Electrical Power and Energy Systems 73 (2015) 34–44

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

A novel random fuzzy neural networks for tackling uncertainties of electric load forecasting

Chin Wang Lou *, Ming Chui Dong

Faculty of Science and Technology, University of Macau, Avenida da Universidade, Taipa, Macau

ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 7 March 2015
Accepted 15 March 2015
Available online 22 April 2015

Keywords:
Load forecasting
Uncertainty
Random fuzzy variable
Random fuzzy neural networks

ABSTRACT

In this paper, fuzzy (inaccuracy, vague, dispersion of individual interpretation) and random (incompleteness, noise and variability) uncertainties of electric load forecasting are modeled by random fuzzy variables (RFVs). Further integrating it into neural networks (NN) to formulate a novel integrated technique–Random Fuzzy NN (RFNN) for load forecasting is presented. The features of this methodology are as follows. (1) It is able to effectively and simultaneously model referred uncertainties occurred in load forecasting by one integrated technique, which existing techniques (e.g. fuzzified NN or Bayesian NN) tackle them separately. (2) Specially, historical data/information containing incompleteness, inaccuracy and vagueness can be modeled by the proposed representations of RFVs. No preprocessing algorithms such as data imputation or discard are required. (3) The proposed RFNN can make NN incorporating both types of uncertainties of inputs and network parameters so as to possess with better tackling uncertainties of load forecasting than other relevant methods. The proposed techniques are applied to electric load forecasting using a real operational data collected from Macau electric utility. Its application is promising in microgrid/small power system or in forecasting curves of individual customer where load curves would present a much higher variability and more noise than global curves of power grid of one country/region.

Introduction

Various techniques of load forecasting had been reported for past decades. In summary, they can be classified as traditional methods based on mathematical models [1–5] and soft-computing techniques [6–8]. Hybrid of the referred methods [9–12] has also been achieving good performances. However, load forecasting is characterized with stochastic and uncertainty, so new techniques have to possess with high-level ability to represent and tackle various uncertainties. Moreover, load curves of microgrid in smart grid or curves of individual customer (measured by smart meter) will present a much higher variability and more noises (Fig. 1). It is because aggregation or smoothing effect is reduced. The uncertainty increases as size of power system/microgrid gets smaller. Thus, load forecasting would face higher challenges due to more uncertainties and variability of load curves. Different techniques [13–15] had been reported to tackle uncertainties in load forecasting application in past.

Artificial neural networks (ANN) have been widely used in load forecasting due to their strong capability to model non-linear mapping relations between inputs and outputs (I/O) [16,17]. However, their drawback is the lack of incorporating uncertainties associated with inputs, parameters of network and I/O relation so as to obstruct their further applications. Thus, Bayesian NN (BNN) [18,19] had been developed to make NN incorporating the ability tackling probability of uncertainty. Fuzzified NN (FNN) [20–23] had also been reported extensively to incorporate fuzzy uncertainty into NN.

However, it is noteworthy that different techniques integrated with fuzzy or probability methods for modeling and tackling uncertainties have following drawbacks:

1. Fuzzy methods can be used to tackle fuzziness on data and knowledge. Probability methods can handle uncertainty caused by random effect. However, individual method cannot fully and effectively tackle various data uncertainties, which hybrid technique is promising and proposed.
2. For load forecasting application, missing or inaccuracy values in historical data would be firstly preprocessed. The simplest technique is to discard those data. Other advanced techniques are

* Corresponding author. Tel.: +853 66806384.
E-mail addresses: jerrylcw2010@macau.ctm.net (C.W. Lou), mcdong@umac.mo (M.C. Dong).

http://dx.doi.org/10.1016/j.ijepes.2015.03.003
0142-0615/

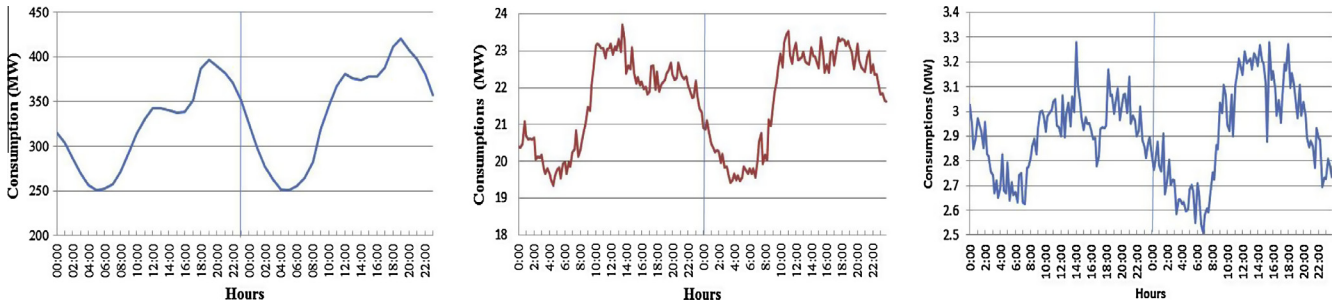

Fig. 1. Global load curve of power system of a country/region (left); load curve of a substation (potential microgrid) (middle); load curve of an individual customer (right). All data in three charts were record in the same period.

data imputation (expectation–maximization and its variants [24], and other algorithms [25–28]). However, their remarkable drawbacks are imputation rate with uncertain accuracy.

This paper proposes a novel technique such that fuzzy and random uncertainties, which occur in load forecasting simultaneously, are modeled by random fuzzy variables (RFVs) and its different representative forms. Furthermore, its integration into NN to formulate a novel integrated technique–Random Fuzzy NN (RFNN) is presented for load forecasting. The features of this methodology are as follows. (1) While existing techniques can model fuzzy and random uncertainties separately, our approach can tackle the referred uncertainties by one integrated technique. (2) In tradition, data imputation and preprocessing are required to handle abnormalities of data. However, special representative forms of proposed RFVs can model missing and inaccurate data or incomplete and vague information directly. (3) The proposed RFNN can make traditional NN incorporating both types of uncertainties of inputs and network's parameters so as to possess with better tackling uncertainties of load forecasting than other relevant methods. The proposed techniques are applied to electric load forecasting using a real operational data collected from Macau electric utility. Test results reveal that the proposed method is superior to other existing relevant methods.

Review of random fuzzy variable [29–31]

Definition

It is well known that a confidence interval is a closed interval in set of real numbers $\Re$ within which possible values of an uncertain variable can be located, i.e.

$$B = [b_1\ b_2] \text{ where } b_1 < b_2 \tag{1}$$

If lower and upper bounds of the closed interval are also uncertain, then Eq. (1) becomes

$$B = [[b_1\ b_2][b_3\ b_4]] \text{ where } b_1 < b_2 < b_3 < b_4 \tag{2}$$

This means that left and right bounds of the confidence interval are not defined by crisp values but are represented by the given intervals. In the way, the bounds vary within these intervals can be accounted by random effect. Under the assumption, it can be stated that internal interval $[b_2\ b_3]$ can represent effect of all unknown but systematic contributions to uncertainty while external interval $[b_1\ b_2]$ and $[b_3\ b_4]$ can represent effect of all random contributions. Therefore, one RFV can be defined as a set of confidence intervals $B_\alpha = [[b_1^\alpha\ b_2^\alpha],\ [b_3^\alpha\ b_4^\alpha]]$, $\alpha \in [0\ 1]$ that obeys the following constraints:

1. $b_1^\alpha \leqslant b_2^\alpha \leqslant b_3^\alpha \leqslant b_4^\alpha \quad \forall \alpha \in [0\ 1]$.
2. Sequence of confidence interval of $[b_1^\alpha\ b_4^\alpha]$ and $[b_2^\alpha\ b_3^\alpha]$ generates two membership functions (MFs) that are normal and convex.
3. $\alpha' > \alpha \Rightarrow \begin{Bmatrix} b_1^{\alpha'} & b_3^{\alpha'} \\ b_2^{\alpha'} & b_4^{\alpha'} \end{Bmatrix} \subset \begin{bmatrix} b_1^{\alpha} & b_3^{\alpha} \\ b_2^{\alpha} & b_4^{\alpha} \end{bmatrix} \quad \forall \alpha, \alpha'$ in range of [0 1].
4. $\begin{bmatrix} b_2^{\alpha=1} & b_3^{\alpha=1} \end{bmatrix} \equiv \begin{bmatrix} b_1^{\alpha=1} & b_4^{\alpha=1} \end{bmatrix}$.

The representation of uncertain effects by one RFV is shown in Fig. 2, in which $\mu_{\mathbf{X}}(x)$ is grade of MF, and $\mathbf{X}$ is set of variables and $x \in \Re$. Bounds of the external and internal interval $[x_1^\alpha\ \ x_4^\alpha]$ and $[x_2^\alpha\ \ x_3^\alpha]$ can define two MFs that represent two distributions of possibility in theory of evidence. As a result, by using appropriate probability–possibility transformation [32–34] and available evidences at hand, one RFV can be built in such a way that its internal MF can represent effect of unknown but systematic contribution and effect of total/partial ignorance while its external MF can model effect of random contribution.

Mathematical operations of RFVs

The defined RFV can be used to solve problem of load forecasting only if its mathematical operations are properly defined. By definition, one RFV is composed by internal and external MFs representing possibility and probability respectively. It follows that when exercising mathematical operation on two RFVs, internal and external MFs should be separately processed in two different mathematical ways in light of different nature of effects they are representing.

The mathematical operation of internal MFs on two RFVs can follow mathematical calculation of fuzzy variables [35]. However, the mathematical operation of external MFs should be processed in accordance with probability theory. However, if different correlations on two random variables are considered, such calculations become very complicated and time consuming [31].

Propose applying RFVs in electric load forecasting

In domain of load forecasting, each data and weather information could be assumed measured independently. Therefore, correlation on random characteristic of two measurements sampled in consecutive intervals could be presumed as zero. Under this

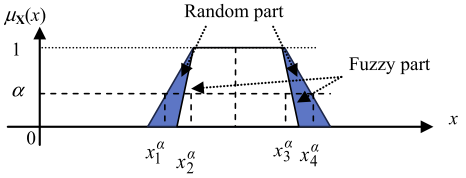

Fig. 2. Representation of uncertainty effects (fuzziness and randomness) by one RFV.

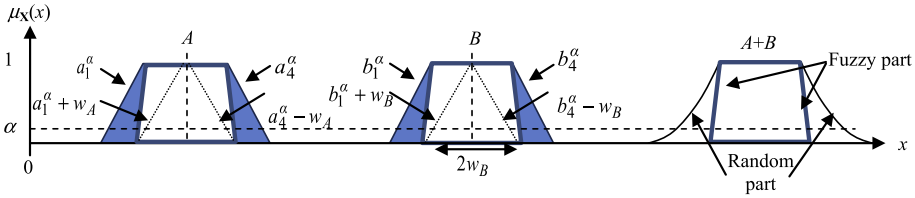

Fig. 3. Demonstration of mathematical 'Sum' operation on two RFVs.

assumption, mathematical operations of RFVs can be simplified so as to apply in realm of load forecasting. As a demonstration, one 'Sum' operation of two RFVs is shown in Fig. 3. It is noted that although shape of external MF of RFVs is trapezoidal, the resulted MF is Gaussian-probability distribution.

Moreover, it should be noteworthy that when the random effect is not accounted, one RFV is degenerated into a pure fuzzy variable shown in Fig. 4a. Therefore, such a special form of RFV can be used to represent vague linguistic terms of human as is shown in Fig. 4b. Moreover, data inaccuracy can be modeled as a special representation of RFV shown in Fig. 5a. Missing data can be modeled by a rectangle-shape possibility function [0 1] shown in Fig. 5b, which simulates complete ignorance on missing value.

In summary, it is discovered that after different representative forms of RFV are utilized, various uncertainties including fuzziness (inaccuracy, vagueness and dispersion of individual interpretation) and randomness (incompleteness, noise and volatility) can be modeled by RFVs and its different representative forms simultaneously and appropriately.

The proposed Random Fuzzy NN

The proposed hybrid system is presented in Fig. 6. Due to the usage of RFVs, all inputs, outputs, weights and biases parameters are represented in forms of RFV shown in Fig. 7, where $\{w_{ij1}, w_{ij2}, w_{ij3}, w_{ij4}\}$, $\{w_{jk1}, w_{jk2}, w_{jk3}, w_{jk4}\}$, $\{\theta_{j1}, \theta_{j2}, \theta_{j3}, \theta_{j4}\}$ and $\{\theta_{k1}, \theta_{k2}, \theta_{k3}, \theta_{k4}\}$ are four base points of respective RFVs. As a result, original mathematical formulas dedicated for learning parameters of traditional NN are required to modify. The complete network relation between layers must be reformulated and processed with mathematical operations using theory of RFV.

The parameters in Fig. 6 are explained as below.

$[X_{pi}]$: Inputs of NN.
$[O_{pi}]$, $[O_{pj}]$, $[O_{pk}]$: Outputs of input, hidden and output layer respectively.
$[W_{ij}]$, $[W_{jk}]$: Weights between input-hidden and hidden-output layers respectively.

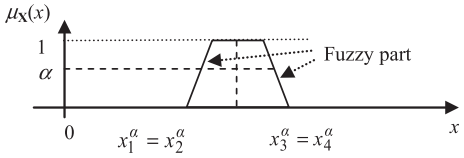

Fig. 4a. One RFV is degenerated into a pure fuzzy variable.

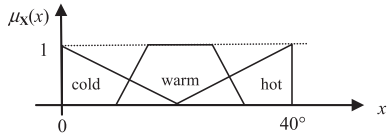

Fig. 4b. A special form of RFV for representing information described in human's vague linguistic terms.

$[\Theta_j]$, $[\Theta_k]$: Biases in hidden and output layers.
m, h, n: Number of NN's inputs, hidden neurons and output neurons.
p: The pth data series inputted to NN.

Computation of input–output relation

Complete I/O relation can be formulated by using theory of RFV (Remark: X are inputs to nodes of input layer, O are outputs at nodes of input, hidden and output layer respectively) as below.

1. Input layer:
Each node in this layer, which represents an input variable to NN, directly transmits each input to next layer as is described below

$$[O_{pi}] = [X_{pi}],\ i = 1,\ldots,m;\ p = 1,\ldots,P \tag{3}$$

where P is number of data series inputted to NN. By using definition of RFV and associated mathematical operations (Appendix 'Mathematical operations of two RFVs'), Eq. (3) is re-written as

$$\begin{aligned}[O_{pi}]^\alpha &= \left[[O_{pi}]_1^\alpha, [O_{pi}]_2^\alpha, [O_{pi}]_3^\alpha, [O_{pi}]_4^\alpha\right] \\ &= \left[[X_{pi}]_1^\alpha, [X_{pi}]_2^\alpha, [X_{pi}]_3^\alpha, [X_{pi}]_4^\alpha\right].\end{aligned} \tag{4}$$

2. Hidden layer:

$$[O_{pj}] = f([Net_{pj}]),\quad j = 1,2,\ldots,h \tag{5}$$

$$[Net_{pj}] = \left(\sum_{i=1}^{m} [W_{ij}] \otimes [O_{pi}]\right) \oplus [\Theta_j] \tag{6}$$

where Σ, $\oplus$, $\otimes$ and $f(\bullet)$ are summation notation, sum and product operation of RFVs, and activation function respectively. Using the definition of RFVs and expanding Eq. (6) can obtain

$$\begin{aligned}&\left[[Net_{pj}]_1^\alpha, [Net_{pj}]_2^\alpha, [Net_{pj}]_3^\alpha, [Net_{pj}]_4^\alpha\right] \\ &= \left[[W_{1j}]_1^\alpha, [W_{1j}]_2^\alpha, [W_{1j}]_3^\alpha, [W_{1j}]_4^\alpha\right] \\ &\quad \otimes \left[[X_{p1}]_1^\alpha, [X_{p1}]_2^\alpha, [X_{p1}]_3^\alpha, [X_{p1}]_4^\alpha\right] \oplus \cdots \\ &\quad \oplus \left[[W_{mj}]_1^\alpha, [W_{mj}]_2^\alpha, [W_{mj}]_3^\alpha, [W_{mj}]_4^\alpha\right] \\ &\quad \otimes \left[[X_{pm}]_1^\alpha, [X_{pm}]_2^\alpha, [X_{pm}]_3^\alpha, [X_{pm}]_4^\alpha\right] \\ &\quad \oplus \left[[\Theta_j]_1^\alpha, [\Theta_j]_2^\alpha, [\Theta_j]_3^\alpha, [\Theta_j]_4^\alpha\right].\end{aligned} \tag{7}$$

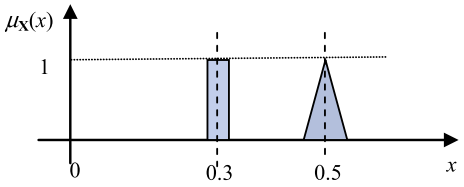

Fig. 5a. The known but inaccurate input is expressed by one form of RFV.

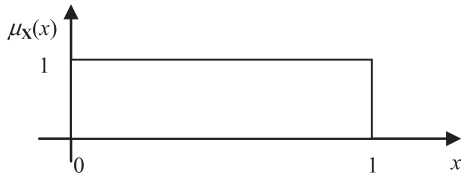

Fig. 5b. The unknown input is expressed by interval representation within the possibility range [0 1], which simulates complete ignorance on missing value

In accordance with Eqs. (A10) and (A11) of product operation and Eqs. (A2) and (A3) of sum operation (Appendix 'Mathematical operations of two RFVs'), these inner parts of $[Net_{pj}]$ can be calculated as below

$$[Net_{pj}]_2^{\alpha} = \sum_{i=1}^{m} [c_{ij}]_2^{\alpha} + [\Theta_j]_2^{\alpha} = \sum_{i=1}^{m+1} [c_{ij}]_2^{\alpha} \tag{8}$$

$$[Net_{pj}]_3^{\alpha} = \sum_{i=1}^{m} [c_{ij}]_3^{\alpha} + [\Theta_j]_3^{\alpha} = \sum_{i=1}^{m+1} [c_{ij}]_3^{\alpha} \tag{9}$$

where

$$[c_{ij}]_2^{\alpha} = \min\left([W_{ij}]_2^{\alpha}[X_{pi}]_2^{\alpha}, [W_{ij}]_2^{\alpha}[X_{pi}]_3^{\alpha}, [W_{ij}]_3^{\alpha}[X_{pi}]_2^{\alpha}, [W_{ij}]_3^{\alpha}[X_{pi}]_3^{\alpha}\right).$$

$$[c_{ij}]_3^{\alpha} = \max\left([W_{ij}]_2^{\alpha}[X_{pi}]_2^{\alpha}, [W_{ij}]_2^{\alpha}[X_{pi}]_3^{\alpha}, [W_{ij}]_3^{\alpha}[X_{pi}]_2^{\alpha}, [W_{ij}]_3^{\alpha}[X_{pi}]_3^{\alpha}\right).$$

$$[c_{(i+1)j}]_2^{\alpha} = [\Theta_j]_2^{\alpha} \text{ and } [c_{(i+1)j}]_3^{\alpha} = [\Theta_j]_3^{\alpha}.$$

The calculation of outer parts of $[Net_{pj}]$ is more complicated. However, they can be carried out by two steps. Firstly, each product term in Eq. (7) can be derived according to Eqs. (A9)–(A11) and (A12) of product operation of RFV

$$[c_{ij}]_1^{\alpha} = [c_{ij}]_2^{\alpha} - 0.5\left([c_{ij}]_{r1}^{\alpha=1} + [c_{ij}]_{r4}^{\alpha=1}\right) + [c_{ij}]_{r1}^{\alpha} \tag{10}$$

$$[c_{ij}]_4^{\alpha} = [c_{ij}]_3^{\alpha} - 0.5\left([c_{ij}]_{r1}^{\alpha=1} + [c_{ij}]_{r4}^{\alpha=1}\right) + [c_{ij}]_{r4}^{\alpha} \tag{11}$$

where

$$[c_{ij}]_{r1}^{\alpha} = \min\left\{[s_{ij}]_{r1}^{\alpha}[t_{ij}]_{r1}^{\alpha}, [s_{ij}]_{r1}^{\alpha}[t_{ij}]_{r4}^{\alpha}, [s_{ij}]_{r4}^{\alpha}[t_{ij}]_{r1}^{\alpha}, [s_{ij}]_{r4}^{\alpha}[t_{ij}]_{r4}^{\alpha}\right\}.$$

$$[c_{ij}]_{r4}^{\alpha} = \max\left\{[s_{ij}]_{r1}^{\alpha}[t_{ij}]_{r1}^{\alpha}, [s_{ij}]_{r1}^{\alpha}[t_{ij}]_{r4}^{\alpha}, [s_{ij}]_{r4}^{\alpha}[t_{ij}]_{r1}^{\alpha}, [s_{ij}]_{r4}^{\alpha}[t_{ij}]_{r4}^{\alpha}\right\}.$$

$$[s_{ij}]_{r1}^{\alpha} = [W_{ij}]_1^{\alpha} + W_{iA}, \quad [s_{ij}]_{r4}^{\alpha} = [W_{ij}]_4^{\alpha} - W_{iA},$$

$$[t_{ij}]_{r1}^{\alpha} = [X_{pi}]_1^{\alpha} + X_{iA} \text{ and } [t_{ij}]_{r4}^{\alpha} = [X_{pi}]_4^{\alpha} - X_{iA}.$$

$$W_{iA} = 0.5\left([W_{ij}]_3^{\alpha=1} - [W_{ij}]_2^{\alpha=1}\right) \text{ and } X_{iA} = 0.5\left([X_{pi}]_3^{\alpha=1} - [X_{pi}]_2^{\alpha=1}\right).$$

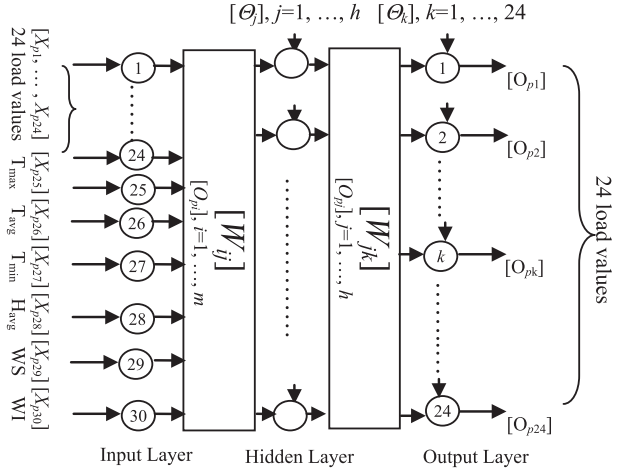

Fig. 6. Architecture of the proposed RFNN.

Secondly, after these m product terms (Eq. (10)) and $[\Theta_j]_1^{\alpha}$ are summed according to Eq. (A1) of sum operation of RFV, the outer part of $[Net_{pj}]$, i.e., $[Net_{pj}]_1^{\alpha}$ is obtained. Similarly, after these m product terms (Eq. (11)) and $[\Theta_j]_4^{\alpha}$ are summed according to Eq. (A4), the $[Net_{pj}]_4^{\alpha}$ is obtained.

3. Output layer:
Each node in this layer are called output node.

$$[O_{pk}] = f([Net_{pk}]), \ k = 1, 2, \ldots, n; \ p = 1, \ldots, P.$$
$$[Net_{pk}] = \left(\sum_{j=1}^{h} [W_{kj}] \otimes [O_{pj}]\right) \oplus [\Theta_k]. \tag{12}$$

Similarly, using the definition of RFV and expanding Eq. (12) can obtain

$$\begin{aligned}&\left[[Net_{pk}]_1^{\alpha}, [Net_{pk}]_2^{\alpha}, [Net_{pk}]_3^{\alpha}, [Net_{pk}]_4^{\alpha}\right]\\ &= [[W_{k1}]_1^{\alpha}, [W_{k1}]_2^{\alpha}, [W_{k1}]_3^{\alpha}, [W_{k1}]_4^{\alpha}]\\ &\quad \otimes \left[[O_{p1}]_1^{\alpha}, [O_{p1}]_2^{\alpha}, [O_{p1}]_3^{\alpha}, [O_{p1}]_4^{\alpha}\right] \oplus \cdots\\ &\quad \oplus [[W_{kh}]_1^{\alpha}, [W_{kh}]_2^{\alpha}, [W_{kh}]_3^{\alpha}, [W_{kh}]_4^{\alpha}]\\ &\quad \otimes \left[[O_{ph}]_1^{\alpha}, [O_{ph}]_2^{\alpha}, [O_{ph}]_3^{\alpha}, [O_{ph}]_4^{\alpha}\right]\\ &\quad \oplus [[\Theta_k]_1^{\alpha}, [\Theta_k]_2^{\alpha}, [\Theta_k]_3^{\alpha}, [\Theta_k]_4^{\alpha}].\end{aligned} \tag{13}$$

Following the similar derivations of Eqs. (8) and (9) can obtain the inner part of $[Net_{pk}]$ (k = 1, …, n) respectively

$$[Net_{pk}]_2^{\alpha} = \sum_{j=1}^{h} [s_{kj}]_2^{\alpha} + [\Theta_k]_2^{\alpha} = \sum_{j=1}^{h+1} [s_{kj}]_2^{\alpha} \tag{14}$$

$$[Net_{pk}]_3^{\alpha} = \sum_{j=1}^{h} [s_{kj}]_3^{\alpha} + [\Theta_k]_3^{\alpha} = \sum_{j=1}^{h+1} [s_{kj}]_3^{\alpha} \tag{15}$$

where $[s_{k(h+1)}]_2^{\alpha} = [\Theta_k]_2^{\alpha}$ and $[s_{k(h+1)}]_3^{\alpha} = [\Theta_k]_3^{\alpha}$.

$$[s_{kj}]_2^{\alpha} = \min\left([W_{kj}]_2^{\alpha}[O_{pj}]_2^{\alpha}, [W_{kj}]_2^{\alpha}[O_{pj}]_3^{\alpha}, [W_{kj}]_3^{\alpha}[O_{pj}]_2^{\alpha}, [W_{kj}]_3^{\alpha}[O_{pj}]_3^{\alpha}\right).$$

$$[s_{kj}]_3^{\alpha} = \max\left([W_{kj}]_2^{\alpha}[O_{pj}]_2^{\alpha}, [W_{kj}]_2^{\alpha}[O_{pj}]_3^{\alpha}, [W_{kj}]_3^{\alpha}[O_{pj}]_2^{\alpha}, [W_{kj}]_3^{\alpha}[O_{pj}]_3^{\alpha}\right).$$

The calculation of outer parts of $[Net_{pk}]$ can follow the similar derivations of Eqs. (10) and (11) to obtain each product term of Eq. (13)

$$[d_{kj}]_1^{\alpha} = [d_{kj}]_2^{\alpha} - 0.5\left([d_{kj}]_{r1}^{\alpha=1} + [d_{kj}]_{r4}^{\alpha=1}\right) + [d_{kj}]_{r1}^{\alpha} \tag{16}$$

$$[d_{kj}]_4^{\alpha} = [d_{kj}]_3^{\alpha} - 0.5\left([d_{kj}]_{r1}^{\alpha=1} + [d_{kj}]_{r4}^{\alpha=1}\right) + [d_{kj}]_{r4}^{\alpha} \tag{17}$$

where

$$[d_{kj}]_1^{\alpha} = \min\left\{[u_{kj}]_{r1}^{\alpha}[v_{pj}]_{r1}^{\alpha}, [u_{kj}]_{r1}^{\alpha}[v_{pj}]_{r4}^{\alpha}, [u_{kj}]_{r4}^{\alpha}[v_{pj}]_{r1}^{\alpha}, [u_{kj}]_{r4}^{\alpha}[v_{pj}]_{r4}^{\alpha}\right\}.$$

$$[d_{kj}]_4^{\alpha} = \max\left\{[u_{kj}]_{r1}^{\alpha}[v_{pj}]_{r1}^{\alpha}, [u_{kj}]_{r1}^{\alpha}[v_{pj}]_{r4}^{\alpha}, [u_{kj}]_{r4}^{\alpha}[v_{pj}]_{r1}^{\alpha}, [u_{kj}]_{r4}^{\alpha}[v_{pj}]_{r4}^{\alpha}\right\}.$$

$$[u_{kj}]_{r1}^{\alpha} = [W_{jk}]_1^{\alpha} + W_{jB} \text{ and } [u_{kj}]_{r4}^{\alpha} = [W_{jk}]_4^{\alpha} - W_{jB},$$

$$[v_{pj}]_{r1}^{\alpha} = [O_{pj}]_1^{\alpha} + O_{jB} \text{ and } [v_{kj}]_{r4}^{\alpha} = [O_{pj}]_4^{\alpha} - O_{jB}.$$

$$W_{jB} = 0.5\left([W_{jk}]_3^{\alpha=1} - [W_{jk}]_2^{\alpha=1}\right) \text{ and } O_{jB} = 0.5\left([O_{pj}]_3^{\alpha=1} - [O_{pj}]_2^{\alpha=1}\right).$$

After these h product terms (Eq. (16)) and $[\Theta_k]_1^{\alpha}$ are summed according to Eq. (A1) of sum operation of RFV, the outer part of $[Net_{pk}]$ is obtained, i.e., $[Net_{pk}]_1^{\alpha}$. Similarly, after these h product terms (Eq. (17)) and $[\Theta_k]_4^{\alpha}$ are summed according to Eq. (A4), the $[Net_{pk}]_4^{\alpha}$ is obtained.

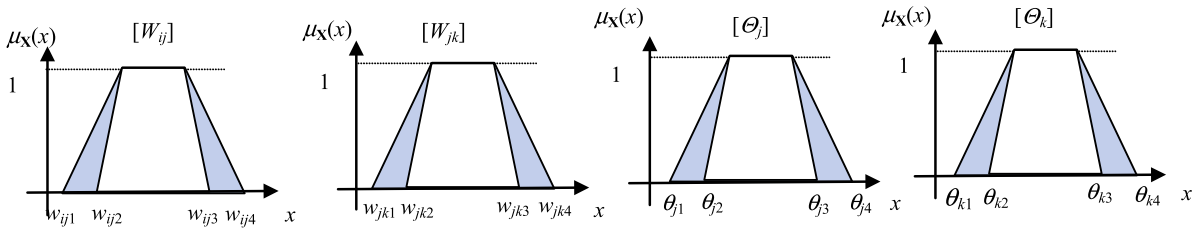

Fig. 7. Parameters of the proposed RFNN are represented in forms of RFV.

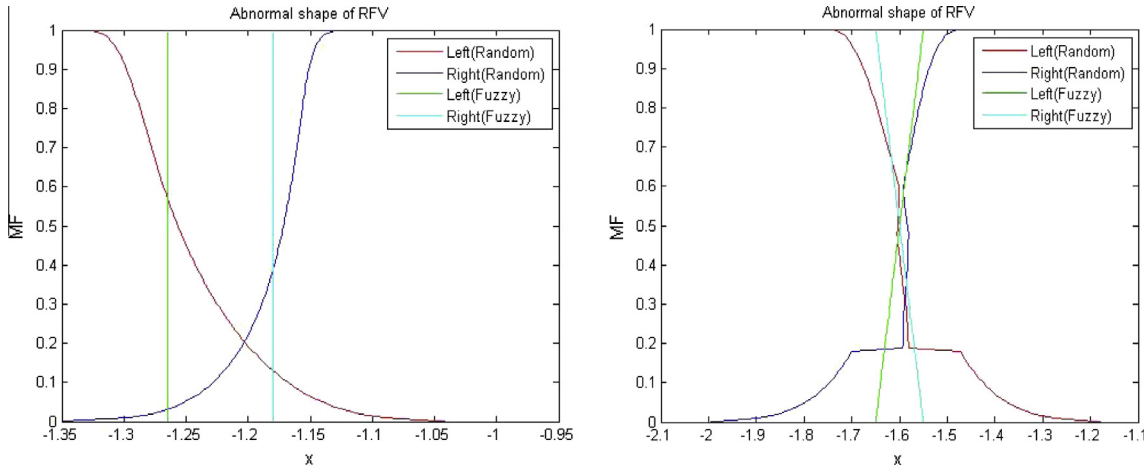

Fig. 8. Examples of incorrect shapes of resultant RFV.

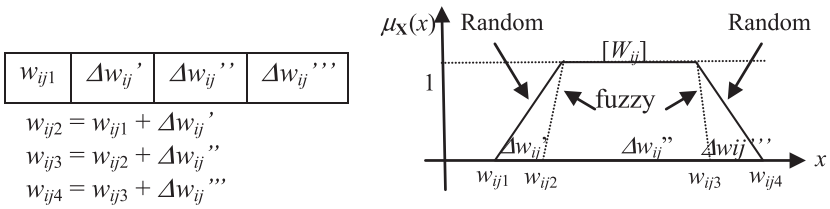

Fig. 9. Demonstration of the proposed encoding technique used in PSO optimization (W_{ij} is served as an example).

Learning algorithm

Given P pairs of input and output data series $(\mathbf{X}_p, \mathbf{T}_p)$, where $\mathbf{X}_p = (X_{p1}, X_{p2}, \ldots, X_{pm})$, $\mathbf{T}_p = (T_{p1}, T_{p2}, \ldots, T_{pn})$ and $p = 1, \ldots, P$. Learning objective is to minimize cost function Eq. (18) to find optimal parameters of RFNN in forms of RFV, i.e., $[W_{ij}]$, $[W_{jk}]$ and $[\Theta_j]$, $[\Theta_k]$.

$$Min\ E = \sum_{p=1}^{P}\sum_{k=1}^{n}[E_{pk}] \tag{18}$$

subject to the following constraints in order to conform correct shapes of RFV

Table 1
Input variables of RFNN system.

Input variable	Description
1–24	24 hourly electric load value
25–27	Maximum, average and minimum daily temperature
28	Average daily humidity value
29	Wind speed
30	Weather information

$$[W_{ij}]^{\alpha_1} \subset [W_{ij}]^{\alpha_2} \text{ and } [W_{jk}]^{\alpha_1} \subset [W_{jk}]^{\alpha_2} \text{ if } \alpha_2 < \alpha_1 \tag{19}$$

$$[\Theta_j]^{\alpha_1} \subset [\Theta_j]^{\alpha_2} \text{ and } [\Theta_k]^{\alpha_1} \subset [\Theta_k]^{\alpha_2} \text{ if } \alpha_2 < \alpha_1 \tag{20}$$

$$w_{ij1} < w_{ij2} < w_{ij3} < w_{ij4} \text{ and } w_{jk1} < w_{jk2} < w_{jk3} < w_{jk4} \tag{21}$$

$$\theta_{j1} < \theta_{j2} < \theta_{j3} < \theta_{j4} \text{ and } \theta_{k1} < \theta_{k2} < \theta_{k3} < \theta_{k4} \tag{22}$$

where $[E_{pk}] = 0.5\left([T_{pk}] \oplus (-[O_{pk}])^2\right)$ is error performance function.

In tradition, parameters of NN are learned and adjusted by differentiability of error performance function, with respect to respective optimized parameters and subject to set of inequality constraints Eqs. (19)–(22) as is shown below.

(a) Equation for updating parameter $[W_{ij}]$

$$[W_{ij}](t+1) = [W_{ij}](t) \oplus \left(-\eta \otimes \frac{\partial [E_{pk}]}{\partial [W_{ij}]}\right) \oplus \left(\beta \otimes \Delta [W_{ij}](t-1)\right)(t) \tag{23}$$

where $\frac{\partial [E_{pk}]}{\partial [W_{ij}]} = [O_{pj}] \otimes (1 \oplus (-[O_{pj}])) \otimes \left[\overline{O}_{pi}\right] \otimes \sum_{k=1}^{n} (\delta_{pk} \otimes [W_{jk}])$ and η, β are learning rates.

(b) Equation for updating parameter $[\Theta_j]$

$$[\Theta_j](t+1) = [\Theta_j](t) \oplus \left(-\eta \otimes \frac{\partial [E_{pk}]}{\partial [\Theta_j]}\right) \oplus (\beta \otimes \Delta[\Theta_j](t-1))(t) \quad (24)$$

where $\frac{\partial [E_{pk}]}{\partial [\Theta_j]} = [O_{pj}] \otimes (1 \oplus (-[O_{pj}])) \otimes \sum_{k=1}^{n} \delta_{pk} \otimes [W_{jk}]$.

(c) Equation for updating parameter $[W_{jk}]$

$$[W_{jk}](t+1) = [W_{jk}](t) \oplus \left(-\eta \otimes \frac{\partial [E_{pk}]}{\partial [W_{jk}]}\right) \oplus (\beta \otimes \Delta[W_{jk}](t-1))(t) \quad (25)$$

where $\frac{\partial [E_{pk}]}{\partial [W_{jk}]} = ([O_{pk}] \oplus (-[T_{pk}])) \otimes [O_{pk}] \otimes (1 \oplus (-[O_{pk}])) \otimes [O_{pj}]$.

(d) Equation for updating parameter $[\Theta_k]$

$$[\Theta_k](t+1) = [\Theta_k](t) \oplus \left(-\eta \otimes \frac{\partial [E_{pk}]}{\partial [\Theta_k]}\right) \oplus (\beta \otimes \Delta[\Theta_k](t-1))(t) \quad (26)$$

where $\frac{\partial [E_{pk}]}{\partial [\Theta_k]} = [O_{pk}] \otimes (1 \oplus (-[T_{pk}])) \otimes [O_{pk}] \otimes (1 \oplus (-[O_{pk}]))$.

However, there are refractory problems delineated as below:

1. The optimized result based on the back-propagation method might be trapped into local minima.
2. Parameter adjustments by derivative method can cause that for example left fuzzy part exceeds right fuzzy part and so on. Therefore, it is hard to let process of optimization computation generate RFV shapes that could fulfill the RFV constraints shown in Eqs. (19) and (20) at all time. Some examples are shown in Fig. 8. Therefore, the derivative method, which involves the mathematic of interval arithmetic, does not work.
3. The derivations and calculations of partial differentiation Eqs. (23)–(26) that are required in back-propagation algorithm based method are very complicated due to usage of RFVs.
4. The inequalities constraints Eqs. (21) and (22) are not easy to meet during optimization process.
5. There is no guarantee of convergence by the derivative method.
6. It requires proof that a derivative method (back-propagation algorithm) is suitable to optimize a function of RFVs.

Therefore, method based on Particle Swarm Optimization (PSO) is proposed to solve these difficulties, which is depicted in the next section.

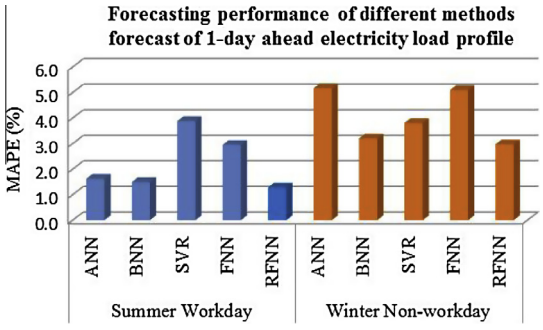

Fig. 10. MAPE performance comparison of different methods.

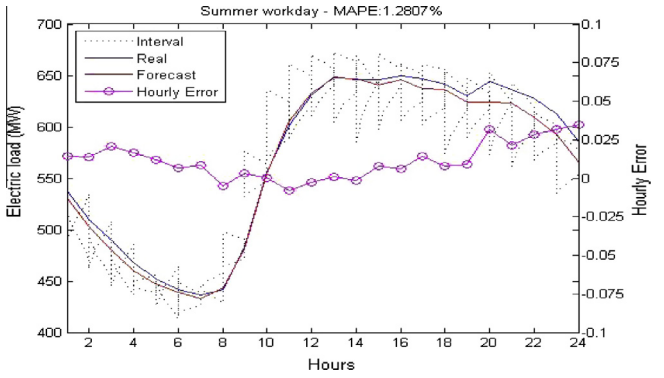

Fig. 11. Forecast of 1-day ahead summer workday load profile (Real: actual load values on 31st of August 2010; Forecast: forecasted values in point-values; Interval: forecasts in forms of RFV; Hourly Error: its hourly forecasting error).

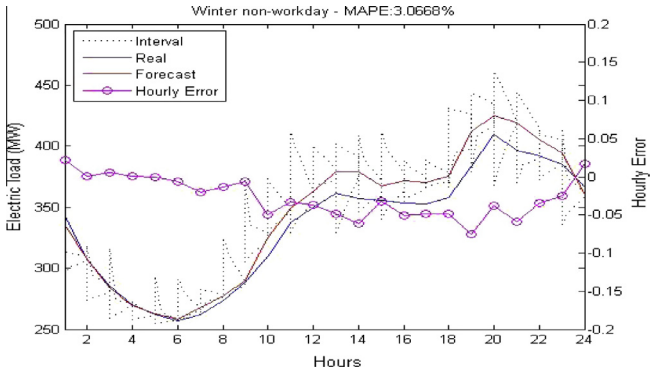

Fig. 12. Forecast of 1-day ahead winter non-workday load profile (Real: actual load values on 30th of January 2011; Forecast: forecasted values in point-values; Interval: forecasts in forms of RFV; Hourly Error: its hourly forecasting error).

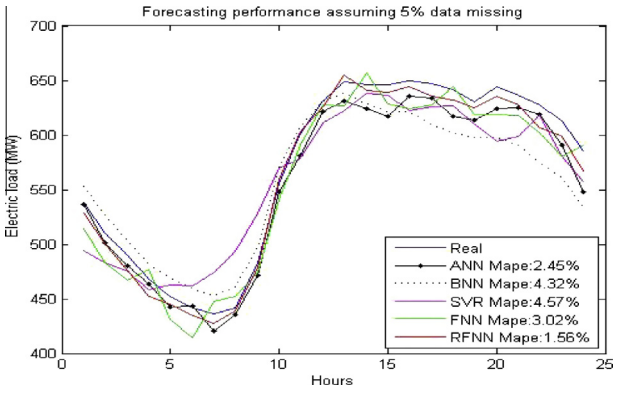

Fig. 13. Forecasting performance on one summer workday by assuming 5% data missing in historical data.

Propose PSO based learning algorithm

The PSO is a population based stochastic optimization technique inspired by social behavior of bird flocking or fish schooling. Its advantages in solving the above-mentioned difficulties are briefed as below:

1. PSO is a global optimization algorithm. It is most likely to locate global minima. Compared with Genetic Algorithms (GA), it has no evolution operators such as crossover and mutation. It has fewer parameters to adjust, so its implementation is not difficult. In addition, its computing cost is much less than that of GA.

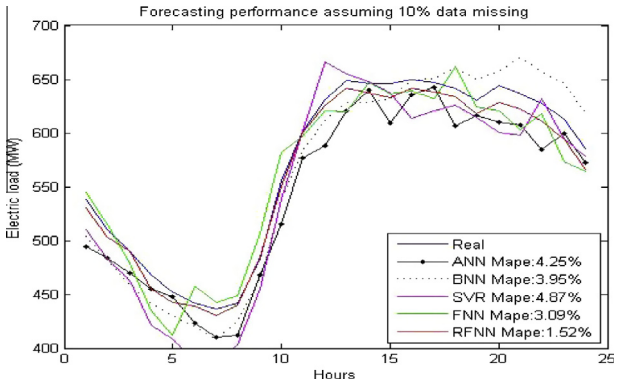

Fig. 14. Forecasting performance on one summer workday by assuming 10% data missing in historical data.

2. RFNN parameters (base points of respective network parameters $[W_{ij}]$, $[W_{jk}]$, $[\theta_j]$ and $[\theta_k]$) are encoded into chromosomes firstly. The whole optimization process is guided by the fitness function Eq. (18) which measures the "goodness" of the candidate solution (chromosome). Differentiability of error performance function that is required in back-propagation algorithm is not required. As a result, equations Eqs. (23)–(26) can be avoided.
3. By using our proposed encoding techniques, chromosomes of PSO can be decoded to generate the correct shapes of network parameters in forms of RFVs shown in Fig. 9, in which outer part represents random feature and inner part represents fuzzy feature. Since calculations of partial derivative are not needed, their corrected shapes can be maintained during optimization process. As a result, inequality constraints Eqs. (19)–(22) can be fulfilled easily at all time.

Implementation and application examples

The proposed technique has been implemented under Matlab programming environment. Its performance is validated using a real operational data collected from Macau electric utility. Macau power system is small. Characteristic of electric load is mainly residential and commercial (hotels, large-scale casino and entertainment complexes) load with minor industry. Macau has subtropical climate with temperatures regulated by its coastal location. It is observed from correlation analysis that temperature and humidity have strong correlation with power load, plus considering available data in observatory, so temperature (max., min. and avg.), humidity (avg.) and 24-hourly load values are selected as input variables. Besides, based on our practical experience on executing load forecasting for years, it is identified that wind speed and weather information have influences too. Therefore, those variables indicated in Table 1 are selected as input variables. The outputs are 24-hourly forecasted load values. As a result, RFNN topology is a 3-layer structure having 30-input variables and 24 outputs as is shown in Fig. 6.

Four other relevant methods were also used for comparison. They are ANN, BNN, FNN and Support Vector Regression (SVR), which are also very active researched techniques in recent years. In the ANN, BNN and SVR, all inputs/outputs and parameters are represented in crisp values. In the FNN system, all inputs/outputs, weights and biases parameters are represented by triangular fuzzy variables. Details of learning algorithms of FNN can be referred to [36–38]. However, in our RFNN system all of them are represented by RFVs in order to fully account for fuzzy and random uncertainties simultaneously. System performances are compared using Mean Absolute Percentage Error (MAPE) shown in Eq. (27).

$$MAPE(\%) = \frac{1}{D}\sum_{i=1}^{D}\sum_{h=1}^{24}\frac{y_a^{hi} - y_o^{hi}}{y_o^{hi}} \times 100\% \tag{27}$$

where y_a is actual hourly load, y_o is output from NN and D is number of forecasted days.

Forecast of one-day ahead electric load profile – data having intrinsic uncertainties

Firstly, test was carried out by using data with intrinsic uncertainties. Its aim is to verify forecasting ability of our proposed method over others. Moreover, to verify its adaptability to different load patterns, two scenarios were tested, namely summer workday and winter non-workday. For the purpose, firstly historical system hourly loads and weather data were collected. For summer season, data from 15th of June to 20th of August 2010 were used to train forecasting system according to our proposed method. Data from 21st of August to 30th of September 2010 were used for testing purpose. For winter season, data from 1st of November 2010 to 25th of January 2011 and data from 26th of January 2011 to 16th of February 2011 were used for training and testing purpose respectively. Secondly, training data were classified into 3 data groups representing 'Mon', 'Tue–Fri', & 'non-workday (Sat, Sun and Holiday)' categories respectively. Since input variables have different values, they are normalized. Then, they were converted into representations of RFV according to Section 'Propose applying RFVs in electric load forecasting' and fed to inputs of RFNN. The PSO was further used to optimize network parameters of the proposed RFNN. Finally, the RFNN was ready. Testing data were used to validate forecasting performance.

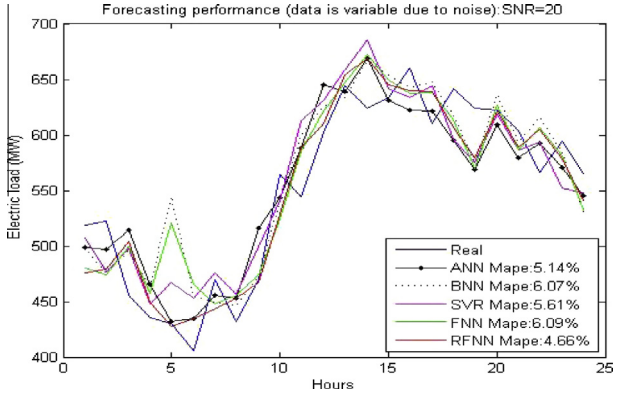

Fig. 15. Forecasting performance on load profile of one summer workday (Real: original data is added the noise of SNR = 20).

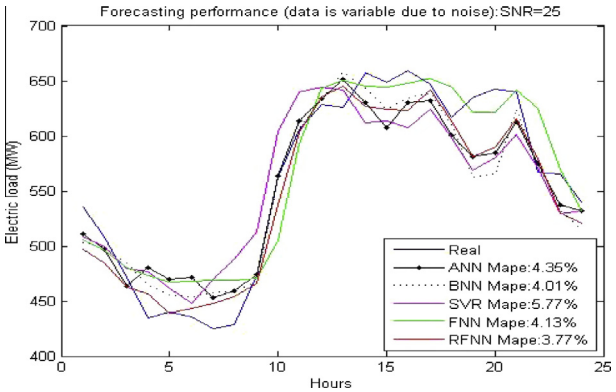

Fig. 16. Forecasting performance on load profile of one summer workday (Real: original data is added the noise of SNR = 20).

Table 2
Forecasting performance of RFNN and other methods.

	SNR = 20					SNR = 25				
	ANN (%)	BNN (%)	SVR (%)	FNN (%)	RFNN (%)	ANN (%)	BNN (%)	SVR (%)	FNN (%)	RFNN (%)
27 August (Fri)	4.56	4.44	4.52	4.09	3.95	4.35	4.01	5.77	4.13	3.77
31 August (Tue)	4.86	3.13	3.95	3.29	2.65	3.51	3.48	5.12	3.62	2.95
1 September (Wed)	4.96	3.94	5.02	2.98	3.12	2.81	3.24	3.27	3.91	2.53
2 September (Thur)	7.60	4.02	5.29	3.62	3.55	4.11	3.75	6.61	3.21	3.37
3 September (Fri)	5.26	3.81	3.29	4.09	3.98	3.19	3.41	3.36	4.50	3.65
7 September (Tue)	6.92	5.28	6.85	5.84	5.26	5.48	5.81	8.21	5.35	5.29
8 September (Wed)	5.64	4.33	3.76	3.09	2.71	3.82	2.33	5.04	3.21	3.37
9 September (Thur)	3.28	3.71	3.92	3.51	2.94	3.63	3.10	3.67	3.98	2.50
10 September (Fri)	4.85	3.21	4.24	2.72	2.78	3.79	3.58	4.16	3.05	2.95
14 September (Tue)	4.16	2.26	2.91	2.28	2.16	2.51	2.24	3.92	3.15	2.09
15 September (Wed)	3.80	5.91	5.73	9.25	5.21	5.95	6.54	7.23	9.43	5.12
Average	5.08	4.00	4.50	4.07	3.48	3.82	3.77	5.04	4.32	3.37

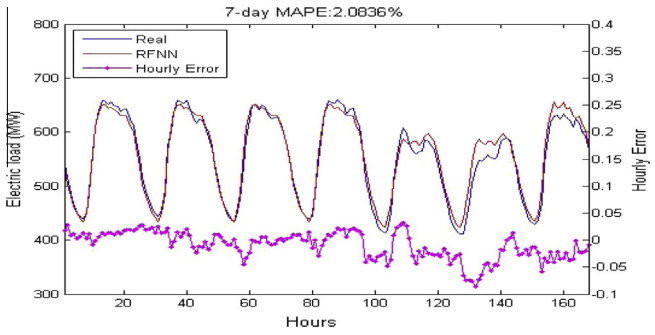

Fig. 17. Forecast of 7-day ahead load profile (Real: actual electric load values; RFNN: forecasted values by the proposed method; Hourly Error: its hourly forecasting error).

Fig. 10 summarizes forecasting performances of these methods. Figs. 11 and 12 demonstrate respectively one-day ahead load curves obtained by RFNN and real values recorded on 31st) of August 2010 and 30th of January 2011 by SCADA system and their respective hourly errors.

Forecast of day ahead electric load profile – containing missing data

The next test is to assume 5% and 10% missing data in historical data. Its aim is to verify its ability handling data incompleteness. The missing data were removed from training data randomly. Figs. 13 and 14 show the performance comparisons.

Forecast of day ahead electric load profile – add noise to make curve more variable

The test is to verify its forecasting ability while data becomes variable due to noise. Two scenarios were tested, i.e. white Gaussian noises with Signal Noise Ratio SNR = 25 (representing volatility) and SNR = 20 (representing more volatility) were added into the historical data. Figs. 15 and 16 show its performances compared with other relevant methods.

We further tested generalization ability of RFNN to forecast other workdays and compared it with other relevant methods. Table 2 shows comparison on their performances of each day and average values.

Forecast of 7-day ahead load profiles

Finally, RFNN is tested to forecast load profiles of one complete week (7-day) ahead, which include workday and non-workday. The forecast is crucial for power utility to evaluate and arrange maintenance of power network equipment and generators. For the purpose, three independent forecasting systems representing load profiles of 'Mon', 'Tue–Fri' and 'non-workday (Sat, Sun and Holidays)' respectively were built by our proposed method. Training data collected from past 2 months were used for 'Tue–Fri' while respective data from past 4 months were used for 'Mon' and 'non-workday'. The obtained forecasting systems produced 1-day (Mon), 4-day (Tue–Fri) and 2-day (non-workday) load forecasting respectively to form a 7-day load forecasting. Fig. 17 shows the forecasting result by RFNN and real values recorded from 24th of August to 30th of August 2010 by SCADA system and its respective hourly errors.

Main advantages of our proposed method can be summarized as below:

1. As is shown in the tests using data having intrinsic uncertainties (Figs. 10–12), forecasting accuracy is improved by RFNN which can model fuzzy and random uncertainties simultaneously and comprehensively.
2. As to system performance on tests using incomplete data, since ANN, BNN and SVR systems cannot handle data with missing values, they were firstly imputed by nearest-neighbor methods. For FNN and RFNN systems, missing data was represented by fuzzy variable and special representation of RFV (Figs. 5a and 5b) respectively. As such, testing results in Figs. 13 and 14 show that both FNN and RFNN have stable performance due to their fuzzy representation of missing data, compared with ANN, BNN and SVR.
3. As to historical data having more volatility and data contaminated by added noises, it is observed from Figs. 15 and 16 that RFNN can generate forecasting results superior to other relevant methods. In addition, Table 2 shows that RFNN obtains better average performances over forecasting 11 workdays than other methods. Therefore, its generalization ability is better.
4. The proposed RFNN taking into account various uncertainties has remarkably better forecasting performance than other relevant methods (see Figs. 11–17).
5. Figs. 11 and 12 also demonstrate the forecasting results in forms of RFV, which can provide additional useful information for power utility to carry out risk assessment and contingency management to take account of forecasting error.

Conclusion

ANN is strong to model non-linear mapping relation between inputs and outputs. Its drawback is the lack of incorporating uncertainties of data and network parameters due to crisp representations. BNN and FNN were developed respectively to incorporate respective probability and fuzziness of uncertainties to enhance

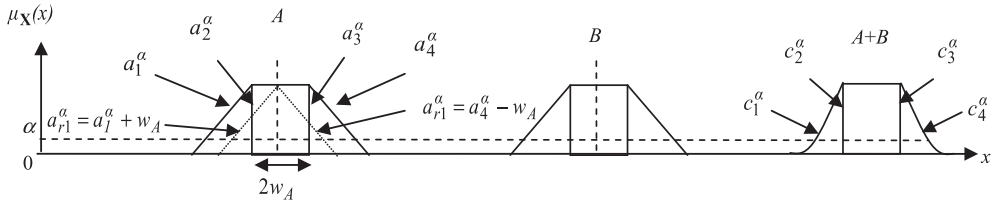

Fig. A1. Operation of two RFVs.

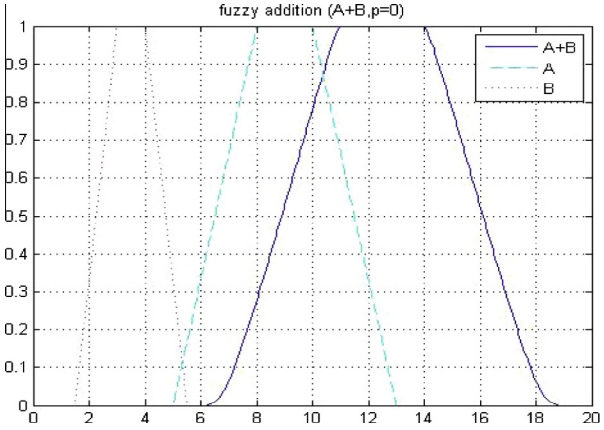

Fig. A2. 'Sum 'operation' of two RFVs.

their abilities such that they can be applied to real-world applications more effectively.

Towards solving the drawbacks, the paper proposes the RFVs, which can model fuzzy and random uncertainties simultaneously. Further integrating the RFVs with NN to form a novel RFNN is presented for load forecasting such that various uncertainties caused by fuzziness and randomness can be modeled and tackled simultaneously by one embedded hybrid technique. Moreover, special representations of RFVs can directly model missing and inaccurate data or incomplete and vague information. Data imputation and/or preprocessing are no longer needed as is required in traditional methods.

Testing results reveal that RFNN has produced in average good forecasting performance. In addition, it can better tackle more volatility on historical data in order to achieve good forecasting performance. Trade-off of the good performances by RFNN algorithm is longer computing time, 30 min (it was tested on a PC with Duo CPU T8100@2.10GHz. 100 generations and 120 population size was used in PSO), compared with 3 min for FNN. It is because mathematical operations on RFVs are required in each epoch in which much computing time is needed, compared with operation of fuzzy variables. However, it is worthwhile that better forecasting performance in average is produced despite of increased complexities of modeling and time of calculation. Future research would be carried out to reduce the computing time. In addition, outputs in forms of RFV give more critical information for power utility to carry out risk analysis because power utility people can evaluate and arrange sufficient generation based on this information to fulfill the load demand due to the possible forecasting error (positive or negative).

Finally, it is rather promising in microgrid/small size of power system or in forecasting curves of individual customer where load curves would present a much higher variability and noise than curves of global power system of country/region.

Acknowledgements

Supported in part by the Research Committee of University of Macau under Grant No. MYRG2014-00060-FST, and in part by the Science and Technology Development Fund (FDCT) of Macau under Grant No. 016/2012/A1, respectively.

Appendix A. Mathematical operations of two RFVs

Assume that $[a_1^\alpha \quad a_4^\alpha]$ & $[b_1^\alpha \quad b_4^\alpha]$ and $[a_2^\alpha \quad a_3^\alpha]$ & $[b_2^\alpha \quad b_3^\alpha]$ are α-cuts of external and internal MFs on two initial RFVs, i.e. A and B as shown in Fig. A1. $[c_1^\alpha \quad c_2^\alpha \quad c_3^\alpha \quad c_4^\alpha]$ denotes α-cuts of final result. Hence, the following definitions are given for four mathematical operations, i.e. sum, difference, product and division.

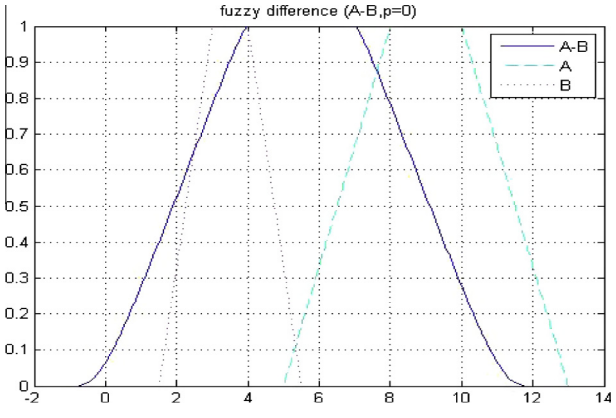

Fig. A3. 'Difference' operation of two RFVs.

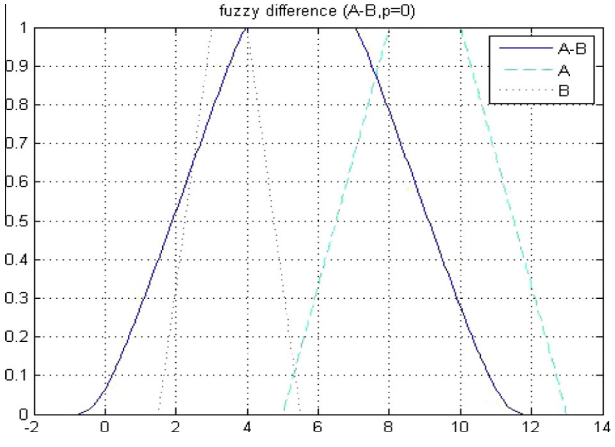

Fig. A4. 'Product' operation of two RFVs.

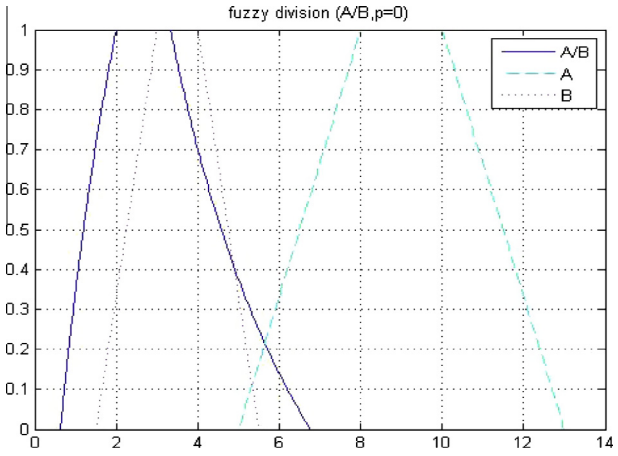

Fig. A5. 'Division' operation of two RFVs.

A1. Sum

(see Fig. A2)

$$c_1^\alpha = c_2^\alpha - \mu_c + K \cdot ext\left(a_{r1}^\alpha + b_{r1}^\alpha, g_1^\alpha\right) + (1-K) \cdot in\left(a_{r1}^\alpha + b_{r1}^\alpha, g_1^\alpha\right) \quad (A1)$$

$$c_2^\alpha = a_2^\alpha + b_2^\alpha \quad (A2)$$

$$c_3^\alpha = a_3^\alpha + b_3^\alpha \quad (A3)$$

$$c_4^\alpha = c_3^\alpha - \mu_c + K \cdot ext\left(a_{r4}^\alpha + b_{r4}^\alpha, g_4^\alpha\right) + (1-K) \cdot in\left(a_{r4}^\alpha + b_{r4}^\alpha, g_4^\alpha\right) \quad (A4)$$

where

$a_{r1}^\alpha = a_1^\alpha + w_A$ and $a_{r4}^\alpha = a_4^\alpha - w_A$.

$b_{r1}^\alpha = b_1^\alpha + w_B$ and $b_{r4}^\alpha = b_4^\alpha - w_B$.

$w_A = 0.5\left(a_3^{\alpha=1} - a_2^{\alpha=1}\right)$ and $w_B = 0.5\left(b_3^{\alpha=1} - b_2^{\alpha=1}\right)$.

$K = 1/\sqrt{2}$.

$\left|g_1^\alpha, g_4^\alpha\right|$ is α-cuts of normal possibility distribution with mean μ_c and standard deviation σ.

$\mu_c = 0.5\left(c_2^{\alpha=1} + c_3^{\alpha=1}\right)$.

$$\sigma = \frac{1}{3}\min\left\{\mu_c - a_{r1}^{\alpha=0} - b_{r1}^{\alpha=0}, a_{r4}^{\alpha=0} + b_{r4}^{\alpha=0} - \mu_c\right\}.$$

"in" and "ext" are functions that provide the value belongings respectively to internal and external MFs.

A2. Difference

See Fig. A3.

$$c_1^\alpha = c_2^\alpha - \mu_c + K \cdot ext\left(a_{r1}^\alpha - b_{r4}^\alpha, g_1^\alpha\right) + (1-K) \cdot in\left(a_{r1}^\alpha - b_{r4}^\alpha, g_1^\alpha\right) \quad (A5)$$

$$c_2^\alpha = a_2^\alpha - b_3^\alpha \quad (A6)$$

$$c_3^\alpha = a_3^\alpha - b_2^\alpha \quad (A7)$$

$$c_4^\alpha = c_3^\alpha - \mu_c + K \cdot ext\left(a_{r4}^\alpha - b_{r1}^\alpha, g_4^\alpha\right) + (1-K) \cdot in\left(a_{r4}^\alpha - b_{r1}^\alpha, g_4^\alpha\right) \quad (A8)$$

A3. Product

See Fig. A4.

$$c_1^\alpha = c_2^\alpha - \mu_r + c_{r1}^\alpha \quad (A9)$$

$$c_2^\alpha = \min\left\{a_2^\alpha b_2^\alpha, a_2^\alpha b_3^\alpha, a_3^\alpha b_2^\alpha, a_3^\alpha b_3^\alpha\right\} \quad (A10)$$

$$c_2^\alpha = \max\left\{a_2^\alpha b_2^\alpha, a_2^\alpha b_3^\alpha, a_3^\alpha b_2^\alpha, a_3^\alpha b_3^\alpha\right\} \quad (A11)$$

$$(c_4^\alpha = c_3^\alpha - \mu_r + c_{r4}^\alpha \quad (A12)$$

where

$c_{r1}^\alpha = \min\left\{a_{r1}^\alpha b_{r1}^\alpha, a_{r1}^\alpha b_{r4}^\alpha, a_{r4}^\alpha b_{r1}^\alpha, a_{r4}^\alpha b_{r4}^\alpha\right\}$.

$c_{r4}^\alpha = \max\left\{a_{r1}^\alpha b_{r1}^\alpha, a_{r1}^\alpha b_{r4}^\alpha, a_{r4}^\alpha b_{r1}^\alpha, a_{r4}^\alpha b_{r4}^\alpha\right\}$.

$\mu_r = 0.5\left(c_{r1}^{\alpha=1} + c_{r4}^{\alpha=1}\right)$.

A4. Division

See Fig. A5.

$$c_1^\alpha = c_2^\alpha - \mu_r + c_{r1}^\alpha \quad (A13)$$

$$c_2^\alpha = \min\left\{a_2^\alpha / b_2^\alpha, a_2^\alpha / b_3^\alpha, a_3^\alpha / b_2^\alpha, a_3^\alpha / b_3^\alpha\right\} \quad (A14)$$

$$c_2^\alpha = \max\left\{a_2^\alpha / b_2^\alpha, a_2^\alpha / b_3^\alpha, a_3^\alpha / b_2^\alpha, a_3^\alpha / b_3^\alpha\right\} \quad (A15)$$

$$c_4^\alpha = c_3^\alpha - \mu_r + c_{r4}^\alpha \quad (A16)$$

where

$c_{r1}^\alpha = \min\left\{a_{r1}^\alpha / b_{r1}^\alpha, a_{r1}^\alpha / b_{r4}^\alpha, a_{r4}^\alpha / b_{r1}^\alpha, a_{r4}^\alpha / b_{r4}^\alpha\right\}$.

$c_{r1}^\alpha = \max\left\{a_{r1}^\alpha / b_{r1}^\alpha, a_{r1}^\alpha / b_{r4}^\alpha, a_{r4}^\alpha / b_{r1}^\alpha, a_{r4}^\alpha / b_{r4}^\alpha\right\}$.

References

[1] Haida T, Muto S. Regression based peak load forecasting using a transformation technique. IEEE Trans Power Syst 1994;9(4):1788–94.
[2] J Huang S, Shih KR. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. IEEE Trans Power Syst 2003;18(2):673–9.
[3] Christiannse WR. Short term load forecasting using general exponential smoothing. IEEE Trans Power Apparatus Syst 1971;90(2):900–10.
[4] Stanton KN. Forecasting by probability methods. IEEE Trans Power Apparatus Syst 1971;90:1183–9.
[5] Uri ND, S Maybee J. Forecasting the load duration curve using Box–Jenkins time series analysis. Eng Optim 1978;3:193–9.
[6] Kandil N, Wamkeue R, Saad M, Georges S. An efficient approach for short term load forecasting using artificial neural networks. Electr Power Energy Syst 2006;28:525–30.
[7] Mori H, Kobayashi H. Optimal fuzzy inference for short-term load forecasting. IEEE Trans Power Syst 1996;11(1):390–6.
[8] Al-Kandari AM, Soliman SA, El-Hawary ME. Fuzzy short-term electric load forecasting. Electr Power Energy Syst 2004;26:111–22.
[9] Alamaniotis M, Ikonomopoulos A, Tsoukalas LH. Evolutionary multiobjective optimization of kernel-based very-short-term load forecasting. IEEE Trans Power Syst 2012;27(3):1477–84.
[10] Elattar EE, Goulermas JY. Generalized locally weighted GMDH for short term load forecasting. IEEE Trans Syst Man Cybern C Appl Rev 2012;42(3):345–6.
[11] Chen T, Wang YC. Long-term load forecasting by a collaborative fuzzy-neural approach. Electr Power Energy Syst 2012;43:454–64.
[12] Akdemir B, Cetinkaya N. Long-term load forecasting based on adaptive neural fuzzy inference system using real energy data. Energy Proc 2012;14:794–9.
[13] Jin M, Zhou X, Zhang ZM, Tentzeris MM. Short-term power load forecasting using grey correlation contest modeling. Expert Syst Appl 2012;39:773–9.
[14] Masataro O, Hiroyuki M. A Gaussian processes technique for short-term load forecasting with consideration of uncertainty. IEEJ Trans Power Energy 2006;126(2):202–8.
[15] Xiao Z, Ye SJ, Zhong B, Sun CX. BP neural network with rough set for short term load forecasting. Expert Syst Appl 2009;36:273–9.
[16] Hippert HS, Pedreira CE, Souza RC. Neural networks for short-term load forecasting: a review and evaluation. IEEE Trans Power Syst 2001;16(1):44–54.
[17] Mandal P, Senjyu T, Funabashi T. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. Energy Convers Manag 2006;47:2128–42.
[18] Lauret P, Fock E, Randrianarivony RN. Bayesian neural network approach to short time load forecasting. Energy Convers Manag 2008;49:1156–66.
[19] Niu DX, Shi HF, Wu DD. Short-term load forecasting using Bayesian neural networks learned by hybrid Monto Carlo algorithm. Appl Soft Comput 2012;12:1822–7.
[20] Liao GC, Tsao TP. Application of fuzzy neural networks and artificial intelligence for load forecasting. Electr Power Syst Res 2004;70:237–44.
[21] Chang PC, Fan CY, Lin JJ. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. Electr Power Energy Syst 2011;33:17–27.
[22] Chen T, Wang YC. Long-term load forecasting by a collaborative fuzzy-neural approach. Electr Power Energy Syst 2012;43:454–64.
[23] Chen T. A collaborative fuzzy-neural approach for long-term load forecasting in Taiwan. Comput Ind Eng 2012;63:663–70.

[24] H Kim S, Yang HJ, Ng KS. Incremental expectation maximization principle component analysis for missing value imputation for coevolving EEG data. J Zhejiang Univ: Sci 2011;12(8):687–97.
[25] Li D, Deogun J, Spaulding W, Shuart B. Dealing with missing data: algorithms based on fuzzy set and rough set theories. Springer-Verlag; 2005.
[26] Liao Z, Lu XJ, Yang T, Wang HG. Missing data imputation: a fuzzy *k*-means clustering algorithm over sliding window. Fuzzy Syst Knowled Discov 2009;14–16:133–7.
[27] Aydilek IB, Arslan A. A novel hybrid approach to estimating missing values in databases using *k*-nearest neighbors and neural networks. Int J Innov Comput I 2012:4705–17.
[28] Nelwamondo FV, Golding D, Marwala T. A dynamic programming approach to missing data estimation using neural networks. Inf Sci 2009.
[29] Ferrero A, Salicone S. The random-fuzzy variables; a new approach for the expression of uncertainty in measurement. In: IMTC 2003 – Instrumentation and measurement technology conference Vail, CO, USA, 20–22 May 2003.
[30] Ferrero A, Gamba R, Salicone S. A method based on random-fuzzy variables for online estimation of the measurement uncertainty of DSP-based instruments. IEEE Trans Instrum Meas 2004;53(5):1362–9.
[31] Ferrero A, Salicone S. Modeling and processing measurement uncertainty within the theory of evidence: mathematics of random-fuzzy variables. IEEE Trans Instrum Meas 2007;56(3):704–16.
[32] Salicone S. Measurement uncertainty. An approach via the mathematical theory of evidence. New York: Springer-Verlag; 2006.
[33] Geer JF, Klir GJ. A mathematic analysis of information preserving transformations between probabilistic and possibilistic formulations of uncertainty. Int J General Syst 1992;20:143–76.
[34] Geer JF, Klir GJ. Probability–possibility transformations: a comparison. Int J General Syst 1992;21:291–310.
[35] Klir GJ, Yuan B. Fuzzy sets and fuzzy logic. Theory and applications. Upper Saddle River (NJ): Prentice-Hall; 1995.
[36] Ishibuchi H, Morioka K, Turksen IB. Learning by fuzzified neural networks. Int J Approx Reason 1995;13:327–58.
[37] Feuring T, Buckley JJ, Hayashi Y. Adjusting fuzzy weights in fuzzy neural nets. In: 1998 2nd Int. conf. on knowledge-based intelligent electronic systems, 21–23 April 1998, Adelaide, Australia.
[38] Li Z, Kecman V, Ichikawa A. Fuzzified neural network based on fuzzy number operations. Fuzzy Sets Syst 2002;130:291–304.