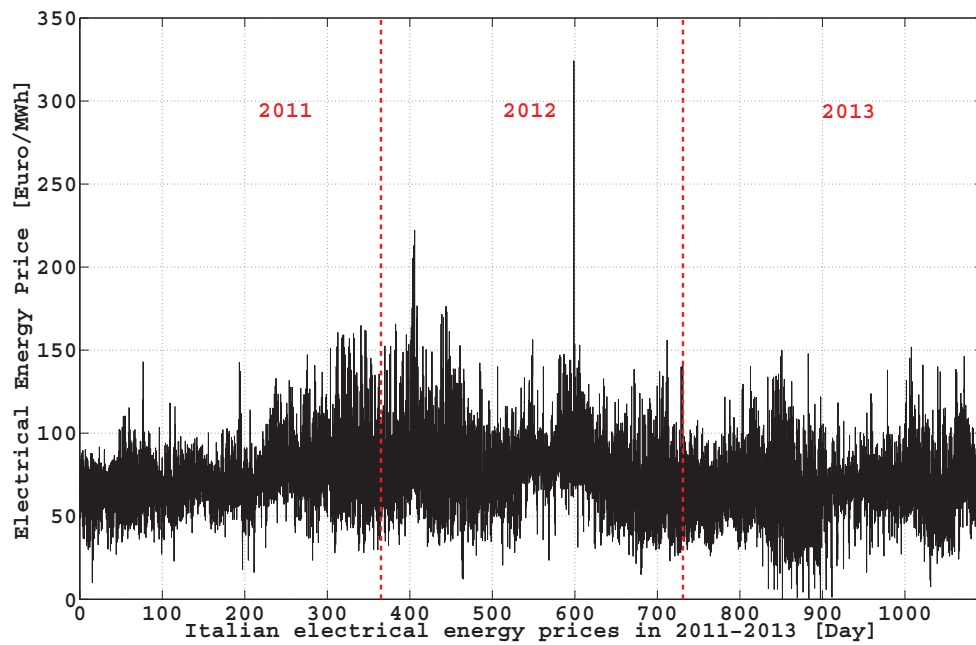


Figure 1: Italian electrical energy prices in years 2011, 2012 and 2013.

### 2.3. Basic algorithms

In this Section we describe the basic formulation of the predictive models considered in this paper.

#### 2.3.1. Day-before forecast

The simplest prediction algorithm assumes that the electrical energy prices of the following day will be the same of the current day. In our notation, such a prediction can be mathematically described as

$$\widehat{\mathbf{p}}_{d+1} = \mathbf{p}_d \quad (2)$$

which makes possible to predict a whatever hourly electrical energy price exactly 24 hours in advance. Note that, according to Equation 2, the price prediction for hour  $h$  of the next day  $d + 1$ , i.e.,  $\widehat{p}_{d+1}(h)$ , is provided as soon as the actual price for the same hour of the present day, i.e.,  $p_d(h)$ , is available, and such prediction for hour  $h$  remains constant for the following 24 hours, i.e. until the actual price for hour  $h$  of day  $d + 1$ , i.e.,  $p_{d+1}(h)$ , becomes available. Also note that the algorithm described by Equation 2 is a simplified version of the general prediction problem (see Equation 1), obtained by neglecting the week correlation. Despite its simplicity, it is worth noticing that such an algorithm is actually applied in some recent works in the Smart Grid literature (see for instance [7]) and that results are quite accurate, as it will be shown in Section 3. The day-before forecast algorithm is used in the following to provide a baseline performance for the more sophisticated methods considered in this paper, such as Kalman Filter (see Section 2.3.3) and Echo State Networks (see Section 2.3.4). Moreover, also other empirical approaches (see Section 2.3.2) aim to outperform the baseline day-before forecast approach.

#### 2.3.2. Empirical Approach

The main drawback of the day-before forecast is that it fails to take into account the difference between electricity price profiles of week-days and week-ends. Under the umbrella of empirical approaches we consider the algorithms that modify the day before prediction algorithm described in the previous section to further consider the current day of the week.

For instance, such an empirical approach has been suggested in the same context in the recent reference [8], as a consequence of a correlation study that revealed that prices are strictly correlated in one day and (in a smaller

way) in a one week interval. On the other hand, correlations at other time scales can be neglected. In [8] hourly prices are predicted according to a linear combination of the day-before (same hour) prices (with an average weight of about 0.8-0.9) and those of the previous week (with a weight never larger than 0.2). The two-days-before data were also used, but only to predict Monday prices (to avoid giving too much importance to the prices of the day-before, being a holiday). In this paper we also adopt a similar strategy for the purpose of comparison, though using different weighting coefficients, as those of [8] were taken from a different context (i.e., electricity prices in 2007-2009 used by Illinois Power Company). In practice, the model of such an approach is

$$\widehat{\mathbf{p}}_{d+1} = \alpha_{1,d+1}\mathbf{p}_d + \alpha_{2,d+1}\mathbf{p}_{d-6} \quad (3)$$

so that the next day prediction is a linear combination of the price of the current day and the price of the previous week (same day of the week). The two scalar coefficients  $\alpha_{1,d+1}$  and  $\alpha_{2,d+1}$  depend on the particular day of the week we want to predict, they are optimally inferred from available historical data and are different for each day of the week. However, they only depend on the particular day of the week, therefore they are periodic with a period of 7 days, and overall only 14 optimal coefficients must be learned from the data. In this paper, the optimal value of the 14 parameters is obtained from historical data by linear regression, according to a least square solution via pseudo-inverse of the data matrix.

Note that as  $\alpha_{1,d+1}$  and  $\alpha_{2,d+1}$  are scalar coefficients, Equation 3 can be also expressed on a hourly basis as  $\widehat{p}_{d+1}(h) = \alpha_{1,d+1}p_d(h) + \alpha_{2,d+1}p_{d-6}(h)$ . From the last expression, it is straightforward to see that also this algorithm can be used to predict the price of electrical energy 24 hours ahead.

### 2.3.3. Kalman Filter

The adoption of the Kalman filter underlies the assumption that the price signal evolves as a dynamic system subject to suitable linear dynamic equations. Noise is used to compensate the difference between such a simplified deterministic model and the real evolution of the price signal. Therefore, the Kalman filter assumes that the electrical energy price evolves according to

$$\begin{cases} \mathbf{p}_{d+1}^{(k)} = \mathbf{p}_d^{(k)} + \mathbf{n}_d \\ \mathbf{y}_d = \mathbf{p}_d^{(k)} + \mathbf{v}_d \end{cases}, \quad (4)$$

where again it is assumed that, apart from some model noise  $\mathbf{n}_d$  and some “measurement noise”  $\mathbf{v}_d$ , the electrical energy price is supposed not to change from one day to the following one. The output  $\mathbf{y}_d$  corresponds to the state of the dynamic system, i.e., electric energy price, apart from the noise  $\mathbf{v}_d$  that together with  $\mathbf{n}_d$  accommodates for the discrepancies from the true value of the price. Notation  $\mathbf{p}_d^{(k)}$  emphasizes that the evolving state is not the real electrical energy price, but its linear approximation required by the Kalman filter. Note that both the state and the output  $\mathbf{y}_d$  are vectors of 24 components (i.e., hourly price). Given the linear dynamical system (4), the next day hourly price can be predicted according to typical Kalman filtering equations, see for instance references [47] and [48], reminded for convenience in the following:

$$\begin{cases} \widehat{\mathbf{p}}_{d+1} &= \mathbf{P}_{d|d} \\ \widehat{\mathbf{P}}_{d+1} &= \mathbf{P}_{d|d} + \mathbf{Q} \end{cases} \quad (5)$$

where  $\mathbf{Q}$  is the covariance matrix of the process noise (assumed known),  $\mathbf{P}_{d|d}$  is the covariance matrix of the posterior distribution of the states (the covariance of the initial distribution  $\mathbf{P}_{0|0}$  is supposed known), and  $\mathbf{p}_{d|d}$  is the *corrected* state. The  $\mathbf{p}_{d|d}$  vector and the covariance matrix  $\mathbf{P}_{d|d}$  are computed at each time step according to the following correction equations of the Kalman filter

$$\begin{cases} \mathbf{K}_{d+1} &= \widehat{\mathbf{P}}_{d+1} \cdot (\mathbf{R} + \widehat{\mathbf{P}}_{d+1})^{-1} \\ \mathbf{p}_{d+1|d+1} &= \widehat{\mathbf{p}}_{d+1} + \mathbf{K}_{d+1}(\mathbf{p}_{d+1} - \widehat{\mathbf{p}}_{d+1}) \\ \mathbf{P}_{d+1|d+1} &= \widehat{\mathbf{P}}_{d+1} - \mathbf{K}_{d+1}\widehat{\mathbf{P}}_{d+1} \end{cases} \quad (6)$$

where  $\mathbf{R}$  is the covariance matrix of the measurement noise (assumed known), and  $\mathbf{K}_{d+1}$  is the Kalman gain. Kalman Filter equations 5 and 6 have a simple form as the underlying model is particularly simple (i.e., the state vector does not evolve in time and corresponds to the output measured vector).

The Kalman filter requires  $O(N^3)$  time and  $O(N^2)$  space per timestep, where  $N$  is the dimension of the state space; however, in presence of very sparse matrices as in the case of interest here, very simple low-rank formulae can be used to speed up the performance. In our case, the prediction of a whole year is of the order of one second, on a personal computer with Intel I5 cpu and using the Matlab software environment.

In our simulations, we later choose  $\mathbf{P}_{0|0} = \mathbf{I}$ ,  $\mathbf{Q} = \{\mathbf{I}, 10 \cdot \mathbf{I}, 100 \cdot \mathbf{I}\}$  and  $\mathbf{R} = \{\mathbf{I}, 10 \cdot \mathbf{I}, 100 \cdot \mathbf{I}\}$ , where  $\mathbf{I}$  is the identity matrix of appropriate size. In particular, as a consequence of a tuning phase for data in year 2010, we specifically chose  $\mathbf{Q} = 10 \cdot \mathbf{I}$ ,  $\mathbf{R} = \mathbf{I}$ , and the sensitivity on the choice of the parameters is discussed later.

In practice, the Kalman Filter assumes that it is more convenient to predict the electrical energy prices using as a starting vector the price of the current day modified to take into account some past history of the price signal. More detailed information about Kalman Filter theory can be found in the original paper [47] and in a more recent tutorial [48].

#### 2.3.4. Echo State Networks

From a dynamical system perspective, ESNs [12, 13] implement discrete-time *non-linear* dynamical systems. In particular, a recurrent *reservoir* component provides the network with a non-linear dynamic memory of the past input history. This allows the network state dynamics to be influenced by a portion of the input history which is not restricted to a fixed-size temporal window, permitting to capture longer term input-output relationships [14]. Therefore, the application of the ESN model to the problem of electricity price forecasting allows us to take into account the benefits of the heuristic approaches outlined in Sections 2.3.1 and 2.3.2, together with a non-linear state-transition modeling of the historical sequence of the prices.

From an architectural point of view, ESNs consist of an input layer with  $N_U$  units, a non-linear, recurrent and sparsely connected *reservoir* layer with  $N_R$  units, and a linear, feed-forward *readout* layer with  $N_Y$  units. In particular, in this paper we take into consideration a variant of the standard ESN model, i.e., the leaky integrator ESN (LI-ESN) [49], characterized by state dynamics which can be suitably used to represent the history of input signals with a speed that can be tuned by the means of a leaky parameter.

In the following, the equations that describe the LI-ESN dynamics are tailored to the problem at hand and to the notation illustrated in Table 1. Therefore, in the specific application context, input and output dimensions of the network architecture are tailored to daily hours according to the basic problem formulation provided by Equation 1, i.e., input dimension is set to  $N_U = 48$  and output dimension is set to  $N_Y = 24$ . Furthermore, the time step granularity of the network computation corresponds to one day for the electricity price forecasting. Based on the general formulation of the prediction problem in Equation 1, for each new day  $d + 1$ , the input for the

LI-ESN model, i.e.,  $\mathbf{u}_{d+1} \in \mathbb{R}^{N_U}$ , is the concatenation of the hourly prices vectors for the current day and for the corresponding day in the previous week:

$$\mathbf{u}_{d+1} = \begin{bmatrix} \mathbf{p}_d \\ \mathbf{p}_{d-6} \end{bmatrix} \quad (7)$$

Given the input  $\mathbf{u}_{d+1}$  and the old state  $\mathbf{x}_d \in \mathbb{R}^{N_R}$ , the reservoir computes the state for day  $d + 1$ , i.e.,  $\mathbf{x}_{d+1}$ , according to the state transition function:

$$\mathbf{x}_{d+1} = (1 - a)\mathbf{x}_d + af(\mathbf{W}_{in}\mathbf{u}_{d+1} + \hat{\mathbf{W}}\mathbf{x}_d) \quad (8)$$

where  $\mathbf{W}_{in} \in \mathbb{R}^{N_R \times N_U}$  is the input-to-reservoir weight matrix (possibly including a bias term),  $\hat{\mathbf{W}} \in \mathbb{R}^{N_R \times N_R}$  is the recurrent reservoir weight matrix,  $f$  is the element-wise applied state activation function (we use  $\tanh$ ) and  $a \in [0, 1]$  is the leaking rate (which controls the speed of reservoir state dynamics). Finally, the readout computes the output of the model as a linear combination of the state vector to provide the model output for day  $d + 1$ , according to the equation:

$$\hat{\mathbf{p}}_{d+1} = \mathbf{W}_{out}\mathbf{x}_{d+1} \quad (9)$$

where  $\hat{\mathbf{p}}_{d+1}$  is the 24-dimensional output vector containing the hourly electricity price predictions for day  $d + 1$  and  $\mathbf{W}_{out} \in \mathbb{R}^{N_Y \times N_R}$  is the reservoir-to-readout weight matrix (possibly including a bias term). Note that according to Equations 7, 8 and 9 the hourly predictions for day  $d + 1$  are computed as soon as the hourly prices for day  $d$  become available.

It is worth to note that the non-linear recurrent reservoir component of the network is left *untrained* after initialization according to the constraints provided by the Echo State Property (ESP) [12, 13], related to a contractive setting of the reservoir state transition function and to the resulting Markovian characterization of the state dynamics [14, 50]. A condition on the spectral radius  $\rho$  of the reservoir weight matrix is typically used for initialization. A resulting initialization procedure consists in selecting the input-to-reservoir weight values in  $\mathbf{W}_{in}$  from a uniform distribution over  $[-scale_{in}, scale_{in}]$ , where  $scale_{in}$  is an input scaling parameter, while the recurrent weight values in  $\hat{\mathbf{W}}$  are initialized from a uniform distribution and then re-scaled to a value of  $\rho < 1$  (see [49, 13] for further details). The linear readout is the only trained part, typically by Moore-Penrose pseudo-inversion or ridge regression [12, 11]. In this regard, in this paper, for our experiments we considered different reservoir hyper-parameterizations corresponding to

a reservoir dimension  $N_R \in \{100, 200, 300, 400, 500, 600, 700, 800\}$ , spectral radius  $\rho \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ , leaky parameter  $a \in \{0.1, 0.5, 1\}$ , input scaling  $scale_{in} = 0.01$  and 20% of connectivity among the reservoir units. For each reservoir hyper-parametrization we independently generated a number of 5 reservoir guesses, and results were averaged over such guesses. ESNs were trained using pseudo-inversion and ridge regression with regularization parameter  $\lambda_r \in \{100, 10, 1, 0.1, 0.01\}$ .

Notice also that the ESN model represents an extremely efficient approach for modeling RNNs. The cost of the encoding process computed by the reservoir scales linearly with both the reservoir dimension and the input length. Indeed the application of the state transition function (equation (8)) to an element of an input sequence requires a number of  $O(RN_R)$  operations, where  $R$  is used to denote the maximum number of connections for every reservoir unit (with smaller  $R$  for sparser reservoirs). Hence, the total cost of the encoding for an input sequence of length  $N_S$  is given by  $O(N_SRN_R)$ . Remarkably, for training and testing, the cost of the encoding process in ESNs is the same, as the parameters of the state transition function are not learned. In this regard, the ESNs compares extremely well with competitive state-of-the-art learning models for sequence domains [11, 14]. Also the cost of output computation scales linearly with both reservoir dimension and the input length. Indeed the computation of the output function (equation (9)) for an input sequence of length  $N_S$  is given by  $O(N_SN_YN_R)$ . Training the readout of an ESN requires to solve a linear regression problem with a computational cost that depends on the algorithm used. This can actually vary from iterative methods, for which the cost of each epoch scales linearly with the length of the input, to direct methods [11, 14]. Notice, however, that the choice of a linear output component involves a generally lower computational training cost with respect to more sophisticated alternatives such as multi-layer perceptrons (MLPs) or SVMs (with kernel).

#### 2.4. Improved Algorithms

In this Section we describe how the previous approaches can be further adapted to the particular scenario of interest. As motivated in Section 2.2, we use a historical window of one year for fitting the internal parameters of predictive models in this application context. Accordingly, the main rationale of the proposed improvements consists in a daily re-adaptation of the parameters of the considered models, on data pertaining to the previous 365 days. In this case, for each predictive model, the price prediction for day  $d+1$ , i.e.,

$\widehat{\mathbf{p}}_{d+1}$ , is obtained after training on the training set  $\mathcal{T}_{d+1} = \{\mathbf{p}_{d-364}, \dots, \mathbf{p}_d\}$ . For instance, the prediction of the hourly electrical prices for December 31<sup>st</sup> 2011 is obtained by coefficients computed using as training set the data from December 31<sup>st</sup> 2010 until December 30<sup>th</sup> 2011, instead of using data from January 1<sup>st</sup> to December 31<sup>st</sup> 2010, as in the basic case. Grounded on such retraining strategy, the following sub-sections 2.4.1, 2.4.2 and 2.4.3 detail the main modifications applied to the EA, KF and ESN approaches. The impact of the proposed improvements on the predictive performance is discussed in Section 3.

#### 2.4.1. Improved Empirical Approach

The EA algorithm illustrated in Section 2.3.2 adopts regression coefficients  $\alpha_{1,d+1}$  and  $\alpha_{2,d+1}$  of Equation 3, learned by minimizing the squared error between the predicted and the real electricity price signals on past historical data belonging to a time window of 1 year, i.e., electricity cost data for year 2010. After training, such values of the regression coefficients are considered fixed, and are used to predict the hourly electrical prices for the following days, i.e. for years 2011, 2012 and 2013. The improved version of this algorithm takes into consideration time varying values for the regression coefficients  $\alpha_{1,d+1}$  and  $\alpha_{2,d+1}$ , which are daily retrained, resulting in a possibly different set of coefficient for every day in the test set.

#### 2.4.2. Improved Kalman Filter

The main drawback of the KF as presented in Section 2.3.3 is that the underlying dynamic model is the day-before prediction, i.e., prices remain constant neglecting the information of the particular day of the week. As a consequence, we now improve the previous KF along two main lines: first, we now assume that the underlying model is the same of empirical approach, i.e., predicted prices are a weighted linear combination of the prices of the day before and of the week before. On purpose, we need to extend the length of the state vector of Equation 4 to include the prices of the whole week (i.e., size  $(24 \times 7) \times 1$ ). Accordingly, the transition matrix has now size  $(24 \times 7) \times (24 \times 7)$ , and consists of 49  $24 \times 24$  diagonal matrices, of which only 14 are non-zero (i.e., as prices of a single day depend on the prices of the before, and on those of the week before, and not on the other five days of the past week). Secondly, we update the parameters every day following the same rationale of the strategy outlined in Section 2.4.1 for the empirical strategy.



As a consequence, the improvements to the basic KF are now twofold: (i) the underlying dynamic model is now more sophisticated as it takes into account the weekly pattern, and (ii) the model parameters are updated every day taking into account the latest information.

#### 2.4.3. Improved ESN

In the version described in Section 2.3.4, ESNs are trained by computing the readout weight values in matrix  $\mathbf{W}_{out}$  of Equation 9 using historical data from the year 2010 as training set. In the improved version of this approach, in analogy to the improvement considered for the empirical algorithm (Section 2.4.1), the readout component of the ESN model undergoes a daily retraining process. This means that for every day  $d + 1$ , a new readout weight matrix  $\mathbf{W}_{out}$  is computed using data pertaining to the previous 365 days as training set. In this context, it is worth noticing that the impact of daily network retraining on the computational cost is limited to the readout part by the nature of the reservoir computation paradigm, whereas the reservoir component is not involved in training processes and its state transition function remains the same as in the standard ESN model, as described by equation 8.

### 3. Experimental Results

In this Section we evaluate the performance of the predictive models illustrated in the previous sections, comparing their ability to accurately predict the hourly electrical energy price, one day in advance. In particular, we take into consideration two experimental settings, corresponding to the basic and the improved settings of the considered predictive models (see Sections 2.3 and 2.4). In the first (basic) experimental setting, the predictive models are trained on the data from the year 2010, and their generalization ability is tested on the year 2011 data (see Section 2.3). In the second (extended) experimental setting, the predictive models are evaluated on data pertaining to years 2011, 2012 and 2013, with a daily retraining on data from the previous 365 days (see Section 2.4).

The performance achieved by the different predictive models is evaluated using hourly averages of the L1-norm and of the L2-norm of the difference between the actual electricity prices and the predicted ones. In this context, such metrics can be respectively computed as hourly Mean Absolute Error

(MAE) and Root Mean Squared Error (RMSE) of the price predictions, according to the following formulae:

$$\begin{aligned}
 MAE &= \frac{1}{24N_{days}} \sum_{d=1, \dots, N_{days}} \sum_{h=1, \dots, 24} |p_d(h) - \hat{p}_d(h)| \\
 RMSE &= \sqrt{\frac{1}{24N_{days}} \sum_{d=1, \dots, N_{days}} \sum_{h=1, \dots, 24} (p_d(h) - \hat{p}_d(h))^2}
 \end{aligned} \tag{10}$$

where  $N_{days}$  is the number of days taken into consideration for the error evaluation. It is worth noticing that the use of MAE and RMSE is interesting in the application context under consideration, as such error measures allow us to assess the deviation between predicted and actual values in the same unit measure in which input and output data are defined, i.e., euros. An additional index of interest, widely used in literature, is the Mean Absolute Percentage Error (MAPE), that takes into account relative errors between the actual electricity prices and the predicted ones. To overcome the difficulty of the occurrence of zero values in the price time-series, the relative error of hour  $h$  of day  $d$  is calculated with respect to the average price of the day  $d$ , denoted as  $\bar{p}_d$  (the daily average price never takes zero values). The MAPE is then defined as:

$$MAPE = \frac{100}{24N_{days}} \sum_{d=1, \dots, N_{days}} \sum_{h=1, \dots, 24} \frac{|p_d(h) - \hat{p}_d(h)|}{\bar{p}_d} \tag{11}$$

$$\bar{p}_d = \frac{1}{24} \sum_{h=1, \dots, 24} p_d(h)$$

In order to assess the statistical significance of the different results of the performance indexes of the algorithms, the signed rank Wilcoxon test [51] was considered, with a significance level of 5%. The test can be used to compare two algorithms at a time, and we used it to assess the statistical significance of the MAE and MAPE, by using all hourly error values involved in the mean calculation. As a result, we obtained that the differences between the performance indexes reported in the following are always statistically significant for every possible choice of the couple of compared algorithms.

As explained in Section 2.4.1, the 14 parameters of the EA algorithms are learnt from the past 365 days, and there are no parameters that need to be tuned by the users. Regarding the KF algorithm, see Section 2.4.2, there are three arbitrary parameters: matrices  $\mathbf{P}_{0|0}$ ,  $\mathbf{Q}$  and  $\mathbf{R}$ . In our experience, the value of matrix  $\mathbf{P}_{0|0}$  has a little impact on the accuracy of the algorithm,

as the initial guess of the state covariance matrix is soon updated after a few iterations of the algorithm. Also, the values of the entries matrices  $\mathbf{Q}$  and  $\mathbf{R}$  are not important if considered alone, but only in a comparative fashion. In fact, by appropriately tuning the entries of the two matrices, one can give more importance either to the history of values (by assuming that the noise of the observations is larger than the noise of the process) or to the last received values (by assuming that the noise of the process is larger than the noise of the observations). In the following simulations, we chose  $\mathbf{P}_{0|0} = \mathbf{I}$ ,  $\mathbf{Q} = 10 \cdot \mathbf{I}$  and  $\mathbf{R} = \mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix of appropriate size, as a consequence of a tuning phase for data in year 2010.

For what concerns the ESN approach, reservoir hyper-parametrization and readout regularization (see Section 2.3.4) were chosen according to a hold-out model selection scheme, on a validation set containing  $\approx 33\%$  of the training data, corresponding to the months September - December of the year 2010. Moreover, in order to produce a single value for the electricity price forecast at each hour, for the results presented in this Section we considered the ensemble of the ESN guesses selected by the model selection process.

Figure 2 shows a comparison among the electricity price values and the predicted ones obtained by EA, KF and ESN. For the sake of graphical representation, such comparison takes into consideration an arbitrarily chosen time period of 10 days, i.e., from March 1<sup>st</sup> 2011 until March 10<sup>th</sup> 2011. As can be seen from Figure 2, the qualitative result of the electricity price prediction achieved by the three predictive models is good, being able to closely follow the trend of the actual electricity price dynamics. The MAE, the RMSE and the MAPE achieved by the basic setting of the proposed predictive models on the whole year 2011 are reported in Table 2, where the results of the day before algorithm are reported as well, for the sake of baseline performance comparison. Table 2 also reports the performance achieved on the 2011 data by the improved versions of the predictive models. Results in Table 2 quantitatively indicate that good performance is obtained by all the EA, KF and ESN approaches, which turned out to outperform the reference performance of the day-before forecast for every setting. Although an exhaustive comparison among the possible predictive models is out of the scopes of this paper, Table 2 also reports the performance achieved by a significant instance in the class of learning models for flat data, i.e., a standard MLP. Experiments on MLPs were conducted according to the same experimental input-output setting adopted for ESNs (see Section 2.3.4), and using gradient descent with momentum backpropagation with regularization and early

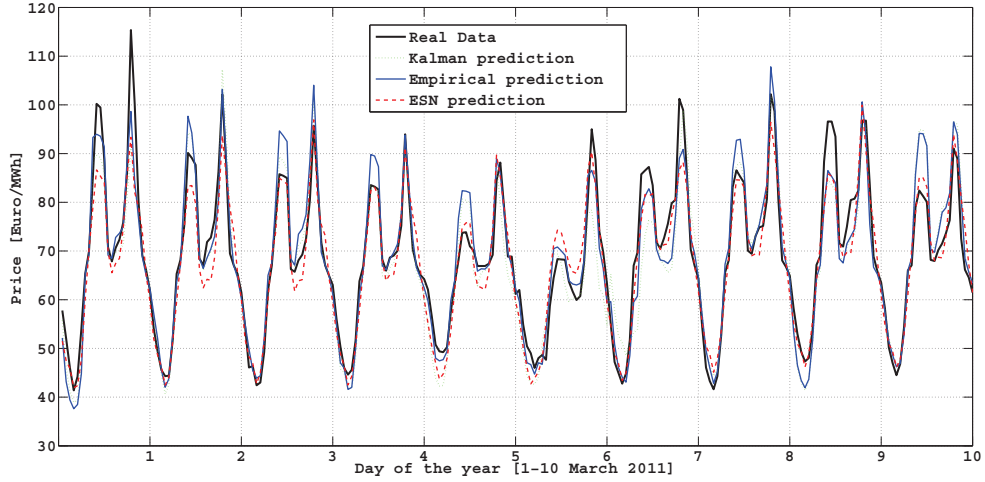


Figure 2: Comparison among the electricity price values and predictions obtained by EA, KF and ESN, over a reference period of 10 days (1-10 March, arbitrarily chosen) in year 2011.

Table 2: Hourly MAE, RMSE and MAPE obtained by the basic and the improved settings of the predictive models on the 2011 data. For performance comparison, the results obtained by the day before algorithm and by MLP are reported as well.

<b>Predictive Model</b>	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>
Day Before	6.83	10.28	9.49
MLP	8.37	11.71	11.44
EA	6.39	9.04	8.72
KF	6.42	9.33	8.88
ESN	5.70	7.98	7.90
EA Improved	5.67	8.34	7.86
KF Improved	6.27	9.15	8.71
ESN Improved	5.25	7.57	7.30

stopping for training. Results reported for MLPs correspond to a number of 30 hidden units, as a result of a model selection procedure implemented using the same validation set adopted for the ESNs experiments. Results in Table 2 show that the MLP does not achieve the same performance level of the other considered learning models for sequential data, empirically suggesting that adopting learning models which are able to process the involved information directly in the form of time series data might be more appropriate for this specific task.

As a general observation for the other methods, it is possible to appreciate the beneficial impact of the effective modeling of the problem (based on both daily and weekly scale dynamics, see Section 2.1), and of the non-linear input-output ESN dynamics. Overall, out of a range of possible values from 10 to 164.80 euros, hourly MAE and RMSE values are not larger than 6.42 and 9.33 euros per *MWh*, respectively, corresponding to a MAPE value of 8.72%, which represents the result obtained by the KF model, whereas the best performance is obtained by the ESN, with MAE and RMSE of 5.70 and 7.98 euros per *MWh*, respectively, and MAPE of 7.90%. From the results in Table 2, the benefit obtained by daily re-training of the models' parameters is also evident, with hourly performance improvements in the order of approximately 20 – 70 cents of euros per *MWh*, corresponding to a MAPE reduction approximately in the range 0.2 – 0.9. In this regard, although the relative improvements of absolute errors are below 10%, it is worth to note that the yearly Italian energy consumption is around 300000*GWh* per year. Accordingly, the availability of accurate electricity price forecasts is of paramount importance for an optimal operation of a microgrid, where an improvement of even a few cents per *MWh* is amplified by volumes of energy of the order of *GWhs*.

Another interesting experimental analysis which can be conducted by referring to the basic experimental setting pertains to the influence that the individual parameters of the learning models have on the final performance. Regarding the Kalman filter approach, Table 3 shows the (minimum-maximum) range of MAE test performance obtained by adopting all the possible values of the matrixes  $\mathbf{Q}$  and  $\mathbf{R}$  considered in section 2.3.3. The lower MAE value is found by taking  $\mathbf{R} = \mathbf{Q}$ , while the larger MAE value is found for  $\mathbf{R} = 10\mathbf{Q}$ . For what concerns the ESN, Table 3 shows the range of MAE obtained by selectively excluding one hyper-parameter from the model selection process and evaluating the performance of the selected models for every possible value of the excluded hyper-parameter. Details

Table 3: Influence of single hyper-parameters on the test performance achieved by Kalman Filter and ESNs in the basic experimental setting.

<b>Kalman filter</b>	<b>Test MAE range</b>
<b>Q/R choice</b>	6.32 – 7.07
<b>ESN Hyper-parameter</b>	<b>Test MAE range</b>
reservoir dimension $N_R$	5.68 – 6.50
spectral radius $\rho$	5.70 – 6.50
leaky parameter $a$	5.70 – 6.73
readout regularization $\lambda_r$	5.70 – 7.34

about the ESN hyper-parameters and the range of values considered in our experiments can be found in Section 2.3.4. From Table 3, it can be observed that the hyper-parameter with a greater influence on the ESN performance is the readout regularization parameter  $\lambda_r$ , while the other hyper-parameters have approximately the same influence on the final result.

Table 4: Hourly MAE, RMSE and MAPE obtained by the predictive EA, KF and ESN models under the daily re-training setting on the 2011, 2012 and 2013 data. For baseline performance comparison, the results obtained by the day before algorithm are reported as well.

	<b>Predictive Model</b>											
	<b>Day before</b>			<b>EA</b>			<b>KF</b>			<b>ESN</b>		
	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>	<b>MAE</b>	<b>RMSE</b>	<b>MAPE</b>
<b>2011</b>	6.83	10.28	9.49	5.67	8.34	7.86	6.27	9.15	8.71	5.25	7.57	7.30
<b>2012</b>	9.55	14.81	13.05	7.99	12.01	10.85	8.54	12.75	11.59	7.59	11.59	10.3
<b>2013</b>	8.94	13.92	15.05	6.87	10.00	11.47	7.71	11.33	12.99	6.51	9.25	10.99
<b>2011-2013</b>	8.44	13.15	12.53	6.84	10.23	10.06	7.51	11.18	11.1	6.45	9.61	9.53

Figure 3 shows a qualitative comparison among the actual electricity prices and the predictions provided by the EA, KF and ESN in the improved setting, with daily re-training over the years 2011, 2012 and 2013. Such results confirm the accuracy of the proposed approaches. Indeed, the performance achieved by EA, KF and ESN clearly outperform the ones obtained by the baseline reference of the day-before forecasting, both if compared on single years results and if compared on the average of the three years consid-

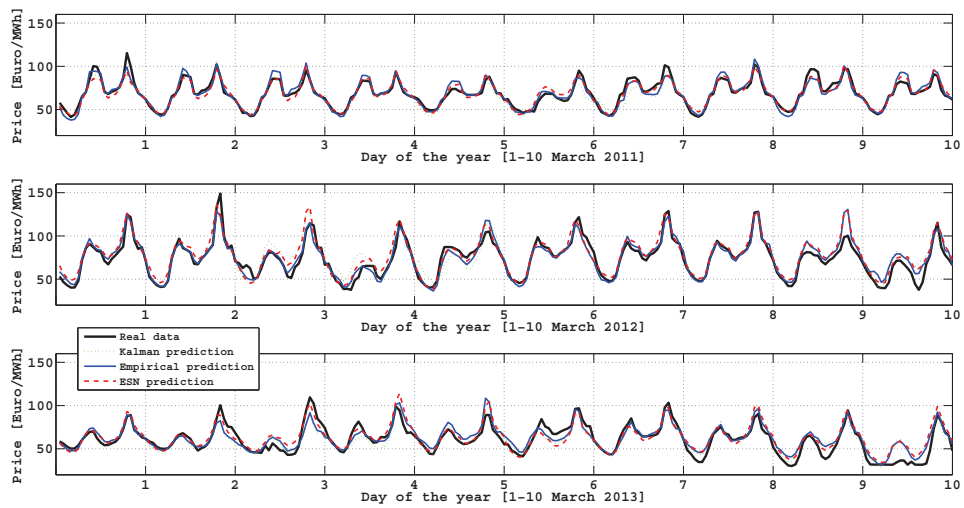


Figure 3: Comparison among the electricity price values and the predictions obtained by EA, KF and ESN with daily retraining over a reference period (1-10 March, arbitrarily chosen) of the years 2011, 2012 and 2013.

ered. In the last case, also, the best results are achieved by the ESN model, which outperforms the EA and the KF approaches. The improvements are about 40 euro cent and 1 euro respectively, based on the MAE index, and are consistent over the 3 years being evaluated. Out of a total range of possible price values from 0 to 324.20 euros, ESNs achieved an hourly MAE of 6.45 euros per *MWh*, an hourly RMSE of 9.61 euros per *MWh* and MAPE of 9.53% on the average of the years 2011, 2012 and 2013 data. Table 4 reports the hourly MAE, RMSE and MAPE for the same models, detailing the results on single years and on the average of the three years 2011-2013.

Apart from the numerical results, at a more general level it is important to remark that another advantage of the ESN approach is given by the fact that it can be easily adapted to modify the horizon of prediction. This is particularly useful if one is interested in predicting prices, say, 48 hours in advance. On the other hand, significant changes to the EA and to the KF algorithms must be introduced to take into account the possibility of changing the prediction horizon to an arbitrary value. At the same time, it is important to remark that the results illustrated in this Section compare the performance obtained by the different predictive models for the electricity price prediction on a next-day scale. However, the day before-based approaches, such as the EA and the KF, can be used to forecast the electricity price for any hour of interest  $h$  with an advance of exactly 24 hours, and the results in Tables 2 and 4 remain unchanged. Indeed the day before-based approaches provide piece-wise constant predictions, such that the prediction for each hour  $h$  is maintained constant for all the following 24 hours (see Section 2.3.1).

It is also noteworthy to point out that the retraining strategy proposed in this paper to improve the electricity price forecasting can be adopted in the actual forecasting practice, as indeed for every considered predictive model the computational time required for daily retraining is quantifiable in the order of seconds.

#### 4. Conclusion

In this paper we have presented different predictive models for the problem of forecasting the electricity price with an advance of one day. Our investigations have originated from the identification of two main periodical patterns in the electricity price dynamics, namely a daily pattern and a weekly pattern. Accordingly, the problem of electrical price forecasting has been formulated as a problem of time series prediction, focusing on predic-



tive models which can directly treat sequential information and taking into consideration a variety of approaches based on empirical strategies common in the smart grid community, Kalman Filters and neural networks. In particular, the Echo State Network, which represents a state-of-the-art neural network approach for efficient sequence processing, has been suitably adapted to model the application context at hand, on the basis of the periodical patterns mentioned above. In addition to the basic formulations, the models proposed for the electricity price prediction have been further tailored to the temporal evolution of the considered phenomena, by including strategies based on daily re-adaptation.

Recent real world data from the Italian market, pertaining to the years 2010-2013, have been used to assess and compare the performance of the proposed predictive models. Results on such data have shown good overall performance, providing experimental evidences that confirm both the hypotheses of the goodness of problem formulation and of the usefulness of the re-training strategy. The best performance has been obtained by the ESN model, which was found able to provide accurate electricity price estimations, with a MAPE of approximately 9.5%, over the whole range of price values.

A natural follow-up of this work consists in extending the current predictive models by including additional information, e.g., the expected load consumption, the expected power generated from renewable sources and weather forecast data. While it is generally acknowledged that all the previous variables do have an impact on the price of electrical energy, it is still questionable whether the accuracy of such predicted variables is enough to improve the current price predictions. In this perspective, the proposed work also represents a benchmark result that can be used to test the accuracy and the usability of such complementary information. Moreover, we also believe that already in the present form, the proposed approaches, due to their efficiency and simplicity, and especially when such extra information are not available or accurate, represent a step towards a more accurate forecast of the electricity price.

## **Acknowledgments**

The authors would like to thank the (anonymous) reviewers whose remarks contributed to improve the quality and the clarity of the final version of this article.

## References

- [1] P. Fox-Penner, *Smart Power: Climate Change, the Smart Grid, and the Future of Electric Utilities*, Washington DC, 2010.
- [2] P. Lombardi, M. Powalko, K. Rudion, Optimal Operation of a Virtual Power Plant, in: *Power & Energy Society General Meeting, Proceedings of the, IEEE*, 2009.
- [3] D. Pudjianto, C. Ramsay, G. Strbac, Virtual power plant and system integration of distributed energy resources, *Renewable Power Generation, IET on 1 (1) (2007) 10–16*.
- [4] Y.A.-R.I. Mohamed, A.A. Radwan, Hierarchical Control System for Robust Microgrid Operation and Seamless Mode Transfer in Active Distribution Systems, *Smart Grid, IEEE Transactions on 2 (2) (2011) 352–362*.
- [5] F. Zaro and M. Abido, *Multi-Objective Optimal Power Flow: Deregulation Perspective*, Lambert Academic Publishing, Saarbrücken, Germany, 2010.
- [6] A. Parisio, L. Glielmo, A mixed integer linear formulation for microgrid economic scheduling, *Smart Grid Communications (SmartGridComm), IEEE International Conference on*, 2011.
- [7] D. Aloini, E. Crisostomi, M. Raugi, R. Rizzo, Optimal Power Scheduling in a Virtual Power Plant, *PES Innovative Smart Grid Technologies (ISGT) Europe, IEEE*, 2011.
- [8] A.-H. Mohsenian-Rad, A. Leon-Garcia, Optimal residential load control with price prediction in real-time electricity pricing environments, *Smart Grid, IEEE Transactions on 1 (2) (2010) 120-133*.
- [9] J. Kolen, S. Kremer (Eds.), *A Field Guide to Dynamical Recurrent Networks*, IEEE Press, 2001.
- [10] D. Verstraeten, B. Schrauwen, M. D’Haene, D. Stroobandt, An experimental unification of reservoir computing methods, *Neural Networks 20 (3) (2007) 391 – 403*.

- [11] M. Lukosevicius, H. Jaeger, Reservoir computing approaches to recurrent neural network training, *Computer Science Review* 3 (3) (2009) 127–149.
- [12] H. Jaeger, H. Haas, Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication, *Science* 304 (5667) (2004) 78–80.
- [13] H. Jaeger, The "echo state" approach to analysing and training recurrent neural networks, Tech. rep., GMD (2001).
- [14] C. Gallicchio, A. Micheli, Architectural and markovian factors of echo state networks, *Neural Networks* 24 (5) (2011) 440 – 456.
- [15] M. D. Skowronski, J. G. Harris, Automatic speech recognition using a predictive echo state network classifier, *Neural networks* 20 (3) (2007) 414–423.
- [16] E. A. Antonelo, B. Schrauwen, D. Stroobandt, Event detection and localization for small mobile robots using reservoir computing, *Neural Networks* 21 (6) (2008) 862–871.
- [17] E. A. Antonelo, B. Schrauwen, J. M. V. Campenhout, Generative modeling of autonomous robots and their environments using reservoir computing, *Neural Processing Letters* 26 (3) (2007) 233–249.
- [18] P. Joshi, W. Maass, Movement generation with circuits of spiking neurons, *Neural Computation* 17 (8) (2005) 1715–1738.
- [19] D. Bacciu, P. Barsocchi, S. Chessa, C. Gallicchio, A. Micheli, An experimental characterization of reservoir computing in ambient assisted living applications, *Neural Computing and Applications* 24 (6) (2014) 1451–1464.
- [20] P. Barsocchi, S. Chessa, A. Micheli, C. Gallicchio, Forecast-driven enhancement of received signal strength (rss)-based localization systems, *ISPRS International Journal of Geo-Information* 2 (4) (2013) 978–995.
- [21] C. Gallicchio, A. Micheli, P. Barsocchi, S. Chessa, User movements forecasting by reservoir computing using signal streams produced by mote-class sensors, in: *Mobile Lightweight Wireless Systems (Mobilight 2011)*,

Vol. 81 of Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Springer Berlin Heidelberg, 2012, pp. 151–168.

- [22] D. Bacciu, C. Gallicchio, A. Micheli, S. Chessa, P. Barsocchi, Predicting user movements in heterogeneous indoor environments by reservoir computing, in: M. Bhatt, H. W. Guesgen, J. C. Augusto (Eds.), Proc. of the IJCAI Workshop on Space, Time and Ambient Intelligence (STAMI) 2011, 2011, pp. 1–6.
- [23] D. Bacciu, S. Chessa, C. Gallicchio, A. Micheli, P. Barsocchi, An experimental evaluation of reservoir computation for ambient assisted living, in: B. Apolloni, S. Bassis, A. Esposito, F. C. Morabito (Eds.), Neural Nets and Surroundings, Vol. 19 of Smart Innovation, Systems and Technologies, Springer Berlin Heidelberg, 2013, pp. 41–50.
- [24] F. Palumbo, P. Barsocchi, C. Gallicchio, S. Chessa, A. Micheli, Multi-sensor data fusion for activity recognition based on reservoir computing, in: J. Botía, J. Alvarez-Garcia, K. Fujinami, P. Barsocchi, T. Riedel (Eds.), Evaluating AAL Systems Through Competitive Benchmarking, Vol. 386 of Communications in Computer and Information Science, Springer Berlin Heidelberg, 2013, pp. 24–35.
- [25] S. Chessa, C. Gallicchio, R. Guzman, A. Micheli, Robot localization by echo state networks using rss, in: S. Bassis, A. Esposito, F. C. Morabito (Eds.), Recent Advances of Neural Network Models and Applications, Vol. 26 of Smart Innovation, Systems and Technologies, Springer International Publishing, 2014, pp. 147–154.
- [26] S. K. Aggarwal, L. M. Saini, A. Kumar, Electricity price forecasting in deregulated markets: A review and evaluation, International Journal of Electrical Power & Energy Systems 31 (1) (2009) 13–22.
- [27] R. C. Garcia, J. Contreras, M. Van Akkeren, J. B. C. Garcia, A garch forecasting model to predict day-ahead electricity prices, Power Systems, IEEE Transactions on 20 (2) (2005) 867–874.
- [28] J. Contreras, R. Espinola, F. J. Nogales, A. J. Conejo, Arima models to predict next-day electricity prices, Power Systems, IEEE Transactions on 18 (3) (2003) 1014–1020.

- [29] A.J. Conejo, M.A. Plazas, R. Espínola, A.B. Molina, Day-Ahead Electricity Price Forecasting Using the Wavelet Transform and ARIMA Models, *Power Systems, IEEE Transactions on* 20 (2) (2005) 1035-1042.
- [30] B. Szkuta, L. Sanabria, T. Dillon, Electricity price short-term forecasting using artificial neural networks, *Power Systems, IEEE Transactions on* 14 (3) (1999) 851–857.
- [31] H. Yamin, S. Shahidehpour, Z. Li, Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets, *International journal of electrical power & energy systems* 26 (8) (2004) 571–581.
- [32] N. Amjady, Day-ahead price forecasting of electricity markets by a new fuzzy neural network, *Power Systems, IEEE Transactions on* 21 (2) (2006) 887–896.
- [33] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, A. K. Srivastava, A novel approach to forecast electricity price for pjm using neural network and similar days method, *Power Systems, IEEE Transactions on* 22 (4) (2007) 2058–2065.
- [34] S. Anbazhagan, N. Kumarappan, Day-ahead deregulated electricity market price forecasting using recurrent neural network, *Systems Journal, IEEE* 7 (4) (2013) 866–872.
- [35] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, K. P. Wong, A hybrid approach for probabilistic forecasting of electricity price, *Smart Grid, IEEE Transactions on* 5 (1) (2014) 463–470.
- [36] C. Gao, E. Bompard, R. Napoli, H. Chenga, Price forecast in the competitive electricity market by support vector machine, *Physica A* 382 (1) (2007) 98–113.
- [37] Z. Xu, Z. Dong, W. Liu, Short-term electricity price forecasting using wavelet and svm techniques, in: *Dynamics of Continuous Discrete and Impulsive Systems-Series B-Applications*, Watam Press, 2003, pp. 372–377.

- [38] J. H. Zhao, Z. Y. Dong, Z. Xu, K. P. Wong, A statistical approach for interval forecasting of the electricity price, *Power Systems, IEEE Transactions on* 23 (2) (2008) 267–276.
- [39] E. Crisostomi, M. Tucci, M. Raugi, SVM Methods for Optimal Management of a Virtual Power Plant, in: B. Apolloni, S. Bassis, A. Esposito, C.F. Morabito (Eds.), *Neural Nets and Surroundings*, Springer-Verlag Berlin Heidelberg, 2012, pp. 271-278.
- [40] X. Lua, Z.Y. Dong, X. Li, Electricity market price spike forecast with data mining techniques, *Electric Power Systems Research* 73 (1) (2005) 19–29.
- [41] R. Faria, P. Moura, J. Delgado, A. de Almeida, Managing the Charging of Electrical Vehicles: Impacts on the Electrical Grid and on the Environment, in *Intelligent Transportation Systems Magazine, IEEE, Special Issue on Electro-Mobility* (2014) 54–65.
- [42] R. Halvgaard, N.K. Poulsen, H. Madsen, J.B. Jørgensen, F. Marra, D.E.M. Bondy, Electric vehicle charge planning using Economic Model Predictive Control, *IEEE International Electric Vehicle Conference (IEVC)*, 2012.
- [43] S. A. Ilić, S. M. Vukmirović, A. M. Erdeljan, F. J. Kulić, Hybrid artificial neural network system for short-term load forecasting, *Thermal Science* 16 (suppl. 1) (2012) 215–224.
- [44] A. Deihimi, H. Showkati, Application of echo state networks in short-term electric load forecasting, *Energy* 39 (1) (2012) 327–340.
- [45] M. J. A. Rabin, M. S. Hossain, M. S. Ahsan, M. A. S. Mollah, A. M. E. Kabir, M. Shahjahan, Electrical load forecasting using echo state network, in: *Computer and Information Technology (ICCIT), 2012 15th International Conference on, IEEE, 2012*, pp. 50–54.
- [46] E. Crisostomi, M. Tucci, M. Raugi, Methods for Energy Price Prediction in the Smart Grid, *PES Innovative Smart Grid Technologies (ISGT) Europe, IEEE, 2012*.

- [47] R.E. Kalman, A New Approach to Linear Filtering and Prediction Problems, ASME - Journal of Basic Engineering, Transactions of the 82 (D) (1960) 35–45.
- [48] R. van der Merwe, A. Doucet, F.F.G. de Freitas, E. Wan, The unscented particle filter, Tech. rep. CUED/F-INFENG/TR 380 (2000).
- [49] H. Jaeger, M. Lukosevicius, D. Popovici, U. Siewert, Optimization and applications of echo state networks with leaky-integrator neurons, Neural Networks 20 (3) (2007) 335–352.
- [50] P. Tiño, M. Cernanský, L. Benusková, Markovian architectural bias of recurrent neural networks, IEEE Transactions on Neural Networks 15 (1) (2004) 6–15.
- [51] N. S. Siegel, J. Castellan Jr, Nonparametric Statistics for The Behavioral Sciences, 2nd Edition, McGraw-Hill, 1988.