A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings

Luis Hernandez, Carlos Baladrón, Javier M. Aguiar, Belén Carro, Antonio J. Sanchez-Esguevillas, Jaime Lloret, and Joaquim Massana

***Abstract*—Recently there has been a significant proliferation in the use of forecasting techniques, mainly due to the increased availability and power of computation systems and, in particular, to the usage of personal computers. This is also true for power network systems, where energy demand forecasting has been an important field in order to allow generation planning and adaptation. Apart from the quantitative progression, there has also been a change in the type of models proposed and used. In the '70s, the usage of non-linear techniques was generally not popular among scientists and engineers. However, in the last two decades they have become very important techniques in solving complex problems which would be very difficult to tackle otherwise. With the recent emergence of smart grids, new environments have appeared capable of integrating demand, generation, and storage. These employ intelligent and adaptive elements that require more advanced techniques for accurate and precise demand and generation forecasting in order to work optimally. This review discusses the most relevant studies on electric demand prediction over the last 40 years, and presents the different models used as well as the future trends. Additionally, it analyzes the latest studies on demand forecasting in the future environments that emerge from the usage of smart grids.**

***Index Terms*—Electric demand forecasting, short-term load forecasting, smart grid, *microgrid*, smart building.**

I. Introduction

THIS introduction aims at setting the background for understanding the demand forecasting problem in modern day and future power networks by taking three different approaches. First, in subsection A, the importance of demand forecasting is explained, presenting the reasons why it is an important step in the energy generation process of utilities, and the role it plays for each of the components of the grids and smart grids of today. Secondly, in subsection B, a brief historical perspective of the evolution of demand forecast is given since its beginning, and third, in subsection C, a justification is given of why demand forecasting will continue being a hot topic in the future energy distribution networks. Then, before moving onto Section II, the global structure of this paper is presented.

A. The importance of demand forecasting

From its onset, the electric system has been based on three levels: generation, transport, and distribution and marketing. With the exception of marketing activities, the rest have focused on solving problems from an electrical perspective. As networks grew bigger and bigger, they started to be very difficult to control and lost efficiency due to electric loss. In addition, traditionally the final consumers have not been taken into account in the system except when the moment comes to pay the bill (large-scale consumers received a separate treatment). The bill, in most cases, is based on estimates and not on real data, and the suppliers have a period of time to issue the invoices.

Recently, new players (*Electric Vehicle* (*EV*), Smart Customers, Renewable Energy) have emerged in the electrical system scene that have caused demand forecasting to gain special interest. Currently, *EVs* are gradually being introduced into the electrical system, and it is expected that the increasing influence of these elements will eventually change overall load profile significantly. This impact comes from the fact that *EVs* are mobile elements that do not only consume power, but can potentially also contribute electricity to the grid. Additionally, it is expected that most of the residential charge operations will take place during the night, in a typical valley period of the demand curve, where the usage of vehicles is reduced. It is believed that *EVs* would then be seized to flatten the consumption curve, taking energy from the grid overnight, and providing energy during peak periods. In any case, it is evident that the massive introduction of these new elements will in any case modify the behaviour of the grid at a very significant scale.

Apart from demand forecasting, electrical generation forecasting models have also received increasing attention, especially when it comes to renewable generation sources, which in turn depend on the forecasting of a particular energy resource (solar radiation, wind, etc.).

Manuscript received April 2, 2013; revised October 16, 2013.

L. Hernandez is with Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Centro de Desarrollo de Energías Renovables (CEDER), Autovía de Navarra A15, salida 56, 42290 Lubia, Soria, (Spain), (e-mail: luis.hernandez@ciemat.es).

C. Baladrón, J. M. Aguiar, B. Carro, and A. J. Sanchez-Esguevillas are with Universidad de Valladolid, E.T.S.I. Telecomunicación, Campus Miguel Delibes, Paseo de Belén 15, 47011 Valladolid (Spain) (e-mail: cbalzor@ribera.tel.uva.es, javagu@tel.uva.es, belcar@tel.uva.es, antsan@tel.uva.es).

J. Lloret is with Universidad Politécnica de Valencia, Departamento de Comunicaciones, Camino Vera, s/n, 46022, Valencia (e-mail: jlloret@dcom.upv.es).

J. Massana is with eXiT Research Group, Institute of Informatics and Applications, University of Girona, Campus Montilivi, Building P4, 17071 Girona (Spain) (e-mail: quim.massana@hotmail.com).

Digital Object Identifier 10.1109/SURV.2014.032014.00094

Next, the main changes occurred in the electrical system, and the actors they have brought in, are explained.

1) Smart Grid (SG): Mantooth (2010) [1] defines *SGs* as hardware and software added to the power system to achieve a more autonomous responsiveness to events that impact the electrical power grid, and optimal day-to-day operational efficiency of electrical power delivery.

The main requirements of a future electrical power grid can be summarized as follows (Ferreira *et al.*, 2010) [2]: (1) flexible: grids must satisfy customers' needs, while at the same time responding to uncertain events (future changes and challenges); (2) accessible: the access of all the users to the grid must be assured, including renewable generation sources and the local units of generation that are more and more frequent; (3) reliable: grids must guarantee a reliable and high-quality delivery in accordance with the requirements of an economy that depends on electrical energy; (4) innovative: from an economical perspective, competition and regularization will foster economic development and evolution of *SGs*. Innovation will be the key to efficient energy management.

The aforementioned requirements lead to the following *SG* goals, according to Sánchez-Jiménez (2006) [3]: development of low-cost and easy-to-deploy technical solutions, as well as adoption of technical standards and protocols that can ensure a wider set of solutions; development of information, computation, and telecommunication systems with business purposes, so as to provide innovative services to the system and its customers; agreement on regulatory and commercial frameworks in different regions for the adopted solutions, as well as on uniform working systems to achieve economies of scale; leverage of the existing infrastructure, synchronizing past and new model designs.

SGs have the distributed intelligence needed to operate efficiently in the aforementioned framework. The importance of reliability and financial considerations, as well as end customers, will make forecasting models essential for the electric utility to harness energy at the lowest cost, and to guarantee optimal quality of the energy supply.

Having real-time information and realistic consumption patterns will allow for more sophisticated forecasting models. Additionally, the fact that information is constantly changing will call for online adaptation of forecasts.

2) Distributed Generation (DG): Distributed Generation (*DG*) builds upon the above framework by providing a new opportunity for renewable energy sources and by placing them in new locations, particularly in the vicinity of consumption points.

It relies on low-power electricity generation technologies being near the end users who consume it. Systems using *DG* resort to multiple small plants, and provide energy to their immediate surroundings. Thus, there is almost no dependence on the distribution and transport network. The power of *DG* will range between one kilowatt (kW) and hundreds of megawatts (MW).

Therefore, renewable generation sources have a new chance thanks to *DG*. Additionally, governments should see *DG* as an opportunity to produce jobs locally. An example of this is the UK, where a sustainable tariff system (feed-in-tariffs) has been in place for years. The future network will have to live together with centralized plants and *DG*, as pointed out by Nissen (2009) [4] and *DG* will certainly be a key ingredient for developers and urban planners if you take into account the mutual impacts of *DG* and current network-configuration, as indicated by Hamedani and Arefifar (2006) [5], which calls for the use of algorithms to optimize control.

Distributed generation makes the control of networks even more uncertain. Demand forecasts will have to be estimated when the network is operating, but taking into account generation under its control. If the *DG* sources are renewable, the random nature of the asset will make models more complicated as results of demand forecasting models will depend on renewable source generation estimates.

3) Virtual Power Plant (VPP): Within *SGs* a new energy production management model emerges—VPP—where the energy plant is no longer a monolithic facility, but a combination of smaller cooperating and intelligent elements, as described by Wille-Haussmann *et al.* (2010) [6]. The aggregation of generators and loads form a single autonomous physical or logical unit, where all elements work coherently. Examples of this could be an industrial park, a residential area with photovoltaic modules or the distributed resources of a utility.

However, managing these *VPP* poses a challenge for *Information and Communication Technologies* (*ICTs*). *VPPs* consist of a number of elements designed to solve local problems, but needing to interact in order to behave as a single unit. Communication protocols and tools are therefore necessary for these elements to communicate, take decisions that affect several entities, and coordinate their behavior so that they can carry our complex tasks, such as the ones presented by Sebastian and Maire (2008) [7]. At the same time, the devices that can be found in a *VPP* are extremely heterogeneous; they are built and supplied by very different organizations with diverse purposes and they are likely to follow different rules when it comes to control and communication.

As showed by Bel *et al.* (2007) [8], the aggregation of small generators and their controlled loads by means of a *VPP* can provide a chance to small production units, and thus reach financially competitive energy production levels, as well as coordinating the elements to offer full services to the network.

In short, *VPPs* imply the coordination and control of distributed low-power generators. In addition, and regardless of the particular offers of the *VPPs* to the electric market, these will have to take into account the demand they are controlling. Thus, the need for electric demand and generation forecasting is again made clear.

4) Microgrid: According to Lasseter *et al.*, (2002) [9], *microgrids* can be defined as an aggregation of loads and microsources operating as a single system providing both power and heat. This new paradigm, a system that aggregates multiple generation technologies and loads, is in a unique position to take the chances created by the liberalized market and, in addition, contributes to an optimized and more rational electrical network.

The definition of *microgrid* provided by the *Consortium for Electric Reliability Technology Solutions* (*CERTS*) does not specify whether it works isolated from the distribution network or connected to it. Anduaga *et al*. (2008) [10] provide a new definition, where a *microgrid* is a system formed by generation sources, storage equipment and electrically connected loads that can be connected to the main system or isolated from it when electrical perturbations occur, and that is controlled by the system operator as an aggregated system without having to plan or manage the generated and consumed power. As can be observed, this definition does not specify if the connection must be at high or medium voltage, nor does it restrict power. These factors depend on the *microgrid* application, the maximum demand to be served, and the location. Therefore, a *microgrid* could consist on the elements present at a building (then, the term nanogrid could be used) or be the size of a small town.

Regardless of its size, for the *microgrid* to work properly and interact with the distribution network, the demand and generation needs in its domain must be known. Hence, the development of applications focusing on forecasting models should be a priority.

5) Smart Building and Smart Environment: Building upon the concept of a *microgrid*, new environments have emerged that deserve consideration. The intelligence deployed at buildings, together with the integration of renewable generation sources, has brought in new concepts, such as *Smart Buildings* and *Smart Environments*. These new environments imply deploying interconnected sensors and intelligent devices in order to obtain energy efficient buildings (*Smart Buildings*) and make the life of the users easier (*Smart Environments*).

In *Smart Buildings*, the integration of renewable generation sources and the interaction of the building itself with the electrical network will call for demand and generation forecasting models. These models, together with other software models, must be integrated within the *Energy Management System* (*EMS*). As for smart environments, it is essential to retrieve information about the "*energy behavior*" (electricity, water, gas, etc.) of the environment inhabitants. Again, this requires forecasting models and applying them directly to the specific variables of the environment to be controlled.

B. Demand Forecasting In Future Environments

As has been explained before, the electrical system model has started to change recently, with the gradual integration of intelligence at the different levels of the electric system: transportation, distribution and, to a lesser extent, the end user level (more recently). This is where *SGs* come into play, since they use *ICTs* to optimize power production and distribution so as to match producer supply with consumer demand. In fact, *SGs* can be defined as "*smart electric power networks*". *SGs* also seek to optimize energy flows and improve fault detection algorithms, thereby improving the network's operation and taking advantage of the sensors and intelligence deployed, as explained by Fang *et al.* (2012) [11].

The goals of the new environments, including smart buildings, smart homes, and smart environments, can be grouped in a new concept called *Smart World*, as explained by Hernández *et al*. (2012) [12]. This new paradigm will enable the creation of new applications, tools and services for all environments, whose goals will be: infrastructure operation, maintenance, and optimization; safety and security; mobility and transport; urban planning; power saving and energy efficiency; sustainability; evolution and environmental control; and quality of life.

In short, according to Werbos (2011) [13], the emergence of new environments in the future will require distributed intelligence and, among other things, new models and applications based on *Artificial Neural Networks* (*ANNs*), not only for demand forecasting but also for the integration of the new actors into the electric system. As Yan *et al.* (2012) [14] point out, apart from distributed intelligence, digital communications and control will have to be carried out safely. Javed *et al.* (2012) [15] show the need for new forecasting models to deal with the requirements set by *Demand Response* (*DR*).

The forecasting models (electric demand and generation) required in the said environments will enable the use of local variables that can have a direct impact on demand or generation behavior. These local variables (climatic, social, economic, those related to habits, etc.) will bring added complexity to the global model but will make estimates more accurate.

However, going back to electric demand, estimates in these new disaggregated environments (*microgrids*, *Smart Building*, *Smart Environment*, etc.) will be more complex. For example, a country's load curve has a much smoother and more predictable profile than that of disaggregated environments (electrical substation, *microgrid*, *Smart Building*, etc.); indeed, the load curves of the latter are more abrupt and sometimes completely atypical, compared to those obtained as more and more load curves are aggregated. Therefore, if models have information that helps to detect this, estimations for that environment will be much more accurate.

Electric demand is clearly dependent on the sector (industrial, commercial, individual, administration, etc.), but disaggregating it and focusing on a particular environment enables a clearer understanding of it by sector. This will allow the implementation of *DR*, since different energy prices could be offered based on the criteria set by the commercializing company. In order to have this knowledge by sectors and to enable forecasting based on them, a clear idea of the consumption habits will be required.

The remaining of this work is organized as follows. Section II presents two different classifications of demand prediction techniques, both from the point of view of the forecasting horizon and according to the aim of the prediction. Section III details the current state of the art for demand prediction divided in linear and non-linear models. Section IV describes the specific demand prediction models available and studied for disaggregated, local environments. Section V provides a comprehensive comparison of all the previously studied solutions. Finally, Section VI summarizes the conclusions.

II. Classification of Demand Forecasting Techniques

This Section classifies demand forecasting models according to two different criteria: the forecasting horizon and the aim of the forecast, with Section III detailing the state of the art divided in linear and non-linear models.

A. According to the forecasting horizon

Hippert *et al.* (2001) [16] indicate that electric demand can be classified following different criteria. Based on the period of time to be predicted, commonly known as forecasting horizon (or forecast horizon), we can find:

- *Very Short-Term Load Forecasting* (*VSTLF*): from seconds or minutes to several hours. These models are generally used to control the flow.
- *Short-Term Load Forecasting* (*STLF*): from hours to weeks. These models are generally used to adjust generation and demand, and to therefore launch offers to the electrical market.
- *Medium-Term and Long-Term Load Forecasting* (*MTLF* and *LTLF*): from months to years. These models are normally used to plan asset utilities.

The most important forecasting horizons are weekly, daily and hourly. Producing an accurate forecast of the next 24 hours is essential for power companies, since it can have a direct impact on the optimal hourly planning of the generation units, as well as on purchase/sale in exchange systems, etc. This is called "*load profile*". Nevertheless, Toyoda *et al.* (1970) [17] present forecasting models for all the aforementioned horizons.

Regardless of the model (all of them will be covered later in this review), the main difference among the three is the scope of the variables used. *VSTLF* models use recent input variables (minutes or hours), *STLF* models use input variables in the range of days, and *MTLF* and *LTLF* use input variables in the range of weeks or even months. Next, some studies on *VSTLF*, *MTLF* and *LTLF* will be analyzed.

Charytoniuk and Chen (2000) [18] present an approach to *VSTLF* based on *ANNs* to model load dynamics in changing environments; they use five different networks to calculate five time intervals with 10-minute spacing. Guan *et al.* (2013) [19] present a model based on wavelet-*ANNs* with pre-data filtering; the *Wavelet Neural Network* (*WNN*) uses as an input the filtering from the wavelet, calendar values (time, month, etc.), and consumption values for the last hour; 12 dedicated *WNNs* are used to forecast the consumption value for the following hour, each *WNN* corresponding to 5 a minute-period.

Doveh *et al.* (1999) [20] present different *ANNs* to obtain *MTLF*, where the forecast horizon is almost a year. The input variables used in the model are heterogeneous, including temperature, day of the week), a fuzzy indicator for the seasonal effect, social and financial variables.

Asber *et al.* (2007) [21] demonstrate how one can use a general load-modeling framework to systematically extract the essence of specific modeling problems in *MTLF* and obtain practical models.

Kandil *et al.* (2002) [22] implement a knowledge-based expert system to support the choice of the most suitable load forecasting model for *M/LTLF* power system planning. The expert system's knowledge base consists on static (histories of load patterns, peak data, power consumers, economic, social and weather) and dynamic (load and energy attributes, loss systems, and estimation errors) facts.

Zhang and Ye (2011) [23] use a regression *ANN* for *LTLF*, including an economic factor, the *Gross Domestic Product* (*GDP*). It is interesting to note that this system needs data from the previous 13 years to train the model, and is able to predict 5 years. The authors claim that *LTLF* models require additional information, apart from load time series.

Daneshi *et al.* (2008) [24] present two models for *LTLF*. The first applies a linear regression method, where the most complex part is to find the equation for the regression model. The second uses *ANN*, applying a fuzzy function to the data set. Regardless of the model, a large number of parameters are used: historical data of curves and weather; time factors; customer information; economic and demographic data and their forecasts; energy prices; regional development of the country; domestic sales of air conditioning and similar equipment; and random perturbations.

Figure 1 summarizes the types of forecasting according to the forecast horizon. The analyzed references have been provided because they were used to detect distinctive features when it comes to the scope of input parameters for each type of horizon. The type of the parameters used has also been included.

B. According to the aim of the forecast

It is also interesting to classify forecasting models based on the number of values to predict. Two main groups exist: the first group is formed by those that forecast only *one value* (next hour's load, next day's peak load, next day's total load, etc.); the second group consists of forecasts with *multiples values*, such as next hours, peak load plus another parameter (for example, aggregated load) or even next day's hourly forecast- the so-called *load profile*.

Next, we analyze some relevant works that fall in the first group because they only seek to forecast a value. This kind of challenge was already taken up in the mid-20th century; Gillies *et al.* (1956) [25] present a manual peak load forecasting model. The process requires a weather prediction that is then converted to the illumination parameter. An illumination vs. peak load table is then used to obtain the increase of the peak load and finally, this increase is applied to curve trends to come up with the forecast.

As far as works dealing with models that predict *multiple values*, Park *et al.* (1991) [26] use three small *ANNs* to forecast next hour's load, next day's peak load and next day's total load. Ho *et al.* (1992) [27] predict next day's peak load, which is later used as an input in an expert system implemented by the same authors (Ho *et al.*, 1990 [28]) to predict next day's *load profile*.

Sometimes, *ANNs* have only an output variable that can be used repeatedly to forecast profile load, as in Drezga and Rahman (1998) [29], and Drezga and Rahman, 1999 [30]. Another alternative is to use a parallel system with 24 *ANNs*, as shown by McMenamin and Monforte (1998) [31].

Lee *et al.* (1992) [32] obtain profile load by dividing the day into three periods and using a different *ANN* for each of them. Lu *et al.* (1993) [33] experiment with three *ANN* models from two utilities and conclude that the systems should be adjusted independently when exchanging them. Papalexopoulos *et al.*

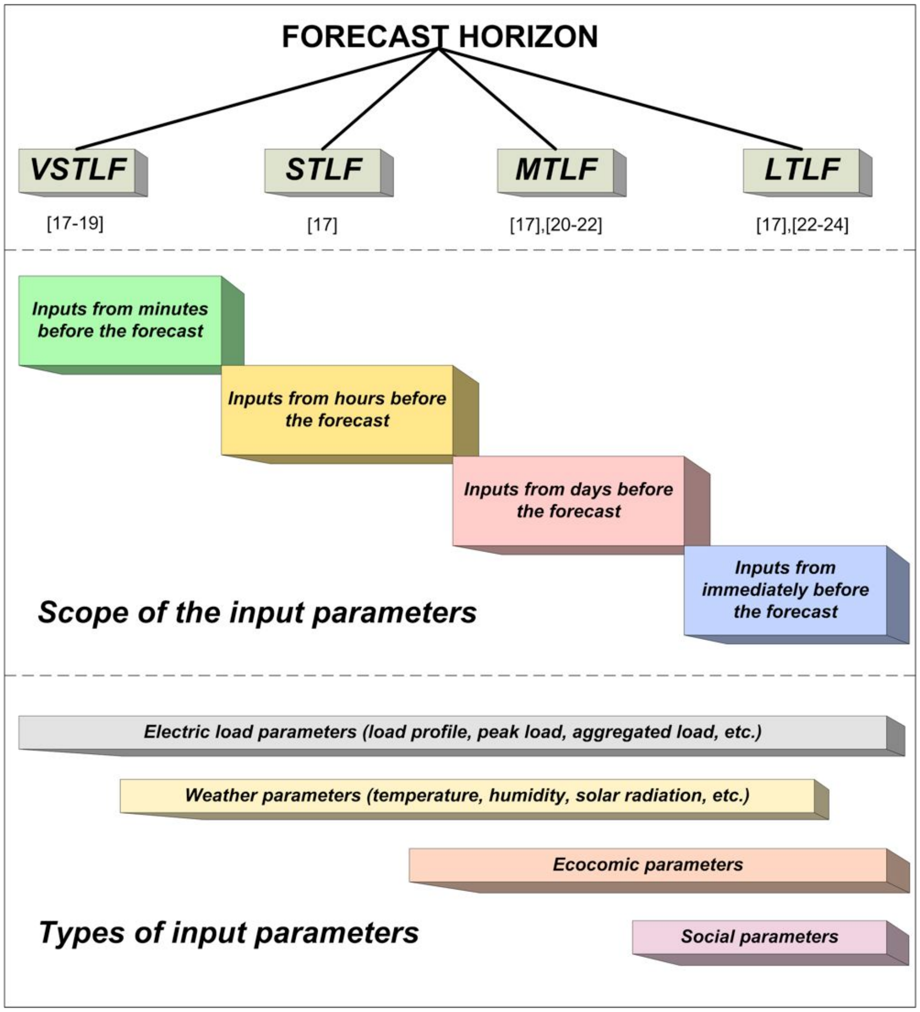


Fig. 1. Types of load forecasting models according to the forecast horizon Identification of the scope and type of input parameters.

(1994) [34] include temperature functions as inputs and also suggest a method to improve forecast for holidays. Bakirtzis *et al.* (1995) [35] present an enhanced model with respect to previous works, focusing on holidays.

As previously mentioned, some of the works suggest systems where multiple *ANNs* work together to obtain the forecast. Alfuhaid *et al.* (1997) [36] use a small *ANN* to preprocess some of the data and produce estimates of the peak load, valley load and next day's peak, which then feed another *ANN*, together with other input data, to obtain next day's *load profile*. Lamedica *et al.* (1996) [37] propose 12 *ANNs*, one for each month of the year, where the *load profiles* are classified using a *Self-Organizing Map* (*SOM*). Mohammed *et al.* (1995) [38] classify hourly loads according to the season, in 7 groups; each group was then modeled by an independent way, forming a large system. Piras *et al.* (1996) [39] use a neural-gas network to sort the data in two groups (summer and winter), and model them with different *ANNs*.

Figure 2 summarizes the types of forecasting according to the aim of the prediction. The analyzed references have been indicated. It can be seen that 1-value forecasts are used for online operation and optimization of load flows, whereas multiple-value forecasts are used for generator scheduling and economic dispatching.

III. Linear and Non-Linear Models

The previous Section has presented possible classifications of demand forecasting, not taking into account the model used. The discussion on the type of model has been considered important enough to deserve a separate section, since it defines the theoretical approach behind each specific solution. All the aforementioned discussion was relevant to the aim of the forecast, but the choice of which model to use is the decision of the researcher.

Next, when it comes to demand forecasting models, the two main types are presented. These types represent two different approaches: *linear* and *non-linear* approach. A third group could consist of models that use a combination of both.

It will later become clear that non-linear models (based on *ANNs*) have gained more and more attention since the second half of the 80's. This evolution is due to the fact that certain researchers achieved great advances on *ANNs*. Among them was James Anderson, who presented a model based on the principle that the connection among neurons are reinforced every time they are activated and named it *Brain-State-in-a-Box* (Anderson *et al.*, 1977 [40]). The same year, at the University of Helsinki, Teuvu Kohonen presented an enhancement of Anderson's model, coming up with *SOM* (Kohonen,

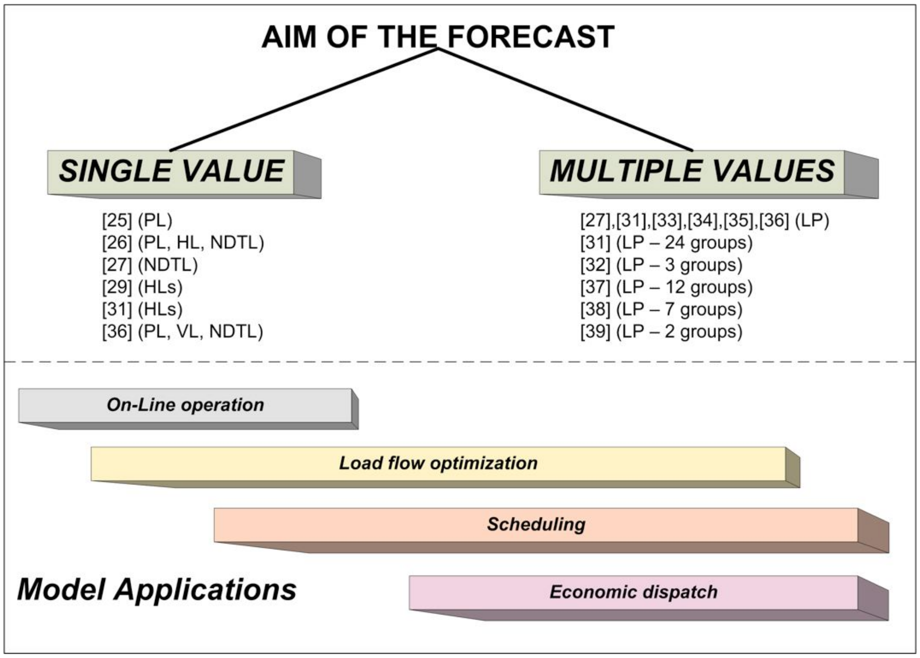


Fig. 2. Types of load forecasting models according to the aim of the prediction. The possible uses of the different models are shown. PL: peak load; VL: valley load; HL: hourly load; NDTL: next day total load; LP: load profile.

1977 [41]). Three years later, in Japan, Kunihiko Fakushima presented a network model that used visual recognition of patterns (Fakushima, 1980 [42]). In 1982, the interest in neural networks re-emerged due to several reasons one of which was the work presented by John Hopfield (Hopfield, 1982 [43]); Hopfield describes clearly and using mathematics, a network named after him and showed how these networks could work. From 1985 on, international conferences on neural networks started to be held all along the planet.

A. Linear models

Before starting, it must be recalled that *linear* models are based on synthesizing all the features of the problem to be solved in more or less complex equations. As we already know, electric demand forecasting is a non-trivial problem, regardless of the forecast horizon or the aim of the prediction. The demand presents a number of *non-linearities* — a great amount of variables and events — which need to be detected and later transferred into the equation/s that model it.

In this Section, the focus is not on the forecast horizon or on the aim of the prediction, but it must be recalled that those models were also grouped depending on whether they predicted *peak load* or *load shape*. The latter are the ones that will set the starting point for the analysis of the different *linear* models used so far.

Load shape models process time series over a specific time interval and among them are *load profile* models, and therefore *STLF* forecasts. *Load shape* and peak load models were sometimes combined, as in Goh *et al.* (1986) [44]. *Load shape* has two variants: *time-of-day* and *dynamic* models.

1) Time-of-day models: Time-of-day models define the load $L(t)$ in each discrete point in time t of a given time series, with prediction period T.

$$\{L(t), t = 1, 2, \ldots, T-1, T\} \tag{1}$$

The model stores T load values based on the previous observations of the said load. Some models store the load curves of several preceding weeks, while others only store that of the previous week. Later, and at the discretion of the operator, the demand has to be forecast by means of formulas. Therefore, the use of expert systems to emulate the operator as well as the application of rules, is common, as shown by Rahman and Bhatnagar (1987) [45]. Expert systems are a branch of artificial intelligence, and are thus close to *ANNs*, but they have been included in this Section because they require a great deal of human participation to create the rules of the knowledge base.

A common form of a *time-of-day* model is:

$$L(t) = \sum_{t=1}^{N} \alpha_i f_i(t) + \nu(t) \tag{2}$$

where the load at the point in time t, $L(t)$ is the sum of a finite number of explicit time functions $f_i(t)$, generally sine waves with a period ranging from 24 to 168 hours, depending on the expected forecast; the coefficients are constant over time, whereas $\nu(t)$ represents the error of the model. When the functions $f(\cdot)$ are selected *a priori* as explicit time functions, such as sine waves, α_i parameters are estimated using simple linear regression or similar techniques, such as the ones found in Quintana *et al.* (1978) [46], and Fankhauser (1984) [47]. The models presented here are simple and their parameters are easily updated.

Within the *time-of-day* models are those based on spectral decomposition, which also have the form presented in (2), where the $f(\cdot)$ represent the corresponding eigenfunctions for the correlation function of the time series, as described in Belik *et al.* (1978) [48]. The results are better than those achieved by the methods previously discussed. Despite their performance compared to other *time-of-day* models, models based on spectral decomposition are used in few applications for utilities. Some examples can be found in Farmer and Potton (1968) [49] and Anelli *et al.* (1978) [50]).

Within the *time-of-day* models are those that use *Ordinary Least Squares* (*OLS*). Meng and Shang (2008) [51] present a *Partial Least-Square* (*PLS*) model for annual forecasting in China. Its main feature is that it requires 12 years for the prediction and it's valid for 5 years.

2) Dynamic models: Dynamic models take into account that the load not only depends on the time of day, but also on weather variables and random inputs. They can be classified in two types: *Auto-Regressive and Moving Average* (*ARMA*) models, and *state-space* models. *ARMA* models have the following form:

$$L(t) = y_p(t) + y(t) \tag{3}$$

where: $y_p(t)$ depends primarily on the time of day and the usual weather patterns for that particular day; this component can be expressed by (2). The term $y(t)$ represents a residual term for the load, accounting for the deviation of the weather patterns due to the random effects in the correlation.

Abou-Hussien *et al.* (1981) [52] describe a *dynamic* model formed by five components in $y_p(t)$, where the main components are the average of the daily load patterns and the average of the weekly load increase; in addition, it includes a bias error based on the hours immediately preceding the forecast and on weather behavior.

Modifications of (3) can be found in the literature, under names such as *Box-Jenkins*, *time series*, *transfer function*, *stochastic*, *ARMA* and *Auto-Regressive Integrated Moving Average* (*ARIMA*), with slight variations among them. They could all be grouped under the *ARMA* model and, specifically, within the *dynamic* models.

Some authors, such as Rajurkar and Nissen (1985) [53] and Schneider *et al.* (1985) [54], choose to represent the periodical load component as in (3). Other authors pre-filter the load data so as to remove the periodical component, as in Abu-El-Magd and Sinha (1981) [55], Ernoult and Mattatia (1984) [56], and Paysti (1984) [57]. After pre-filtering, the new load process is:

$$L'(t) = L(t) - L(t - t_p) \tag{4}$$

where t_p represents the period of the time of day of the component, and $L'(t)$ represents the result of the process, now without a periodical term and following equations similar to (3) (ARMA).

Hagan and Behr (1987) [58] describe some limitations of the *Box-Jenkins* model. The most important of them is precisely linearity, which is not the best option when dealing with a demand forecasting problem that is certainly *not linear*. They propose an *ARIMA* model, where a temperature variable is added, modified by a *non-linear* transformation function.

Only some of the models use weather factors as inputs for the *ARMA* models because when weather variables change very quickly, it is not possible to update the parameters of the model automatically. Some of the *ARMA* models include the effects of climatic variables as additional inputs, as in Schneider *et al.* (1985) [54], Paysti (1984) [57], Keyhani and Miri (1983) [59], and Vemuri *et al.* (1986) [60], whereas others rely on a heuristic model and the load is corrected based on the temperature, before applying the *ARMA* model, as in Ernoult and Mattatia (1983) [56]. As indicated by Gross and Galiana (1987) [61], the inputs of the model are temperature values. Its main feature is the *non-linear* behavior.

Identifying the parameters of *non-linear* models is a complex and tedious task. To do that in *ARMA* models, a recursive system can be used which solves the Yule-Walker equation, as described in Vermudi *et al.* (1986) [60].

As shown by Ljung and Soderstrom (1983) [62], the *state-space* models can be easily transformed to *ARMA* and vice-versa. They are therefore especially interesting, because there are no big differences between them. However, *state-space* models are also presented because they show a structure that is not present in *ARMA*. In these models, the load is represented by:

$$L(t) = \mathbf{c^t} x(t) \tag{5}$$

where:

$$x(t+1) = \mathbf{Ax}(t) + \mathbf{Bu}(t) + w(t) \tag{6}$$

$x(t)$ is the state vector at the point in time t, $u(t)$ is the weather variable input vector, and $w(t)$ is the white noise input vector. Matrices $\mathbf{A}$, $\mathbf{B}$ and vector $\mathbf{c}$ are assumed constant. Many variations exist from (5) and (6) but in all models the difference with *ARMA* is that in *state-space* models the parameters defining the load's periodical component are random processes.

In some of the models, $\mathbf{A}$ and $\mathbf{B}$ do not need to be known, as in Irisarri *et al.* (1982) [63], and Campo and Ruíz (1987) [64]. However, Abu-El-Magd and Sinha (1981) [55] claim that the matrices need to be completely identified. The advantages of *state-space* models over *ARMA* are not clear, but Gross and Galiana (1987) [61] point towards an application of the former in bus load forecasting, where bus load presents high correlation.

Some works that focus on *ARMA* can still be found in the early years of this century, but *ANNs* are more frequently studied. However, it must be noted that nowadays many *Transmission System Operators* (*TSOs*) and *Distribution System Operator* (*DSOs*) use these models, as shown in Fan and McDonald (1994) [65] and Yang *et al.* (2005) [66].

From 2000 on, new *ARMA* models can be found that use innovative parameters or are combined with the aforementioned *ANNs*. Some authors have compared *ARMAs* and *ANNs*, emphasizing the advantages of the former. Some examples follow.

Huang and Shin (1997) [67] describe an *ARMA* model combined with a *non-Gaussian* process for a better identification of the parameters, which slightly outperforms a tested model. Chen *et al.* (2006) [68] present a generalization of the *ARMA* model called *AutoRegressive Conditional Heteroscedasticity* (*ARCH*), which outperforms previous *ARMA* models but not models. Baharudi and Kamel (2008) [69] propose an enhancement of the *AutoRegressive* (*AR*) model called *Modified COVariance* (*MCOV*), which achieves better

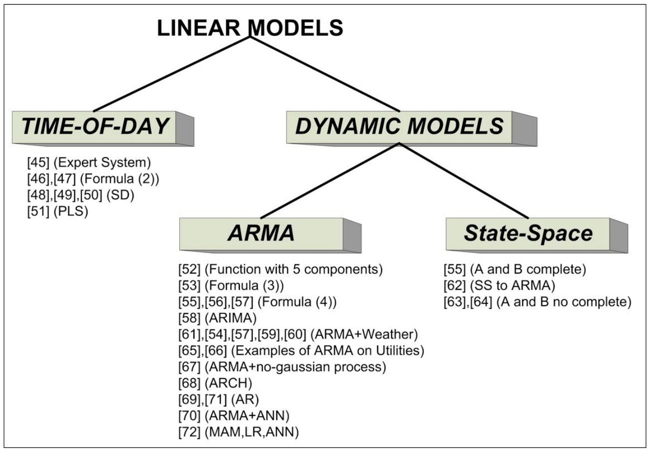


Fig. 3. Summary of linear models presented in this Section. SD: Spectral Decomposition; PLS: partial Least-Square; ARIMA: Auto-Regressive Integrated Moving Average; ARCH: AutoRegressive Conditional Heteroscedasticity; AR: Auto-Regressive; ARMA: Auto-Regressive and Moving Average; MAM: Moving Average Model; LR: Linear Regression; SS: state-space.

results than the *Durbin* (*ARMA*) and the *Burg* (*AR*) methods. The forecast is validated for a few days, and therefore the effect of seasonal variations or holidays cannot be assessed. Wang *et al.* (2009) [70] combine an *ARMA* model with an *ANN* trained with *backpropagation*, obtaining better results than the 1-stage *ARMA*. Fadhila *et al.* (2009) [71] compares the best results of an *Adaptive NeuroFuzzy Inference System* (*ANFIS*) to those of an *AR* when weekends are ruled out, obtaining similar results to those achieved by an *ARMA* model. The authors conclude that non-linear models are required to detect the peak in the series, which happen in weekends and holidays, so as to smooth out some features of the series and improve the results. Kamaev *et al.* (2012) [72] compare the *Moving Average Model* (*MAM*), *Linear Regression* (*LR*) and models. They point out that not having an internal knowledge of the model is a disadvantage for *ANNs*.

3) Linear models: a summary: Figure 3 summarizes the models presented in this Section. After the discussion, it can be concluded that the challenge for *linear* models is translating knowledge of the problem to the equations defining the model. Demand prediction is a complex problem due to its non-linearity, and it is therefore a big challenge to translate it into *linear* models.

IV. NON-LINEAR MODELS

From 1985 on, researchers such as Bunn and Farmer (1985) [73] started to realize that *non-linear* models accurately described the relation between periodical and residual components. In addition, Hagan and Behr (1987) [58] emphasized the limitations of *linear* models in comparison to *non-linear* models based mostly on *ANNs*.

4) Types of Artificial Neural Networks: ANNs are able to learn from experience, generalize to new cases based on the past, obtain essential features from input variables representing relevant information, etc. They have therefore certain advantages and have gained increasing attention over time. Some of these advantages are: adaptive learning, self-organization, fault tolerance, real-time operation, and ease of integration in existing technology.

It is precisely their ability to generalize and, above all, their ability to discover non-linear relation in complex environments that makes *ANNs* especially attractive for use in problems such as demand forecasting, where the limitations suffered by linear techniques have become apparent (and big efforts are required from researchers to overcome those).

ANNs are formed of neurons arranged in different layers (input, hidden, and output layers). For the so-called "*standard neuron*" presented by McClelland and Rumelhart, among others, and described in McClelland and Rumelhart (1986) [74], and Rumelhart and McClelland (1986) [75], the propagation rule is the weighted sum of the input values, and the output function is the identity. After applying the propagation rule, the output for the neuron is obtained from the input values. The neuron's inputs are affected by some weights that are defined once the training stage is complete.

ANNs operate in two modes: learning or training mode, and execution or recall mode. The learning mode is of particular interest, since it is one of the keys of *ANNs*, making them trainable, able to carry out certain processing tasks that they learn from a set of patterns and then reproduce in the execution mode with unknown patterns. The training involves updating synaptic weights following learning or training rules, which are built by optimizing an error function. This function measures how efficient the operation of the network is. If $w_{ij}(t)$ is the weight of the $j-th$ pre-synaptic neuron with respect to the $i-th$ post-synaptic neuron at iteration t, the training algorithm that depends on the input signals received at the point in time t from the environment will yield a value $Deltaw_{ij}(t)$. This value represents the change to be made in the corresponding weight, and is obtained as follows:

$$\Delta w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t) \quad (7)$$

The two main training modes are the supervised and the unsupervised, the difference between them being the way they deal with patterns. Supervised training gives information about the functions to be estimated by providing a distribution of the types or labels of the output patterns. Unsupervised training, on the contrary, does not provide such information.

Different models of *ANNs* exist, according to the neuron model, the architecture or type of wiring, and the learning algorithm. The main features of the models are the type of learning and the network architecture. Figure 4 classifies *ANNs* based on these two criteria.

Since the emergence of *ANNs*, multiple authors have analyzed whether these models outperform those based on *ARMA* and, of course, the early models based on historical average measures. Sharda and Patil (1990) [76] show that *Multi-Layer Perceptron* (*MLP*) achieves the same results as *Box-Jenkins* models. Tang *et al.* (1991) [77] show that in the case of long time series, the three models they analyze perform similarly to *Box-Jenkins* model. However, when time series are short, *ANNs* outperform the latter. Based on this, the authors suggest *ANNs* as a future alternative to *ARMA* and similar models. The same conclusion when it comes to long and short series can be found in Badri (1996) [78]; the author also finds *ANNs* extremely interesting when the series are short.

The following are the conclusions that can be inferred from

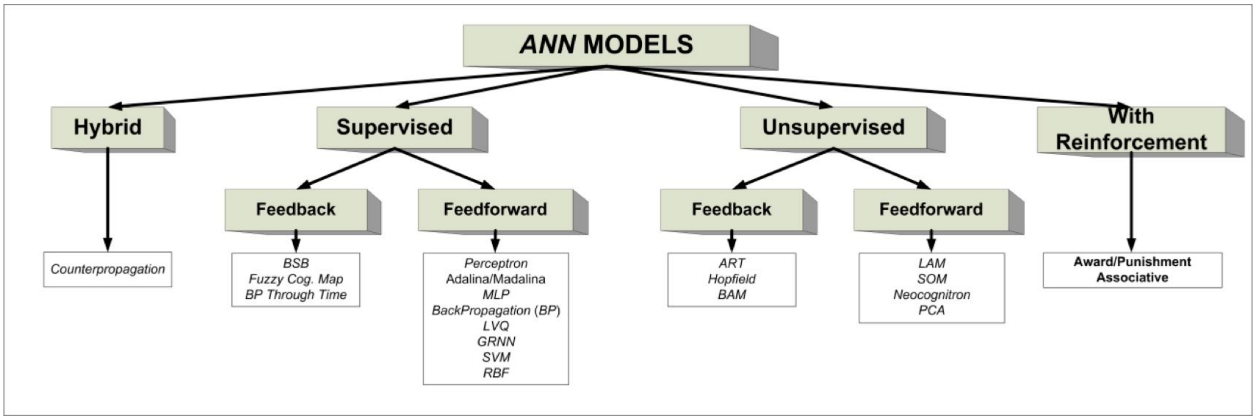


Fig. 4. Classification of *ANNs* according to the type of learning and the architecture. *RBF*: Radial Basis Function; *BSB*: Brain State in a Box; *MLP*: Multi-Layer Perceptron; *LVQ*: Learning Vector Quantization; *GRNN*: General Regression Neural Network; *SVM*: Support Vector Machine; *ART*: Adaptive Resonance Theory; *BAM*: Bi-directional Associative Memory; *LAM*. Linear Associative Memory; *PCA*: Principal Component Analysis.

the previous discussion: first of all, since the last decade of the 20th century, models based on *ANNs* that outperform linear models can be found in the literature; secondly, for short time series, *ANNs* are able to come up with more efficient forecasts, due to their ability to detect non-linear features, which produces a smaller margin of error when the forecast is made for a small number of patterns. With the deployment of measurements in *SGs* (and other disaggregated environments), and bearing in mind the fast social and economic changes that society is undergoing, updated and regular consumption data will be increasingly available, thus making or *hybrid* models the most suitable solution (*hybrid* models will be described later).

Hippert *et al.* (2001) [16] review the models used to forecast demand over the previous 10 years and highlight the following concepts: data pre-processing, design, implementation and validation. On the other hand, for these authors it is still not clear how overfitting and overparameterization affect the process. They conclude that in the light of recent works, models are the ones to be used in the future.

The issue posed by overfitting was later solved thanks to the algorithm presented by Wu and Liu (2009) [79], among others. Also van Wyk and Engelbrecht (2010) [80] prevent overfitting using an unbounded activation function. Corrales and Auron (2000) [81] address the compensation phenomenon caused by overparameterization in the *Nonlinear AutoRegressive with eXogenous* inputs (*NARX*) algorithm when it is applied to event-related potentials.

Based on the classification shown in Figure 4, only the most relevant works carried out in the last 30 years will be discussed, since the literature on the topic is extremely vast. Some of the works presented use *ANNs* exclusively, while others combine them with other methods. These are called hybrid models. Among the latter, some of them integrate other models based on *ANNs*, or even *linear* or other (fuzzy logic) models. Within this classification, and as described in Section II, some of the models predict a single value while others forecast multiple values.

5) Supervised feedforward ANN models: Within this category, since models started to be used to forecast electric demand, and regardless of the forecast horizon and aim of the prediction, the potential of models based on the *back-propagation* learning algorithm was clear. The reason for that was that this algorithm shows excellent results in the adjustment of functions, and is able to easily generalize as well. Some of the works presented in Section II have been developed using *MLP* ([18], [20], [24], [26-36] and [38]). García-Ascanio and Mate (2009) [82] propose a *MLP* model for *STLF* where data are previously arranged according to an analysis criterion. Then, they compare the model with a *Vector AutoRegressive* (*VAR*) model, obtaining similar results. Hernández *et al.* (2013) [83] present an *MLP*-based model for *STLF* in a disaggregated environment. The monthly results are satisfactory and, additionally, a study of the average operational error vs. the number of patterns used in the training stage is presented. They propose the possibility of updating the data for training purposes based on a given number of patterns. Indeed, the model's performance does not improve if more patterns are added beyond that number.

Other models have also been proposed as an alternative to *MLP*. Nair and Joshi (2010) [84] carry out a *VSTLF* (half an hour and one hour) using a *Probabilistic Neural Network* (*PNN*) and they obtain good results compared to similar models. They also emphasize the ability of *ANNs* to learn the non-linearities associated with the demand. Mori and Iwashita (2005) [85] compare a *RBFN* for *STLF* with a *MLP* model and others; the results obtained are very similar to *MLP*, but only 79 days have been used for the forecast, therefore preventing us from observing the seasonal effect and, in addition, special days were also excluded.

Another very important paradigm within supervised models is *Support Vector Machine* (*SVM*). Li *et al.* (2012) [86] propose a variation of the *SVM* model called *Least Squares Support Vector Machine* (*LSSVM*), with good results compared to similar *ANN*-based models. Hong *et al.* (2011) [87] explore *Support Vector Regression* (*SVR*), based on *SVM*, to model seasonal effect and the cyclic nature of the demand. Aung *et al.* (2012) [88] claim their proposed least-squares version of *SVR* with online learning strategy gives superior performance over *MLP* (or similar) demand forecasting model.

6) Supervised feedback ANN models: As far as supervised feedback models go, Sun *et al.* (2005) [89] study the entropy of variables to select the most relevant ones and then carry out *STLF* using an *Elman* model (with a *feedback ANN*), and the results obtained are comparable to those of a *backpropagation* neural network. Su and Zhang (2007) [90] use a filter to remove the noise of the data and then an *Elman* model for *STLF*. However, seasonal effects cannot be observed, because the days used for the prediction are not enough. Yongchum (2010) [91] uses an *Elman* model, but instead of using all the times of the previous day as an input, it uses the load from the last 12 hours. Therefore, the model has 15 inputs, 12 corresponding to the demand and 3 to weather variables.

7) Unsupervised ANN models: Some works based on unsupervised models have proposed an alternative to supervised models. Some of them are based on cascading models to improve the results. Marín *et al.* (2002) [92] apply *SOM* to divide a vast area of central Spain in different groups and then train *Elman* models for each group. The disadvantage of this method is that it requires manual grouping in supergroups, after applying *SOM*. Hernández *et al.* (2012) [93], on the contrary, use *SOM* with pattern recognition purposes, and then use a *k-means* algorithm to form similar-curve clusters in an industrial park without human intervention, so as to carry out *STLF*. Additionally, a *Principal Component Analysis* (*PCA*) is completed to detect technical outliers present in the historical load curves. Chicco *et al.* (2004) [94] propose two pattern recognition techniques applied to sort electrical customers using *SOM* and a *modified follow-the-leader* algorithm. This sorting can then be used to apply different fees depending on the customer's profile and the forecast of the electric demand. Carpinteiro and Reis (2005) [95] forecast electric demand in Brazil using two cascaded *SOMs*; this can also be considered a *hybrid* system. Tsekouras *et al.* (2007) [96] perform a two-stage sorting using *SOMs*. The first forecasts customer's load and is used to apply different fees, whereas the second provides valuable information to the utility and is used in power markets. Zhang *et al.* (2012) [97], similarly to [93], propose a load-profile sorting algorithm. They use *SOM* for sorting and clustering (*k-means*). They propose sorting systems as the first stage in demand forecasting.

8) Hybrid ANN models: Some *hybrid* systems focus on fuzzy strategies. Wang (2006) [98] forecasts electric demand in China using *SOM* and *fuzzy-rough*. Duan *et al.* (2011) [99] combine a clustering stage with *fuzzy c-means* (*FCM*) and *SVR* techniques. Che *et al.* (2012) [100] propose a *hybrid* system called *adaptive fuzzy combination model* (*AFCM*) whose goal is to iteratively combine the solution of different subgroups, after applying *SOM* and *SVR*, as well as finding fuzzy functions for each homogeneous subgroup. Other *hybrid* systems that combine fuzzy logic and other techniques are Nadimi *et al.* (2010) [101], and Lou and Dong (2012) [102].

As for *hybrid* systems that use *RBF*, Mori and Kanaoka (2006) [103] predict temperature and then carry out *STLF*. To do that, they perform a clustering using the *Deterministic Annealing* (*DA*) algorithm and they use a *Radial Basis Function Network* (*RBFN*) model for each cluster, obtaining high efficiency compared to other tested models (*MLP* and *RBFN*). These authors based their work on a previous study (Mori and Yuihara, 2001 [104]), where the approach was similar but using *MLP* instead of *RBF*.

Hybrid systems have been used and verified while comparing them to different *ARMA* models. The conclusion has been that, even if the architecture of the system is more complex, the solution is better compared to the use of *linear* functions to carry out the forecasting. Amjady (2007) [105] proposes a *hybrid* method which consists of a linear forecasting stage called *Forecast-Aided State Estimator* (*FASE*) followed by *MLP*, and compares the results with *MLP*, *FASE*, and *AR* models. Meng *et al.* (2011) [106] suggest using a *grey model* to mitigate the stochastic effect existing in the electric load trend, and then use *RBF* to forecast monthly demand.

It should be noted that the complexity inherent to *hybrid* systems involves a non-negligible increase of the time needed to solve the problem. Che *et al.* (2011) [107] approach demand forecasting using *SVR* in an attempt to solve the computing time issue (with N^2 time complexity). This can be replicated in systems that use cascading models, which must be assessed in advance to know the order of computation complexity required by the solution.

Hybrid systems have also been used to model the parameters of the forecasting model. For example, the following works used chaotic algorithms to optimize *SVR*, with good performance compared to other regression models and even to *ARMA* models (Hong, 2010 [108]; Hong, 2011 [109]; Zhang *et al.*, 2012 [110]; Hong *et al.*, 2013 [111]).

Also using chaotic algorithms and genetic algorithms in *hybrid* systems, Deihimi and Showkati (2012) [112] propose a recurrent *Echo State Network* (*ESN*) for *STLF*. Genetic algorithms are used at the first stage to select the most suitable input variables for the model (load parameters and weather variables). The final performance of the model is good, but computational cost is increased. Li *et al.* (2013) [113] propose a variation of *RBF*, called *Generalized Regression Neural Network* (*GRNN*) to forecast annual demand; this hybrid system uses the *Fly Optimization Algorithm* (*FOA*) for a suitable selection of the *GRNN* parameters.

9) Types of ANN models: a summary: To finish this Section, the works discussed will be represented and summarized in a simple figure. As has been seen, *non-linear* models used in demand forecasting are mostly based on *ANN* techniques or a combination of these with other models. Same as *linear* models, they have been used for a wide range of forecast horizons, and also with a great variety of forecast aims. Their main difference with respect to *linear* models is that they do not require coming up with a clear mathematical model representing the dependence on different parameters to solve the demand forecasting problem, which saves on time; conversely, *linear* models require an accurate definition of the equations that define them. All *ANN* models need a proper parameterization of the model's input variables, as well as a topological definition of the network (number of layers and neurons in each, training functions, etc.). However, this is a reasonably simple task (a simple script) compared to the parameterization of the functions that define *linear* models. Nonetheless, if the system includes additional stages for the forecast using more models, these will involve an increase of the global computational cost, for both the training and the operation (forecast) stages.

In short, a forecast model based on a single *ANN* will have a

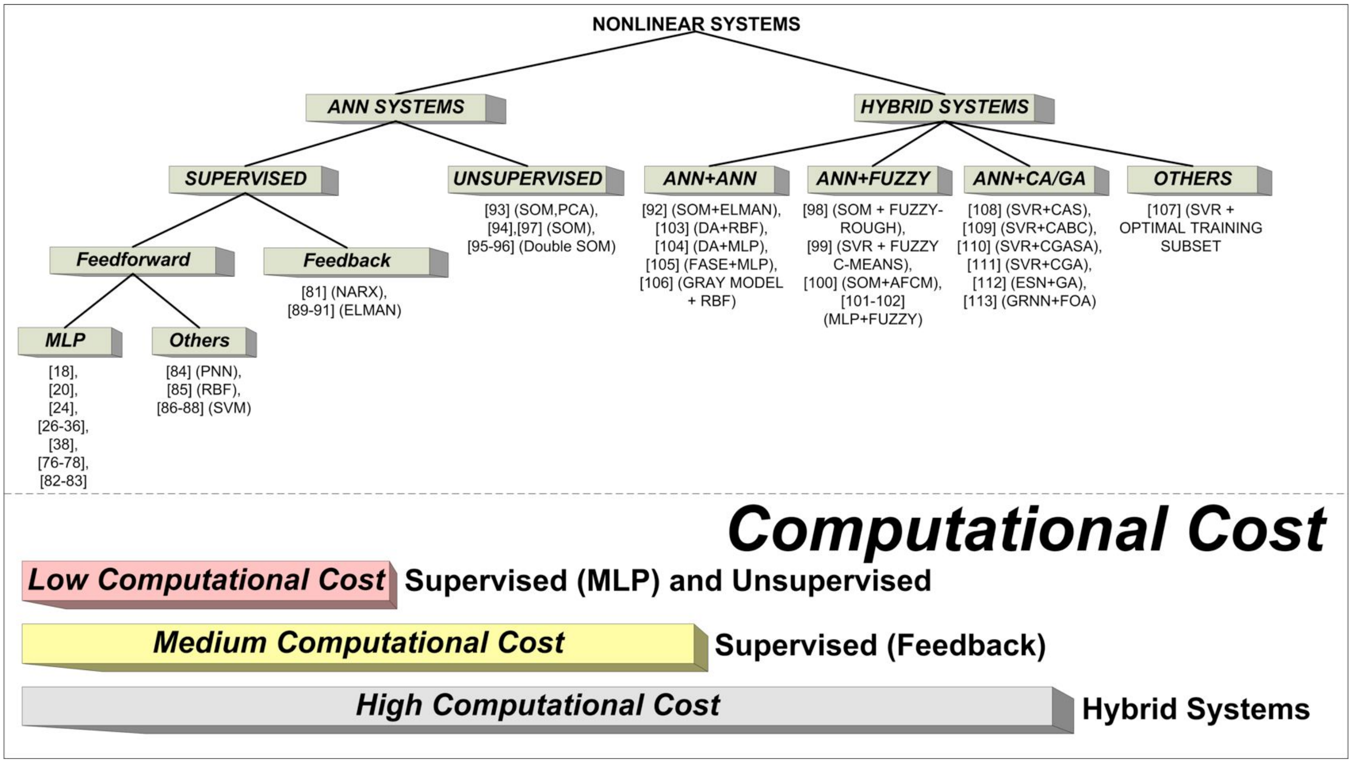


Fig. 5. Summary of the non-linear models discussed in this Section NARX: Nonlinear AutoRegressive with eXogenous; PNN: Probabilistic Neural Network; RBF: Radial Basis Function; MLP: Multi-Layer Perceptron; GRNN: General Regression Neural Network; SVM: Support Vector Machine; SVR: Support Vector Regression; PCA: Principal Component Analysis; SOM: Self-Organizing Map; DA: Deterministic Annealing; FASE: Forecast-Aided State Estimator; AFCM: Adaptive Fuzzy Combination Model; CAS: Chaotic Ant Swarm; CABC: Chaotic Artificial Bee Colony; CGASA: Chaotic genetic Algorithm Simulated Annealing; CGA: Chaotic Genetic Algorithm; GA: Genetic Algorithm; CA: Chaotic Algorithm; FOA: Fly Optimization Algorithm.

much lower computational cost than a system that uses several stages or additional algorithms, such as a genetic algorithm. Even if a single *ANN* model is used, the computational cost of the learning stage could be also high, depending on the model chosen. Figure 5 shows the works discussed in this Section, taking into account that forecasting systems can be of a single type or *hybrid*. For the latter, the base model is indicated. Note that in this case the term hybrid is used with a different meaning than in Figure 4. In addition, an estimate of the global computational cost is provided.

V. Disaggregated Environments

As explained in Section I, the emergence of new disaggregated operation and control spaces (substations, *microgrid*s, *Smart Buildings*, etc.), together with the deployment of measurement systems and intelligence in these places, will allow for the development of planning tools, which will in turn depend on demand and generation estimates.

Regarding demand forecasting, with the exception of works such as the one presented by Sargunaraj *et al.* (1997) [114], which addresses a disaggregated substation, all works discussed (regardless of their model and the aim of their prediction) seek to predict for a broad geographical area, as in Hsu and Yang (1991) [115], Taylor and Buizza (2002) [116], Chu *et al.* (2011) [117], Nose-Filho *et al.* (2011) [118], Rejc and Partos (2011) [119] and Wang *et al.* (2011) [120].

In recent years the attention is shifting to new spaces for the integration of generation and demand which emphasize the need for prediction models (most of them based on artificial intelligence) for demand and generation forecasting that enable operation as shown by Chaouachi *et al.* (2013) [121].

For correct operation and management, the *microgrid*, considered as a single and autonomous entity will require demand and generation prediction tools (as will its interaction with the *SG*), in order to be able to adapt its behaviour accordingly. The *microgrid*'s external price indicators and storing situation will also need to be considered. Dimeas and Hatziargyriou (2005) [122] present the operation of a *Multi-Agent System* (*MAS*), and claim that despite the advances in forecasting techniques, knowledge is focused on large interconnected systems and hourly prediction horizons. They suggest the need for models applicable to disaggregated environments and forecast horizons shorter than several hours. For the *microgrid* to operate correctly, Khodayar *et al.* (2012) [123] present an economic model of the *microgrid* operation. In it, a prediction of the demand is necessary so that there is a balance between the available DG energy and the energy stored. Sugaya *et al.* (2012) [124] indicate that for a stable control of a *microgrid* it is necessary to build an efficient schedule of the energy plants existing in the *microgrid*, which will in turn depend on the demand estimation. As expected, the EMSs of these disaggregated environments must know in advance the generation and the demand they are in charge of, as shown in Guan *et al.* (2010) [125], Chen *et al.* (2011) [126], and Kanchev *et al.* (2011) [127].

The works on forecasting and classification of the demand in *microgrid*s show the critical need to adapt existing forecasting models to the specific features of these environments with disaggregated demands and abrupt *load profiles*. This makes forecasting more complicated, as seen in Amjady *et al.* (2010) [128], and Chan *et al.* (2011) [129].

Techniques for classification and clustering of the load curves of industrial parks via *SOM+k-means* for a posterior forecasting of the electric demand, as shown in Hernández *et al.* (2012) [93], can help the aggregator, among others, to obtain better prices and realize *DR*. As presented by Kamaev *et al.* (2012) [72], demand forecasting in shopping centers (*microgrid*s) is deemed essential. The manager of these environments must estimate the electric power consumption with the greatest possible accuracy in order to place energy purchase orders to the utilities. Shopping centers have a schedule which depends on the opening hours, the turnout and seasonal factors. Better predictions are obtained by introducing the previous variables, as well as by using next day's forecast, as done in Hernández *et al.* (2013) [83]. This study tests linear and non-linear models and concludes that *ANN*-based models yield optimal results at the expense of losing the possibility to interpret the problem, since the reference of the selected variables' aim with respect to demand forecasting is lost. This does not happen when using linear models, where the researcher's work in creating the functions provides this insight.

In *microgrids*, having these prediction tools will not only lead to energy saving by enabling optimal power purchase, as reported in [72], but it will also allow for the integration of renewable generation sources and electric storage, as shown in Hida *et al.* (2010) [130]. The responsibility for demand and generation forecasting, the correct operation of the *microgrid* and the optimal economic dispatch is assumed by the *MicroGrid Central Controller* (*MGCC*) (Lopes, 2009 [131]). Dimeas and Hatziargyriou (2005) [122] and Sechilariu *et al.* (2013) [132] present the optimal economic dispatch for a *microgrid*, where the possibility of the maximum penetration of renewable energy sources and electric storage relies on an accurate demand prediction.

If we further atomize the problem, more disaggregated spaces with associated intelligence emerge, such as *Smart Buildings*, etc.; back in 1987, Skrzypczak (1987) [133] already forecast the intelligent home of year 2010, where the main challenge would be the integration of areas such as mechanics, electrics, electronics, informatics and ergonomics. Monitorization and control should try to adjust electric parameters for a maximum comfort, on the basis of environmental and behavioral measurements. Constanzo *et al.* (2012) [134] propose a *Demand Side Management* (*DSM*) for a building, where a load forecaster will be a fundamental element. In Harris (2012) [135] several definitions of smart building are provided, among others: "*A smart building should have the ability to integrate automated building controls and optimise operations to lowering both cost and energy usage compared to the conventional buildings*"; in order to automate and control the energy, it will be necessary to predict its consumption. Sechilariu *et al.* (2013) [132] show the need for communication between the *SG* and the building control so that the *DG* and storage installed in the latter can operate correctly. Buildings will require generation and demand predictions to operate and make decisions. Abreu *et al.* (2012) [136] utilize pattern recognition techniques to identify behavioral patterns in residential consumption, thus allowing the forecast of demand in disaggregated environments such as smart buildings.

In addition, these disaggregated environments will allow us to carry out more in-depth and thorough studies on the influence of weather variables on the electric demand. The conclusions could then be applied to demand forecasting models in *Smart World* frameworks, as in Hernández *et al.* (2012) [12].

Rathore (2012) [137] proposes a new hybrid grid based residential utility interfaced smart energy system; based on pricing, renewable energy output, load demand, storage, and forecasting, decisions are made and communicated by a central intelligent unit.

Borges *et al.* (2011) [138] present a *STLF* for a Campus (University of Deusto, a *microgrid*); they present a comparative between 2 forecasting methods (*SVM* and *MLP*) and 4 forms of post-processing their results. Wai *et al.* (2011) [139] investigate a fuzzy neural network with varied learning rates for *STLF* in a Taiwan Campus (*microgrid*), and compare its better forecasting performance with a conventional *ANN*; the proposed model *STLF* could be modified and extended to *MTLF* and *LTLF*. Llanos *et al.* (2012) [140] present two methods of *STLF* in a *microgrid*; first method is proposed to be used before the implementation of the *microgrid* in the design state, and it includes a household classifier based on a SOM; the second method is used after the implementation of the *microgrid*, in the operation state, and consists of a *ANN* with on-line learning for the load forecasting. Shimoda *et al.* (2012) [141] show an electrical load forecasting method and an optimized operation planning method for *microgrid*; for the electrical load prediction, the authors use an operation planning of the heat sources which run according to heat load prediction (with a *MLP* network).

In summary, the emergence of these future environments, which are a consequence of the *SGs*, will require electric demand forecasting models. Some works addressing this issue have been discussed, but more are bound to emerge from the increasing research on related topics (*microgrid*, *Smart Building*, etc.). It should be kept in mind that conceptually, within these environments, the integration of renewable generation sources and/or power storage are crucial; therefore, demand and generation forecasting techniques will be key to the success of renewable energy sources.

VI. Comparison of the Available Solutions

The aim of this Section is to compare *linear* and *non-linear* models for electric demand forecasting. First, the two big groups will be compared, with the goal of analyzing the defining features of each and their contribution to the problem of demand forecasting.

Then, a classification of the presented works will be provided, grouping works according to their common features and in line with that taxonomy. The grouping will be:

- Models: *linear*; *non-linear*.

TABLE I
COMPARISON BETWEEN THE MAIN FEATURES OF LINEAR AND NON-LINEAR MODELS FOR ELECTRIC DEMAND FORECASTING.

	Linear models	***Non-linear* models**
Forecast horizon	*VSTLF/STLF/MTLF/LTLF*	*VSTLF/STLF/MTLF/LTLF*
Aim of the prediction	Single/multiple value(s)	Single/multiple value(s)
Knowledge base loss	No	Sometimes
Ease of non-linearities inherent to the model	No	Yes
Ease of parameterization	No	Relatively easy
Need for initial Adjustment	Yes	Yes
Computational cost	Medium/High (depends on model complexity)	Low/Medium/High (depends on the model)
Possibility to implement hybrid syst.	Yes	Yes
Response to high number of hist. patterns	High	High
Response to low number of hist. patterns	Average	High
Works before 1990	Many	Few
Works 1990-2000	Many	Increasing
Works after 2000	Few	Many

- Area: country, region and city; *small city*, *microgrid* and *Smart Building*.
- Horizon: *VSTLF*; *STLF*; *MTLF*; *LTLF*.
- Aim: *one value*; *several values*.
- Variables: load (with calendar data); load (with calendar data)+weather; load (with calendar data)+weather+other; load (with calendar data)+other.

Before addressing the global comparative study of the two big model groups (i.e., *linear* and *non-linear*) for demand forecasting, let us reflect upon the previous analysis. As stated in Section IV, forecasting models based on *ANNs* require a smaller intellectual effort from the researcher, as the non-linearity detection between the various variables and the demand is intrinsically handled by the network. A researcher with average knowledge of electric demand will be able to propose different alternatives for demand forecasting, focusing only on the development of a complete testing methodology and its controls for a posterior comparative study. Nevertheless, as shown in [72], this is a double-edged sword: in exchange for the ease of demand prediction given by *ANNs*, there is an inherent reduction in the knowledge base of the problem.

On the other hand, the availability of computers with high processing capability and enhanced algorithms is also in favour of the use of *ANN*-based forecasting models.

Additionally, the minimization of the objective function in the learning stage (Rying *et al.*, 2002 [142]), together with the enhanced training algorithms for *ANNs* (Han and Qiao, 2012 [143]), their reduced computation complexity (Bragatto *et al.*, 2008 [144]); also preliminarily studied by Orponen (1994) [145] and lastly, the automatic detection of *ANN* architectures (Islam *et al.*, 2009 [146]), make *ANNs* the most promising alternative to the traditional linear-model approach. This is the reason why in the first decade of the 21st century, the number of works on forecasting models that focus on *ANNs* greatly outnumbers those focusing on linear models.

Table 1 shows a comparison between *linear* and *non-linear* models. The most critical and distinctive features discussed are shown, namely: forecast horizon, aim of prediction, loss of knowledge base, ease of non-linearity detection inherent to model, easy of parameterization, need for initial adjustment, computational cost, possibility to implement hybrid systems, response when using a great/small number of historical patterns, works published before 1990, between 1990 and 2000, and from 2000 on.

Table 1 can be used to decide between *linear* and *non-linear* demand forecasting models. The main factors to consider when choosing one of them are:

- The choice of *non-linear* models could involve a reduction in the knowledge base of the problem, since forecasting models based on *ANNs* intrinsically lay on the networks the responsibility of detecting the defining features of the problem. Therefore, the researcher might not spend the required time to understand the problem to be solved. However, the need to suggest optimal input variables to the system will require a minimum/medium study of the electric demand problem.
- For *linear* models, computational cost depends on the complexity of the functions used, as well as on the use of *hybrid* systems. In the case of *ANNs*, computational cost will be defined at the learning stage of the chosen model, and similarly will depend on whether *hybrid* systems are used or not. Further discussion on computational cost and model performance can be found in Kaushik *et al.* (2010) [147] for *ANNs* and in Chow and Tan (1998) [148] for *ARMA*.
- The deployment of intelligence in the future environments (*SG*, *microgrid*, *Smart Building*, etc.) will enable a more frequent update of models based on available data. Electric demand is clearly dependent on some external factors (social, economic, etc.); therefore, the availability of recent data which could carry information on possible changes in these parameters (and thus on electric demand) is very promising. *Non-linear* models show a better response than linear models when a small amount of patterns is used at the learning stage. These two reasons lead us to think that the models that will prevail in the future will be those based on *ANNs*, especially in the aforementioned environments.

This seems also clear judging by that fact that most of the works on demand forecasting published since 2000 have focused on *non-linear* models, particularly in the last 5 years.

Next, Tables 2 and 3 summarize and highlight the main features of the works presented in this review. Table 3 is used to put together similar references into groups identified with a number (linear models) or a character (non-linear models). These groups are represented in Table 2 to study their main features. *Linear* models have been grouped in two main categories, those based on *ARMA* models (1-6) and those that use *other* models (7-9). *Non-linear* models have also been sorted in two categories: *ANN*-based (A-K) and *hybrid* systems (L-Y). For all groups, the main features have been indicated, namely: area under control (country or region and *microgrid* or *Smart Building*), forecast horizon, aim of the prediction, and main variables used.

Table 3 shows each of the groups, indicating the corresponding references. Rows represent each of the groups and columns correspond to the main two types of model, *linear and non-linear*.

It is easy to notice that only a small number of categories are focused on applying demand forecasting techniques in disaggregated environments (*microgrid*, *Smart Building*, small town, industrial park, etc.). In particular, the groups modeling these environments are: 5, 6, C, H, N and Q. Two of them correspond to *linear* models and three of them to *non-linear* models. The 2 groups under *linear* models consist of 3 works, which represents only 3% of the total works presented here (95 works on Table 3). As for C, H, N y Q, they account for a total of 7 works, which is only 7.36% of the works discussed in this review. Out of the 10 works, 3 date back to the late-90s and the remaining 7 were published in 2012 and 2013. The deployment of disaggregated environments has started in the last 5 years, which leads us to think that more and more works studying electric demand and generation forecasting and focusing on these types of environments will be published in the near future.

Regarding the forecast horizon, 6% of the groups analyzed fall into *VSTLF*, 58% carry out *STLF*, 20% deal with *MTLF* and finally, 16% focus on *LTLF*. Among *linear* models, 7 correspond to *STLF*, 1 to *MTLF*, and 1 to *LTLF*. *VSTLF* has not been analyzed.

As far as the aim of the prediction goes, 15 groups forecast a *single value* whereas 19 forecast *multiple values*. The difference comes from *linear* models, where only 1 group focuses on a *single value* and the rest forecast *multiple values*. In the case of *non-linear* models, both alternatives are evenly distributed.

As has been explained, the chosen forecast horizon and aim of the forecast depend on the problem that the researcher needs to solve. Therefore, future trends will follow a similar path, and works on all possible alternatives regarding forecast horizon and aim of prediction will be published.

With respect to the variables used, 47% of the groups use models that only use load values (together with calendar variables). 32% of the groups combine the use of load and weather variables. 9% of the groups use load variables, and other types of variables. Finally, 12% of the groups use load and weather variables, as well as other types of variables.

Finally, Table 4 presents an in depth analysis of the main features of the most meaningful solutions studied in order to facilitate comparison among them. The table includes information about the core technologies used for forecasting, input variables employed, accuracy/error, and a qualitative analysis of the most interesting advantages of each approach.

Before the appearance of the new smart environments (*SG*, *microgrid*, etc.) utilities were already using demand forecasting. Regarding geographical location, these forecasting models have been traditionally focused on entire countries or large regions of a country. In Table 4, it is possible to see that for these environments, there is a variety of solutions available: manual and *linear* models [25,51], expert systems [22,28], models based on *ANN* [18-19,21,23,26-27,29-30,32,35,78,82,84-86,88,91,95,106,112,115-117,126]; and hybrid models with *ANN* [20,24,37,39,71,87,89,92,94,96-98,100-102,104,107-111,119-120].

Linear models and expert systems require large amounts of data for learning phase, with an average of 7 years in the papers presented. Otherwise, data from the training phase for the *ANN* or hybrid models are based on 24 months average, with the exception of long term solutions [23], [24], [51], [82], [86], [92], [101], [106] and [108] which use 16, 22, 7, 33, 12, 10, 27 and 19 years respectively. As for performance predictions, for manual and *linear* models an average error of 2.04 % is obtained, 3.20% for *ANN*-based models and 3.14% for hybrid *ANN*-based models. Therefore, in these environments the best performances have been obtained by *linear* techniques, mainly due to the large amount of data used and precise adjustment formulas of the models for discovering nonlinearity patterns in the demand curve.

With the emergence of the *SGs* and *Smart Cities* (*SCs*) demand forecasting models have continued to evolve in order to seize the peculiarities and abundance of information available. In Table 4, there are hybrid or linear models based in linear models [68,70], and hybrid models based in *ANN* and different optimization algorithms [90,99,103,113]. All these models are defined by the large amount of data needed for the learning phase (two years on average except [113] which requires 24 years because the one year forecasting horizon). The average error for the linear models (including hybrids) is 4.71%, whereas for *ANN*-based hybrid models performance and different optimization algorithms is 2.28 %, improving the previous models of the same type for the electric demand forecasts of the country or region.

Finally, Table 4 shows the *microgrid*, substation and *Smart Buildings* type locations. For this kind of environments classifiers and evolutionary algorithms [105.128-129] and *ANN*-based models or hybrid *ANN*-based models [72,83,93,114,118,121,138-141], trying to capture the increased variability of less aggregated scenarios. There are few differences regarding the volume of date employed, presenting an average of 22 months for all of them. Regarding the performance of the models, an average error of 4.82% is obtained, which is a worse performance than that obtained in the previous locations. This deterioration in performance is due to the inherent increased variability in the demand curve, and therefore the need to use new variables in the prediction models, such as data which anticipate the operation of the main loads that control the *Microgrid*/Substation/*Smart*

TABLE II

SUMMARY OF THE WORKS PRESENTED IN THIS REVIEW L: LOAD (WITH CALENDAR DATA); L+W: LOAD (WITH CALENDAR DATA) + WEATHER; L+W+O: LOAD (WITH CALENDAR DATA) + WEATHER + OTHER; L+O: LOAD (WITH CALENDAR DATA) + OTHER.

		LINEAR MODELS									*NON-LINEAR MODELS*																									
		1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y
Area	COUNTRY, REGION, CITY	✓	✓	✓	✓			✓	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
	SMALL CITY, MICROGRID, SMART BUILDING					✓	✓						✓					✓							✓			✓								
Horizon	VSTLF										✓											✓														
	STLF	✓	✓	✓		✓	✓	✓	✓			✓	✓			✓	✓	✓			✓		✓	✓	✓	✓	✓	✓								✓
	MTLF				✓														✓	✓									✓	✓			✓		✓	
	LTLT									✓				✓	✓																✓	✓		✓		
Aim	ONE VALUE				✓				✓	✓	✓	✓			✓					✓		✓	✓	✓					✓	✓	✓			✓	✓	
	SEVERAL VALUES	✓	✓	✓		✓	✓	✓					✓			✓	✓	✓	✓		✓				✓	✓	✓	✓				✓	✓			✓
Variables	L	✓			✓	✓					✓	✓		✓	✓	✓		✓	✓			✓					✓	✓	✓			✓	✓			
	L+W		✓				✓		✓	✓							✓			✓				✓	✓	✓				✓						✓
	L+W+O			✓									✓										✓								✓					
	L+O																				✓													✓	✓	

TABLE III

CATEGORIZED GROUPS, INDICATING THEIR TYPE OF MODEL AND THE CORRESPONDING REFERENCES.

GROUP	*LINEAR MODEL*	*NON-LINEAR MODEL*
1	[48], [54-55], [67-70]	—
2	[52], [57-58], [60], [65-66]	—
3	[56], [61]	—
4	[71]	—
5	[53]	—
6	[59], [72]	—
7	[45-47], [50], [63-64]	—
8	[49]	—
9	[51]	—
A	—	[18], [84]
B	—	[31]
C	—	[72], [88]
D	—	[23]
E	—	[86]
F	—	[32], [76-78], [82], [94-97]
G	—	[26], [29-30], [33-35], [85], [89-91], [115], [117]
H	—	[83], [93], [138]
I	—	[87]
J	—	[116]
K	—	[118]
L	—	[19]
M	—	[27-28]
N	—	[36]
Ñ	—	[136], [140-141]
O	—	[38-39], [102-104], [112], [119-120], [129]
P	—	[70], [92], [107]
Q	—	[114]
R	—	[98], [109], [111], [139]
S	—	[20], [110]
T	—	[22]
U	—	[108]
V	—	[37], [100]
W	—	[101]
X	—	[116]
Y	—	[24], [113]

Bulding. Furthermore, it can be concluded that in these heavily non-linear scenarios, *ANN*-based models or hybrid *ANN*-based models are used.

Independently of the location, of all the papers presented in Table 4, the 7.20% correspond to linear models and the remaining 92.80% non-linear models (expert systems 2.80%, non-linear models based on *ANN* 40% and hybrid *ANN*-based models 50%). Regardless of whether the non-linear model is used alone or as part of a hybrid model, 40% of the works have used *MLP* and 15.71% *SVR*. Within *Micro-grid*/Substation/*Smart Building* locations the 61.53% of the papers presented in Table 4 have been used *MLP*, either as single model or as part of a hybrid model.

VII. CONCLUSION

Since the early electrical systems were in place, there has been interest in having an insight on demand behaviour, sometimes just for the sake of understanding it but most of the time in order to plan generation units that supply consumers. As described, about a century ago simple systems such as punch cards with information on the average past consumption were used in order to get a rough estimate of the future behavior of demand. In the 1940s, *linear* models started to be used to forecast demand, and these evolved into the *ARMA* model and its variations, which were the undisputed favourites until the 80s. The main difficulty with these models is having

TABLE IV
ANALYSIS OF THE MAIN FEATURES OF THE MOST MEANINGFUL SOLUTIONS STUDIED.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[18]	2000	MLP	Predicts 8 values of load for the time leads from 20-90 minutes in 10-minute increments	Power consumption and climate variables	24 hourly consumption values	VSTLF	Power Utility US	MAPE: 0.66%	—	No information about dataset. Flow of the load forecasting process is too long.	Implementation in a power utility in US confirmed its good performance and reliability.
[19]	2013	WNN	12 dedicated WNN are used to perform moving forecasting every 5 minutes	Power consumption	24 hourly consumption values	VSTLF	Region (New England)	MAPE: 0.09%-0.45%	November 2007 to January 2008. Training phase: November 2007 to December 2007. Validation phase: January 2008	Few data for validation of anomalous days	This method can capture the load components at different frequencies. Daubechies-4 with two-level decomposition is the best configuration, which balances the decomposed level, the filter length, and the minimum padding length for decomposition.
[20]	1999	MLP+fuzzy indicator of season.	Seasonal variables are used for describing the four seasons of the year to build a fuzzy indicator of season. MLP with 10 hidden neurons	Power consumption and climate variables	1 year	MTLF	Country (Israel)	—	April 1992 to March 1996. Training phase: April 1992-March 1995. Validation phase: April 1995-March 1996	The climatic variables process is complicated.	—
[21]	2007	MTLF and STLF	They use two functions to optimize.	Power consumption and climate variables	24 hours-200 hours	STLF and MTLF	Hydro-Quebec feeder area	MAPE: 3.9%	—	The models are compared with other models	Forecasts at all system load points, as well as feeder forecasts, improving the quality of distribution state estimation.
[22]	2002	Knowledge-based Expert System (ES)	ES is implemented to support the choice of the most suitable load forecasting model for LTLF. The structure of the ES for LTLF is: - Knowledge base. - Inference Engine. -Solution Strategy. - User Interface	Power consumption, climate variables and social variables	Annual peak load	LTLF	Country (Egipto)	MAPE: 1.06%	1994 to 1998 and 2001. Training phase: 1994-1998, Validation phase: 2001	The model requires large amount of data in the training phase. Flow of the load forecasting process is too long	The developed ES can serve as a valuable assistant to system planners and for training purposes
[23]	2011	SVR	The relationship between Gross Domestic Product (GDP) and load is a nonlinear function, then, taking the historical data as the learning samples, implement the learning of SVR	Power consumption and economic factors	Annual	LTLF	Country (China)	—	1995 to 2011. Training phase: 1995-2008, Validation phase: 2009-2011,	The model requires large amount of data in the training phase.	—

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[24]	2008	MLP +fuzzy logic	LFLF model is a combination of MLP and fuzzy logic that enhances the ability of MLP to adapt through training to complex relationships and uses fuzzy logic to simulate uncertainties in real data. MLP: 12 outputs (12 months) and 37 neurons in the input layer. The fuzzy variables are: - Population growth. - Weather. - Number of households (residential load forecasting). - Number of employees (industrial load forecasting)	Power consumption, climate variables and social variables	Annual.	LTLF	Region (New England)	MAPE: 5.73%-11.10%	1980 to 2002. Training phase: 1980-1999, Validation phase: 2000-2002	The model requires large amount of data in the training phase. Flow of the load forecasting process is too long	The model shows better results than a MLP or a regression model
[25]	1956	Manual	The authors present a manual peak load forecasting, the process requires a weather prediction that is then converted to a critical parameter, illumination.	Power consumption and climate variables	Peak load.	STLF	Region (Ontario Hydro)	MAPE: 1.00%	10 months	Manual tables	—
[26]	1991	MLP	Peak load: Input neurons 3, hidden neurons 5 and output neuron 2. Total load: input neurons 3, hidden neurons 5 and output neurons 1. Hourly load: Input neurons 6, hidden neurons 10 and output neurons 1 (x24 STLF)	Power consumption and climate variables	Peak load, total load and 24 hourly consumption values	STLF	Region (Tacoma - US)	MAPE peak load: 2.60%. MAPE Total load: 3.39%. MAPE Hourly load: 1.64%	Nov. 1988 - Jan. 1989.	Big errors on days when people have specific start-up activities such as Monday	The results shows that the ANN is suitable to interpolate among the load and temperature pattern data of training sets to provide the future load pattern
[27]	1992	MLP	46 input nodes, 60 hidden nodes and one output node	Power consumption	Peak Load	STLF and MTLF	Country (Taiwan)	MAPE peak load: 1.19%	—	No information about the datase	—
[28]	1990	ES	Expert System	Power consumption	24 hourly consumption values	STLF	Country (Taiwan)	MAPE 2.52%	5 years.	A characteristic feature of the knowledge-based expert system is that it is not easy to add new information and new rules to the knowledge base	To benefit from the expert knowledge and experience of the system operator, eleven different shapes, each with different means of load calculations, were established. With these load shapes at hand, some peculiar load characteristics pertaining to the Taiwan Power Company can be taken into account
[29]	1998	MLP	15 input variables. MLP have only an output variable that can be used repeatedly to forecast profile load.	Power consumption, climate variables and social variables	24 hourly consumption values	STLF	—	MAPE: 1%-4.63%.	—	—	Novel input variable identification for ANN-based STLF.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[30]	1999	MLP	The model used two identical ANNs, input data randomly split into two disjoint training sets of the same size, cross-validation with early stopping, and simple averaging of individual forecasts.	Power consumption, climate variables and social variables	24 hourly consumption values	STLF	Two utilities	MAPE 24-hour: 2.05%-2.47%. MAPE Peak Load: 2.12%-2.38%.	2 years	Flow of the load forecasting process is too long	The implemented forecasting technique was able to reduce the size of the training set, further improve forecasting accuracy for both hourly and peak-load forecasts and reduce maximum errors in STLF
[32]	1992	MLP	Two different ANN-based methods are presented in STLF. Method 1 is a static approach which forecasts the 24-hour load simultaneously, while method 2 is a dynamic approach in the sense that the 24-hour load is forecasted sequentially using the previous-time forecasts. Method 1: Input neurons 48, first hidden layer neurons 90, second hidden layer 24 and output neurons 24. Method 2: Input neurons 8, hidden neurons 10, output layer 1 (x24 STLF)	Power consumption	24 hourly consumption values	STLF	Country (Korea)	MAPE Method 1: 1.457%-2.511%. MAPE Method 2: 1.575%-2.374%	-	No information about the amount of data (training and validation)	The dynamic approach (Method 2) performs better than the static approach (Method 1) in the sense that it uses less neurons and weights, is trained faster, and gives better results, especially for the peaks
[35]	1995	MLP	Input neurons 63, hidden neurons 24 and output neurons 24	Power consumption and climate variables	24 hourly consumption vaues	STLF	Country (Grecia)	MAPE: 2.34%–2.55\%.	2 years.	—	—
[37]	1996	SOM+MLP	Normal Days: a MLP. Anomalous days (vacation periods, holidays and weekends): SOM+MLP. MLP: variables in input layer are 51, neurons in hidden layer are 31 and variables in output layer are 24. SOM: the input pattern is a daily load vector composed by 24 values and the classification has been performed through an 8x8 SOM classifier	Power consumption	24 hourly consumption values	STLF	Country (Italia)	MAPE normal day: 2.6%. MAPE anomalous days: 2.7% (April) and 7.7% (Dec.)	1991–1993	The training set for anomalous days is normally small due to their inherent rarity	Increased accuracy at the cost of identification of anomalous days.
[39]	1996	Neural-gas+MLP	They use a neural-gas network to sort the data in two groups (summer and winter) and model them with different MLP	jPower consumption and climate variables	24 hourly consumption values	STLF	Region (Switzerland)	MAPE: 2.07%–3.82%.	1992-1993	The test only clusters data in two groups. It would be desirable to test the neural gas model with more advanced classifications	The classification process leads to models which increase accuracy inside the two groups of data.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[51]	2008	Partial Least-squares Regression (PLS)	No information	Power consumption	Total anual consumption	LTLF	Country (China)	MAPE: 3.07%	—	Complicated definition of the model equations. The model requires large amount of data in the training phase	PLS method could use few data to simulate the relationship between annual electric power consumption and influent factors. It avoids the number restriction and multi-collinearity demand in OLS. The modeling results show that it is more suitable than neural network in annual electric power consumption forecasting
[68]	2006	Auto-Regressive Conditional Heteroscedasticity	—	Power consumption	24 hourly consumption values	STLF	City (Nanjing)	MAPE: 0.14%–8.50%	January 2002 to June 2004. Training phase: Jan. 2002-May 2004. Validation phase: June 2004	Complicated definition of the model equations. The model is not compared with ANN	The model improves the classical ARMA model
[70]	2009	ARMA+ MLP	The ARMA model is used to forecast the linear component of time series. Afterwards, the error can be seen as the nonlinear component of the time series. An MLP is used to forecast this nonlinear error.	Power consumption	24 hourly consumption values	STLF	City (Jiang-Men)	MAPE: 0.92\%	2005. Validation phase: December 2005	Complicated definition of the model equations	The model (ARMA+MLP) is better than a single ARMA or MLP
[71]	2001	Adaptive Neuro-Fuzzy Inference System (ANFIS).	Besides statistical time series ARMA model, the artificial intelligent model was also considered such as Adaptive Neuro-Fuzzy Inference System (ANFIS) to capture some features in time series	Power consumption	Peak load	STLF	Country (Malasya)	MAPE: 7.26\% (including weekends) and 1.95\% (excluding weekends)	2004-2005. Training phase: 370 data points; Validation phase: 30 data points	The model fails to include weekends	—
[72]	2012	Model Comparison: MAM, LRM, and NN	No information	Energy, usage profile of the building, building status, temperature, external disturbance	24 hourly consumption values	STLF	Mall (micro-grid)	MAPE (MAM)=25.4 % MAPE (LRM)=17.96 % MAPE (NN)=11.06%	1 año anterior.	The disadvantage of the model for practical usage is low interpretability (with the aim of explaining the forecasted values obtained)	The model of artificial neural networks turns out the most accurate among the studied ones. This model could be employed to design energy-saving control loop. problem for electric energy in shopping center with a year of sample size. One is provided by linear regression model with building usage profile
[78]	1996	MLP	A three layered network was used for all models	Power consumption	24 hourly consumption values	STLF	Country (United Arab Emirates)	MAPE 16.47%-19.22%	—	High error figures	Suitable for short time series

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[82]	2009	Interval time series (ITS)+Vector auto-regressive (VAR), multi-layer perceptron model adapted to interval data (iMLP).	VAR Model: Yt=C+fip(B)Yt+epsilont; epsilont is the Kx1 vector of disturbances of the system, C is the Kx1 vector of constants, Yt es a k-dimensional stationary vector time series, fip is the autoregressive KxK matrix polynomial of order p and B is the backshift operator. iMPL: iMLP is comprised of a hidden and an output layer. See Muñoz San Roque et al. 2007 (iMLP: applying multilayer perceptron to interval-valued data)	Annual electricity consumption	Monthly consumption hours	LTLF 2 years (2006 to 2007)	Country(España)	RMSSE (iMLP) = 0.2757–0.4844% RMSSE (ITS-VAR) =0.1952 – 0.3877%	January 1, 2000 to December 12, 2007	The model requires a large amount of data in the training phase	High prediction accuracy
[83]	2013	MLP	Input layer: 29 neurons (Periodic variables are supplied to the network in the form of values of sines and cosines, as it has been demonstrated that this transformation significantly improves the performance of the ANN). Hidden layer: 16. Output layer: 24	Power consumption	24 hourly consumption values	STLF	Microgrid (small city)	MAPE: 2.47%.	3 years.	—	Not only a prediction algorithm, but a complete ANN-based system for forecasting electric load in microgrids. Real world electric load data from a microgrid-sized area.
[84]	2010	Probabilistic Neural Network (PNN)	- The input layer has m-units, to which m-dimensional vector is applied. - The hidden layer has one pattern unit for each pattern exemplar; - The second hidden layer contains one summation unit for each class. Each summation unit is to realize a sum of the type, from the outputs of the previous layer. - The output layer is the decision layer, selects the class with maximum posteriori probability	Power consumption	Half an hour ahead electric load forecasting	VSTLF	Region (Australia).	MAPE: 1.0419%	2004	It is relatively slow to classify, and that it requieres large amounts of memory	PNNs learning speed is better comparatively.
[85]	2005	RBF	Learning rate: 0.04%. Moment Coefficient: 0.3 Nº of Hidden Units: 13. Nº of Output Units: 24. N. of Input units	Power consumption and climate variables	24 hourly consumption values	STLF	Region (Japan).	MAPE: 0.2845%-82.41%	2000-2002	Only 79 days have been used for the forecast: no seasonal effect observed. Special days not considered.	—
[86]	2011	Variation of the SVM: LSSSVM	The model combines LSSV and FFO	Power consumption	Annual load	Next annual	Country (China)	MAPE – 3.00% –+ 3.00%	1978 - 2011	LSSVM+FFO flowchart is long	Artificial intelligence forecasting models have good non-linear fitting capacity, improving accuracy

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[87]	2011	SVR+ chaotic immune algorithm.	The idea of the chaotic algorithm is to transform the problem variables from the solution space to the chaos space and then perform searches to find out the solution based on three characteristics (randomicity, ergodicity and regularity) of the chaotic variables. The chaotic immune algorithm is employed to determine the values of three parameters in a SVR model	Power consumption and climate variables	Monthly load	Next month	Region (Northeast China).	MAPE: 1.766%.	2004-2009	Chaotic immune algorithm flowchart is long	Innovative method combining SVR with chaotic immune algorithm and seasonal adjustment for forecasting monthly electric loads.
[88]	2012	SVR	SVR using least-squares	Power consumption and climate variables	Peak load	Next day	Two datasets: Smart Meter Deployed in a household (Lower Saxony and North Rhine-Westphalia)	MAPE: 1.4%-6.20%	2009. Training phase: Feb. 2009-June 2009. Validation phase: July 2009-Dec. 2009	Small dataset. The model has not been validated for months January to May	The proposed method can be used for utilities in load forecasting for consumer entities at any granularity level
[89]	2005	Information Entropy+Elman ANN	Information entropy theory is used to select relevant variables from all influential factors; the results are used as inputs of neural network. Secondly, according to the features of power load, the typical historical load data samples were selected as the training set which have the same weather characteristic as the certain forecasting day by using data mining theory. Finally, Elman neural network forecasting model is constructed combining the reduced factors and typical training set	Maximum temperature, minimum temperature, average temperature, quantity of rainfall, humidity content and day type	24 hourly consumption values	STLF	Hebei province of China	MSE is about 1.56%	2002	—	The benefit of the hybrid structure was to reduce the training time and improve convergence speed. Improved forecasting accuracy
[90]	2007	H∞ Filter+Elman	Hi filter is introduced to overcome the unknown disturbance and noise in the linear part of the systems during forecasting, and then, Elman dynamic neural network is applied to implement the nonlinear loads prediction	Temperature, humidity, thunderstorms, wind speed, rain, fog, cloud cover or sun shine, hour-of-day, day-of-week, and month-of -year, weekend and holiday effects and electric load of the day before	12 load values on the forecasting day (every 2 hours)	STLF	Guangzhou city of China	—	10 to 21 June 2005	Very small dataset	Compared with normal Kalman filter and BP neural networks, the proposed method possesses high learning efficiency, strong adaptability, high forecasting accuracy and good forecasting behavior, and is very suitable to short-term load forecasting for power system

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[91]	2010	Elman ANN	An Elman neural network with a weather component is proposed for the power load forecasting. Elman neural network can meet nonlinear recognition and process predition of the dynamic system, and make power system having the ability to adapt to time-varying characteristics in mechanism	Power load, the highest temperature, the lowest temperature and the weather characteristic	12 load values on the forecasting day (every 2 hours)	STLF	Hebei province of China.	The maximum difference is about 0.095	2/7/2008 to 11/7/2008	Very small dataset	This model has a good performance in increasing forecasting accuracy because of its inherent dynamic behavior and memory behavior. The forecasting ability of Elman neural networ are better than BP neural network
[92]	2002	SOM+Elman ANN	The prediction process consists of three phases. In the first phase, using historical data, a SOM classifies the days according to their load profile; this classification is performed only once over the life of the package. The second phase involves finding and training the network associated with each class, Elman recurrent networks have been selected. During the third phase, the building ANNs provide the prediction. The SOM is used with a toroidal two-dimension network (10x10 size) where both top and bottom and left and right sides are attached. Elman: the number of hidden layer neurons was selected by experimentation, varying between 30 (class 7) and 100 (class 12), depending on the size of training patterns	Power consumption and climate variables	24 hourly-consumption values	STLF	Country (España).	MAPE: 1.26%–1.91%	1989 to 1999	Elman network requires great time for training. Grouping after SOM is performed by manually selecting the parameters of each class	Good accuracy, human understandable classes (due to manual classification), validation with long database including seasonality
[93]	2012	SOM+k-means	Classification of days in different groups, each of them presenting a distinct load curve pattern and a set of meaningful features. The system is comprised by a SOM to classify the load curves, followed by a clustering with the k-means algorithm. Historical data+data pre-processing+outliear ditection+SOM+clustering+graphic output	Day number, day of the week, month, year, workability and the load curves (24 values) for each day	Load curve: 24 values of electric load for a day (hour)	STLF	"Las Casas" industrial park located in Soria (Castilla y León, Spain)	—	1 January 2008 to 31 December 2010	—	Tested with real world data of a \emph{microgrid}. The validation results show that the system adequately finds different behavior patterns.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[94]	2004	Follow-the-leader algorithm and SOM	The focus is on adequacy assessment of customer classification tools in terms of their accuracy for the classification process. The classification process can be performed by using various clustering tools: Follow-the-leader unsupervised clustering: distance threshold 2.266 and 16 clusters. SOM: 34x16 map and k-means clustering procedure	Power consumption	Classification of electricity customers	Clustering of electricity customers	Country (Rumanía)	—	A measurement campaign that encompassed 234 customers spread all over the country over three-week time intervals	Classification only. Forecast is not provided.	Provides customer classification. The follow-the-leader algorithm makes it effective for applications to large customer sets and for easily isolating uncommon load patterns, while the easy visualization of the SOM makes it useful especially for tutorial and training purposes
[95]	2005	Hierarchical neural model (SOM)	Two different hierarchical neural models operate in parallel. The first one is required to forecast in the interval from the first to the sixth hour. The training of the two SOMs of this model takes place in two phases — coarse-mapping and finetuning. In the coarse-mapping phase, the learning rate and the radius of the neighborhood are reduced linearly whereas in the fine-tuning phase, they are kept constant. The forecasting of the remaining time — seventh to twenty-fourth hour — is addressed by the second model. The same training process is applied	Load curves (24 values) for each day	24 hourly consumption values	STLF	Brazilian electric utility	MAPE (%)=2.33-2.03	3 months, February, 1995	Not tested with weather information or large number of samples	The results are better than those obtained by MLP on equal data.
[96]	2007	Pattern recognition methods: k-means, SOM, fuzzy k-means and hierarchical clustering	The classification of customers is achieved by applying a pattern recognition methodology consisting of two stages. Firs stage: - Data and features selection. - Customers' clustering using a priori indices. - data preprocessing. - Typical load curves clustering for each customer. - Selection of typical chronological load curves for customers. Second stage: Clustering of customers	Power consumption	Classification of electricity customers	STLF	Country (Greek)	No information	Ten months in the year 2003	Flow of the load forecasting process is long. Only provides clustering.	Provides typical daily load curves for each customer.
[97]	2012	k-means, Fuzzy and SOM	This approach clusters load curves of customers : k-means, fuzzy c-means and the self-organizing map.	Load curves (48 values) for each day	48 values of electric load for a day (1/2 hour)	STLF	Medium voltage distribution system in Northern China	—	August 2009	Different normalization methods may cause different clustering results, so the effects of normalization methods on clustering results should be studied further	The main advantage of work from previous studies is in the proposal of a stability index to characterize the performance of the clustering methods and the proposal of a priority index to rank the number of clusters.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[98]	2006	SOM+Fuzzy+CBR	This study presents an STLF model that applies SOM and CBR: The fuzzy-rough attributes reduction method is first used to select the most significant variables as case components; then some clusters are derived from all cases by SOM clustering; afterwards each case is compared to the small set of cluster centroid values	Weather, day type and load.	One day ahead maximum load forecasting.	STLF	Hang Zhou Electric Power Company (HZEPC) in China	MAPE (%)=0.05	Year 2000 to 2001	—	SOM-CBR approach alleviates the computational complexity by use of cluster-centroid values obtained from SOM process. It yields superior results compared to ANN network, similar day method and conventional CBR method.
[99]	2011	Fuzzy c-means clustering based on particle swarm optimization SVR model (FCM Based PSO-SVR) and particle swarm optimization SVR model (PSO-SVR).	Flow chart of the forecasting process: - Input and pre-processing of the historical load data - Creation of the basic sample set and prediction of input samples. - Forecasting from t=1 to t=24: a) Optimize the FCM clusters to form the basic sample set in period t; b) Calculate the distance from the forecast input vector and the cluster center in period t; c) Select the minimum distance of the corresponding subclass as the PSO-SVR prediction model sample set; d) Determine the SVR model parameters by using PSO optimal algorithm and the selected sample; e) Establish the objective function, and form the optimal regression equation according to the selected training samples and the SVR parameters; f) Obtain the load forecasting value at period t	Power consumption	Electric load in a specific time.	STLF	City (Chongqing).	MAPE (FCM Based PSO-SVR)=1.066 % MAPE in a day (PSO-SVR)=1.443 % MAPE in a week (FCM Based PSO-SVR)=1.218 % MAPE in a week (PSO-SVR)=1.542 %	Not specified.	Flow of the load forecasting process is long. Computing power-greedy	This method increases the accuracy and reduces the number of samples needed
[100]	2012	AFCM, optimal selection approach of SVR	AFCM is affected by many parameters, pertaining to SVR and SOM. The parameters used in the PSO-BR are: a) The first set related to BP neural network, input layer dimension 2, hidden layer dimension 3 and output layer dimension 1; b) The second set related to PSO, maximum iteration number 300, number of particles 400, length of particles 3 and weight 2.	Half an hour ahead electric load	Half an hour ahead electric load	2 forecast: 1 day and 7 days	Region (New South Wales)	MAPE (SVR)=11.6955 – 12.8765% MAPE (PSO-SVR)=11.4189 – 13.503% MAPE (PSO-P)=10.9094 – 12.2384% MAPE (AFCM)=9.9524 – 11.1019%	Training phase: 5 days; Validation phase: 1 day. Training phase: 23 days; Validation phase: 7 days.	Very small dataset for validation	The adaptive fuzzy combination model can effectively count for electric load forecasting with good accuracy and interpretability at the same time.
[101]	2010	NN: Adaptive-network-based fuzzy inference system (ANFIS)	The process model is: - Determine the inputs of the model. - Collect data set in all available previous periods for each of input variables and output variable. - Divide the data into two sets. - The value of input variables in coming periods are predicted by using Auto Regressive model. - The yielded values are fed to the selected ANFIS - finally, the value of electric consumption in each coming year is calculated by selected ANFIS and values of input variables that are fed to ANFIS	Gross Domestic Product (GDP) and Population (POP)	Annual load	LTLF (1 year). The next 8 years	Countries G7	MAPE: U.S.A=0.005696% Italy=0.014929% United Kingdom=0.00446% Japan=0.007985% France=0.012971% Germany=0.013136% Canada=0.011739%	1980 to 2007. Training phase: 1980-2003. Validation phase: 2004-2007.	Very large amount of data required in the training phase. The process model is complex..	The ANFIS algorithm is capable of dealing both with data complexity and ambiguity due to its mechanism, providing high precision.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[102]	2012	Type-2 Self-Developing and Self-Adaptive Fuzzy Neural Networks (T2SDSA-FNN).	Layer 1-input nodes: each node in this layer, wich represents an input variable to FLS, directly transmits input vector to next layer. - Layer 2-Fuzzification nodes: each node in layer 2 represents one linguistic load of one input variable, the output of this node specifies the degree to which a T2 input variable intersects with k-th T2 fuzzy set associated with j-th input variable in layer 1. - Layer 3-pre-condition node: each node in this layer represents one fuzzy rule and performs antecedent matching of this rule, the output of each node in this layer represents the firing strength of corresponding fuzzy rule due to activation of its antecedent by the observed input. - Layer 4-Consequent node: node in this layer is called a consequent node, the output of this node is T2 fuzzy set defined by membership function. - layer 5-Join node: node in this layer joins outputs of M rules and then generates one T2 fuzzy set with membership function. - Layer 6-Output node: node in this layer converts T2 fuzzy set in layer 5 into crisp number by processing two steps	Power consumption and temperature	Electric load in a day and a time interval	Tres: 1 day, 7 days y 14 days	Region (Macao).	MAPE (1 day) = 2.0428% MAPE (7 days) = 3.4096% MAPE (14 days) = 2.552%	January to December 2010	Flow of the load forecasting process is long	Improved computation time. Simulation and test results reveal that it has superior accuracy performance on electric forecasting problem than other techniques shown in existing literatures
[103]	2006	DA-preconditioned RBFN with DA clustering	As a precondition technique, Deterministic Annealing (DA) is used to classify input data into some clusters. The RBFN is employed as ANN at each cluster so that one-step ahead temperature is evaluated precisely.	Average temperature, maximum temperature, minimum temperature, average humidity, minimum humidity, wind direction, maximum wind speed, daylight and isolation	One day ahead maximum temperature forecasting	STLF	Tokyo	MAPE (%) = 2.17	1999 to 2001	—	The proposed method was successfully applied to real data of one-step ahead daily maximum temperature. The simulation results have shown that the proposed method outperforms others. The proposed method allows the power system operators to deal with short term temperature forecasting appropriately
[104]	2001	DA+MLP	Hybrid method of ANN and the DA clustering for short-term load forecasting. As a data-portioning technique, the DA clustering plays a significant role to classify input data into clusters efficiently so that MLP has input data smaller variance and enhances the prediction error at each cluster	Daily maximum load	Daily maximum load forecasting	STLF	Japan	MAPE (%) = 1.394%	5 years of summer days	Weather data are not used	The proposed method alleviates the nonlinearity of time series with data partitioning. Classified data has smaller variance in each cluster. The proposed method is capable of handling special days or unexpected events although the accumulation of historical data is needed to some extent

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[105]	2007	NN+EA	A new bilevel prediction strategy is proposed. The low level includes a feature selection method and hybrid forecast engine. The first step is to identify factors that affect load signal. The set of candidate inputs should be first refined so that a minimum subset of the most informative features is selected and the other unimportant candidates are filtered out	Power consumption and temperature	24 hourly consumption values	STLF	Microgrid	MAPE: Sep. 2008 2.85%; Dec. 2008 1.94%; Mar. 2009 2.044%; Jun. 2009 2.77%	Training Pphase: - Validation phase: September 2008, December 2008, March 2009 and June 2009	Flow of the load forecasting process is long. Small dataset for validation	Better forecast accuracy and stability of the proposed bilevel prediction strategy is shown for STLF of a \emph{microgrid} compared with several other forecast methods
[106]	2011	NN: Grey model and RBF	Forecasting model involves the following steps: first the historical data are decomposed by DWT and then reconstructed by each frequency; second, the rising trend is modeled using Grey model (1,1) and other periodical waves are determined using the RBF; third, the forecasting results of monthly electric energy consumption are obtained by adding the forecasting values of the previous models	Load month consumption	Monthly total consumption	2 years 2005 to 2006	Country (China)	MAPE = 2.67%	1990 to 2006	The model requires large amount of data in the training phase	Reduction of the maximum forecasting error; second, the application of DWT in forecasting may result in a simultaneous reduction of the MAPE and maximum error, and third, when GM and DWT are used simultaneously, the MAPE and maximum error are both further reduced
[107]	2012	Optimal training subset+SVR	Pre-processing variables + SVR	The half an hour ahead electric load	The half an hour ahead electric load	3 forecasts: 1 day, 15 days y 1 mounth	Region (New South Wales)	MAPE = 12.0965 – 16.1032%	Three datasets: 3 days, 15 days and 1 month. 48 values per day. Validation phase: 1 day	The model needs a preprocessing phase (ad-hoc) of the input variables. Extremely small dataset for validation (1 day).	Reduced computation time in the SVR algorithm
[108]	2010	Chaotic ant swarm optimization (CAS)+SVR	The model is: - SVR. - CAS optimization in parameters selection for the SVR model. - Chaotic ant swarm optimization	Annual load	Annual load	LTLF 4 years. (1997 to 2000)	Region (Taiwan)	MAPE (SVR-CAS) = 1.3083 –2.2262 % MAPE (SVRCGA) = 1.3558-2.5695% MAPE (SVR-CPSO) = 1.3187-2.1860% MAPE (ANN) = 1.06-3.62% MAPE (Regression) = 2.45-8.52%	1981 to 2000. Training phase: 1981-1996. Validation phase: 1997-2000	The model requires large amount of data in the training phase. Flow of the load forecasting process is long	Solves premature convergence

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[109]	2011	Hybrid chaotic sequence and evolutionary algorithms +SVR	- SVR - CABC algorithm for parameters determination of the SVR model - chaotic artificial bee colony algorithm - Recurrent learning mechanism framework in an SVR model - Seasonal adjustment	Mounthly load	Mounthly load	LTLF (Prediction of 7 months of consumption)	Region (North of China)	MAPE (ARIMA) = 6.044% MAPE (TF-ε-SVR-SA) =3 .799% MAPE (SSVR-CABC) = 3.056% MAPE (SRSVR-CABC) =2.387%	2004 to 2009. Training phase: 32 months. Validation phase: 7 months	Flow of the load forecasting process is too long	Prevents premature convergence (avoid local minimum) and has a higher prediction performance. The forecasting results indicate that the proposed model yields more accurate forecasting results than ARIMA and TF-ε-SVR-SA models. Therefore, the SRSVRCABC model is a promising alternative for electric load forecasting
[110]	2012	Chaotic genetic algorithm-simulated annealing algorithm	The SSVRCGASA model includes the formulation of SVR, the CGASA algorithm and the seasonal adjustment process. CGASA is used to optimize and select well parameter combination of an SVR model. Deo and Hurvich's approach is used to adjust the cyclic effects (seasonal), namely seasonal mechanism	Mounthly load	Mounthly load	LTLF (Prediction of 7 months of consumption)	Region (North of China)	MAPE (ARIMA) = 6.044% MAPE (TF-\checkmark-SVR-SA) = 3.799% MAPE (SVRC-GASA) = 3.731% MAPE (SSVRC-GASA) = 1.901%	December 2004 - September 2008. Training phase: 39 months; Validation phase: 14 months	Flow of the load forecasting process is long	Novel evolutionary algorithms (hybrid genetic algorithm with simulated annealing algorithm) to enhance the superior capabilities of the SVR model mentioned above
[111]	2013	Chaotic genetic algorithm to determine the parameters of SVR (SSVR-CGA)	The SSVRCGA model includes the formulation of SVR and the CGA algorithm	Mounthly load	Mounthly load	LTLF (Prediction of 7 months of consumption)	Region (North of China)	MAPE (ARIMA) = 6.044% MAPE (TF-e-SVR-SA) = 3.799% MAPE (SVRCGA) = 3.382% MAPE (SSVRCGA) = 2.659%	December 2004–September 2008. Training phase: 39 months; Validation phase: 14 months	Flow of the load forecasting process is long	Experiment results indicate that the proposed SSVRCGA model has significant superiority among other alternatives in terms of forecasting accuracy.
[112]	2012	NN: ESN	24 ESNs are considered where each one operates as a 24-h ahead forecaster at a specific hour of the day	Power consumption , temperature, day of the week	Daily electric load	STLF	Nation (Utility in USA)	MAPE (1h) = 1.048% MAPE (24h) = 2.1174%	15.000 hours (45 months)	If the predicted value of an exogenous variable (temperature) at the target time is selected as an input, its poor prediction can lead to large errors in load forecasting. Then, It is necessary to predict the temperature in a separate model, complicating the global model.	The trend of temperature errors on the ESNs load forecasts shows more sensitive to positive deviations than to negative ones

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[113]	2013	FOA+ GRNN with PSO-GRNN, SALSSVM, OLS-LR	FOA to automatically select the spread parameter accuracy in the annual power load forecasting	Annual consumption	Annual consumption	5 years (2006-2010)	Beijing city and China	Beijing-FOA-GRNN = 1.149% Beijing-GRNN = 2.392% Beijing-PSOGRNN = 1.857% Beijing-SALSSVM = 1.367% Beijing-OLS-LR = 2.744% China-FOAGRNN = 1.252% China-GRNN = 2.746% China-PSOGRNN = 2.533% China-SALSSVM = 3.232% China-OLS-LR = 2.061%	1978 to 2010. Training phase: 1981-2005. Validation phase: 2006-2010	The model requires large amount of data in the training phase. Flow of the load forecasting process is long	Improves the performance of the model, because the methodology selects the best parameters for the GRNN
[114]	1997	Kalman filtering algorithm with "fading memory"	In the first stage, 168 separate filters are applied in parallel to 168 separate hour-of-the-week load sequences. In stage 2, hourly corrections are carried out. Forecast errors increase if there is excessive noise present in the data, and usually some smoothing technique is used to increase the signal to noise ratio for STLF.	Power consumption and climate variables	24 hourly consumption values	STLF	Substation	MAPE: 8.4%-11.5%.	2 years.	—	—
[115]	1991	MLP	Two hidden layers with 46 hidden units for each layer. In the input layer they have 46 nodes, each receiving an input signal from the external world. There is only one output, which gives either peak load or valley load. For the problem of peak and valley load forecasting, they know in advance the output (peak and valley load) for each input pattern (weather data and others) in the training set	Annual consumption and three temperatures of Taiwan (north, center and south)	24 hourly consumption values	STLF	Taiwan (Region of China)	MAPE: 0.014%-2.4%	2 months in 1987. Validation of model with 4 days	Very complex structure with two hidden layers is not necessary Very small dateset for validation	It is concluded from the study results that accurate forecasting of hourly loads can be achieved by the neural network in a very efficient manner
[116]	2002	ANN	This study investigates the use of weather ensemble predictions in the application of NNs to load forecasting for lead times from 1 to 10 days ahead. A weather ensemble prediction consists of multiple scenarios for a weather variable. We use these scenarios to produce multiple scenarios for load	Electric load, spot temperature, wind speed and cloud cover	Daily maximum load forecasting	1 to 10 days ahead.	England and Wales	MAPE = 1.7-2.6%	November 1998 to 30 June 2000	—	The results show that the average of the load scenarios is more accurate than that produced using traditional weather forecasts

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[117]	2009	ANN	This paper proposes a multiregion short-term load-forecasting methodology, taking into account the Heat Index to improve load-forecasting accuracy in Taiwan Power Company's (Taipower's) system	Electric load and Heat Index (HI)	24 values of electric load for a day (hour)	STLF	Taiwan Power Company's (Taipower's) system	MAPE (%) = 1.46	2005 and 2006	HI leads to more accurate forecasting during summer only when the temperature is higher [27°C (80°F)]	The numerical results demonstrate a positive improvement on load forecasting for Taipower's system by applying the multiregion model and using HI as an ANN parameter
[118]	2011	Modified general regression NN	This paper uses two methodologies for short-term multinodal load forecasting. The first individually forecasts the local loads and the second forecasts the global load and the load participation factors to estimate the local loads. For the forecasts, a modified general regression neural network and a procedure to automatically reduce the number of inputs of the artificial neural networks are proposed	Date information (day/month/year, day of the week, daylight saving time, holidays, and sample number) and half-hourly active load from each of the nine electrical substations	48 values of electric load for a day (1/2 hour)	STLF	New Zealand distribution subsystem (nine electrical substations)	MAPE (%) = 2.93-3.21	January 2007 to March 2009	In most of the cases, daily peak values were not predicted correctly	Good generalization ability, stability, and training, with the ability to provide faster forecasting and reduce the number of inputs, avoiding redundancies that may compromise the results in some cases
[119]	2011	MLR+MA+ANN+Fuzzy	The weather-sensitive ANN, multiple linear regression approach, and the non-weather-sensitive moving average are used to forecast the load and generation. To minimize the forecast errors and increase the robustness, a fuzzy-based grouping of the forecasting-method results is employed	Temperature, cloud cover, precipitation, wind speed, load and generation data	24 values of electric load for a day (hour)	STLF	Slovenian power system	MAPE (%) = 2.02	1 year	—	The use of combined forecasts gives the lowest transmission-loss forecast MAPE and standard deviation, indicating improved accuracy compared to the use of individual forecasting methods.
[120]	2011	ANLR+SVM+QP	A two-stage hybrid forecasting method. In the first stage, daily load is forecasted by time-series methods; in the second stage, the deviation caused by time-series methods is forecasted considering the impact of relative factors, and then is added to the result of the first stage. On the basis of this analysis, an adaptive algorithm is used to perform the second stage which can be used to choose the most appropriate algorithm among linear regression (ANLR), quadratic programming (QP) and support vector machine (SVM) according to the characteristic of historical data	Temperature, humidity, load and day type	96 values of electric load for a day (1/4 hour)	STLF	Two typical Chinese power grids (Beijing and Jiangxi)	MAPE (%) = 4.0088 - 5.2129	June 8, 2005 to September 6, 2005	Small dataset for validation	The SFP has already been embedded as an independent module into existing STLF software to amend the widely used time-series methods.

No.	Year	Model	Architecture	Variables	Forecast	Horizon	Application Area	MAPE or RMSE	Data	Drawbacks	Advantages
[121]	2013	Neural Network Ensemble (NNE)	A robust NNE based on Regularized Negative Correlation Learning (RNCL) is proposed for an enhanced learning model with improved generalization ability and reduced variance and performances instability with a relative simple implementation	Power consumption and climate variables	24 values of electric load	STLF	Building (Tokyo)	MAPE: 3.04%-3,.1%.	No information	—	The experimental results showed that the NNE achieved a higher forecasting accuracy that conventional MLPNN, RBFNN and RNN
[126]	2010	ANN	NN is incorporated to forecast power generation of a PV energy source	Wind speed and direction, solar irradiation and circumstance temperature, solar power and day type	24 values of power generation of the PV instalation for a day (hour)	STLF	Wuhan city of the province of Hubei (China).	MAPE (%) = 16.47-11.59	19-25 December 2006	Some further researches can be performed to perfect this method by concerning more relevant factors, such as the industrial and commercial profiles of a city or region	The results show that the forecasting model is able to predict hourly power generation according to the weather forecasting inputs
[128]	2010	EDE	The proposed strategy is composed of a feature selection technique and a forecast engine (including neural network and evolutionary algorithm) in the lower level as the forecaster and an enhanced differential evolution algorithm in the upper level for optimizing the performance of the forecaster	Weather factors, calendar factors and load	24 values of electric load for a day (hour)	STLF	Load data of the University of Calgary (Canada)	Weekly Mean Error (WME) = 2.40%	Four seasons (first week of September 2008, December 2008, March 2009, and June 2009)	Prediction of realized demand (load demand minus generation of renewable resources), which is another important issue for the operation of microgrids, will be considered in future works	Better forecast accuracy and stability of the proposed bilevel prediction strategy is shown for STLF of a microgrid compared with several other forecast methods
[129]	2011	Multiple Classifier System (MCS)	The base classifier can be different types or trained differently, such as by different algorithms and data sets. A fusion method is used to combine the decisions of base classifiers	Power consumption and climate variables	24 values of electric load for a day (hour)	STLF	Microgrid	MAPE: 15.12%–15.66%	September 2008 to August 2010, Validation phase: August 2009 to August 2010	Big errors	MCS can adapt the highly fluctuation of load series in Microgrid
[138]	2011	SVM and MLP	MLP: one hidden layer with 10 neurons. v-SRV using a RBF as kernel and parameters: v=0.9, epsilon=0.001	Power consumption	24 values of electric load for a day (hour)	STLF	Microgrid (Campus)	MAPE: MLP 13.26%; SVM 7.92%	1 year	—	—
[139]	2011	Fuzzy Neural Network	—	Power consumption	24 values of electric load for a day (hour)	STLF	Microgrid (Campus)	MAPE: 3.69%-4.04%	4 months	Reduced dataset. The model is not validated for every month of the year	Application to \emph{microgrid}. The proposed model STLF could be modified and extended to mid-term and long-term load forecasting
[140]	2012	SOM and MLP	SOM: 3x3, gaussian function, hexagonal. MLP: neurons of the input layer 96, neurons of the hidden layer 8, neurons of the output layer 24,backpropagation	Power consumption and climate variables	24 values of electric load for a day (hour)	STLF	Microgrid	MAPE MLP: 12%-14%	6 months	Reduced dataset. The model is not validated for every month of the year	Application to real world \emph{microgrid} data.
[141]	2012	MLP and Formules	MLP: MLP is used for Thermal load forecasting. Input layer neurons 13, Hidden layer neurons 10 and Output layer neurons 1. Flow chart of the load forecasting and operation planning in \emph{microgrid}: thermal load prediction, operation planning of the heat sources, estimation of heat sources power, estimation of power consumption, operation planning of DGs and ESSs, adjustment of DGs and ESSs outputs	Power consumption and climate variables	24 values of electric load for a day (hour)	STLF	Microgrid	MAPE Heat load: 11.8%. MAPE Electrical load: 6.7%	—	No information available about the dataset	The paper showed the electrical load forecasting method and the optimized operation planning method for microgrid. For the electrical load prediction, the paper used an operation planning of the heat sources which run according to heat load prediction

to interpret the random nature of demand and represent it using equations in the model.

Even though McCulloch and Pitts (1943) [149] already suggested transferring neural behavior into computational environments in 1943, it was only in the late 80s when *ANN*-based models started to be applied to demand forecasting. The improvement in the training stage and, above all, the ability to generalize and to detect non-linearities inherent to electric demand made *ANNs* triumph over *ARMA* and similar models in the late 20th and early 21st centuries. Nowadays, these models are much more widely used by the scientific community and the industry due to their high efficiency, the relatively small amount of time needed to set up the system, and their performance. In these systems, the researcher must focus on feeding the model with suitable inputs and the network handles the task of finding complex associations inherent in the problem, with impressive results.

Additionally, the 21st century has brought up the conceptualization of new generation, transportation, distribution and consumption environments for electric power, with *SGs* as the core. Similarly, the increasing presence of *DG* in power grids and the need for the system to operate more efficiently in order to avoid unnecessary distribution and transportation loss, has brought the emergence of disaggregated environments (*VPP*, *microgrid*, *Smart Building*, among others). In these new environments, a balance must be found between consumption and storage, which requires information and knowledge about what happens around these elements. This disaggregation and, therefore, the emergence of new intelligent spaces, will involve the need to come up with demand and generation forecasting models which, regardless of the model, will enable the introduction of new, more suitable variables (weather variables at a highly local scale, behavior and stay of users, use of facilities, etc.), so as to use them in operation management systems and plan these in the future. For these environments, next hour's load, next day's peak load, next day's total load and *load profile* will continue to be of great interest. For example, *VPPs* pose a challenge for demand forecasting and generation, and a possible approach is to use a management model which takes into account the multiple elements that are part of it, making them cooperate to obtain a demand forecast via *ANN* in disaggregated environments, as shown by Hernández *et al.* (2013) [150]. Intelligence will involve the deployment of measurements. Therefore, more recent data will be available and the past behavioural data will be more recent, thus allowing the models to be adjusted (or trained, in the case of *ANNs*) more regularly. This, in turn, will provide more information on the changes inherent to the demand (such as the influence of social or economical changes in the area of interest), which will also be more and more frequent in this increasingly changing world.

Regarding the future evolution of the field, it seems clear that there is a main direction that research and development is taking: the integration of forecasting technologies as a key input element to support the increasing intelligence of the *SG*. Therefore, it is expected to see an increasing interest in the usage of predictions to automate infrastructure operations such as energy balancing, resource planning and auto-diagnosis. Some works have been already opening this direction, specifically in the field of intelligent scheduling.

Intelligent scheduling of generation sources and loads is essential to the operation of a *microgrid*, in order to allow the integration of highly dynamic *Distributed Energy Resources* (*DERs*) while ensuring stability and reliability. A major limitation of intelligent scheduling implies the unrealistic requirement of constant human intervention during the scheduling of loads and generators, motivating the need for algorithms that automatically modulate the generation or consumption levels with uncertain renewable power availability. Traditional generation scheduling paradigms rely on perfect prediction of future electricity supply and demand. Narayanaswamy *et al.* (2012) [151] present for instance an online algorithm designed to compensate for the renewable resource when they are not available, taking also into account physical generator constraints; This work also includes online algorithms that intelligently leverage available information about the future, such as predictions of wind intensity and loads, and show that they can be used to guarantee near optimal performance under mild assumptions. Lu *et al.* (2013) [152] study online algorithms for the *microgrid* generation scheduling problem with intermittent renewable energy sources and co-generation, with the goal of maximizing the cost-savings with local generation and local loads; the authors propose a class of competitive online algorithms, called *CHASE* (*Competitive Heuristic Algorithm for Scheduling Energy-generation*), that track the offline optimal in an online fashion. Alvarez *et al.* (2009) [153] present a new online dispatch algorithm for both electrical active power and heat in the same installation (*microgrid*); its objective is to minimize running cost, adjust the generation of active power at the point of demand (while in grid connected mode), as well as to cover heat demand; the new algorithm uses a heuristic approach, based on electrical generation cost functions of microsources, to improve the results obtained using state-of-the-art nonlinear programming optimization methods.

Additionally, [154] predicts an increase in demand, generation, electricity prices, and emissions from the utilities created by the introduction of *EVs* and *Plug-in Hybrid Electric Vehicles* (*PHEVs*); it also suggests that by 2030 almost all regions (in US) will need to add capacity to provide for charging *PHEVs*, mostly in the scenario where *PHEVs* are charged at 6 kW in the evenings; to avoid these problems the *utilities* in the regions would expand their capacity, increase their imports, or establish *DR* programs. This effect should be taken into account by the predictive models. As shown by Pérez *et al.* [155], without coordination of the charging, the vehicles are charged instantaneously when they are plugged in or after a fixed start delay; this uncoordinated power consumption on a local scale can lead to grid problems therefore, coordinated charging is proposed to minimize the power losses and to maximize the main grid load factor; the optimal charging profile of the *PHEVs* and *EVs* is computed by minimizing the power losses and it is required load forecast.

On the other hand, while the literature shows an enormous interest in the *microgrid* domain in general, and the application of these forecasting techniques in particular, the vast majority of the available works deal with theoretical or simulated results. It is detected a lack of field experiences and real world deployment of these technologies, and therefore the availability of experimental evidence is low. However, it is expected that this kind of works will start to appear progressively as *microgrid* technologies are rolled out towards the general public.

In any case, it is worth noting that *microgrid* technologies are very promising as one of the driving forces behind the energy grid of the future, and they rely on the availability of forecasts in order to be able to plan, manage and automate their operations. And luckily, as this work shows, the literature is rich in studies showing how to apply different techniques in order to obtain accurate and effective predictions.

References

[1] H. A. Mantooth, "Smart Grid: Definition, Road map and Breadth of this Grand Challenge," Connected Planet, (Accessed 12/20/2011), 2010.

[2] H. C. Ferreira, L. Lampe, J. Newbury, and T. G. Swart editors, "Power Line Communications. Theory and Applications for Narrowband and Broadband Communications over Power Lines," Ed. John Wiley and Sons, second edition, 2010.

[3] M. Sánchez-Jimínez, "European SmartGrids Technology Platform. Vision and Strategy for Europe's Electricity Networks of the Future," European Commission: Directorate-General for Research. (Accessed 01/10/2012), 2006.

[4] M. B. Nissen, "High performance development as distributed generation," IEEE Potentials, vol. 28, no. 6, pp. 25-31, 2009.

[5] M. E. Hamedani and S. A. Arefifar, "Distributed generation, reactive sources and network-configuration planning for power and energy-loss reduction," IEE Proceedings-Generation, Transmission & Distribution, vol. 153, no. (2), pp. 127-136, 2006.

[6] B. Wille-Haussmann, T. Erge, and C. Wittner, "Decentralized optimisation of cogeneration in virtual power plant," Solar Energy, vol. 84, no. 4, pp. 604-611, 2010.

[7] M. Sebastian and J. Maire, "Towards Virtual Power Plant integration into the Transmission and Distribution grid tools," Bulleting Editor Michael Durstewitz ISET, pp. 1-2, 2008.

[8] I. Bel, A. Valenti, J. Maire, J. M. Corera, and P. Lang, "Innovative Operation with Aggregated Distributed Generation," in proc., 19th International Conference on Electricity Distribution, May 2007, pp. 1-4.

[9] R. Lasseter, A. Akhil, C. Mamy, J. Stephens, J. Dagle, R. Guttromson, A. S. Meliopoulous, R. Yinger, and J. Eto, "Integration of Distributed Energy Resources. The CERTS Microgrid Concept," White Paper. (Accessed 01/01/2013). http://certs.lbl.gov/pdf/50829.pdf, 2002.

[10] J. Anduaga, M. Boyra, I. Cobelo, E. García, A. Gil, J. Jimeno, I. Laresgoiti, J. M. Oyarzabal, R. Rodríguez, E. Sánchez, F. J. Santiago, E. Turienzo, E. Zabala, and E. Perea, "La microrred, una alternativa de futuro para un suministro energético integral," TECNALIA, Corporación Tecnológica, 2008.

[11] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart Grid – The New and Improved Power Grid: A Survey," IEEE Communications Surveys & Tutorials, vol. 14, no. 4, pp. 944-980, 2012.

[12] L. Hernández, C. Baladrón, J. M. Aguiar, L. Calavia, B. Carro, A. Sánchez-Esguevillas, D. J. Cook, D. Chinarro, and J. J. Gómez, "A Study of the Relationship between Weather Variables and Electric Power Demand inside a Smart Grid/Smart World Framework," Sensors, vol. 12, no. 9, pp. 11571-11591, 2012.

[13] P. J. Werbos, "Computational Intelligence for the Smart Grid-History, Challenges, and Opportunities," IEEE Computational Intelligence Magazine, vol. 6, no. 3; pp. 14-21, 2011.

[14] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A Survey on Cyber Security for Smart Grid Communications," IEEE Commun. Surveys & Tutorials, vol. 14, no. 4, pp. 998-1010, 2012.

[15] F. Javed, N. Arshad, F. Wallin, I. Vassileva, and E. Dahlquist, "Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting," Applied Energy, vol. 96, pp. 150-160, 2012.

[16] H. S. Hippert, C. E. Pedreira, and C. R. Souza, "Neural Networks for Short-Term Load Forecasting: A review and Evaluation," IEEE Trans. Power Syst., vol. 16, no. 1, pp. 44-51, 2001.

[17] J. Toyoda, M.-S. Chen, Y. and Inoue, "An Application of State Estimation to Short-Term Load Forecasting, Part I: Forecasting Modeling," IEEE Trans. Power App. Syst., vol. PAS-89, no. 7, pp. 1678-1682, 1970.

[18] W. Charytoniuk and M.-S. Chen, "Very Short-Term Load Forecasting Using Artificial Neural Networks," IEEE Trans. Power Syst., vol. 15, no. 1, pp. 263-268, 2000.

[19] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, "Very Short-Term Load Forecasting: Wavelet Neural Networks With Data Pre-Filtering," IEEE Trans. Power Syst., vol. 28, no. 1, 2013.

[20] E. Doveh, P. Feigin, D. Greig, and L. Hyams, "Experience with FNN Models for Medium Term Power Demand Predictions," IEEE Trans. Power Syst., vol. 14, no. 2, 1999.

[21] D. Asber, S. Lefebvre, M. Saad, and C. Desbiens, "Modeling of Distribution Loads for Short and Medium-Term Load Forecasting," in proc., IEEE Power Engineering Society General Meeting, 2007, June 2007, pp. 1-5.

[22] M. S. Kandil, S. M. El-Debeiky, and N. E. Hasanien, "Long-Term Load Forecasting for Fast developing Utility Using a Knowledge-Based Expert System," IEEE Trans. Power Syst., vol. 17, no. 2, 2002.

[23] Z. Zhang and S. Ye, "Long Term Load Forecasting and Recommendations for China Based on Support Vector Regression," in proc., 2011 International Conference on Information Management, Innovation Managemenet and Industrial Engineering, Nov. 2011, vol. III, pp. 597-602.

[24] H. Daneshi, M. Shahidehpour, and A. L. Choobbari, "Long-Term Load Forecasting in Electricity Market," in proc., 2008 IEEE International Conference on Electro/Information Technology, May 2008, pp. 395-400.

[25] D. K. A. Gillies, B. Bernholtz, and P. J. Sandiford, "A New Approach to Forecasting Daily Peak Loads," Transactions of the Ametrican Institute of Electrical Engineers, Power Apparatus and Systems, Part III, vol. 75, no. 3, pp. 382-387, 1956.

[26] D. C. Park, M. A. El-Sharkawi, R. J. Marks II, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," IEEE Trans. Power Syst., vol. 6, no. 2, pp. 442-449, 1991.

[27] K.-L. Ho, Y.-Y. Hsu, and C.-C. Yang, "Short term load forecasting using a multilayer neural network with and adaptive learning algorithm," IEEE Trans. Power Syst., vol. 7, no. 1, pp. 141-149, 1992.

[28] K.-L. Ho, Y.-Y. Hsu, C.-F. Chen, T.-E. Lee, C.-C. Liang, T.-S. Lai, T-S., and K.-K. Chen, "Short term load forecasting of Taiwan power system using a knowledge-based expert system," IEEE Trans. Power Syst., vol. 5, no. 4, pp. 1214-1221, 1990.

[29] I. Drezga and S. Rahman, "Input variable selection for ANN-based short-term load forecasting," IEEE Trans. Power Syst., vol. 13, no. 4, pp. 1238-1244, 1988.

[30] I. Drezga and S. Rahman, "Short-term load forecasting with local ANN predictors," IEEE Trans. Power Syst., vol. 14, no. 3, pp. 844-850, 1999.

[31] J. S. McMenamin and F. A. Monforte, "Short-term energy forecasting with neural networks," The Energy Journal, vol. 19, no. 4, pp. 43-61, 1998.

[32] K. Y. Lee, Y. T. Cha, and J. H. Park, "Short-term load forecasting using an artificial neural network," IEEE Trans. Power Syst., vol. 7, no. 1, pp. 124-132, 1992.

[33] C.-N. Lu, H.-T. Wu, and S. Vemuri, "Neural network based short term load forecasting," IEEE Trans. Power Syst., vol. 8, no. 1, pp. 336-342, 1993.

[34] A. D. Papalexopoulos, S. Hao, and T.-M. Peng, "An implementation of a neural network based load forecasting models for the EMS," IEEE Trans. Power Syst., vol. 9, no. 4, pp. 1956-1962, 1994.

[35] A. G. Bakirtzis, V. Petridis, S. J. Kiartzis, M. C. Alexiadis, and A. H. Maissis, "A neural network short term load forecasting model for the Greek power system," IEEE Trans. Power Syst., vol. 11, no. 2, pp. 858-863, 1995.

[36] A. S. Alfuhaid, M. A. El-Sayed, and M. S. Mahmoud, "Cascaded artificial neural networks for short-term load forecasting," IEEE Trans. Power Syst., vol. 12, no. 4, pp. 1524-1529, 1997.

[37] R. Lamedica, A. Prudenzi, M. Sforma, M. Caciotta, and V. O. Cencelli, "A neural network based technique for short-term forecasting of anomalous load periods," IEEE Trans. Power Syst., vol. 11, no. 4, pp. 1749-1756, 1996.

[38] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J. Farah, and C. Drake, "Practical experiences with an adaptive neural network short-term load forecasting system," IEEE Trans. Power Syst., vol. 10, no. 1, pp. 254-265, 1995.

[39] A. Piras, A. Germond, B. Buchenel, K. Imhof, and Y. Jaccard, "Heterogeneous artificial neural network for short term electrical load forecasting," IEEE Trans. Power Syst., vol. 11, no. 1, pp. 397-402, 1996.

[40] J. Anderson, J. Silverstein, S. Ritz, and R. Jones, "Distinctive features, categorical perception and probability learning: some applications on a neural model," Psychological Review, vol. 84, pp. 413-451, 1977.

[41] T. Kohonen, "Associative Memory. A system theoretical approach," Ed. Springer-Verlag, 1977.

[42] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biological Cybernetics, vol. 36, no. 4, pp. 192-202, 1980.

[43] J. J. Hopfield, "Neural Networks and physical systems with emergent collective computational abilities," Proceedings of the National Academy of Sciences of the United States of America, vol. 79, no. 8, pp. 2554-2558, 1982.

[44] T. N. Goh, H. L. Ong, and Y. O. Lee, "A new approach to statistical forecasting of daily peak power demand," Electric Power Systems Research, vol. 10, no. 2, pp. 145-148, 1986.

[45] S. Rahman and R. Bhatnagar, "An expert system based algorithm for short term load forecast," IEEE Trans. Power Syst., vol. 3, no. 2, pp. 392-399, 1987.

[46] A. Quintana, J. Gómez, and N. D. Reppen, "Integration of an Adaptive ShortTerm Load Forecast Procedure into a New Energy Control Center," in proc., 7th PSCC, Aug. 1978, pp.426-431.

[47] H. R. Fankhauser, "A novel approach to on-line short and intermediate-term load forecasting," in proc., 8th Power Systems Computation Conference, Aug.1984, pages 376-380.

[48] D. D. Belik, D. J. Nelson, and D. W. Olive, "Use of the Karhunen-Loeve expansion to analyze hourly load requirements for a power utility," IEEE Power Engineering Society Winter Meeting, vol. A78, pp. 225-230, 1978.

[49] E. D. Farmer and M. J. Potton, "Development of online load-prediction techniques with results from trials in the south-west region of the CEGB," Proc. Institution of Electrical Engineers, vol. 115, no. 10, pp. 1549-1558, 1968.

[50] R. Anelli, U. D. Caprio, V. Marchese, and S. Pozzi, "Short Term Predictionof Stationary Load Processes with a Correlation Function FiniteSum of Exponentials," in proc., 7th PSCC, Aug. 1978, pp. 401-408.

[51] M. Meng and W. Shang, "Research on Annual Electric Power Consumption Forecasting Based on Partial Least-Squares Regression," in proc., 2008 International Seminar on Business and Information Management, Dec. 2008, vol. 2, pp. 125-127.

[52] M. S. Abou-Hussien, M. S. Kandlil, M. A. Tantawy, and S. A. Farghal, "An Accurate Model for Short-Term Load Forecasting," IEEE Trans. Power App. Syst., vol. PAS-100, no. 9, pp. 4158-4165, 1981.

[53] K. P. Rajurkar and J. L. Nissen, "Data-dependent systems approach to short-term load forecasting," IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-15, no. 4, pp. 532-536, 1985.

[54] A. Schneider, T. Takenawa, and D. A. Schiffman, "24-hour electric utility load forecasting," in proc., Comparative Models for Electrical Load Forecasting, D. Bunn and E. Farmer, Eds. Wiley, 1985.

[55] M. A. Abu-El-Magd and N. K. Sinha, "Two New Algorithms for On-Line Modelling and Forecasting of the Load Demand of a Multinode Power System," IEEE Trans. Power App. Syst., vol. PAS-100, no. 7, pp. 3246-3252, 1981.

[56] M. Ernoult and R. Mattatia, "Short term load forecasting: New developments at E.D.F," in proc., 8th Power Systems Computation Conference, Aug.1984, pp. 369-375.

[57] K. Paysti, "Box-Jenkins method in short-term forecasting of grid load in Finland," in proc., 8th Power Systems Computation Conference, Aug. 1984, pp. 357-369.

[58] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short Term Load Forecasting," IEEE Trans. Power Syst., vol. 2, no. 3, pp. 785-791, 1987.

[59] A. Keyhani and S. M. Miri, "On-line weather sensitive and industrial group bus load forecasting for microprocessor-based applications," IEEE Trans. Power App. Syst., vol. PAS-102, vol. 12, pp. 3868-3876, 1983.

[60] S. Vemuri, B. Hovelda, and S. Mohebbi, "Short-term load forecasting based on weather load models," in proc., IFAC Conf. on Power Systems and Power Plant Control, Aug. 1986, pp. 565-570.

[61] G. Gross and F. D. Galiana, "Short-Term Load Forecasting," Proc. IEEE, vol. 75, no. 12, pp. 1558-1573, 1987.

[62] L. Ljung and T. Soderstrom, "Theory and Practice of Recursive Identification," IEEE Trans. Autom. Control, vol. AC-30, no. 9, pp. 1054-1056, 1983.

[63] G. D. Irisarri, S. E. Widergren, and P. D. Yehsakul, "On-line Load Forecasting for Energy Contron Center Application," IEEE Trans. Power App. Syst., vol. PAS-101, no. 1, pp. 71-78, 1982.

[64] R. Campo and P. Ruíz, "Adaptive Weather-Sensitive Short Term Load Forecast," IEEE Trans. Power Syst., vol. 2, no. 3, pp. 592-598, 1987.

[65] J. Y. Fan and J. D. McDonald, "A real-time implementation of short-term load forecasting for distribution power systems," IEEE Trans. Power Syst., vol. 9, no. 2, pp. 988-994, 1994.

[66] H. T. Yang, S. C. Chen, and P. C. Peng, "Genetic k-means-algorithm-based classification of direct load-control curves," IEEE Proc. Generation, Transmission and Distribution, vol. 152, no. (4), pp. 489-495, 2005.

[67] S.-J. Huang and K.-R. Shin, "Short-term load forecasting using threshold autoregressive models," IEE Proceedings Generation, Transmission and Distribution, vol. 144, no. 5, pp. 477-481, 1997.

[68] H. Chen, J. Wu, and S. Gao, "A Study of AutoRegressive Conditional Heteroscedasticity Model in Load Forecasting," in proc., 2006 International Conference on Power System Technology, Oct. 2006, pp. 1-8.

[69] Z. Baharudi and N. Kamel, "Autoregressive Method in Short Term Load Forecast," in proc., 2nd IEEE International Conference on Power and Energy (PECon 08), Dec. 2008, pp. 1603-1608.

[70] J.-J. Wang, D.-X. Niu, and L. Li, "An ARMA cooperate with Artificial Neural Network approach in Short-Term," in proc., 2009 Fifth International Conference on Natural Computation Load Forecasting, Aug. 2009, pp. 60-64.

[71] A. R. Fadhilah, S. Suriawati, H. H. Amir, Z. A. Izham, and S. Mahendran, "Malaysian day-type load forecasting," in proc., 3rd International Conference on Energy and Environment, 2009 (ICEE 2009), Dec. 2009, pp. 409-411.

[72] V. A. Kamaev, M. V. Shcherbakov, D. P. Panchenko, N. L. Shcherbakova, and A. Brebels, "Using Connectionist Systems for Electric Energy Consumption Forecasting in Shopping Centers," Automation and Remote Control, vol. 73, no. 6, pp. 1075-1084, 2012.

[73] D. W. Bunn and E. D. Farmer, "Comparative models for electrical load forecasting," New York: Wiley, pp. 13-30, 1985.

[74] J. L. McClelland and D. E. Rumelhart, "Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol 2: Psychological and biological models," Ed. MIT Press, 1986.

[75] D. E. Rumelhart and J. L. McClelland, "Parallel Distributed Processing. Vol 1: Foundations," Ed. MIT Press, 1986.

[76] R. Sharda and R. B. Patil, "Neural networks as forecasting experts: an empirical test," Proc. International Joint Conference on Neural Networks, June 1990, pp. 491-494.

[77] Z. Tang, C. de Almeida, and P. Fishwick, "Time series forecasting using neural networks vs. Box Jenkins methodology," Transactions of The Society for Modeling and Simulation International - SIMULATION, vol. 57, no. 5, pp. 303-310, 1991.

[78] M. A. Badri, "Neural networks of combination of forecasts for data with long memory pattern," in proc., IEEE International Conference on Neural Networks, June 1996, pp. 359-364.

[79] X.-X. Wu and J.-G. Liu, "A New Early Stopping Algorithm for Improving Neural Network Generalization," in proc., Second International Conference on Intelligent Computation Technology and Automation, 2009. ICICTA 09, Oct. 2009, pp. 15-18.

[80] A. B. van Wyk, and A. P. Engelbrecht, "Overfitting by PSO trained feedforward neural networks," in proc., 2010 IEEE Congress on Evolutionary Computation (CEC), July 2010, pp. 1-8.

[81] T. A. Corrales and J. J. Auron, "Nonlinear system identification and overparameterization effects in multisensory evoked potential studies," IEEE Trans. Biomed. Eng., vol. 47, no. 4, pp. 472-486, 2000.

[82] C. García-Ascanio and C. Mate, "Electric power demand forecasting using interval time series: A comparasion between VAR and iMLP," Energy Policy, vol. 38, pp. 715-725, 2009.

[83] L. Hernández, C. Baladrón, J. M. Aguiar, B. Carro, A. J. Sánchez-Esguevillas, and J. Lloret, "Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks," Energies, vol. 6, no. 3, pp. 1385-1408, 2013.

[84] A. Nair and S. K. Joshi, "Short Term Load Forecasting Using Probabilistic Neural Network Based Algorithm," in proc., 2010 International Conference on Computational Intelligence and Communication Networks, Nov. 2010, pp. 128-132.

[85] H. Mori and D. Iwashita, "ANN-based risk assessment for short-term load forecasting," in Proc. 13th International Conference on Intelligent Systems Application to Power Systems 2005, Arlington, VA, 6–10 Nov. 2005, pp. 446-451. DOI:10.1109/ISAP.2005.1599305.

[86] H. Li, S. Guo, H. Zhao, C. Su., and B. Wang, "Annual Electric Load Forecasting by a Least Squares Support Vector Machine with a Fruit Fly Optimization Algorithm," Energies, vol. 5, no. 11, pp. 4430-4445, 2012.

[87] W.-C. Hong, Y. Dong, C.-Y. Lai, L.-Y. Chen, and S.-Y. Wei, "SVR with Hybrid Chaotic Immune Algorithm for Seasonal Load Demand Forecasting," Energies, vol. 4, no. 6, pp. 960-977, 2011.

[88] Z. Aung, M. Toukhy, J. R. Williams, A. Sánchez, and S. Herrero, “Towards Accurate Electricity Load Forecasting in Smart Grids,” in proc., the Foyrth International Conference on Advances in Databases, Knowledge, and Data Applications, Feb. -March 2012, pp. 51-57.
[89] W. Sun, J. Lu, and Y. He, “Information Entropy Based Neural Network Model for Short-Term Load Forecasting,” in proc., 2005 IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific, Aug. 2005, pp. 1-5.
[90] H. Su and Y. Zhang, “Short-term load forecasting using H∞ filter and Elman neural network,” in proc., IEEE International Conference on Control and Automation (ICCA), May 2007, pp. 1868-1872, 2007.
[91] L. Yongchum, “Application of Elman Neural Network in Short-Term Load Forecasting,” in proc., 2010 International Conference on Artificial Intelligence and Computational Intelligence (AICI), Oct. 2010, pp. 141-144.
[92] F. J. Marín, F. Garcia-Lagos, G. Joya, and F. Sandoval, “Global model for short-term load forecasting using artificial neural networks,” IEE Proceedings-Generation, Transmission and Distribution, vol. 149, no. 2, pp. 121-125, 2002.
[93] L. Hernández, C. Baladrón, J. M. Aguiar, B. Carro, and A. Sánchez-Esguevillas, “Classification and Clustering of Electricity Demand Patterns in Industrial Parks,” Energies, vol. 5, no. 12, pp. 5215-5228, 2012.
[94] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, “Load pattern-based classification of electricity customers,” IEEE Trans. Power Syst., vol. 19, no. 2, pp. 1232-1239, 2004.
[95] O. A. S. Carpinteiro and A. J. R. Reis, “A SOM-based hierarchical model to short-term load forecasting,” in proc., 2005 IEEE Russia Power Tech, Jun. 2005, pp. 1-6.
[96] G. J. Tsekouras, N. D. Hatziargyriou, and E. N. Dialynas, “Two-Stage Pattern Recognition of Load Curves for Classification of Electricity Customers,” IEEE Trans. Power Syst., vol. 22, no. 3, pp. 1120-1128, 2007.
[97] T. Zhang, G. Zhang, J. Lu, X. Feng, and W. Yang, “A New Index and Classification Approach for Load Pattern Analysis of Large Electricity Customers,” IEEE Trans. Power Syst., vol. 27, no. 1, pp. 153-160, 2012.
[98] Z. Y. Wang, “Development Case-based Reasoning System for Short-term Load Forecasting,” in proc., 2006 IEEE Russia Power Engineering Society General Meeting, Oct. 2006, pp. 1-6.
[99] P. Duan, K. Xie, T. Guo, and X. Huang, “Short-Term Load Forecasting for Electric Power Systems Using the PSO-SVR and FCM Clustering Techniques,” Energies, vol. 4, no. 1, pp. 173-184, 2011.
[100] J. Che, J. Wang, and G. Wang, “An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting,” Energy, vol. 37, no. 1, 657-664, 2012.
[101] V. Nadimi, A. Azadeh, P. Pazhoheshfar, and M. Saberi, “An Adaptive-Network-Based Fuzzy Inference System for Long-Term Electric Consumption Forecasting (2008-2015): A Case Study of the Group of Seven (G7) Industrialized Nations: U.S.A, Canada, Germany, United Kingdom, Japan, France and Italy,” in proc., 2010 Fourth UKSim European Symposium on Computer Modeling and Simulation (EMS), Nov. 2010, pp. 301-305.
[102] C.-W. Lou and M.-C. Dong, “Modeling data uncertainty on electric load forecasting based on Type-2 fuzzy logic set theory,” Engineering Applications of Artificial Intelligence, vol. 25, no. 8, pp. 1567-1576, 2012.
[103] H. Mori and D. Kanaoka, “Application of Preconditioned RBFN to Temperature Forecasting for Short-term Load Forecasting,” in proc., 2006 IEEE region 10 Conference (TENCON 2006), Nov. 2006, pp. 1-4.
[104] H. Mori and A. Yuihara, “Deterministic Annealing Clustering for ANN-BasedShort-Term Load Forecasting,” IEEE Trans. Power Syst., vol. 16, no. 3, pp. 545-551, 2001.
[105] N. Amjady, “Short-Term Bus Load Forecasting of PowerSystems by a New Hybrid Method,” IEEE Trans. Power Syst., vol. 22, no. 1, pp. 333-341, 2007.
[106] M. Meng, D. Niu, and W. Sun, “Forecasting Monthly Electric Energy Consumption using Feature extraction,” Energies, vol. 4, no. 10, pp. 1495-1507, 2011.
[107] J. Che, J. Wang, and Y. Tang, “Optimal training subset in a support vector regression electric load forecasting model,” Applied Soft Computing, vol. 12, no. 5, 1523-1531, 2012.
[108] W.-C. Hong, “Application of chaotic and swarm optimization in electric load forecasting,” Energy Policy, vol. 38, no. 10, pp. 5830-5839, 2010.
[109] W.-C. Hong, “Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm,” Energy, vol. 36, no. 9, pp. 5568-5578, 2011.
[110] W. Y. Zhang, W.-C. Hong, Y. Dong, G. Tsai, J.-T. Sung, and G.-F. Fan, “Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting,” Energy, vol. 45, no. 1, pp. 850-858, 2012.
[111] W.-C. Hong, Y. Dong, W. Y. Zhang, L.-Y. Chen, and B. K. Panigrahi, “Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm,” Electrical Power and Energy Systems, vol. 44, no. 1, pp. 604-614, 2013.
[112] A. Deihimi and H. Showkati, “Application of echo state networks in short-term electric load forecasting,” Energy, vol. 39, no. 1, pp. 327-340, 2012.
[113] H.-Z. Li, S. Guo, C.-J. Li, and J.-Q. Sun, “A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm,” Knowledge-Based Systems, vol. 37, pp. 378-387, 2013.
[114] S. Sargunaraj, D. P. Sen Gupta, and S. Devi, “Short-term load forecasting for demand side management,” IEE Proceedings-Generation, Transmission and Distribution, vol. 144, no. 1, pp. 68-74, 1997.
[115] Y.-Y. Hsu and C.-C. Yang, “Design of artificial neural networks for short-term load forecasting. Part II: Multilayer feedforward networks for peak load and valley load forecasting,” IEE Proceedings C Generation, Transmission and Distribution, vol. 138, no. 5, pp. 414-418, 1991.
[116] J. W. Taylor and R. Buizza, “Neural Network Load Forecasting With Weather Ensemble Predictions,” IEEE Trans. Power Syst., vol. 17, no. 3, pp. 626-632, 2002.
[117] W.-C. Chu, Y.-P. Chen, Z.-W. Xu, and W.-J. Lee, “Multiregion Short-Term Load Forecasting in Consideration of HI and Load/Weather Diversity,” IEEE Trans. Ind. Appl., vol. 47, no. 1, pp. 232-237, 2011.
[118] K. Nose-Filho, A. D. Plasencia, and C. R. Minossi, “Short-Term Multinodal Load Forecasting Using a Modified General Regression Neural Network,” IEEE Trans. Power Del., vol. 26, no. 4, pp. 2862-2869, 2011.
[119] M. Rejc and M. Partos, “Short-Term Transmission-Loss Forecast for the Slovenian Transmission Power System Based on a Fuzzy-Logic Decision Approach,” IEEE Trans. Power Syst., vol. 26, no. 3, pp. 1511-1521, 2011.
[120] Y. Wang, Q. Xia, and C. Kang, “Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting,” IEEE Trans. Power Syst., vol. 26, no. 2, pp. 500-507, 2011.
[121] A. Chaouachi, R. M. Kamel, R. Andoulsi, and K. Nagasaka, “Multiobjetive Intelligent Energy Management for a Microgrid,” IEEE Transactions on Industrial Electronics, vol. 60, no. 4, pp. 1688-1699, 2013.
[122] A. L. Dimeas and N. D. Hatziargyriou, “Operation of a Multiagent Systemfor Microgrid Control,” IEEE Trans. Power Syst., vol. 20, no. 3, pp. 1447-1445, 2005.
[123] M. E. Khodayar, M. Barati, and M. Shahidehpour, “Integration of High Reliability Distribution System inMicrogrid Operation,” IEEE Trans. Smart Grid, vol. 3, no. 4, pp. 1997-2006, 2012.
[124] Y. Sugaya, S. Omachi, A. Takeuchi, and Y. Nozaki, “A Statistical Analysis on Operation Scheduling for an Energy Network Project,” IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 9, pp. 1583-1592, 2012.
[125] X. Guan, Z. Xu, and Q.-S. Jia, “Energy-Efficient Buildings Facilitated by Microgrid,” IEEE Trans. Smart Grid, vol. 1, no. 3, pp. 243-252, 2010.
[126] C. Chen, S. Duan, T. Cai, B. Liu, and G. Hu, “Smart energy management system for optimal Microgrid economic operation,” IET Renewable Power Generation, vol. 5, no. 3, pp. 258-267, 2011.
[127] H. Kanchev, D. Lu, F. Colas, V. Lazarov, and B. Francois, “Energy Management and Operational Planning of a Microgrid With a PV-Based Active Generator for Smart Grid Applications,” IEEE Trans. Ind. Electron., vol. 58, no. 10, pp. 4583-4592, 2011.
[128] N. Amjady, F. Keynia, and H. Zareipour, “Short-Term Load Forecast of Microgrids by a New Bilevel Prediction Strategy,” IEEE Trans. Smart Grid, vol. 1, no. 3, pp. 286-294, 2010.
[129] P. P. K. Chan, W.-C. Chen, W. W. Y. Ng, and D. S. Yeung, “Multiple classifier system for short term load forecast of Microgrid,”in proc., 2011 International Conference on Machine Learning and Cybernetics (ICMLC), July 2011, vol. 3, pp. 1268-1273.
[130] Y. Hida, Y. Ito, R. Yokoyama, and K. Iba, “A study of optimal capacity of PV and battery energy storage system distributed in demand side,” in proc., 2010 45th International Universities Power Engineering Conference (UPEC), Aug. 2010, pp. 1-5.
[131] J. A. P. Lopes, “Advanced MicroGrids as a component for active management of distribution networks,” in proc., International Conferenceon Power Engineering, Energy and Electrical Drives, 2009 (POWERENG 09), March 2009, pp. 7-8.

[132] M. Sechilariu, W. Baochao, and F. Locment, "Building Integrated Photovoltaic System With Energy Storage and Smart Grid Communication," IEEE Transactions on Industrial Electronics, vol. 60, no. 4, pp. 1607-1618, 2013.

[133] C. Skrzypczak, "The intelligent home of 2010," IEEE Commun. Mag., vol. 25, no. 12, pp. 81-84, 1987.

[134] G. T. Costanzo, G. Zhu, M. F. Anjos, and G. Savard, "A System Architecture for Autonomous Demand SideLoad Management in Smart Buildings," IEEE Trans. Smart Grid, vol. 3, no. 4, pp. 2157-2165, 2012.

[135] A. Harris, "Smart buildings," Engineering & Technology, vol. 7, no. 6, pp. 52-54, 2012.

[136] J. M. Abreu, F. Camara, and P. Ferrao, "Using pattern recognition to identify habitual behavior in residential electricity consumption," Energy and Buildings, vol. 49, pp. 479-487, 2012.

[137] A. K. Rathore, "Hybrid micro-grid (μG) Based residential utility interfaced smart energy system: Applications for green data centers and commercial buildings," in proc., 2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), July 2012, pp. 2063-2068.

[138] C. E. Borges, Y. K. Penya, and I. Fernández, "Optimal combined short-term building load forecasting", in proc., 2011 IEEE PES, Innovative Smart Grid Technologies Asia (ISGT), Nov. 2011, pp. 1-7.

[139] R.-J. Wai, Y.-C. Chen, and Y.-R. Chang, "Short-term load forecasting via fuzzy neural network with varied learning rates," in proc., 2011 IEEE International Conference on Fuzzy Systems (FUZZ), June 2011, pp. 2426-2431.

[140] J. Llanos, D. Saez, R. Palma-Behnke, A. Nunez, and G. Jiménez-Estevez, "Load profile generator and load forecasting for a renewable based *microgrid* using Self Organizing Maps and neural networks," in proc., The 2012 International Joint Conference on Neural Networks (IJCNN), June 2012, pp. 1-8.

[141] E. Shimoda, S. Numata, J. Baba, and T. Nitta, "Operation planning and load prediction for microgrid using thermal demand estimation," in proc., 2012 IEEE Power and Energy Society General Meeting, July 2012, pp. 1-7.

[142] E. A. Rying, G. L. Bilbro, and J.-C. Lu, "Focused local learning with wavelet neural networks," IEEE Trans. Neural Netw., vol. 13, no. 2, pp. 304-319, 2002.

[143] H.-G. Han and J.-F. Qiao, "Adaptive Computation Algorithm for RBF Neural Network," IEEE Trans. Neural Netw. Learning Syst., vol. 23, no. 2, pp. 342-347, 2012.

[144] T. A. C. Bragatto, G. S. I. Ruas, and M. V. Lamar, "Real-time hand postures recognition using low computational complexity Artificial Neural Networks and Support Vector Machines," in proc., 3rd International Symposium on Communications, Control and Signal Processing, 2008 (ISCCSP 2008), March 2008, pp. 1530-1535.

[145] P. Orponen, "Computational Complexity of Neural Networks: A Survey," NeuroCOLT Technical Report Series, 1994.

[146] M. M. Islam, M. A. Sattar, M. F. Amin, Y. Xin, and K. Murase, "A New Constructive Algorithm for Architectural and Functional Adaptation of Artificial Neural Networks," IEEE Trans. Syst. Man Cybern. B, Cybern., vol. 39, no. 6, pp. 1590-1605, 2009.

[147] B. Kaushik, N. Kaur, and A. K. Kohli, "A computational study on the design of overall reliability measures of optimized ANN for computer networks with fixed and varying link reliabilities," in proc., 2010 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Dec. 2010, pp. 1-6.

[148] T. W. S. Chow and H.-Z. Tan, "Semiblind identification of nonminimum-phase ARMA models via order recursion with higher order cumulants," IEEE Trans. Ind. Electron., vol. 45, no. 4, pp. 663-671, 1988.

[149] W. S. McCulloch and W. A. Pits, "A logical calculus of the ideas immanent in nervous activity," The Bulletin of Mathematics and Biophysics, vol. 5, no. 4, pp. 115-133, 1943.

[150] L. Hernández, C. Baladrón, J. M. Aguiar, B. Carro, A. Sánchez-Esguevillas, J. Lloret, D. Chinarro, J. J. Gómez-Sanz, and D. Cook, "A Multi-Agent System Architecture for Smart Grid Management and Forecasting of Energy Demand in Virtual Power Plants," IEEE Commun. Mag., vol. 51, no. 1, pp. 106-113, 2013.

[151] B. Narayanaswamy, V. K. Garg, and T. S. Jayram, "Online optimization for the smart (micro) grid," in proc., 2012 Third International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy), May 2012, pp. 1-10.

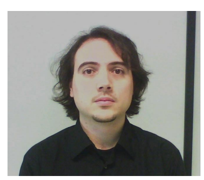
[152] L. Lu, J. Tu, C.-K. Chau, M. Chen, and X. Lin, "Online energy generation scheduling for microgrids with intermittent energy sources and co-generation," in proc., ACM SIGMETRICS/international conference on Measurement and modeling of computer systems, June 2013, pp. 53-66.

[153] E. Alvarez, J. Gomez-Aleixandre, N. de Abajo, and A. M. Campos Lopez, "Algorithm for microgrid on-line central dispatch of electrical power and heat," in proc., 2009 Proc. 44th International Universities Power Engineering Conference (UPEC), Sep. 2009, pp. 1-5.

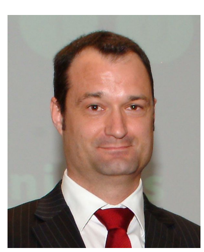
[154] S. W. Hadley and A. Tsvetkova, "Potential Impacts of Plug-in Hybrid Electric Vehicles on regional Power Generation," White Paper. (Accessed 01/08/2013). http://web.ornl.gov/info/ornlreview/v41_1_08/regional_phev_analysis.pdf, 2008.

[155] K. Clement-Nyns, E. Haesen, and J. Driesen, "The Impact of Charging Plug-In Hybrid Electric Vehicles on a Resilient Distribution Grid," IEEE Trans. Power Syst., vol. 25, no. 1, pp. 371-380, 2010.

Luis Hernandez (luis.hernandez@ciemat.es) has a degree in electrical engineering, a degree in computer engineering, a Master European Higher Education Area (EHEA) in communication, network, and content management and Ph.D. telecommunications engineering. He is responsible for communications systems in CEDER-CIEMAT and is project coordinator for Smart Grid/Smart Metering/Microgrid in CEDER-CIEMAT. LUIS HERNANDEZ is Member of the Steering Group FUTURED (Spanish Platform for Electricity Networks).



Carlos Baladrón holds a Ph.D. on Information and Communication Technologies, and a M. Eng. Degree in Telecommunications Engineering from the University of Valladolid, Spain, where he works as a researcher. He has been involved in several National and European projects, including IST FP6 SATSIX, and has been the Technical Coordinator of the IST FP6 OPUCE project. His research activities include applications of Artificial Intelligence and Machine Learning, Context-aware services, Semantics and Next Generation Networks.

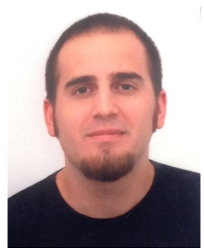


Javier M. Aguiar holds a PhD in Telecommunications and Telecommunications Engineering from the University of Valladolid, Spain, where he is a Professor in the Higher Technical School of Telecommunications Engineering. His research is focused on next-generation networks and services. He has managed and participated in technical activities in several national and European research projects, as well as cooperation with relevant companies of the telecommunication sector. He has contributed in the standardisation field as expert at the European Telecommunications Standards Institute. He serves as editor, guest editor and reviewer, and also author of numerous papers, in several international journals, books and conferences. Furthermore he has been involved and he is participating as reviewer and rapporteur in several international research initiatives including FP7 and H2020 European programmes.

Belén Carro received the Ph.D. degree in the field of broadband access networks from the University of Valladolid, Valladolid, Spain, in 2001. She is a Professor at the Department of Signal Theory and Communications and Telematics Engineering, University of Valladolid. She is the Director of the Communications Systems and Networks (SRC) laboratory, working as a Technical Researcher and Research Manager in several Europe-wide projects in the areas of service engineering and SOA systems, IP broadband communications, NGN/IMS and voice over IP and quality of service. Recently, she has been Technical Director of the OPUCE (Open Platform for User-centric service Creation and Execution) a 10 million project based on Internet protocol (IP) multimedia subsystem (IMS). She has supervised several Ph.D. students on topics related to personal communications, IMS, and session initiation protocol (SIP). She has extensive research publications experience, as an author, a reviewer, and an editor.

Antonio Sanchez-Esguevillas (Senior Member, IEEE) received the Ph.D. degree (with honors) in the field of QoS for real time multimedia services over IP networks from the University of Valladolid, Valladolid, Spain, in 2004. He has been managing innovation at Telefonica (both at Telefonica I+D-Services line and at Telefonica Corporation), Madrid, Spain. He has also been Adjunct Professor and Honorary Collaborator at the University of Valladolid, supervising several Ph.D. students. He has coordinated very large (in excess of 100 million) international R&D projects in the field of personal communication services, particularly related to voice over IP (VoIP) and Internet protocol (IP) multimedia subsystem (IMS). He has more than 50 international publications and several patents. His current research interests are in the area of digital services. Dr. Snchez-Esguevillas is a member of the Editorial Board of IEEE COMMUNICATIONS MAGAZINE among others.

Jaime Lloret received his M.Sc. in Physics in 1997, his M.Sc. in electronic Engineering in 2003 and his Ph.D. in telecommunication engineering (Dr. Ing.) in 2006. He is a Cisco Certified Network Professional Instructor. He worked as a network designer and administrator in several enterprises. He is currently Associate Professor in the Polytechnic University of Valencia. He is the head of the research group ”communications and remote sensing” of the Integrated Management Coastal Research Institute and he is the head of the ”Active and collaborative techniques and use of technologic resources in the education (EITACURTE)” Innovation Group. He is the director of the University Expert Certificate Redes y Comunicaciones de Ordenadores, the University Expert Certificate Tecnologas Web y Comercio Electrnico, and the University Master ”Digital Post Production”. He is currently Chair of the Internet Technical Committee (IEEE Communications Society and Internet society). He has authored 12 books and has more than 240 research papers published in national and international conferences, international journals (more than 80 with ISI Thomson Impact Factor). He has been the co-editor of 15 conference proceedings and guest editor of several international books and journals. He is editor-in-chief of the international journal ”Networks Protocols and Algorithms”, IARIA Journals Board Chair (8 Journals) and he is associate editor of several international journals. He has been involved in more than 200 Program committees of international conferences and in many organization and steering committees. He led many national and international projects. He is currently the chair of the Working Group of the Standard IEEE 1907.1. He has been general chair (or co-chair) of 22 International workshops and conferences. He is IEEE Senior and IARIA Fellow.



Joaquim Massana is a PhD student in Artificial Intelligence at the University of Girona. He is an Electronic Engineer (2006, University of Girona) and Industrial Engineer (2009, University of Girona). In 2010 he received a MSc. in Industrial, Automation, and Computer Systems at the University of Girona and is currently doing his PhD on prediction and pattern discovery. His research interests include artificial intelligence, neural networks, machine learning, energy forecasting, smart cities and pattern discovery.