Expert Systems With Applications 119 (2019) 128–141

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Prediction interval methodology based on fuzzy numbers and its extension to fuzzy systems and neural networks

Luis G. Marín[a], Nicolás Cruz[a], Doris Sáez[a], Mark Sumner[b], Alfredo Núñez[c,*]

[a] Department of Electrical Engineering, University of Chile, Santiago, Chile
[b] Department of Electrical and Electronic Engineering, University of Nottingham, Nottingham NG7 2RD, UK
[c] Section of Railway Engineering, Delft University of Technology, Delft, The Netherlands

ARTICLE INFO

Article history:
Received 6 April 2018
Revised 8 September 2018
Accepted 20 October 2018
Available online 24 October 2018

Keywords:
Prediction interval
Fuzzy number
Neuronal network
Renewable energy
Microgrid

ABSTRACT

Prediction interval modelling has been proposed in the literature to characterize uncertain phenomena and provide useful information from a decision-making point of view. In most of the reported studies, assumptions about the data distribution are made and/or the models are trained at one step ahead, which can decrease the quality of the interval in terms of the information about the uncertainty modelled for a higher prediction horizon. In this paper, a new prediction interval modelling methodology based on fuzzy numbers is proposed to solve the abovementioned drawbacks. Fuzzy and neural network prediction interval models are developed based on this proposed methodology by minimizing a novel criterion that includes the coverage probability and normalized average width. The fuzzy number concept is considered because the affine combination of fuzzy numbers generates, by definition, prediction intervals that can handle uncertainty without requiring assumptions about the data distribution. The developed models are compared with a covariance-based prediction interval method, and high-quality intervals are obtained, as determined by the narrower interval width of the proposed method. Additionally, the proposed prediction intervals are tested by forecasting up to two days ahead of the load of the Huatacondo microgrid in the north of Chile and the consumption of the residential dwellings in the town of Loughborough, UK. The results show that the proposed models are suitable alternatives to electrical consumption forecasting because they obtain the minimum interval widths that characterize the uncertainty of this type of stochastic process. Furthermore, the information provided by the obtained prediction interval could be used to develop robust energy management systems that, for example, consider the worst-case scenario.

1. Introduction

Nonlinear models can provide excellent insight into complex real-world processes and systems. Such models represent the relationships among variables and are useful in planning and operational stages as well as in the analysis of measured data (Rencher & Schaalje, 2008). In general, the aim of predictive models is to obtain a reliable representation of the target system (Ghanbari, Hadavandi, & Abbasian-Naghneh, 2010). In recent decades, several methodologies have been proposed to solve nonlinear model identification problems that use a finite number of measured data and consider an optimality criterion (Škrjanc, 2011). Many studies have examined methods for improving the accuracy of these approaches to obtain higher precision in expected value prediction (Khodayar, Wang, & Manthouri, 2018; Kroll & Schulte, 2014).

Neural networks and fuzzy systems are efficient for nonlinear modelling because they have a high fitting accuracy for nonlinear systems (Veltman, Marín, Sáez, Gutierrez, & Nuñez, 2015; Xu, Zhang, Zhu, & He, 2017). Although computational intelligence methods exhibit adequate performance in estimation and prediction, uncertainty is not typically quantified by these modelling approaches, and only expected value is obtained. However, information on the dispersion of the output of the model provides more information about the phenomena modelled with uncertainty and more useful information from a decision-making point of view than the models with only expected value (Kabir, Khosravi, Hosen, & Nahavandi, 2018; Shrivastava, Lohia, & Panigrahi, 2016).

Confidence intervals and prediction intervals have been proposed to model the uncertainties of a system. Confidence intervals are used to capture uncertainties in the unknown parameters of a model. Confidence intervals are usually associated with parameters rather than with observations. Prediction intervals are used

* Corresponding author.
E-mail addresses: luis.marin@ing.uchile.cl (L.G. Marín), nicolas.cruz@ing.uchile.cl (N. Cruz), dsaez@ing.uchile.cl (D. Sáez), mark.sumner@nottingham.ac.uk (M. Sumner), a.a.nunezvicencio@tudelft.nl (A. Núñez).

https://doi.org/10.1016/j.eswa.2018.10.043
0957-4174/

to capture uncertainties in random variables yet to be observed and provide a probability that the random variable will be within a given interval (Dybowski & Roberts, 2001; Heskes, 1997; Ramezani, Bashiri, & Atkinson, 2011; Rencher & Schaalje, 2008). Prediction intervals consider more sources of uncertainty than do confidence intervals; these additional sources of uncertainty include model error and noise variance. The predicted outputs are intervals that represent (with a given coverage probability) the most likely region defined by the upper and lower bounds of the interval to which the output of the uncertain phenomena will belong.

In this paper, prediction interval models are used to represent both nonlinear behaviour and uncertainty derived from nonconventional energy sources and electrical demand. The uncertainties associated with wind and photovoltaic power are due to the stochastic intermittency of the primary input (wind speed and solar radiation), and the uncertainty of the demand profiles in energy communities (microgrids) is due to minor load variations, which can generate large changes in the total profile (Parhizi, Lotfi, Khodaei, & Bahramirad, 2015). Moreover, for control of microgrids, the uncertainty associated with intermittent power sources and load is typically handled using a robust model predictive control (Robust MPC) in the formulation of the energy management system (EMS) (Sáez, Ávila, Olivares, Cañizares, & Marín, 2015). The EMS is a control strategy that allows coordination of the energy resources to supply the demand, guaranteeing an economic and reliable operation. Robust EMS has become popular because it can be applied using uncertainty sets rather than probabilistic models, reducing the difficulties related to PDF identification (Lara, Olivares, & Cañizares, 2018).

In the work of Valencia, Collado, Sáez, and Marín (2016), a wind-based energy source was modelled by a fuzzy prediction interval based on the method reported in the work of Škrjanc, Blažič, and Agamennoni (2005), and a Robust EMS was achieved using the convex sum of the lower (worst case) and upper (best case) bounds of the available wind energy. In a similar way, in the works of Valencia et al. (2016) and Xiang, Liu, and Liu (2016), prediction interval models of the solar power, wind power, and electrical demand of a microgrid were generated to formulate a scenario-based Robust EMS. In Valencia et al. (2016), the combination of all the lower and upper bounds of the prediction intervals allowed the various scenarios for Robust EMS to be defined, and the solution was obtained using a second-order cone optimization problem. In Xiang et al. (2016), scenarios were generated via Taguchi's orthogonal array testing method using the prediction intervals of the uncertain variables modelled, and the optimization problem was solved using a search strategy based on an orthogonal array. The results of the previous studies showed that a more secure and reliable operation is achieved with Robust EMS than with EMS without uncertainty. However, the performance of a Robust EMS depends on the quality of the prediction interval models over the future time horizon; therefore, improved prediction interval model designs are required (Marín, Valencia, & Sáez, 2016; Parhizi et al., 2015; Sáez et al., 2015).

Several approaches have been proposed that use neural networks and fuzzy systems to generate prediction interval models (see Section 2). In many cases, these approaches carry high computational costs and/or require making assumptions about the data. Additionally, in several of the reported approaches, the prediction interval models are tuned only one step ahead, which could decrease the quality of the interval in terms of the information about the uncertainty modelled for a higher prediction horizon.

This paper presents a new methodology for developing prediction interval models using a novel criterion that includes the coverage probability and the normalized average width of the interval as metrics for training models at future steps. Thus, the prediction interval models aim to achieve the desired coverage probability with the sharpest interval possible. Note that the narrower the width of the prediction intervals, the more accurate the information about the uncertainty phenomena. However, a width that is too narrow might compromise the coverage probability.

The main contribution of this work is a new modelling methodology for constructing prediction intervals based on fuzzy numbers and its extension to fuzzy and neural network prediction interval models. The fuzzy number concept is used because the affine combination of interval fuzzy numbers generates, by definition, prediction interval models that can handle uncertainty without requiring assumptions to be made about the data and the noise distribution. The proposed prediction interval is developed in two stages. First, model identification is performed to tune the parameters necessary for obtaining the expected value. Then, the spreads of the parameters of the prediction interval are found for the future step. The proposed methodology can be used to describe a large family of uncertain nonlinear functions, such as the electrical demand in small communities.

The remainder of this paper is organized as follows: Section 2 presents a literature review of prediction interval modelling. Section 3 introduces the problem statement. Section 4 presents prediction interval models based on interval fuzzy numbers and the extension to fuzzy and neural network models. Section 5 describes the proposed method for developing prediction interval models. Section 6 presents the results of a benchmark test and two case studies involving load forecasting for residential dwellings in the town of Loughborough, UK, and for the isolated microgrid installed at Huatacondo in northern Chile. The last section provides the main conclusions and the focus of future work.

2. Literature review for prediction interval modelling

In the specialized literature, several methods based on neural networks and fuzzy systems have been proposed to obtain prediction intervals. In the work of Khosravi, Nahavandi, Creighton, and Atiya (2011a), traditional methods based on neural networks were analysed, including the delta, Bayesian, mean-variance estimation, and bootstrap methods. These approaches are computationally expensive and/or make assumptions about noise. For instance, the delta method assumes that noise is homogeneous, and the calculation of Jacobian matrices is required. The Bayesian method assumes that the parameters are a random set of variables with a distribution defined a priori, and it requires the computation of the Hessian matrix. The mean-variance estimation method considers that model errors are normally distributed around the true target; therefore, the method requires the known mean and variances. This approach assumes that the corresponding neural network accurately estimates the true targets, which is not always true, leading to a low coverage probability. The bootstrap method assumes that a high-precision estimate of the true targets will be produced by a group of neural networks. This method is the most computationally demanding in the development stage because several neural networks are necessary to estimate variance. However, after the models are trained offline, online computations are simple and do not require the evaluation of complex matrices or derivatives.

In the works of Škrjanc (2011) and Škrjanc et al. (2005), two prediction interval methods based on type-1 fuzzy systems were proposed. In Škrjanc (2011), the upper and lower bounds that define the interval are constructed based on the error covariance of each rule of the fuzzy model. However, in this approach, regionally dependent noise with a normal distribution, with zero mean value and variance, is an a priori assumption. In Škrjanc et al. (2005), an optimization procedure was used to find the lower- and upper-bound parameters of a fuzzy model. This method for prediction interval construction is considered a parametric method. The interval bounds are obtained as the result, and the optimization problem does not impose a desired value for coverage probability or interval width. In the work of Sáez et al. (2015), one-day-ahead fuzzy prediction interval models were developed, supported by the covariance method derived in Škrjanc (2011). This prediction interval model was validated using renewable generation (photovoltaic and wind) and demand data from an installed microgrid in Huatacondo, Chile. Khosravi, Nahavandi, and Creighton (2011) used the delta method to develop an adaptive neuro-fuzzy (ANFIS) prediction interval model. The parameters of the ANFIS prediction interval were found by minimizing a nonlinear cost function that includes the coverage probability and sharpness of the interval. A simulated annealing algorithm was used as a solution method for the optimization problem.

Khosravi, Nahavandi, Creighton, and Atiya (2011b) proposed the lower upper bound estimation (LUBE) method for building neural network prediction intervals as an alternative to solving the problem regarding the assumptions about the data and/or the intensive computational burden. Several studies have used the LUBE method to develop prediction interval models of distributed energy resources and electric consumption. For instance, Khosravi and Nahavandi (2013) and Khosravi, Nahavandi, and Creighton (2013) developed prediction intervals for forecasting wind farm power generation. In a similar way, Quan, Srinivasan, and Khosravi (2014) used both electrical load and wind power generation data to construct prediction interval models, and Wang, and Jia (2015) used photovoltaic power data to construct prediction intervals based on a radial basis function (RBF) neural network. However, in these approaches, neural network training is based on a cost function that includes the combinational coverage width-based criterion (CWC). The problem with this criterion, as established in Shrivastava, Khosravi, and Panigrahi (2015), Wan, Xu, Østergaard, Dong, and Wong (2014), Pinson and Tastu (2014), and Khosravi, and Nahavandi (2014b), occurs when an extremely narrow interval width is obtained: the entire term (CWC) becomes zero irrespective of the prediction interval coverage probability; therefore, the coverage probability can be very low. Additionally, in Khosravi, and Nahavandi (2014a) an interval type-2 fuzzy system was proposed for construction of prediction intervals. The left and right points of the type-reduced set were defined as the lower and upper bounds of the prediction interval. However, the parameters of the system were obtained using the same CWC criteria.

Other approaches regarding the prediction intervals of renewable resources, the price of energy, and the electricity demand have been reported (Hu, Hu, Yue, Zhang, & Hu, 2017; Li et al., 2018; Shrivastava et al., 2015, 2016; Voyant et al., 2018). In the works of Shrivastava et al. (2016) and Shrivastava et al. (2015), methodologies were proposed based on the support vector machine (SVM) to generate the prediction intervals for wind speed and electricity costs. In Shrivastava et al. (2016), a multi-objective differential evolution algorithm was used to tune model parameters such that multiple opposing objectives were achieved to generate Pareto-optimal solutions. In Shrivastava et al. (2015), using particle swarm optimization (PSO), the optimal model parameters were obtained by minimizing the interval width while a desired coverage probability was achieved. In both studies, SVMs were used to generate the upper and lower bounds of the prediction interval. However, the upper and lower values for the training process were unknown; they were artificially generated by modifying the training values within a given percentage. Hu et al.(2017) used the kernel extreme learning machine (KELM) method to develop the prediction interval for wind power using data from two wind farms. The artificial bee colony algorithm was used to find the parameters necessary for the KELM models. The optimization was performed using a cost function that included the coverage probability, the sharpness of the interval and the average deviation of the data from the prediction interval as metrics. In the work of Voyant et al. (2018), prediction interval models of the global horizontal irradiation using regression tree methods were presented. Several prediction models were tested, including classic, pruned, bagged and boosted regression tree and classic and smart persistence models. Several predictors based on subsets of the training data were used to build the prediction interval. Then, a cumulative distribution function (CDF) was constructed based on predicted values from each regression tree model developed. In the work of Li et al. (2018), an improved bootstrap method was proposed for constructing prediction intervals using extreme gradient boosting (XGB) as the base model. The approach was compared with traditional bootstrap, LUBE and SVR-2D using solar power data. The proposed method in this study achieved the best performance in terms of the quality of the prediction interval. Although the prediction interval performed well in previous studies, they considered only short prediction horizons (a few hours ahead), and for some applications, for instance, for EMSs in microgrid operations based on receding horizon control, a higher prediction horizon could be necessary.

Several types of Takagi–Sugeno fuzzy systems that model the uncertainty in the antecedent, in the consequent or in both parts of the fuzzy rules have been discussed in the literature. Type-2 fuzzy sets are used to model the uncertainty in the antecedent, and fuzzy numbers are used to model the uncertainty in the consequents (Khosravi, Nahavandi, Creighton, & Srinivasan, 2012; Mendel, 2017). Most reported studies have demonstrated that these kinds of systems are excellent tools for handling uncertainty and have a degree of accuracy superior to traditional type-1 fuzzy systems. For instance, in the work of Jafarzadeh, Fadali, and Yaman (2013) type-1 and interval type-2 fuzzy systems were proposed for the prediction of solar power. Type-2 systems with type-2 antecedents and crisp consequents provided the more precise output. Khosravi et al. (2012) compared neural networks, traditional type-1 fuzzy systems and interval type-2 fuzzy systems for electrical demand prediction. The results show that the interval type-2 fuzzy model showed improved the prediction accuracy compared to the other approaches due to its additional degrees of freedom. In a similar way, in the work of Khosravi, and Nahavandi (2014c), an interval type-2 fuzzy system was proposed for one-day-ahead load prediction. An optimal type reducer based on a neural network was proposed to improve prediction performance without increasing computational burden compared to traditional type reduction. In Begian, Melek, and Mendel (2008) a novel inference engine for a type-2 fuzzy system was presented. This approach uses a closed form for inference instead of the type-reduction process. The results showed that the proposed inference mechanism outperforms the type-1 fuzzy systems.

Additionally, several studies have demonstrated that the use of complex fuzzy sets and logic in intelligent systems can improve the prediction of future observations in a time series (Yazdanbakhsh & Dick, 2018). In the work of Chen, Aghakhani, Man, and Dick (2011), an adaptive neuro-complex fuzzy inferential system (ANCFIS) was proposed. This system was applied to time-series prediction for synthetic and real-world datasets. The results showed that complex fuzzy sets are a useful tool in intelligent systems design. ANCFIS achieved good performance using a maximum of three rules for all experiments; in contrast, the best ANFIS network used 128 rules for the synthetic dataset. Yazdanbakhsh, and Dick (2017) proposed an extension of the ANCFIS to the multivariable time-series prediction. The proposed approach was compared with the results showed in the work of Li, and Chiang (2013) for the NASDAQ dataset. The results showed that ANCFIS has superior performance regarding the complex neuro-fuzzy autoregressive integrated moving average (CNFS-ARIMA) approach. The work of Yazdanbaksh, Krahn, and Dick (2013) compared three approaches:

ANFIS, radial basis function network (RBFN) and complex fuzzy logic (ANCFIS) for photovoltaic power prediction. ANCFIS was more accurate than the other approaches regarding one-step-ahead prediction.

However, the previous studies have focused only on improving the precision of the expected value rather than obtaining the prediction interval. Because these kinds of the systems can naturally provide a prediction interval, with the type-1 fuzzy set at the consequent, some studies have included the dispersion of the output in the design of the system to obtain the prediction interval in an active way. For instance, Veltman et al. (2015) developed a fuzzy prediction interval model for an electric load application. The fuzzy interval was obtained by including only the uncertainty in the parameters of the consequences. All parameters (premises and consequences) of the fuzzy prediction interval model were found using an improved teaching-learning-based optimization algorithm (ITLBO) to minimize a multi-objective cost function. Marín et al. (2016) characterized uncertainty in wind power and electric load using type-2 fuzzy prediction interval models. As Mendel (2017) mentioned, this work is arguably the first paper to use the type-reduced set in an active way (rather than using it only as a means to obtain the expected value) during parameter identification. In a manner similar to that reported in the work of Veltman et al. (2015), all parameters of the type-2 fuzzy system were tuned using some optimality criteria of the prediction interval, such as the coverage probability, interval width and prediction error. Consequently, many parameters must be tuned in these approaches using an optimization algorithm. The prediction interval models in Veltman et al. (2015) and Marín et al. (2016) were validated for predictions up to two days ahead using data on the energy resources of an isolated microgrid installed in Chile. Although the desired coverage probability in previous studies is fixed during the training process of the prediction interval, this coverage probability could decrease as the number of future steps increases because these models are trained one step ahead. Therefore, prediction intervals that generate the most information in terms of the relationship between the width of the interval and the coverage probability over future steps are required.

Deep learning has achieved state-of-the-art results in several areas, such as computer vision. However, only recently have deep learning-based algorithms become a popular solution for time-series forecasting (Chen, Zeng, Zhou, Du, & Lu, 2018; Qiu, Zhang, Ren, Suganthan, & Amaratunga, 2014; Rodrigues, Markou, & Pereira, 2018). Several deep learning methods exist for this purpose. The most popular is the long short-term memory network (LSTM), which is a form of recurrent neural network (RNN) with the ability to model complex patterns in time series due to its specialized cell architecture. LSTM networks are not affected by the exploding gradient problem that is common in regular RNN when trained to predict values at future steps. These characteristics have made LSTM networks popular for time-series forecasting (Bao, Yue, & Rao, 2017; Bui, Le, & Cha, 2018; Liu, Wang, Yang, & Zhang, 2017). Most reported works in deep learning for time-series forecasting aim to predict the expected value rather than a prediction interval. Some development has also been made for the estimation of uncertainty using deep learning models. In Gal, and Ghahramani (2016), a Monte Carlo dropout approach for representing model uncertainty was presented. The same approach was used in Zhu, and Laptev (2017), where a LSTM encoder-decoder architecture was used to predict the daily completed trips processed by the Uber platform using uncertainty estimation based on the Monte Carlo dropout of hidden units. However, these approaches make several assumptions on the process distribution. Additionally, although LSTM networks are currently used for time-series forecasting, sometimes the improvement of the prediction does not compensate for the higher complexity of the LSTM network.

3. Problem statement

In this paper, prediction interval modelling based on fuzzy numbers provides a systematic framework for representing uncertainty and nonlinear dynamics, which makes it useful for forecasting the uncertainty associated with stochastic variables, such as renewable energy-based generation variables. Next, the general interval modelling problem is detailed.

The result of mapping an input vector $Z(k)$ onto a nonlinear real continuous function g can be written as follows:

$$y(k) = g(Z(k), w) + \varepsilon(k) \qquad k = 1, \ldots, N \tag{1}$$

where $Z(k)=\{z_1(k), z_2(k), \ldots, z_p(k)\}$ represents the input vector of all measurements at time k, $y(k)$ is the output obtained from the set of measured data at time k, w is the true parameter set, and random variable $\varepsilon(k)$ is noise. The aim of the model is to find a real function $g \in \mathcal{G}$ that belongs to the model class $\mathcal{G}$, such that g is the best representation of the system (Shrivastava et al., 2016). The condition for selecting the model is $\| y(k) - \hat{y}(k) \| \leq \varepsilon_0 \;\; k = 1, \ldots, N,$ where $\hat{y}(k) = g(Z(k), \hat{w})$ is the model output at time k, $\hat{w}$ are the estimated parameters and ε_0 is the desired error model. The error may be due to unknown or unobserved variables that affect the model output $\hat{y}(k)$ (Heskes, 1997; Rencher & Schaalje, 2008). When the nonlinear real function g is an uncertain function, it can be assumed that it is a member of the following family of functions (Škrjanc, 2011; Škrjanc et al., 2005):

$$\mathcal{G} = \left\{ g : S \to \mathbb{R}^1 | g(Z(k)) = g_{nom}(Z(k)) + \Delta_g(Z(k)) \right\} \tag{2}$$

where g_{nom} represents the nominal function and Δ_g models the uncertainty and satisfies $\sup_{Z \in S} |\Delta_g(Z)| \leq c, c \in \mathbb{R}$. According to Eq. (2), the function $g \in \mathcal{G}$ can be used to predict a new observation, and its uncertainty based on observed data. This type of function ($g \in \mathcal{G}$) is called a prediction interval model. The goal of prediction interval modelling is to find the lower function $\hat{y}_L$ and the upper function $\hat{y}_U$ that satisfy:

$$\hat{y}_L(k) \leq g(Z(k), w) \leq \hat{y}_U(k) \quad \forall Z(k) \in S \tag{3}$$

In this respect, a function g from the class $\mathcal{G}$ can be found in the band defined by the upper and lower functions. The prediction intervals are developed with a certain coverage probability $(1-\alpha)\%$ that future observations of the uncertain phenomena belong to the interval defined by the lower $\hat{y}_L$ and upper $\hat{y}_U$ bounds (Ak et al., 2013):

$$P\left\{ \hat{y}_L(k) \leq y(k) \leq \hat{y}_U(k) \right\} \geq (1-\alpha)\% \tag{4}$$

As in the works of Veltman et al. (2015), Marín et al. (2016), Shrivastava et al. (2016), Khosravi et al. (2011a,2011b), and Khosravi, Nahavandi, and Creighton (2010), in this methodology, the prediction interval coverage probability *(PICP)* and the prediction interval normalized average width *(PINAW)* are the metrics to be incorporated in the identification process of prediction intervals. *PICP* is used to quantify the number of measured values that fall within the interval defined by the model, and *PINAW* is used to measure the width of the interval.

In this paper, new prediction interval models based on the concept of fuzzy numbers are derived such that the width defined by the upper $\hat{y}_U(k)$ and lower $\hat{y}_L(k)$ values of the interval is as narrow as possible while the interval contains a certain percentage of measured data $y(k)$. This condition implies that, to generate prediction intervals, the average width measured by *PINAW* must be minimized while considering a certain desired coverage probability measured by *PICP*. In the next section, the proposed prediction interval models based on fuzzy and neural network modelling are presented.

4. Prediction interval models based on fuzzy numbers

In this section, a new approach to developing prediction intervals based on fuzzy and neural network models is derived. In general, the models consider a set of p inputs ($z_1(k) \in Z_1, \ldots, z_p(k) \in Z_p$) that represent the input measurement data at time step k.

When an affine linear model is used, the model output $\hat{y}(k)$ at time k is defined as follows:

$$\hat{y}(k) = \theta_o + \theta_1 z_1(k) + \cdots + \theta_p z_p(k), \tag{5}$$

where θ_i ($i = 0, 1, \ldots, p$) are the regression coefficients. In this paper, to include uncertainty, the coefficients θ_i are defined as interval fuzzy numbers (Lee, 2005; Mendel, 2017). Therefore, the parameters are expressed as a fuzzy set that defines a fuzzy interval for representing the value of θ_i.

Thus, the parameters θ_i (interval fuzzy numbers) are characterized by a mean (m) and spread (s). The uncertainty distribution regarding the expected value is characterized using various spread values, i.e., $\theta_i = [m_i - \underline{s}_i, m_i + \bar{s}_i]$. The lower bound ($\hat{y}_L$) and upper bound ($\hat{y}_U$) that define the prediction interval are defined based on the theorem of the affine combination of type-1 interval fuzzy numbers (see Karnik and Mendel (2001) and Mendel (2017) for details on this theorem):

$$\hat{y}_L(k) = \sum_{i=1}^{p} m_i z_i(k) + m_0 - \sum_{i=1}^{p} |z_i(k)| \underline{s}_i \tag{6}$$

$$\hat{y}_U(k) = \sum_{i=1}^{p} m_i z_i(k) + m_0 + \sum_{i=1}^{p} |z_i(k)| \bar{s}_i \tag{7}$$

Based on Eqs. (6) and (7), the expected value is characterized by the mean (m_i). The last term in both equations is associated with the prediction interval, and it is characterized by the parameters ($\bar{s}_i, \underline{s}_i$).

In this interval modelling approach, the parameters associated with spread ($\bar{s}_i, \underline{s}_i$) are obtained to assure the desired coverage probability $(1-\alpha)$% with the smallest interval width at the defined future prediction horizons. The proposed method for identifying these parameters (spreads) is described in Section 5. The models thus provide the values of the upper ($\hat{y}_U$) and lower ($\hat{y}_L$) bounds given a coverage probability and the expected value $\hat{y}(k)$.

The proposed method is used to characterize uncertainty. Uncertainty corresponds to the fitting error between the prediction $\hat{y}(k)$ and the actual output $y(k)$; thus, uncertainty is defined by the interval $[\hat{y}_L, \hat{y}_U]$ to which the predicted value could belong. In the next section, both fuzzy and neural network prediction interval models based on fuzzy numbers are presented.

4.1. Fuzzy prediction interval modelling

Mathematically, a fuzzy system is defined by a set of p inputs ($z_1(k) \in Z_1, \ldots, z_p(k) \in Z_p$), a set of rules, and an output $\hat{y}^j(k)$ related to each rule at time k. The rules of the Takagi–Sugeno models are expressed as follows:

$$R^j : if\ z_1(k)\ is\ F_1^j\ and \cdots and\ z_p(k)\ is\ F_p^j\ then \tag{8}$$
$$\hat{y}^j(k) = \theta_o^j + \theta_1^j z_1(k) + \cdots + \theta_p^j z_p(k)$$

$j = 1, \ldots, M$, where M is the number of rules. Let $F^j(Z(k)) = \prod_{i=1}^{p} \mu_{F_i^j}(z_i(k))$ be the activation degree of each rule. Then, the normalized activation degree $\beta^j(Z(k))$ is defined as follows:

$$\beta^j(Z(k)) = \frac{F^j(Z(k))}{\sum_{j=1}^{M} F^j(Z(k))} \tag{9}$$

In this paper, singleton fuzzification, Gaussian membership functions (F_i^j), and the t-norm product are used to provide the output of the fuzzy system:

$$\hat{y}(k) = \sum_{j=1}^{M} \beta^j(Z(k)) \hat{y}^j(k) \tag{10}$$

Considering the proposed interval modelling framework, in the fuzzy prediction interval models, the consequence parameters (θ_i^j) of each rule (Eq. (8)) can be considered as interval fuzzy numbers with their corresponding means (m_i^j) and spreads ($\bar{s}_i^j, \underline{s}_i^j$). Thus, the local interval output for each rule (j) is calculated as follows:

$$\hat{y}_L^j(k) = \sum_{i=1}^{p} m_i^j z_i(k) + m_0^j - \sum_{i=1}^{p} |z_i(k)| \underline{s}_i^j \tag{11}$$

$$\hat{y}_U^j(k) = \sum_{i=1}^{p} m_i^j z_i(k) + m_0^j + \sum_{i=1}^{p} |z_i(k)| \bar{s}_i^j \tag{12}$$

Finally, the lower $\hat{y}_L(k)$ and upper $\hat{y}_U(k)$ bounds are calculated considering the activation degrees Eq. (9)) and the local outputs of each rule (Eqs. (11) and ((12)) as follows:

$$\hat{y}_L(k) = \sum_{j=1}^{M} \beta^j(Z(k)) \hat{y}_L^j(k) \tag{13}$$

$$\hat{y}_U(k) = \sum_{j=1}^{M} \beta^j(Z(k)) \hat{y}_U^j(k) \tag{14}$$

In this paper, a fuzzy clustering method is considered for defining the rule numbers and the parameters (centre and standard deviation) of the Gaussian membership functions (F_i^j). The means (m_i^j) of the consequences are estimated by the minimum least-squares optimization method (Babuška, 1998). The method for tuning the spreads ($\bar{s}_i^j, \underline{s}_i^j$) is explained in Section 5.

The lower and upper bounds of the fuzzy prediction model for forecasting the output of future steps are defined as follows:

$$\hat{y}_L(k+h) = f^{fuzzy}(Z(k+h), \beta^j(Z(k+h)), m_i^j, \underline{s}_i^j(k+h)) \quad \forall h = 1, \ldots, N_p \tag{15}$$

$$\hat{y}_U(k+h) = f^{fuzzy}(Z(k+h), \beta^j(Z(k+h)), m_i^j, \bar{s}_i^j(k+h)) \quad \forall h = 1, \ldots, N_p \tag{16}$$

where $j = 1, \ldots, M$ is the rule number, $i = 1, \ldots, p$ is the input number, and N_p is the prediction horizon. Note that the parameters $\underline{s}_i^j(k+h)$ and $\bar{s}_i^j(k+h)$ are the spreads tuned $\tilde{h}$ steps ahead, where $\tilde{h} \in \{1, \ldots, N_p\}$, using experimental data with certain coverage probability at the future steps. After the tuning process is completed and the prediction interval is obtained, these parameters are held constant through horizon prediction, i.e., $\underline{s}_i^j(k+h) = \underline{s}_i^j(k+\tilde{h})$ and $\bar{s}_i^j(k+h) = \bar{s}_i^j(k+\tilde{h})$ for $h = 1, \ldots, N_p$. More details about the tuning method are provided in Section 5.

4.2. Neural network prediction interval modelling

Mathematically, a neural network system is defined by a set of p inputs ($z_1(k) \in Z_1, \ldots, z_p(k) \in Z_p$), a set of weights ($w$) and biases ($b$) per layer, and an activation function per layer. If the neural network uses a hyperbolic tangent activation function for the hidden layer and a linear activation function for the output layer, the output of the neural network at time k is defined as follows:

$$\hat{y}_l(k) = \sum_{j=1}^{L} w_{j,l}^0 \left(\tanh\left(\sum_{i=1}^{p} w_{j,i}^h z_i(k) + b_j^h \right) \right) + b_l^0 \tag{17}$$

$j=1,\ldots,L$, where L is the number of hidden layer units and l is the number of output units; in this paper, $l=1$. The hidden weights, hidden bias, output weights and output bias are $w_{j,i}^h$, b_j^h, $w_{j,l}^0$ and b_l^0 respectively. The neural network in Eq. (17) can be written as follows:

$$\hat{y}(k) = \sum_{j=1}^{L} w_j^0 \tilde{Z}_j(k) + b^0 \tag{18}$$

where:

$$\tilde{Z}_j(k) = \tanh\left(\sum_{i=1}^{p} w_{j,i}^h z_i(k) + b_j^h\right) \tag{19}$$

In this paper, Bayesian regularization is used to train the neural network. Bayesian regularization consists of a paradigm designed to minimize overfitting of neural networks. The method provides a Bayesian criterion for terminating training, thus generating better results for the test dataset (Gençay & Qi, 2001).

In this approach, the neural network prediction interval is developed such that the output weights (w_j^0) are considered interval fuzzy numbers with their means (m_j) and spreads ($\underline{s}_j, \bar{s}_j$). The lower and upper bounds of the prediction interval can be calculated as follows:

$$\hat{y}_L(k) = \sum_{j=1}^{L} m_j \tilde{Z}_j(k) + b^0 - \sum_{j=1}^{L} \left|\tilde{Z}_j(k)\right| \underline{s}_j \tag{20}$$

$$\hat{y}_U(k) = \sum_{j=1}^{L} m_j \tilde{Z}_j(k) + b^0 + \sum_{j=1}^{L} \left|\tilde{Z}_j(k)\right| \bar{s}_j \tag{21}$$

The neural network can be defined as a neural network whose outputs are the upper and lower bounds and the target prediction. As fuzzy prediction interval models, neural network prediction interval models are used to forecast the output of future steps as follows:

$$\hat{y}_L(k+h) = f^{NN}(\tilde{Z}_j(k+h), m_j, \underline{s}_j(k+h)) \qquad \forall h = 1, \ldots, N_p \tag{22}$$

$$\hat{y}_U(k+h) = f^{NN}(\tilde{Z}_j(k+h), m_j, \bar{s}_j(k+h)) \qquad \forall h = 1, \ldots, N_p \tag{23}$$

where $j=1,\ldots,L$ is the number of hidden layer units and N_p is the prediction horizon. Note that the parameters $\underline{s}_j(k+h)$ and $\bar{s}_j(k+h)$ are the spreads tuned $\tilde{h}$ steps ahead, where $\tilde{h} \in \{1, \ldots, N_p\}$, using experimental data with a certain coverage probability at the future steps. After the tuning process is completed and the prediction interval is obtained, these parameters are held constant through horizon prediction.

Next, the method for identifying the parameters of the prediction interval based on fuzzy systems and neural networks is explained.

5. Proposed method for developing prediction intervals based on fuzzy systems and neural networks

The identification procedure for deriving the prediction interval models is shown in Fig. 1. The first part of this procedure corresponds to the identification method of the fuzzy and neural network models for obtaining the expected value, and the second part is the method for prediction interval parameter (spreads) identification.

Regarding model identification (Fig. 1), the first step involves data collection for training, validation and testing; sufficient information is collected to represent the various operational points of the process to be modelled. The training dataset is used to obtain the model parameters. The validation dataset is not directly used in the training process; however, it allows the model generalization capacity given by the model behaviour to be evaluated under a new dataset. Finally, the test dataset is used to evaluate the performance of the obtained model.

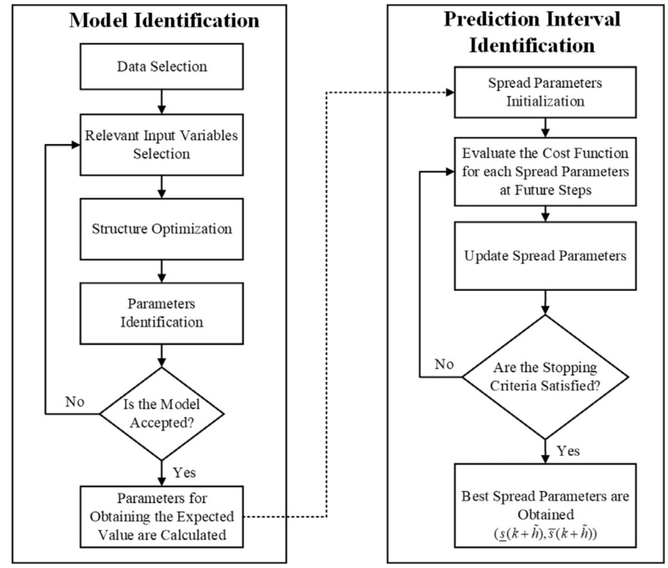

Fig. 1. Methodology for developing prediction intervals.

In this procedure, a structural optimization is made. The structural optimization of fuzzy and neural network models consists of proposing several structures. Specifically, several fuzzy models are obtained when the number of clusters (rules) is modified, and several neural network models are obtained by modifying the hidden neuron number. Then, relevant input variables are selected via sensitivity analysis, and a structural optimization is made. Finally, the parameters necessary for obtaining the expected value are calculated using the relevant input variables, the optimal structure and the training dataset. As proposed in Sáez, and Zuñiga (2004), the best structure is defined when the validation error is either increased or stabilized in comparison with the training error when the structure of the model increases in complexity.

For the fuzzy models, the Gustafson–Kessel clustering algorithm is used to obtain the premise parameters, and the consequence parameters are estimated by the minimum least-squares optimization method. Bayesian regularization is used to obtain the parameters of the neural network models. Finally, the model is evaluated using a test dataset to verify model performance. Then, if the performance of the model is not suitable, the model identification procedure in previous steps must be reviewed; otherwise, this procedure is completed (Sáez & Zuñiga, 2004).

After the model identification procedure, the parameters associated with providing the expected value are obtained. In fuzzy models, the standard deviation and centre of the Gaussian functions (F_i^j) of the fuzzy model are found, where p ($i=0,1,\ldots,p$) are the relevant inputs identified and M ($j=1,\ldots,M$) is the rules number. These parameters are necessary for obtaining the normalized activation degree ($\beta^j(Z(k))$) of the premises. The identified consequence parameters (θ_i^j) (see Eq. (8)) are assigned to the mean values ($m_i^j = \theta_i^j$) required in Eqs. (11) and (12), and the expected value (Eq. (10)) can be obtained.

Regarding neural network models, hidden weights ($w_{j,i}^h$) and hidden biases (b_j^h) are found. With these parameters, the term $\tilde{Z}_j(k)$ in Eq. (19) is calculated, where p ($i=1,\ldots,p$) are the relevant inputs identified and L ($j=1,\ldots,L$) is the number of hidden layer units. Additionally, the output weights (w_j^0) and output bias (b^0) are identified. Finally, the output weights are used to obtain the expected value, where $m_j = w_j^0$ (used in Eqs. (20) and (21)). After the model identification stage, the spreads of the parameters for developing the prediction interval at future steps must be identified (see Fig. 1). This method is described in the following section.

5.1. Parameters identification for prediction intervals

This identification method stage obtains the parameters (spreads) of the prediction interval models such that the upper and lower values of the interval are as narrow as possible and the interval contains a certain percentage of measured data. The prediction interval models derived in this paper can include endogenous $y(k)$ and exogenous variables $u(k)$, where $Z(k)=[y(k-1),\dots, y(k-q_1),u(k-1),u(k-2),\dots,u(k-q_2)]^T$ is the vector of regressors associated with the output and input variables. Then, the prediction interval is a function of the real and/or prediction data, depending on the number of future steps (Sáez et al., 2015). In this paper, the spreads for developing the prediction interval are tuned according to the required steps ahead. Then, based on the formulation described in the previous sections for developing the prediction interval models and the metrics for evaluating the performance of the prediction interval, the spread identification procedure consists of the solution to the following optimization problem (24):

$$\min_{\underline{s}(k+\tilde{h}),\bar{s}(k+\tilde{h})} PINAW$$
$$st.\ PICP = 1-\alpha \tag{24}$$

where $\tilde{h} \in \{1,\dots,N_p\}$ is the number of steps ahead and $(1-\alpha)\%$ is the desired coverage probability. The prediction interval normalized average width *(PINAW)* and the prediction interval coverage probability *(PICP)* for N_p steps ahead are defined as follows:

$$PICP = \frac{1}{N}\sum_{k=1}^{N}\delta_{k+h} \times 100\% \tag{25}$$

$$PINAW = \frac{1}{N\cdot R}\sum_{k=1}^{N}\left(\hat{y}_U(k+h)-\hat{y}_L(k+h)\right) \times 100\%$$
$$\forall h = 1,\dots,N_p \tag{26}$$

where $\delta_{k+h}=1$ if $y(k+h) \in [\hat{y}_L(k+h), \hat{y}_U(k+h)]$; otherwise, $\delta_{k+h}=0$. The parameters $(\underline{s}(k+\tilde{h}), \bar{s}(k+\tilde{h}))$ are the decision variables in the optimization problem, and the dimensionalities of these parameters depend on the model selected. For fuzzy models, $2pM$ parameters that correspond to the spreads $(\underline{s}_i^j(k+\tilde{h}), \bar{s}_i^j(k+\tilde{h}))$ and $2L$ parameters for the neural network model that corresponds to the spreads $(\underline{s}_j(k+\tilde{h}), \bar{s}_j(k+\tilde{h}))$ should be identified. Then, the parameters $(\underline{s}(k+\tilde{h}), \bar{s}(k+\tilde{h}))$ (Eq. (24)) must be computed such that *i)* *PICP* is greater than or equal to the desired coverage probability $(1-\alpha)\%$ and *ii)* *PINAW* is as small as possible at future steps. The equality constraint $PICP=(1-\alpha)\%$ in Eq. (24) is a hard constraint and is therefore included in the optimization problem as a barrier function to relax this constraint. Therefore, the solution of the minimization problem (24) is computed following the procedure for the unconstrained minimization problem:

$$\min_{\underline{s}(k+\tilde{h}),\bar{s}(k+\tilde{h})} J = \eta_1 PINAW + \exp^{-\eta_2(PICP-(1-\alpha))} \tag{27}$$

In (27), η_1 is a weighting factor and η_2 is a penalty factor. These parameters are chosen such that, if *PICP* is less than $(1-\alpha)\%$, the term $\exp^{-\eta_2(PICP-(1-\alpha))}$ is the dominant term in the cost function; otherwise, *PINAW* is dominant. Finally, the solution to the nonlinear optimization problem (27) is computed using particle swarm optimization (PSO), as outlined in the next section.

5.2. Solution method

To solve the nonlinear optimization problem in Eq. (27), traditional algorithms, such as gradient descent methods, are not adequate. These methods entail a risk of falling into a local optimum when solving non-convex optimization problems. Therefore, other optimization methods are needed (Quan et al., 2014). In this paper, PSO is used to solve the problem because it generally outperforms other algorithms in terms of success rate and solution quality, as reported in the work of Elbeltagi, Hegazy, and Grierson (2005). In PSO, the generated solutions are called particles, and each particle has a position vector with an associated velocity vector (Tran, Wu & Nguyen, 2013). The first step in the algorithm consists of the initialization of particle positions $x_{i,j}$ and velocities $v_{i,j}$ for the *j-th* dimension of the *i-th* particle. In this paper, the particle positions are all the spread parameters $(\underline{s}(k+\tilde{h}), \bar{s}(k+\tilde{h}))$ required to develop the prediction interval model, as explained in Section 4.

The velocity $v_{i,j}$ and position $x_{i,j}$ in the *j-th* dimension of every *i-th* particle are updated according to the following relations:

$$v_{i,j}(t+1) = Wv_{i,j}(t) + c_1 rand()(Pbest_{i,j}(t)-x_{i,j}(t)) + c_2 rand()(gbest_j(t)-x_{i,j}(t))$$
$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \tag{28}$$

$i=1, 2,\dots,NP$ where NP is the number of particles, and $j=1,2,\dots,N_0$ is the total number of parameters to be identified, which depends on the type of model used (fuzzy or neural). W is an inertia factor, *Pbest* is the best previous solution of the particle, and *gbest* is the best solution of the swarm up to the current step. The terms c_1 and c_2 are called the cognitive and social acceleration constants, and *rand*() is a random number between 0 and 1. The training termination criterion is set when a minimum error or a defined maximum number of iterations is achieved. Once the training process terminates, the *gbest* value is chosen as the spread parameter to generate the prediction interval model.

PINAW and *PICP* are used as metrics for the evaluation of the quality of the interval. Additionally, the root mean square error *(RMSE)* and the mean absolute error *(MAE)* are included as performance indices to evaluate the accuracy of the prediction model associated with the expected value. All indices are evaluated for several prediction horizons with the test dataset. In this paper, the prediction interval models based on fuzzy systems and neural networks are used to represent the nonlinear behaviour and uncertainty derived from electricity demand; however, the proposed methodology can be used to describe a large family of uncertainty nonlinear functions.

6. Results

A comparative analysis between the proposed prediction interval models based on interval fuzzy numbers (PI-IFN) and covariance prediction interval models is presented following the definition presented in Rencher and Schaalje (2008) and Škrjanc (2011). The prediction interval based on the covariance establishes the interval based on the error between the observed data $y(k)$ and the model output $\hat{y}(k)$. This method is based on the assumption that the noise is normally distributed with a zero mean value and variance σ^2 that is expressed as $e=N(0, \sigma^2)$ (Škrjanc, 2011).

As indicated in Eq. (18), the neural network model is a linear model of the parameters. Therefore, the prediction interval based on the covariance method can be developed using Eqs. (29) and (30), following the definition presented in Rencher and Schaalje (2008):

$$\hat{y}_U = \tilde{Z}^{*T}W^0 + b^0 + t_\alpha\sigma_e\left(1+\tilde{Z}^{*T}\left(\tilde{Z}^T\tilde{Z}\right)^{-1}\tilde{Z}^*\right)^{1/2} \tag{29}$$

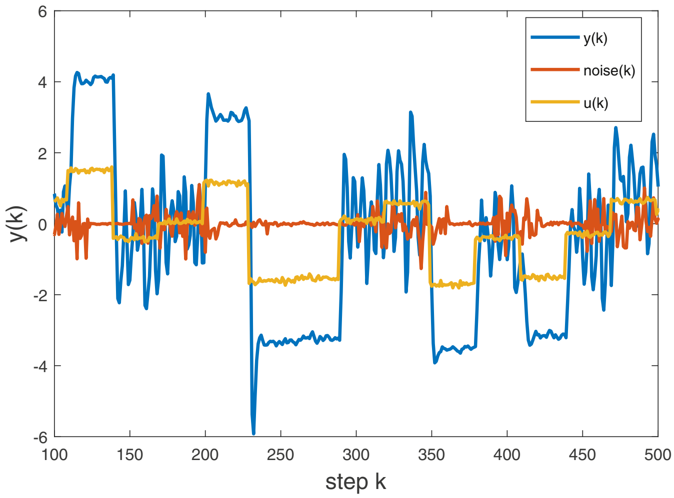

Fig. 2. Modified Chen series for 400 training data.

$$\hat{y}_L = \tilde{Z}^{*T} W^0 + b^0 - t_\alpha \sigma_e \left(1 + \tilde{Z}^{*T}\left(\tilde{Z}^T \tilde{Z}\right)^{-1} \tilde{Z}^*\right)^{1/2} \quad (30)$$

where σ_e is the variance of the error, t_α is the parameter related to the interval width, $\tilde{Z}^*$ is the new datum used to predict the future observation and $\tilde{Z}$ is the matrix that considers all data used in the training process in which the output weights and output bias were determined.

Finally, the fuzzy prediction interval model based on covariance proposed in Škrjanc (2011) is used to obtain the upper and lower bounds of the local linear model as follows:

$$\hat{y}_U^j = \psi_j^{*T} \theta_j + t_\alpha \sigma_j \left(1 + \psi_j^{*T}\left(\psi_j \psi_j^T\right)^{-1} \psi_j^*\right)^{1/2} \quad j = 1, \ldots, M \quad (31)$$

$$\hat{y}_L^j = \psi_j^{*T} \theta_j - t_\alpha \sigma_j \left(1 + \psi_j^{*T}(\psi_j \psi_j^T)^{-1} \psi_j^*\right)^{1/2} \quad j = 1, \ldots, M \quad (32)$$

$j = 1,\ldots,M$, where M is the total number of rules composing the fuzzy system, σ_j is the local variance of the error, and $\psi_j^T = \beta^j(Z)[1\ Z^T]$ is the matrix that considers all the values used in the training process.

For all the models (fuzzy and neural), t_α is tuned using experimental data to achieve the desired coverage probability $(1-\alpha)$%, as explained in Sáez et al. (2015). In the next section, the results of a benchmark and the load forecast with the proposed prediction interval models are presented.

6.1. Benchmark

In this paper, the original Chen series in Chen, Billings, and Grant (1990) is modified and used to evaluate the prediction interval models:

$$\begin{aligned} y(k) = {} & (0.8 - 0.5\exp(-y^2(k-1)))y(k-1) \\ & - (0.3 + 0.9\exp(-y^2(k-1)))y(k-2) + u(k-1) \\ & + 0.2u(k-2) + 0.1u(k-1)u(k-2) + e(k) \end{aligned} \quad (33)$$

where the system noise $e(k) = 0.5\exp(-y^2(k-1))\gamma(k)$ depends on the previous state of the output model and $\gamma(k)$ is white noise. The system input $u(k)$ is band-limited Gaussian white noise. The system is simulated, and 10,000 data points are generated. The data are divided into training, validation and testing sets accounting for 55%, 25% and 20% of the total dataset, respectively. Fig. 2 shows the input, output and noise of the modified Chen series simulation for 400 training data. As shown in Fig. 2, the noise level is high when the output $y(k)$ is close to zero.

The regressors $u(k-1)$, $u(k-2)$, $y(k-1)$ and $y(k-2)$ are defined as the inputs for deriving the prediction interval models. Regarding the structure of the fuzzy model, five rules are defined, whereas eight hidden layer units are defined for the neural network model. With these structures defined, the parameters associated with providing the expected value are obtained as explained in Section 5. The PSO algorithm is used to identify the spread parameters for generating the prediction interval at future steps. The desired coverage probability $(1-\alpha) = 90\%$, the weighting factor $\eta_1 = 250$ and the penalty factor $\eta_2 = 150$ in Eq. (27) are defined. A particle size of 50 and the parameters $c_1 = 2.5$ and $c_2 = 1.5$ are used. Finally, W runs from 0.9 to 0.3 during offline optimization. The number of iterations for PSO is set to 5000, the optimizations are executed several times, and the best solution is selected. The cost function value (J) in Eq. (27) and the developed metrics are reported in Table 1 for the test dataset using various numbers of steps ahead.

As shown in Table 1, the fuzzy and neural network models provide cost function values (J) lower than those of the linear model, which is consistent with the nonlinear benchmark structure. These results are expected because of the ability of the fuzzy and neural network models to better fit the dynamics and nonlinearities of the systems, which are more notable for longer future-step predictions. Additionally, it can be observed that the cost function of the proposed method (Eq. (27)) is lower than that of the covariance method for all models (i.e., linear, fuzzy and neural network). The *RMSE* and *MAE* values are equal in the proposed and covariance methods because the identification method is the same. However, the prediction error increases for a larger horizon prediction because the accumulative error of the model is larger when the steps of the horizon increase, as shown in Table 1.

Furthermore, it can be observed that the *PICP* term is close to 90% because the interval models are trained to maintain *PICP* near the desired value for various steps ahead. In terms of prediction intervals, the proposed method (PI-IFN) provides narrower intervals for all step-ahead forecasts. While the covariance method maintains a constant width for the interval (see Figs. 3(a), 4(a) and 5(a)), the proposed method achieves a narrower interval in states with little noise and an interval with a width similar to that of the covariance method in states with high noise.

Importantly, the information level delivered by a prediction interval is directly related to its width (Marín et al., 2016; Xu et al., 2017); thus, the proposed method yields a better information level regarding the covariance method (smaller widths). Wider intervals could produce a higher *PICP*, but these intervals provide less useful information about the uncertainty of the modelled phenomena. In this respect, the neural network models exhibit sharper prediction intervals than the linear and fuzzy models.

Figs. 3–5 show sixteen-step-ahead forecasts of the linear, fuzzy and neural network prediction interval models. The figures show that nearly all the data are included in the interval; only the outliers of the time series are left outside the region constructed by the prediction interval. Additionally, the intervals produced by the proposed method (PI-IFN) are narrower than those obtained by the method used for comparison, as shown in the figures.

6.2. Application for load forecasting

In this section, two case studies involving the implementation of the prediction interval models are presented to address the uncertainty associated with a load. The first case is from the isolated microgrid in the village of Huatacondo in Chile, and the second is from 20 residential dwellings in the town of Loughborough, UK. Both fuzzy and neural network prediction interval models based

Table 1
Performance indices.

Prediction horizon	Performance indices	Linear models		Fuzzy models		Neural models	
		Covariance	PI-IFN	Covariance	PI-IFN	Covariance	PI-IFN
One step ahead	J	52.23	32.62	26.26	23.60	20.93	18.20
	RMSE	0.4312	0.4312	0.3517	0.3517	0.2372	0.2372
	MAE	0.2883	0.2883	0.2253	0.2253	0.1241	0.1241
	PINAW (%)	20.89	12.40	10.09	9.37	7.90	6.85
	PICP (%)	98.15	89.68	89.98	91.18	89.89	89.95
Four steps ahead	J	57.71	47.47	40.27	35.56	36.58	24.36
	RMSE	0.6124	0.6124	0.5266	0.5266	0.4076	0.4076
	MAE	0.5010	0.5010	0.3765	0.3765	0.2446	0.2446
	PINAW (%)	23.05	17.35	15.14	13.73	14.18	9.18
	PICP (%)	91.62	89.06	89.41	89.86	89.92	89.77
Eight steps ahead	J	57.98	49.32	43.39	39.48	39.94	32.62
	RMSE	0.6761	0.6761	0.5621	0.5621	0.5105	0.5105
	MAE	0.4647	0.4647	0.4015	0.4015	0.3127	0.3127
	PINAW (%)	23.19	18.16	17.08	15.29	15.13	11.72
	PICP (%)	93.26	89.09	90.25	89.85	89.50	89.20
Sixteen steps ahead	J	58.16	50.75	45.21	41.40	45.93	36.15
	RMSE	0.6814	0.6814	0.5839	0.5839	0.5937	0.5937
	MAE	0.4717	0.4717	0.4167	0.4167	0.3685	0.3685
	PINAW (%)	23.26	18.97	17.85	15.98	18.20	14.10
	PICP (%)	93.19	89.20	90.36	89.75	90.57	90.07

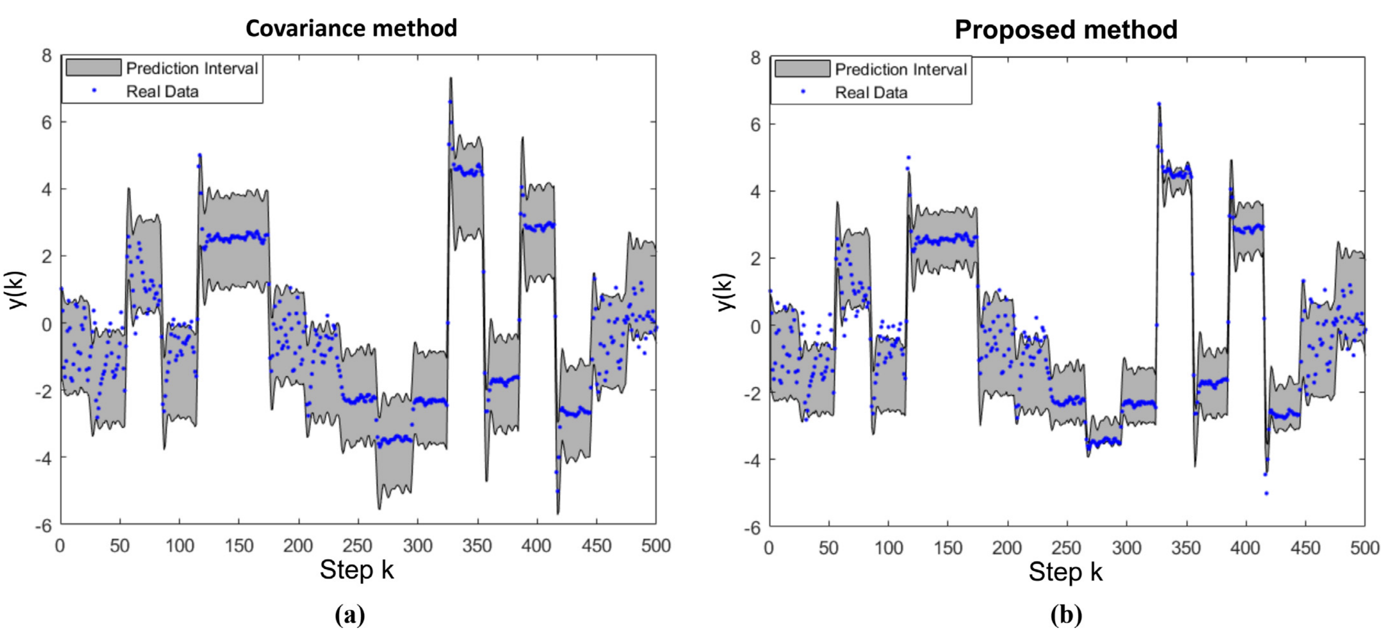

Fig. 3. Sixteen-step-ahead linear prediction interval model: (a) covariance and (b) proposed methods.

on the concept of interval fuzzy numbers (PI-IFN) were used to develop the load forecasting models supported by the results obtained with the benchmark.

6.2.1. Huatacondo microgrid

In this section, the proposed prediction interval model based on the concept of interval fuzzy numbers (PI-IFN) is identified using data from an isolated microgrid in the village of Huatacondo in the Atacama Desert, Chile. The dataset used corresponds to a period of 147 days, spanning November 1st of 2012 to March 27th of 2013, with a sampling time of 15 min, and the peak power load is 27.54 kW. The training dataset corresponds to the period spanning November 1st of 2012 to January 19th of 2013, the validation dataset corresponds to the period spanning January 20th to February 26th of 2013, and the test dataset corresponds to the period spanning February 27th to March 27th of 2013.

Based on the obtained data and using the identification procedure described in Section 5, an optimal structure consisting of three rules and nine regressors is obtained for the fuzzy model:

$$\hat{p}_L(k) = f^{fuzzy}(p_L(k-1), p_L(k-2), p_L(k-3), p_L(k-4), p_L(k-92), \ldots \\ p_L(k-93), p_L(k-95), p_L(k-96), p_L(k-100)) \tag{34}$$

Similarly, eight neurons in the hidden layer and ten regressors are obtained for the neural network model as an optimal structure:

$$\hat{p}_L(k) = f^{NN}(p_L(k-1), p_L(k-2), p_L(k-3), p_L(k-4), p_L(k-92), \ldots \\ p_L(k-93), p_L(k-95), p_L(k-96), p_L(k-97), p_L(k-100)) \tag{35}$$

Note that exogenous variables are not included in the models used to represent the behaviour of the load. During the model identification stage, a prediction horizon of $N_p = 1$ is considered, as explained in Section 5. However, the spread parameters for generating the prediction interval models are identified using PSO by considering various steps ahead with a desired coverage probability of $(1-\alpha)\% = 90\%$. The parameters used for the PSO algorithm

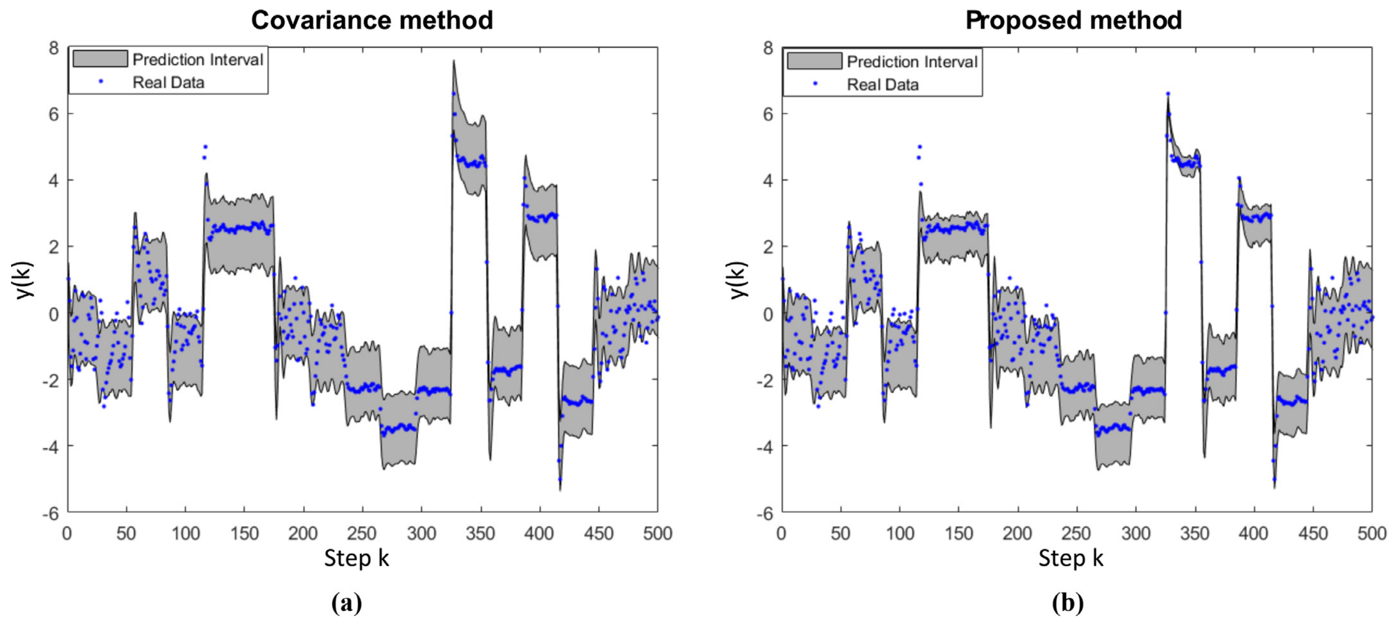

Fig. 4. Sixteen-step-ahead fuzzy prediction interval model: (a) covariance and (b) proposed methods.

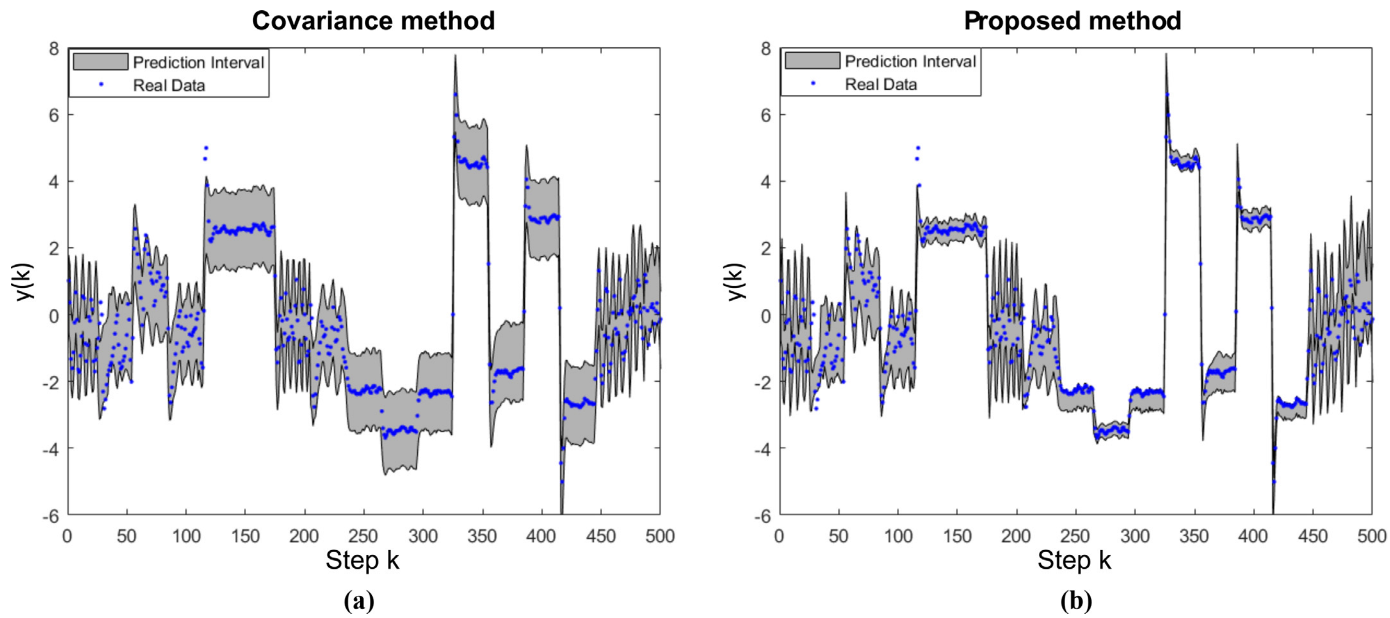

Fig. 5. Sixteen-step-ahead neural network prediction interval model: (a) covariance and (b) proposed methods.

are identical to those in the benchmark problem. To minimizing the random initialization of the parameters (spreads), the optimization is performed several times, and the model with the lowest cost function value is selected.

The performance of the prediction interval models is evaluated by considering 192-step-ahead (two-day-ahead) predictions. This prediction horizon of the load is selected because, for instance, Palma-Behnke et al. (2013) implemented an EMS that minimizes the operational costs of an isolated microgrid by considering a two-day-ahead prediction of renewable resources (wind and solar) and electricity demand.

Table 2 presents *RMSE* and *MAE* for one-step-ahead (15 min), one-hour-ahead, one-day-ahead and two-day-ahead predictions of electricity demand using test data. In the neural network and fuzzy models, the prediction errors increase when the prediction horizon increases. However, the neural network model produces *RMSE* and *MAE* values lower than those of the fuzzy model.

The prediction intervals are evaluated and compared using *PINAW* and *PICP*. Both models can maintain a coverage probability close to 90% because the spread parameters for the prediction intervals are trained to remain within the bounds of the desired coverage probability for various steps ahead. The neural network model generates a narrower interval for all prediction steps, but it also yields a slightly lower *PICP* than that of the fuzzy model. Note that the J values for the neural network model (see Table 2) are higher than those of the fuzzy model because a small variation in *PICP* modifies the total cost function (J) (see Eq. (27)).

Fig. 6 shows the one-day-ahead prediction interval of the fuzzy and neural network models tuned with 90% coverage probability using a receding horizon strategy. In these figures, five days of the test dataset are presented. In small communities (microgrids), the

Table 2
Performance indices.

Prediction horizon	Performance indices	Fuzzy model	Neural model
One step ahead	J	50.42	60.02
	RMSE (kW)	1.5110	1.4929
	MAE (kW)	1.0458	1.0318
	PINAW (%)	17.40	15.73
	PICP (%)	88.71	87.98
One hour ahead	J	52.58	60.57
	RMSE (kW)	1.7468	1.6966
	MAE (kW)	1.1873	1.1638
	PINAW (%)	19.94	17.91
	PICP (%)	89.33	88.16
One day ahead	J	57.49	73.48
	RMSE (kW)	1.9140	1.8936
	MAE (kW)	1.2874	1.2740
	PINAW (%)	20.31	19.48
	PICP (%)	88.73	87.86
Two days ahead	J	71.18	88.05
	RMSE (kW)	2.1780	2.1329
	MAE (kW)	1.4305	1.4232
	PINAW (%)	21.13	20.36
	PICP (%)	88.06	87.59

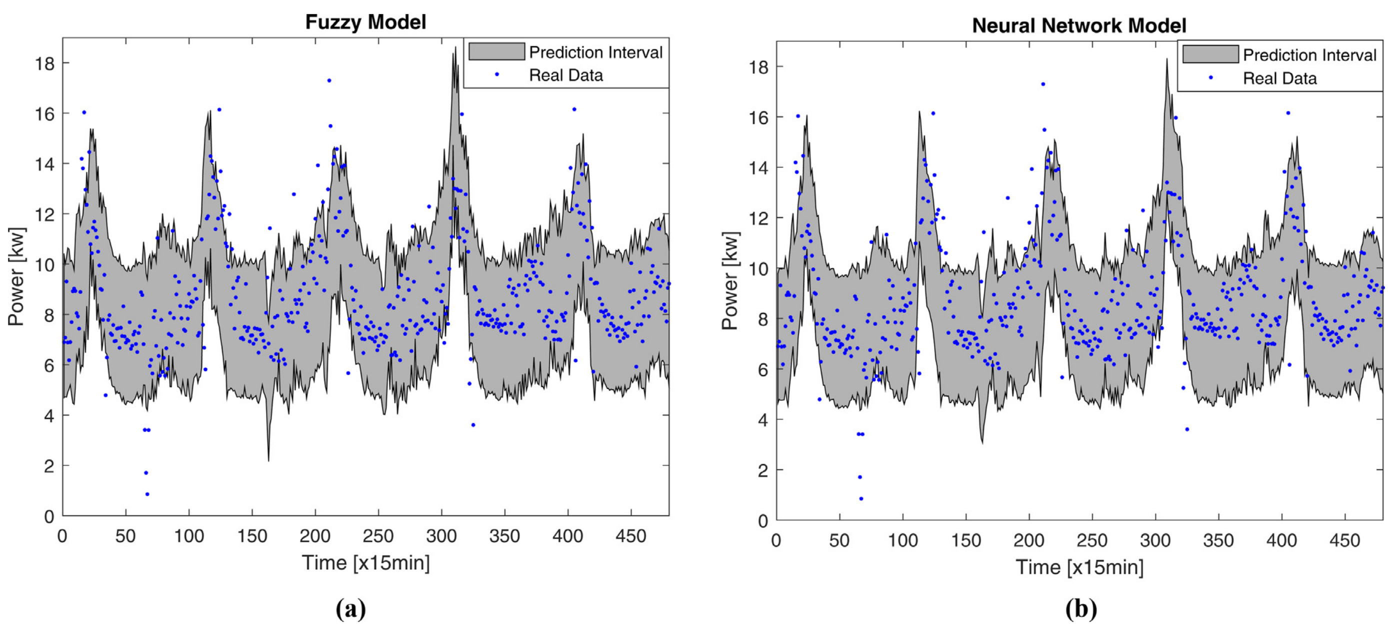

Fig. 6. One-day-ahead prediction interval: (a) fuzzy model and (b) neural network model.

variability of the load is due to the significant effects of each slight change in the load on the total electricity demand. These findings confirm that prediction interval modelling of the load is an important task for microgrid operation when the prediction of future quantities is considered for designing controllers.

6.2.2. Residential dwellings in Loughborough

The load data from 20 dwellings in the town of Loughborough, UK (Richardson & Thomson, 2010), are used to develop fuzzy and neural network prediction interval models using the proposed method (PI-IFN). The available load data correspond to the year 2008; however, only summer data are used to develop the prediction interval. Therefore, a period of 94 days is used, which is divided into 52 days for training, 23 days for validation and 19 days for test data, corresponding to 55%, 25% and 20% of the total dataset, respectively. The maximum electric load is 29.54 kW with a sample time of 15 min.

Eqs. (36) and (37) show the relevant regressors obtained during the model identification process for the fuzzy and neural network models, respectively:

$$\begin{aligned}\hat{p}_L(k) = f^{fuzzy}(&p_L(k-1), p_L(k-2), p_L(k-3), p_L(k-5), p_L(k-6), p_L(k-8), \dots\\ &p_L(k-90), p_L(k-91), p_L(k-92), p_L(k-93), p_L(k-94), \dots\\ &p_L(k-95), p_L(k-96), p_L(k-97), p_L(k-98))\end{aligned} \quad (36)$$

$$\begin{aligned}\hat{p}_L(k) = f^{NN}(&p_L(k-1), p_L(k-2), p_L(k-3), p_L(k-4), p_L(k-5), p_L(k-91), \dots\\ &p_L(k-92), p_L(k-95), p_L(k-96), p_L(k-97), p_L(k-100))\end{aligned} \quad (37)$$

Three rules and nine neurons in the hidden layer correspond to the optimal structure for the fuzzy and neural network models, respectively. Note that exogenous variables are not included in the models. Specifically, in this study, one-step-ahead (15 min), one-hour-ahead, one-day-ahead, and two-day-ahead prediction horizons are considered. For both models (fuzzy and neural network), the performance indices are computed based on the method described in Section 5 for the prediction horizons considered.

As shown in Table 3, the fuzzy and neural models provide similar performances in terms of *RMSE* and *MAE* for the test dataset. The maximum *RMSE* is 2.6054 kW for a peak load of 29.54 kW, corresponding to the fuzzy model at two days ahead. The coverage

Table 3
Performance indices.

Prediction horizon	Performance indices	Fuzzy model	Neural model
One step ahead	J	93.07	92.33
	RMSE (kW)	2.0678	2.0613
	MAE (kW)	1.5225	1.5147
	PINAW (%)	36.68	36.44
	PICP (%)	89.79	89.86
One hour ahead	J	110.64	110.17
	RMSE (kW)	2.2942	2.2910
	MAE (kW)	1.7269	1.7139
	PINAW (%)	44.25	44.05
	PICP (%)	92.73	92.12
One day ahead	J	118.85	119.83
	RMSE (kW)	2.3429	2.2984
	MAE (kW)	1.8076	1.7403
	PINAW (%)	47.54	47.93
	PICP (%)	93.86	94.06
Two days ahead	J	128.23	128.05
	RMSE (kW)	2.6054	2.5532
	MAE (kW)	2.0665	1.9422
	PINAW (%)	51.29	51.22
	PICP (%)	94.63	94.22

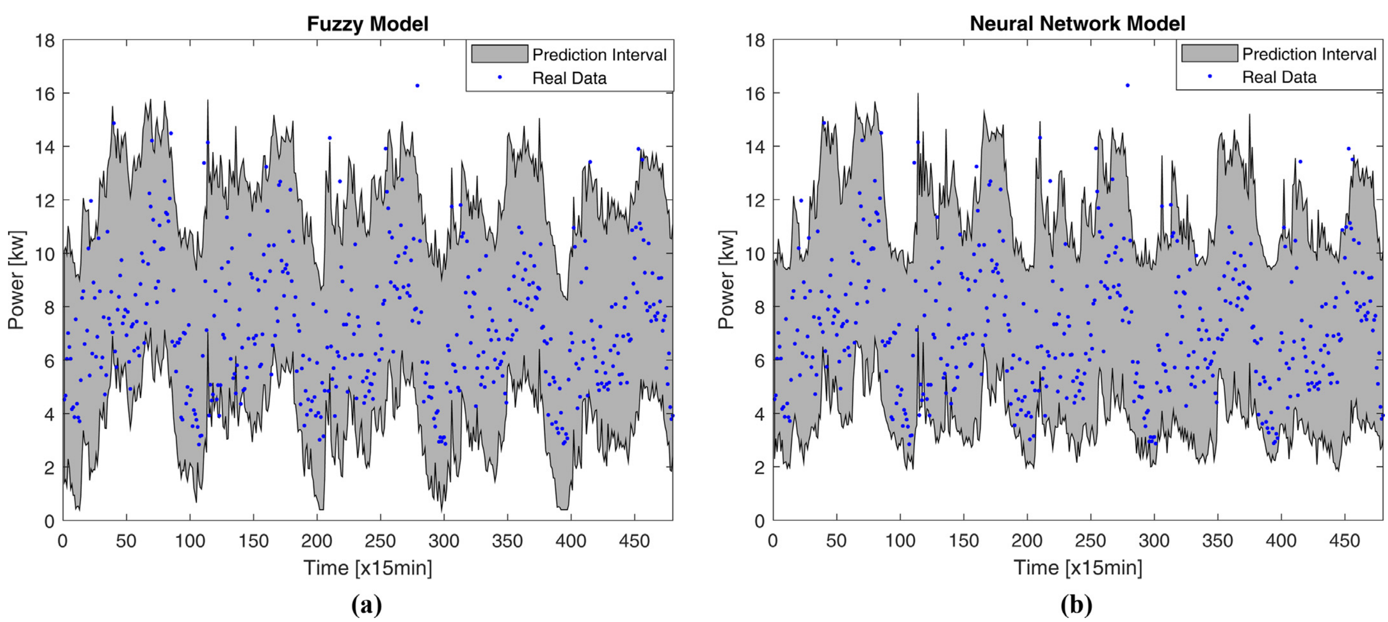

Fig. 7. One-day-ahead Prediction Interval: (a) fuzzy model and (b) neural network model.

probability *(PICP)* for all prediction horizons with respect to the training data is in accordance with the desired coverage probability.

These results suggest that the prediction interval is tuned appropriately to 90% of the desired *PICP*. However, Table 3 shows that the *PICP* values for the test data are higher than the desired coverage probability as the prediction horizon increases. For instance, the *PICP* values are 93.86% and 94.06% at one-day-ahead for the fuzzy and neural network models, respectively. These results are obtained because of the high variability of the data used in this study case, as shown in Fig. 7.

As explained in Section 5, to calculate the spread parameters of all prediction interval models, the coverage probability is fixed in the optimization problem (see Eq. (24)). The interval width finding therefore corresponds to the minimum width at the step ahead defined for characterizing the uncertainty of the modelled demand, given the desired *PICP*. The interval width *(PINAW)* increases with the prediction horizon.

Fig. 7 shows the one-day-ahead prediction intervals of the fuzzy and neural network models, with interval widths of 47.54% and 47.93%. Despite the widths of these intervals, the relationship between the coverage probability and the interval width for this study case provides sufficient information about the uncertainty modelled, and this information could be useful, for instance, for the design of a robust energy management system.

7. Conclusions

In this paper, a new prediction interval modelling framework based on the concept of interval fuzzy numbers was proposed to represent nonlinear dynamics and uncertainties. These models provide the upper and lower bounds of the predicted values given a coverage probability with the minimum interval width at future prediction horizons. This prediction interval modelling was extended to fuzzy systems and neural networks to describe a large family of uncertain nonlinear functions. In this paper, the fuzzy number concept was used because the affine combination of interval fuzzy numbers generates, by definition, interval models that can address the uncertainty of the modelled phenomena without requiring assumptions to be made about the data or the noise distribution. In this methodology, the spreads of the prediction interval models were tuned at future steps based on a novel criterion that minimizes the width of the interval given a desired coverage probability.

Based on a benchmark problem, the proposed method was compared with a covariance prediction interval method. The results show that the proposed prediction interval models generated a narrow interval width and retained the desired coverage probability. In this sense, narrow width prediction intervals provide more information about the uncertainty phenomena modelled. Furthermore, the proposed method was used to represent the future load uncertainty of a microgrid installed in Chile and residential dwellings in the town of Loughborough, UK for several prediction horizons. The results indicated that the proposed method for developing prediction intervals is suitable for load forecasting in applications of energy communities.

Future work will focus on the development of a robust energy management system based on the prediction interval modelling developed in this paper. Additionally, other evaluation metrics for the prediction interval could be included in the optimization problem, and a Pareto analysis could be included in the multi-objective cost function to obtain the best compromise solution.

Acknowledgements

This study was partially supported by the Complex Engineering Systems Institute (CONICYT–PIA–FB0816), the Solar Energy Research Centre SERC-Chile (CONICYT/FONDAP Project under Grant no. 15110019), and FONDECYT Chile (Grant no. 1170683; “Robust Distributed Predictive Control Strategies for the Coordination of Hybrid AC and DC Microgrids”). Luis G. Marín has also been supported by a Ph.D. scholarship from CONICYT-PCHA/Doctorado Nacional para extranjeros/2014-63140093, by a scholarship from COLCIENCIAS-Colombia and by a short-term research scholarship from the Graduate Department, Vice-Presidency of Academic Affairs of the University of Chile.

References

Ak, R., Li, Y., Vitelli, V., Zio, E., López Droguett, E., & Magno Couto Jacinto, C. (2013). NSGA-II-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment. *Expert Systems with Applications, 40*(4), 1205–1212. doi:10.1016/j.eswa.2012.08.018.

Babuška, R. (1998). *Fuzzy modeling for control* (1st ed.). Boston: Kluwer Academic Publishers. doi:10.1007/978-94-011-4868-9.

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long- short term memory. *Plos One*, 1–24. doi:10.6084/m9.figshare.5028110.

Begian, M., Melek, W., & Mendel, J. M. (2008). Parametric design of stable type-2 TSK fuzzy systems. In *Proceedings of the 2008 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS)* (pp. 1–6).

Bui, T. C., Le, V. D., & Cha, S. K. *A deep learning approach for forecasting air pollution in South Korea using LSTM*. http://arxiv.org/abs/1804.07891.

Chen, J., Zeng, G. Q., Zhou, W., Du, W., & Lu, K. D. (2018). Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. *Energy Conversion and Management, 165*, 681–695. doi:10.1016/j.enconman.2018.03.098.

Chen, S., Billings, S. A., & Grant, P. M. (1990). Non-linear system identification using neural networks. *International Journal of Control, 51*(6), 1191–1214. doi:10.1080/00207179008934126.

Chen, Z., Aghakhani, S., Man, J., & Dick, S. (2011). ANCFIS : A neurofuzzy architecture employing complex fuzzy sets. *IEEE Transactions on Fuzzy Systems, 19*(2), 305–322. doi:10.1109/TFUZZ.2010.2096469.

Dybowski, R., & Roberts, S. (2001). Confidence intervals and prediction intervals for feed-forward neural networks. In R. Dybowski, & V. Gant (Eds.), *Clinical applications of artificial neural networks* (pp. 298–326). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511543494.013.

Elbeltagi, E., Hegazy, T., & Grierson, D. (2005). Comparison among five evolutionary-based optimization algorithms. *Advanced Engineering Informatics, 19*(1), 43–53. doi:10.1016/j.aei.2005.01.004.

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the thirty-third international conference on machine learning: 48*. New York, NYUSA. doi:10.1109/TKDE.2015.2507132.

Gençay, R., & Qi, M. (2001). Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging. *IEEE Transactions on Neural Networks, 12*(4), 726–734. doi:10.1109/72.935086.

Ghanbari, A., Hadavandi, E., & Abbasian-Naghneh, S. (2010). Comparison of artificial intelligence based techniques for short term load forecasting. In *Proceedings of the third international conference on business intelligence and financial engineering* (pp. 6–10). Hong Kong: IEEE. doi:10.1109/BIFE.2010.12.

Heskes, T. (1997). Practical confidence and prediction intervals. *Advances in Neural Information Processing Systems, 9*, 176–182.

Hu, M., Hu, Z., Yue, J., Zhang, M., & Hu, M. (2017). A novel multi-objective optimal approach for wind power interval prediction. *Energies MDPI, 10*(4), 419–433. doi:10.3390/en10040419.

Jafarzadeh, S., Fadali, M. S., & Yaman, C. (2013). Solar power prediction using interval type-2 TSK modeling. *IEEE Transactions on Sustainable Energy, 4*(2), 333–339. doi:10.1109/TSTE.2012.2224893.

Kabir, H. M. D., Khosravi, A., Hosen, M. A., & Nahavandi, S. (2018). Neural network-based uncertainty quantification: A survey of methodologies and applications. *IEEE Access, 6*, 36218–36234. doi:10.1109/ACCESS.2018.2836917.

Karnik, N. N., & Mendel, J. M. (2001). Operations on type-2 fuzzy sets. *Fuzzy Sets and Systems, 122*(2), 327–348. doi:10.1016/S0165-0114(00)00079-8.

Khodayar, M., Wang, J., & Manthouri, M. (2018). Interval deep generative neural network for wind speed forecasting. *IEEE Transactions on Smart Grid*. doi:10.1109/TSG.2018.2847223.

Khosravi, A., & Nahavandi, S. (2013). Combined nonparametric prediction intervals for wind power generation. *IEEE Transactions on Sustainable Energy, 4*(4), 849–856. doi:10.1109/TSTE.2013.2253140.

Khosravi, A., & Nahavandi, S. (2014a). An interval type-2 fuzzy logic system-based method for prediction interval construction pdf. *Applied Soft Computing, 24*, 222–231. doi:10.1016/j.asoc.2014.06.039.

Khosravi, A., & Nahavandi, S. (2014b). Closure to the discussion of “Prediction intervals for short-term wind farm generation forecasts” and “Combined nonparametric prediction intervals for wind power generation” and the discussion of "Combined nonparametric prediction intervals for wind power. *IEEE Transactions on Sustainable Energy, 5*(3), 1022–1023. doi:10.1109/TSTE.2014.2323852.

Khosravi, A., & Nahavandi, S. (2014c). Load forecasting using interval type-2 fuzzy logic systems: Optimal type reduction. *IEEE Transactions on Industrial Informatics, 10*(2), 1055–1063. doi:10.1109/TII.2013.2285650.

Khosravi, A., Nahavandi, S., & Creighton, D. (2010). A prediction interval-based approach to determine optimal structures of neural network metamodels. *Expert Systems with Applications, 37*(3), 2377–2387. doi:10.1016/j.eswa.2009.07.059.

Khosravi, A., Nahavandi, S., & Creighton, D. (2011). Prediction interval construction and optimization for adaptive neurofuzzy inference systems. *IEEE Transactions on Fuzzy Systems, 19*(5), 983–988. doi:10.1109/TFUZZ.2011.2130529.

Khosravi, A., Nahavandi, S., & Creighton, D. (2013). Prediction intervals for short-term wind farm power generation forecasts. *IEEE Transactions on Sustainable Energy, 4*(3), 602–610. doi:10.1109/TSTE.2012.2232944.

Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011a). Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on Neural Networks, 22*(9), 1341–1356. doi:10.1109/TNN.2011.2162110.

Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011b). Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Transactions on Neural Networks, 22*(3), 337–346. doi:10.1109/TNN.2010.2096824.

Khosravi, A., Nahavandi, S., Creighton, D., & Srinivasan, D. (2012). Interval type-2 fuzzy logic systems for load forecasting : A comparative study. *IEEE Transactions on Power Systems, 27*(3), 1274–1282. doi:10.1109/TPWRS.2011.2181981.

Kroll, A., & Schulte, H. (2014). Benchmark problems for nonlinear system identification and control using soft computing methods: Need and overview. *Applied Soft Computing, 25*, 496–513. doi:10.1016/j.asoc.2014.08.034.

Lara, J. D., Olivares, D. E., & Cañizares, C. A. (2018). Robust energy management of isolated microgrids. *IEEE Systems Journal*, 1–12. doi:10.1109/JSYST.2018.2828838.

Lee, K. H. (2005). *First course on fuzzy theory and applications* (1st ed.). Berlin: Springer-Verlag Chapter 5. doi:10.1007/3-540-32366-X_5.

Li, C., & Chiang, T. (2013). Complex neurofuzzy ARIMA forecasting–A new approach using complex fuzzy sets. *IEEE Transactions on Fuzzy Systems, 21*(3), 567–584. doi:10.1109/TFUZZ.2012.2226890.

Li, K., Wang, R., Lei, H., Zhang, T., Liu, Y., & Zheng, X. (2018). Interval prediction of solar power using an improved bootstrap method. *Solar Energy, 159*, 97–112. doi:10.1016/j.solener.2017.10.051.

Liu, Y., Wang, Y., Yang, X., & Zhang, L. (2017). Short-term travel time prediction by deep learning: A comparison of different LSTM-DNN models. In *Proceedings of the IEEE twentieth international conference on intelligent transportation systems (ITSC)*. doi:10.1109/ITSC.2017.8317886.

Marín, L. G., Valencia, F., & Sáez, D. (2016). Prediction interval based on type-2 fuzzy systems for wind power generation and loads in microgrid control design. In *Proceedings of the IEEE international conference on fuzzy systems (FUZZ-IEEE)* (pp. 328–335). Vancouver, BC. doi:10.1109/FUZZ-IEEE.2016.7737705.

Mendel, J. M. (2017). *Uncertain rule-based fuzzy logic systems: Introduction and new directions* (2nd ed.). Springer International Publishing. doi:10.1007/978-3-319-51370-6.

Palma-Behnke, R., Benavides, C., Lanas, F., Severino, B., Reyes, L., Llanos, J., et al. (2013). A microgrid energy management system based on the rolling horizon strategy. *IEEE Transactions on Smart Grid, 4*(2), 996–1006. doi:10.1109/TSG.2012.2231440.

Parhizi, S., Lotfi, H., Khodaei, A., & Bahramirad, S. (2015). State of the art in research on microgrids: A review. *IEEE Access, 3*, 890–925. doi:10.1109/ACCESS.2015.2443119.

Pinson, P., & Tastu, J. (2014). Discussions of "prediction intervals for short-term wind farm generation forecasts" and "combined nonparametric prediction intervals for wind power generation". *IEEE Transactions on Sustainable Energy, 5*(3), 1019–1020. doi:10.1109/TSTE.2014.2323851.

Qiu, X., Zhang, L., Ren, Y., Suganthan, P., & Amaratunga, G. (2014). Ensemble deep learning for regression and time series forecasting. In *Proceedings of the* IEEE symposium on computational intelligence in ensemble learning. doi:10.1109/CIEL.2014.7015739.

Quan, H., Srinivasan, D., & Khosravi, A. (2014). Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE Transactions on Neural Networks and Learning System, 25*(2), 303–315. doi:10.1109/TNNLS.2013.2276053.

Ramezani, M., Bashiri, M., & Atkinson, A. C. (2011). A goal programming-TOPSIS approach to multiple response optimization using the concepts of non-dominated solutions and prediction intervals. *Expert Systems with Applications, 38*(8), 9557–9563. doi:10.1016/j.eswa.2011.01.139.

Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics* (2nd ed.). John Wiley & Sons, Inc. doi:10.1002/9780470192610.

Richardson, I., & Thomson, M. (2010). *One-minute resolution domestic electricity use data, 2008–2009*. Colchester, Essex: UK Data Archive [Distributor]. doi:10.5255/UKDA-SN-6583-1.

Rodrigues, F., Markou, I., & Pereira, F. C. (2018). Combining time-series and textual data for taxi demand prediction in event areas: A deep learning approach. *Information Fusion*, 1–20. doi:10.1016/j.inffus.2018.07.007.

Sáez, D., Ávila, E., Olivares, D., Cañizares, C., & Marín, L. G. (2015). Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids. *IEEE Transactions on Smart Grid, 6*(2), 548–556. doi:10.1109/TSG.2014.2377178.

Sáez, D., & Zuñiga, R. (2004). Cluster optimization for Takagi & Sugeno fuzzy models and its application to a combined cycle power plant boiler. In *Proceedings of the 2004 American control conference: 2* (pp. 1776–1781). Boston, MAUSA.

Shrivastava, N. A., Khosravi, A., & Panigrahi, B. K. (2015). Prediction interval estimation of electricity prices using PSO tuned support vector machines. *IEEE Transactions on Industrial Informatics, 11*(2), 322–331. doi:10.1109/TII.2015.2389625.

Shrivastava, N. A., Lohia, K., & Panigrahi, B. K. (2016). A multiobjective framework for wind speed prediction interval forecasts. *Renewable Energy, 87*(Part 2), 903–910. doi:10.1016/j.renene.2015.08.038.

Škrjanc, I. (2011). Fuzzy confidence interval for pH titration curve. *Applied Mathematical Modelling, 35*(8), 4083–4090. doi:10.1016/j.apm.2011.02.033.

Škrjanc, I., Blažič, S., & Agamennoni, O. (2005). Interval fuzzy model identification using l -Norm. *IEEE Transactions on Fuzzy Systems, 13*(5), 561–568. doi:10.1109/TFUZZ.2005.856567.

Tran, D. C., Wu, Z., & Nguyen, V. X. (2013). A new approach based on enhanced PSO with neighborhood search for data clustering. In *Proceedings of the international conference on soft computing and pattern recognition (SoCPaR)* (pp. 98–104). Hanoi. doi:10.1109/SOCPAR.2013.7054109.

Valencia, F., Collado, J., Sáez, D., & Marín, L. G. (2016). Robust energy management system for a microgrid based on a fuzzy prediction interval model. *IEEE Transactions on Smart Grid, 7*(3), 1486–1494. doi:10.1109/TSG.2015.2463079.

Valencia, F., Sáez, D., Collado, J., Ávila, F., Marquez, A., & Espinosa, J. (2016). Robust energy management system based on interval fuzzy models. *IEEE Transactions on Control Systems Technology, 24*(1), 140–157. doi:10.1109/TCST.2015.2421334.

Veltman, F., Marín, L. G., Sáez, D., Gutierrez, L., & Nuñez, A. (2015). Prediction interval modelling tuned by an improved teaching learning algorithm applied to load forecasting in microgrids. In *Proceedings of the IEEE symposium series on computational intelligence* (pp. 651–658). Cape Town. doi:10.1109/SSCI.2015.100.

Voyant, C., Motte, F., Notton, G., Fouilloy, A., Nivet, M. L., & Duchaud, J. L. (2018). Prediction intervals for global solar irradiation forecasting using regression trees methods. *Renewable Energy, 126*, 332–340. doi:10.1016/j.renene.2018.03.055.

Wan, C., Xu, Z., Østergaard, J., Dong, Z. Y., & Wong, K. P. (2014). Discussion of combined nonparametric prediction intervals for wind power generation. *IEEE Transactions on Sustainable Energy, 5*(3), 1021. doi:10.1109/TSTE.2014.2323836.

Wang, S., & Jia, C. (2015). Prediction intervals for short-term photovoltaic generation forecasts. In *Proceedings of the fifth international conference on instrumentation and measurement, computer, communication and control (IMCCC)* (pp. 459–463). Qinhuangdao. doi:10.1109/IMCCC.2015.103.

Xiang, Y., Liu, J., & Liu, Y. (2016). Robust energy management of microgrid with uncertain renewable generation and load. *IEEE Transactions on Smart Grid, 7*(2), 1034–1043. doi:10.1109/TSG.2014.2385801.

Xu, Y., Zhang, M., Zhu, Q., & He, Y. (2017). An improved multi-kernel RVM integrated with CEEMD for high-quality intervals prediction construction and its intelligent modeling application. *Chemometrics and Intelligent Laboratory Systems, 171*, 151–160. doi:10.1016/j.chemolab.2017.10.019.

Yazdanbakhsh, O., & Dick, S. (2017). Forecasting of multivariate time series via complex fuzzy logic. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47*(8), 2160–2171. doi:10.1109/TSMC.2016.2630668.

Yazdanbakhsh, O., & Dick, S. (2018). A systematic review of complex fuzzy sets and logic. *Fuzzy Sets and Systems, 338*, 1–22. doi:10.1016/j.fss.2017.01.010.

Yazdanbaksh, O., Krahn, A., & Dick, S. (2013). Predicting solar power output using complex fuzzy logic. In *Proceedings of the joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS)* (pp. 1243–1248). Edmonton, AB. doi:10.1109/IFSA-NAFIPS.2013.6608579.

Zhu, L., & Laptev, N. (2017). Deep and confident prediction for time series at Uber. In *Proceedings of the IEEE international conference on data mining workshops (ICDMW)*. doi:10.1109/ICDMW.2017.19.