A fuzzy theory-based machine learning method for workdays and weekends short-term load forecasting

Li, C.

Energy & Buildings 245 (2021) 111072

Contents lists available at ScienceDirect

Energy & Buildings

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

A fuzzy theory-based machine learning method for workdays and weekends short-term load forecasting

Chen Li

Institute of Environmental Sciences (CML), Leiden University, P.O. Box 9518, 2300 RA Leiden, the Netherlands

ARTICLE INFO

Article history:
Received 9 April 2020
Revised 20 April 2021
Accepted 27 April 2021
Available online 4 May 2021

Keywords:
Short-term load forecasting
Workdays and weekends
Fuzzy time series
Multi-objective optimization

ABSTRACT

Countries around the globe have introduced renewable energies (RE) and minimized the dependency of fossil resources in power systems to address extensive environmental risks. However, such large-scale energy transitions pose a great challenge to power systems due to the volatility of RE. Meanwhile, power demand is increasing over time and it shows temporal characteristics, such as seasonal and peak-valley patterns. Whether the future power system with a larger proportion of RE can meet the surging but fluctuated electricity demand remains problematic. Previous studies on short-term load forecasting focused more on forecasting accuracy than stability. Further, there is a relative paucity of research into temporal patterns. In order to fill in these research gaps, this paper proposes a fuzzy theory-based machine learning model for workdays and weekends short-term load forecasting. Fuzzy time series (FTS) is applied for data mining and back propagation (BP) neural network is used as the main predictor for short-term load forecasting. To exploit the trade-offs between forecasting stability and accuracy, multi-objective optimization is applied to modify the parameters of BP. Moreover, an interval forecasting architecture with several statistical tests is constructed to address forecasting uncertainties. Short-term load data from Victoria in Australia is selected as a case study. Results demonstrate that the proposed method can significantly boost forecasting stability and accuracy, and help strategy making in the field of energy and electricity system management and planning.

1. Introduction

Transitions on power systems are happening as countries around the globe are introducing RE and minimizing the dependency of fossil resources. Integration of large-scale RE such as wind and solar energy into the electricity grids has been increasing. However, this integration poses a great challenge that hampers the stable operation of the electricity systems as RE's unstable power output. Meanwhile, electricity demand is prone to be highly temporal: it varies between workdays and weekends, also shows seasonal and festival variation.

Research on short-term load forecasting has been conducted for decades. In the early stage, classic arithmetic has been widely deployed with its simple mathematical principles and assumptions. Classic arithmetic is mainly based on statistical models, such as regression-based models [1,2], Box-Jenkins models [3,4] and Bayesian models [5]. However, this arithmetic is highly dependent on the quantity of historic data and strict statistical assumptions. Also, it cannot achieve high forecasting accuracy when dealing with non-linear time series.

Owing to the shortcomings of classic arithmetic, intelligent arithmetic has been developed that comprises of artificial neural networks (ANNs) [6,7], ANNs simulate the human brain and can yield satisfactory training results when dealing with multi-structural and non-linear problems. In the year of 1991, Park et. al [8] used ANN for power forecasting for the first time, which proved the good forecasting performance of ANN. Based on this pioneer work, various types of ANNs have been developed and used to short-term load time series forecasting [9,10,11]. However, intelligent arithmetic has some limitations: a) they are easy to fall into local optima due to their slow self-learning convergence rates [12]; b) it is difficult to determine parameters such as layer and neuron numbers in a network structure [13].

Given the limitations of intelligent arithmetic, hybrid arithmetic has developed in these years and gradually become the mainstream in short-term forecasting fields [14]. The basic principle of hybrid arithmetic is to integrate the outputs of different individual models in terms of utilizing certain weights and narrowing the value ranges [15]. To improve the extrapolation ability and reduce the learning time of ANNs, ANNs have been combined with some methods [16,17]. For example, Azimi et al. [18] developed a hybrid short-term load forecasting model based on ANNs and

E-mail address: c.li@cml.leidenuniv.nl

https://doi.org/10.1016/j.enbuild.2021.111072
0378-7788/

autoregressive moving average (ARMA). Lu et al. [19] integrated a growing hierarchical self-organizing map (SOM) with support vector machines (SVMs); Okumus et al. [20] proposed an adaptive neuro-fuzzy inference system that combines of ANNs. Elvira et al. [21] used several prediction models to predict the electricity demand in the southeastern region of Oklahoma. The choice of a combination of models depends largely on the characteristics of the research problem and on error evaluation. To optimize the parameters of ANNs, heuristic algorithms are employed to integrate with ANNs. For example, enlightened by the biological evolution, Pandian et al. [22] and Pai et al. [23] proposed a set of optimization algorithms to combine with ANNs for electricity forecasting. Goudarzi et al. developed a hybrid model optimized by the particle swarm optimization (PSO) for the optimal configuration of building-wide energy dissemination policies [24]. Moreover, differential evolution optimization is used to combine with the SVMs for half-hourly and daily electricity consumption [25].

Hybrid arithmetic significantly improves short-term load forecasting accuracy but stability is too large extent missing. Also, there is a relative paucity of research into dynamicity and volatility of short-term electricity in terms of multiply temporal patterns. Important temporal characteristics, such as seasonal and weekday-weekend patterns are also not well addressed in current model settings. To fill in these gaps, this paper proposes a hybrid short-term load forecasting model based on data de-noising, the fuzzy time series (FTS) and ANNs with multi-objective optimization. Workdays and weekends short-term load data in four seasons from Victoria in Australia is chosen as a case study. The proposed model is concluded as following steps: 1) The original short-term load data are filtered by using a de-noising method; 2) A modified FTS model based on fuzzy sets further mines hidden features of the pre-processed data; 3) the fuzzified data is imported to BP neural network; 4) the parameters of BP are optimized by a multi-objective optimization algorithm and 5) the forecasted results are exported by BP and finally generated by defuzzification (See Fig. 1). This paper provides separate forecasting results in regard to seasonal and workday-weekend patterns. An interval forecasting is also used to provide the possible intervals of forecasting results. Moreover, multiply tests including algorithm tests, statistical tests and error measurement metrics are employed to validate the proposed model.

The novelties of the proposed model are summarized as follows:

- A data cleaning is conducted to eliminate the noise and further mine hidden characteristics through data de-nosing methods and the FTS, respectively.
- To boost the forecasting stability and accuracy simultaneously, a multi-objective optimization algorithm is used to optimize ANNs.
- Seasonal and workday-weekend temporal patterns are considered. Separate short-term load forecasting, varying in seasons, in workdays and weekends, is provided.
- This paper is organized as follows: The methodology is introduced in Section 2. Section 3 describes the datasets and forecasting results. Section 4 shows the discussion and Section 5 presents the conclusion.

2. Methodology

2.1. Data de-noising method—ICEEMDAN

The improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) is considered as one of state-of-the-art data de-noising methods. This paper employs the ICEEMDAN to mitigate noise and chaos of the original short-term load time series. The ICEEMDAN decomposes the original signal into several intrinsic mode functions (IMFs) [26,27] and overcomes cumulative noise neutralization and residual noise modification problems [28]. The ICEEMDAN has been successfully applied for wind speed forecasting [29], air quality early-warning [30], and electricity price forecasting [31].

2.2. Further mine hidden characteristics- fuzzy time series (FTS)

The FTS is applied for short-term forecasting with its capability of dealing with non-linear problems. It has been successfully used for predicting nonlinear and dynamic datasets in various fields, such as offshore wind energy [32], electricity consumption [33], the stock market [34], subject enrollment [35], and environmental materials [36]. In this paper, the FTS is modified with fuzzy sets and employed to fuzzify the preprocessed short-term load time series. The fuzzified short-term load series later is then imported into input layers of ANNs. The final forecasted results are obtained by defuzzification. The basic theory of the FTS is clarified in Appendix C.

2.3. Back propagation (BP) neural network

The BP is one of the commonly used ANNs [37] and it can learn and store a number of input–output map relations through feed-forward back propagation. Here the BP is selected as the main predictor for short-term load forecasting. The preprocessed and fuzzified load data is fed into BP through the input layer, which communicates to hidden layers. And the output of the output layer is considered as forecasting values. The mechanism of the BP relies on the steepest descent approach, in which the neural network continuously amends the weights and thresholds until it reaches the minimized sum of squared errors [38].

2.4. Multi-objective optimization algorithm

The dragonfly algorithm (DA) is one of the novel swarm intelligence algorithms. To address multi-objective optimization problems, Seyedali et al. developed a multi-objective version of DA, e.g., multi-objective dragonfly algorithm (MODA) [39]. The detailed description of the DA and MODA models can be found in [39,40] and the pseudo-code of the MODA is shown in Appendix D. In this paper, the parameters (e.g. the weights and thresholds) of the BP are optimized by the MODA to achieve higher forecasting accuracy and stability simultaneously. In doing so, this paper uses Pareto optimal strategy to determine the balance of two objects [41]. An introduction of Pareto optimal strategy can be found in Appendix E.

2.5. The hybrid forecasting model

A hybrid model that comprises of several components is proposed in this paper to achieve reliable short-term load forecasting results. To eliminate negative noise and extract inner characteristics of the raw short-term load, a data de-noising method ICEEMDAN and the FTS are performed. The preprocessed and fuzzified data is then imported into the BP. To achieve high forecasting stability and accuracy simultaneously, the MODA is applied to optimize the parameters of the BP. Based on the bias-variance system [42], the fitness function of the MODA has denoted both accuracy and stability. The calculation of the bias-variance framework is expressed as Eqn 1.

$$E(\hat{x}-x)^2 = E[\hat{x}-E(\hat{x})+E(\hat{x})-E(x)]^2 \\ = E[\hat{x}-E(\hat{x})]^2 + [E(\hat{x})-E(x)]^2 = Var(\hat{x}) + Bias^2(\hat{x}) \quad (1)$$

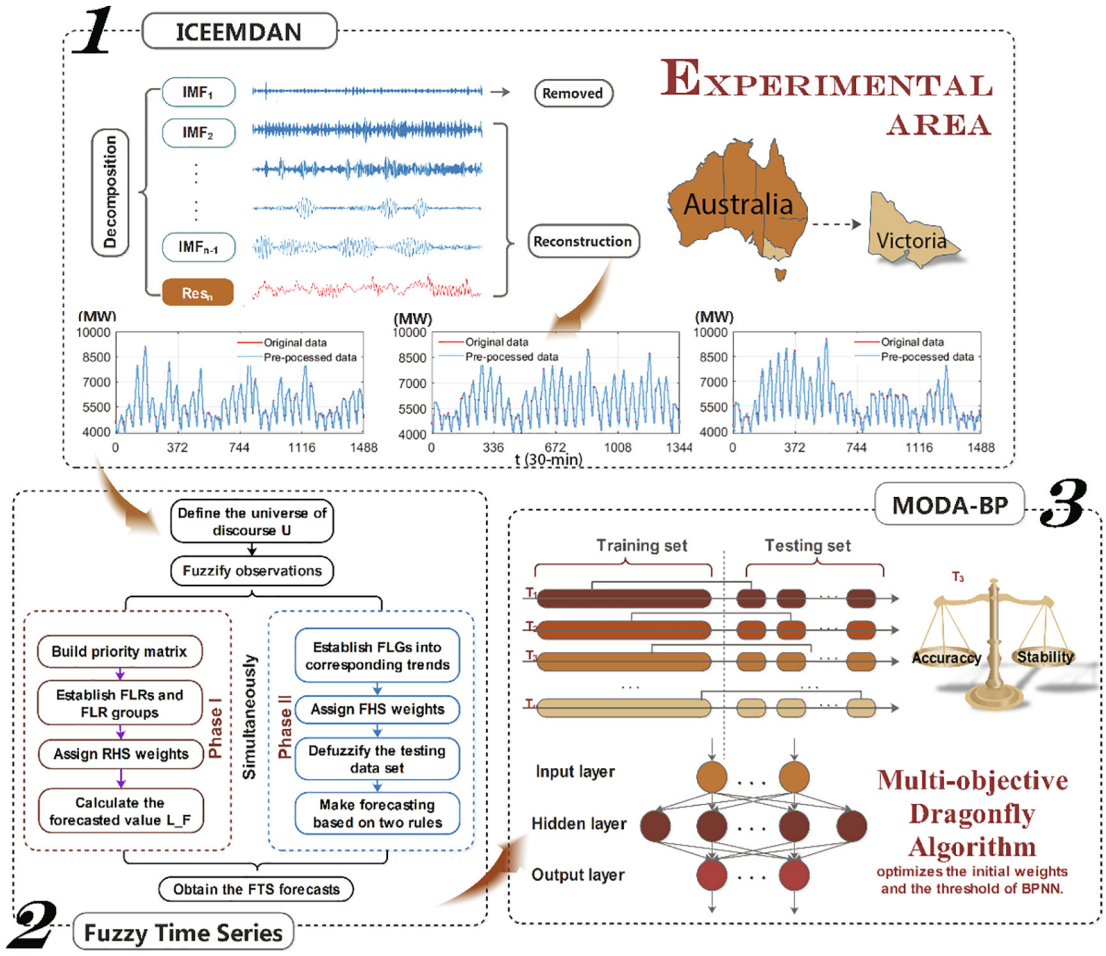

Fig 1. Flowchart of the proposed hybrid forecasting model.

where E represents the mathematical expectation, x and $\hat{x}$ denote the actual value and the forecasted value, respectively. The bias measures the average difference between the actual and forecasted values, and the variance represents the forecasting volatility. In this paper, the fitness function for accuracy and stability is designed as Eqn 2:

$$\min\begin{cases} fitness_a = |Bias(\hat{x})| \\ fitness_b = Std(x-\hat{x}) \end{cases} \quad (2)$$

where the first objective is the absolute bias and the second objective is the standard deviation. The final forecasted results are exported by BP and finally generated by defuzzification.

2.6. Interval forecasting theory

To provide the lower and upper intervals of forecasting results, an interval forecasting (IF) based on point forecasting is used in this paper. In the IF, predictive range and confidence level of forecasting are provided [43,44]. With a significance level α, the probability formula with the interval limits ($\boldsymbol{I}_{min}$ and $\boldsymbol{I}_{max}$) and the observed value $\boldsymbol{Y}_t$ can be calculated as follows:

$$P(I_{min} \leqslant Y_t \leqslant I_{max}) = 1 - 2\alpha \quad (3)$$

Lognormal distribution is considered as a desirable distribution to fit the input short-term load data. In this paper, the distribution of the output data is assumed the same of the input data. To do the IF, the output layer of the BP is set to two to indicate the minimum and maximum of intervals.

3. Experimental setting and results

3.1. Database

Short-term load data (30-min) from the Victoria in Australia is used as a case study in this paper. To consider seasonal and weekday-weekend patterns, the selected database is divided into multiply subsets, i.e., quarter I, quarter II, quarter III, quarter IV, workday and non-workday. A descriptive statistic of these subsets is shown in Table 1, where the train-test ratio of each dataset is set to 4:1 by the 'trial and error' rule [45,46]. Unsurprisingly, workdays have more electricity consumption than non-workdays. Among four quarters, quarter I has larger variety but less complexity than the other three quarters.

3.2. Experiment setup

The simulation in this paper is performed on Windows 8 with a 64-bit 3.30 GHz Intel Core i5 4590HQ CPU and 8 GB of RAM on the platform of MATLAB R2018a environment. To validate the proposed short-term load forecasting model, two experiments are carried out, namely Experiment I: point forecasting on workdays and non-workdays, and Experiment II: interval forecasting on workdays and non-workdays.

The proposed forecasting model is named as ICE-FTS-MODA-BP in this paper. To testify the effectiveness and superiority of the ICE-FTS-MODA-BP, four types of comparisons, i.e., comparison to Naïve forecasting models, traditional statistic models, artificial

Table 1
Descriptive statistical analysis of datasets.

Quarter	Date	Dataset	Length	Min	Max	Std.	Complexity
Quarter I	Workday	All samples	3072	3865.07	9587.51	1110.914	0.2849
		Training set	2458	3865.07	9138.75	1084.544	0.2991
		Testing set	614	4028.46	9587.51	1156.986	0.2882
	Non-workday	All samples	1248	3833.48	7843.03	782.3815	0.4258
		Training set	998	3874.07	7843.03	820.4033	0.3783
		Testing set	250	3833.48	5553.79	391.4460	0.6051
Quarter II	Workday	All samples	3120	3979.31	7813.35	820.1743	0.3132
		Training set	2496	3979.31	7481.25	767.2383	0.3500
		Testing set	624	4310.93	7813.35	846.7359	0.3677
	Non-workday	All samples	1248	3839.88	6860.55	590.1071	0.4913
		Training set	998	3839.88	6548.54	562.3212	0.4606
		Testing set	250	4396.26	6860.55	581.6237	0.5835
Quarter III	Workday	All samples	3168	3848.86	7699.93	806.7898	0.3718
		Training set	2534	4007.27	7699.93	813.6822	0.3859
		Testing set	634	3848.86	6856.18	686.4078	0.4173
	Non-workday	All samples	1248	3705.93	6857.28	603.1894	0.4304
		Training set	998	3705.93	6857.28	614.8645	0.4167
		Testing set	250	3943.52	5855.58	437.6508	0.6699
Quarter IV	Workday	All samples	3168	3748.16	9007.52	742.0011	0.3495
		Training set	2534	3798.98	7408.14	652.4790	0.3919
		Testing set	634	3748.16	9007.52	878.8699	0.3627
	Non-workday	All samples	1248	3551.60	5926.84	437.4109	0.4632
		Training set	998	3587.77	5629.87	416.6170	0.5099
		Testing set	250	3551.60	5926.84	530.1535	0.6267

intelligence (AI) models, and different optimization algorithm-based models are taken into consideration. For traditional statistic models, the typical time series model AR and ARIMA models, and the FTS are implemented as the benchmarks; for AI models, the BPNN and ELMAN model are selected; For different decomposition approach-based models, the EEMD method is chosen, which derives from EMD family and has developed the new form—ICEEMDAN; and the PSO and the MODA algorithms are employed to optimize the BPNN model (i.e., PSO-BP, and MODA-BP), which are selected as different optimization algorithm based models. To guarantee the experimental fairness, the neuron numbers of each ANN model are all supposed to set to optimal ones. As there is no uniform theory to determine the exact best neuron number of ANNs, this paper also applied the trial-and-error [47,48] for parameter settings of ANNs. In doing so, the parameters of ANNs are determined by a number of experiments. Each AI model was repeated 10 times to assure the reliability and independency in ultimate results, especially about the initial random weight values and the optimization algorithm. The experimental parameters employed in this paper are shown in Appendix H.

3.3. Experiment I: point forecasting on workdays and non-workdays

Experiment I is aimed at short-term load point forecasting on workdays and non-workdays. To comprehensively assess the accuracy and stability of the proposed forecasting model, this paper employs multiple error criteria (shown in Appendix F). Tables 2 and 3 show the forecasting results and the values in bold indicate the best values for each criterion among all the benchmarks. Forecasting on workdays shows poorer performance than that of on non-workdays in terms of both accuracy and stability. Interestingly, according to most criteria the forecasting performance on quarter I is better than the other three quarters. Fig. 2 shows the results of the first day in each testing sample, which are selected as examples to reveal the superiority of the proposed forecasting model in terms of various criterion. The proposed hybrid model possesses the smallest values in terms of the majority of criteria and a detailed analysis is presented as follows:

(a) Data preprocessing makes large contributions to enhance the forecasting performance. Compared the hybrid models based on the ICEEMDAN or the FTS to other individual models, it is clearly shown that the data preprocessing-based models significantly outperform individual models. Further, the ICEEMDAN-FTS data pre-processing is also shown better performance than the simple ICEEMDAN or FTS. Tables 2 and 3 show the forecasting results of the FTS, ICEEMDAN-FTS, MODA-BP and ICE-FTS-MODA-BP model. According to the listed results, it is concluded that the proposed ICE-FTS-MODA-BP model outperforms three other models in various forecasting horizons.

(b) Compared the BP to classical statistical model (e.g., ARIMA) and AI model (e.g., ELMAN), the results show that the BP is superior to the other two models in terms of forecasting accuracy and stability. For example, on workdays, the BP achieves the least MAPE values of 1.4416%, 1.4374%, 1.5562% and 1.5844% in four quarters, respectively. However, the ELMAN has comparatively larger MAPE values of 1.5743%, 1.5494%, 1.5806% and 1.9741% in four quarters, respectively. The ARIMA performs well in the second quarter but shows large bias in other quarters. In terms of forecasting stability, the proposed model achieves better results than individual models regarding to DA, FB and TIC values in most situations.

(c) Compared the proposed model ICE-FTS-MODA-BP to other hybrid models, it is found that the proposed model has a positive influence on improving the forecasting accuracy and stability. For example, in comparing the proposed model with the ICE-FTS-MODA-BP and EEMD-MODA-BP, the original EEMD can also enhance the forecasting performance but the contribution to forecasting accuracy is comparatively limited. The forecasting stability is also further improved when in comparison with simple hybrid models, e.g., PSO-BP and MODA-BP.

3.4. Experiment II: Interval forecasting on workdays and non-workdays

In this section, an interval forecasting was conducted to establish forecasting intervals. Point forecasting in Experiment I pro-

Table 2
The proposed model in comparison with benchmark models on workdays.

DATASET	MODEL	AE	MAE	RMSE	MAPE	DA	FB(−)	TIC(−)
	Naïve	16.1269	233.9330	354.3321	4.4031	0.7619	0.0033	0.0017
	AR	33.6785	465.3611	707.9355	8.4154	0.6464	−0.0049	0.0023
Quarter I	**ARIMA**	26.6856	417.0460	673.9117	6.9712	0.6828	−0.0045	0.0022
	ELMAN	1.1692	88.3886	121.1396	1.5743	0.7116	−0.0002	0.0001
	BPNN	−2.5901	81.1295	110.2048	1.4416	0.7202	0.0004	0.0002
	PSO-BP	−2.0167	80.9192	103.7989	1.4384	0.7385	0.0004	0.0002
	MODA-BP	−2.2314	77.1690	100.4598	1.4024	0.7736	0.0002	0.0001
	EEMD-FTS-MODA-BP	0.8732	69.9264	89.6573	1.3243	**0.8212**	0.0002	0.0001
	ICE-FTS-MODA-BP	**0.2005**	**54.5045**	**75.2808**	**0.9580**	0.8058	**0.0000**	**0.0000**
	Naïve	−10.6500	80.9413	137.5630	1.7036	0.7827	0.0577	0.0301
	AR	−0.2160	77.9150	111.8873	1.5164	0.7660	0.0385	0.0606
Quarter II	**ARIMA**	**−0.1759**	75.2649	96.6758	1.2774	0.8007	**0.0000**	**0.0000**
	ELMAN	−26.3098	93.1895	125.4139	1.5494	0.7941	0.0043	0.0022
	BPNN	−29.2107	87.2830	117.8140	1.4374	0.7842	0.0048	0.0024
	PSO-BP	−30.2352	82.4159	116.7411	1.4361	0.7991	0.0052	0.0021
	MODA-BP	−32.3276	81.2479	111.4140	1.4293	0.8178	0.0046	0.0020
	EEMD-FTS-MODA-BP	−17.4184	67.3829	89.5283	1.1963	0.8865	0.0037	0.0015
	ICE-FTS-MODA-BP	−19.4186	**61.8046**	**83.5457**	**1.0130**	**0.9014**	0.0032	0.0016
	Naïve	33.7407	132.0464	273.0754	2.5385	0.7084	−0.0028	0.0026
	AR	22.2995	257.3217	301.2899	4.9387	0.6943	−0.0055	0.0027
Quarter III	**ARIMA**	19.3877	226.4415	282.0416	4.2734	0.7239	−0.0036	0.0018
	ELMAN	28.9227	83.4954	126.6667	1.5806	0.7579	−0.0053	0.0026
	BPNN	10.0044	83.6873	116.2518	1.5562	0.7408	−0.0018	0.0009
	PSO-BP	9.7047	82.1706	114.6825	1.5457	0.7427	−0.0018	0.0008
	MODA-BP	9.4829	83.6923	113.4370	1.5225	0.7747	−0.0014	0.0008
	EEMD-FTS-MODA-BP	3.3299	70.6198	98.0094	1.2739	0.8364	−0.0005	0.0003
	ICE-FTS-MODA-BP	**2.3158**	**64.4928**	**91.9450**	**1.2251**	**0.8909**	**−0.0004**	**0.0002**
	Naïve	23.9068	190.1425	384.5889	6.7303	0.7104	−0.0322	0.0044
	AR	80.2634	555.9149	980.6115	12.0061	0.7098	−0.2208	0.0070
Quarter IV	**ARIMA**	74.6988	542.2105	837.6239	9.7783	0.7051	−0.0140	0.0069
	ELMAN	**1.9311**	111.8817	216.4933	1.9741	0.7255	**−0.0004**	**0.0002**
	BPNN	2.3071	87.2307	141.6930	1.5844	0.7279	**−0.0004**	**0.0002**
	PSO-BP	2.2877	86.3005	139.1835	1.5549	0.7390	−0.0010	0.0003
	MODA-BP	2.0986	82.7770	140.4068	1.5401	0.7307	−0.0008	0.0003
	EEMD-FTS-MODA-BP	−8.8232	62.8767	122.8767	1.1021	**0.8344**	0.0020	0.0009
	ICE-FTS-MODA-BP	−13.7248	**60.6163**	**118.0424**	**1.0780**	0.8284	0.0026	0.0013

vides deterministic forecasting results while interval forecasting provides the forecasting range under a certain confidence level. The interval forecasting highly relies on point forecasting accuracy and stability in Experiment I. The IFCP and IFAW (introduced in Appendix F) are used to measure the capability of interval forecasting in this paper. The interval forecasting results are shown in Table 4.

Given a certain significance level α, a desirable situation for interval forecasting is that the predictive range covers most of the observed data. Meanwhile, the smaller predictive range the better. However, there is a trade-off between high IFCP and low IFAW values. This paper employs interval forecasting under α = 0.30, where the IFCP and IFAW achieve comparatively good results. Take quarter I as an example, under α = 0.30, the IFCP and IFAW values are 0.7679 and 0.1533, respectively. As for the non-workdays, the values of IFCP and IFAW are 0.9184 and 0.4778, respectively. The performance of the interval forecasting is shown in Figs. 3 and 4.

4. Discussion

4.1. Significance of the forecasting performance

This paper employed the DM test and the FE (details can be found in Appendix G) to verify the outperformance of the proposed forecasting model over comparison models. It is shown in Table 5 that the proposed model is significantly superior to comparison models, e.g., the BPNN, GRNN, ELMAN and ARIMA models, with the DM test values greater than critical values under 1% significance level in quarters I, II and III. In quarter IV, the DM test results for the BPNN and ELMAN is 10% and 5% significant, respectively. From the FE test results we can see the 1st-order forecasting effectiveness provided by the proposed forecasting framework is greater than 0.94, whereas the 2nd-order values are greater than 0.90.

4.2. Effectiveness and improvement of each component

To quantify the contribution of each component that is embedded in the proposed ICE-FTS-MODA-BP model, the reduced relative error (RRE) of the MAPE is used. The REE results can be found in Table 6 and the formula of the RRE is shown as follows:Table D1.

$$RRE_{MAPE_{ij}} = \left| \frac{MAPE_{\ mod\ el_i} - MAPE_{\ mod\ el_j}}{MAPE_{\ mod\ el_i}} \right| \tag{4}$$

From the contribution analysis we can conclude that data cleaning and mining framework, including ICEEMADAN and FTS, can boost forecasting performance significantly. Their contributions to forecasting performance are larger than optimization algorithms.

4.3. Practical applications and limitations

The proposed short-term load forecasting model could effectively reduce risks of power generation caused by variability. Results show that the proposed can achieve accurate and stable forecasting. However, parameters in the proposed model do not continuously update in an online implementation. In general,

Table 3
The proposed model in comparison with benchmark models on non-workdays.

DATASET	MODEL	AE	MAE	RMSE	MAPE	DA	FB(−)	TIC(−)
	Naïve	3.1742	50.1982	76.1645	1.6795	0.6511	0.0045	0.0016
	AR	2.9058	44.6449	57.2942	0.9517	0.8045	0.0076	0.0026
Quarter I	**ARIMA**	**1.2021**	40.8306	53.4681	0.8602	0.8458	**−0.0003**	**0.0001**
	ELMAN	5.5077	68.1617	94.1227	1.4447	0.7089	−0.0011	0.0006
	BPNN	2.4138	63.5111	91.8614	1.3377	0.7173	−0.0005	0.0003
	PSO-BP	2.5687	62.6704	88.4742	1.3042	**0.8520**	−0.0005	0.0003
	MODA-BP	2.6513	60.9577	86.1611	1.2731	0.8234	−0.0004	0.0003
	EEMD-FTS-MODA-BP	2.9024	36.5225	50.6454	0.8967	0.8391	−0.0007	0.0002
	ICE-FTS-MODA-BP	3.9974	**29.2106**	**41.7554**	**0.6231**	0.8517	−0.0008	0.0004
	Naïve	−27.8890	81.4261	128.8215	1.5647	0.8736	0.0028	0.0041
	AR	−20.2634	68.9149	77.0061	1.2315	0.9031	0.0033	0.0013
Quarter II	**ARIMA**	−14.6946	56.3361	**68.9003**	0.9994	0.9322	0.0027	0.0013
	ELMAN	−30.9636	77.3054	117.0300	1.3729	0.8143	0.0057	0.0028
	BPNN	−11.4325	72.4860	110.4069	1.2867	0.8312	0.0021	0.0010
	PSO-BP	−17.1471	71.4402	108.4957	1.2963	0.8484	0.0018	0.0010
	MODA-BP	−13.9109	69.4179	106.9065	1.2531	0.8511	0.0017	0.0004
	EEMD-FTS-MODA-BP	−2.7874	58.4439	77.5950	1.1253	0.9399	0.0004	0.0001
	ICE-FTS-MODA-BP	**−0.1682**	**55.4968**	70.1216	**0.9982**	**0.9417**	**0.0000**	**0.0000**
	Naïve	27.1711	86.6645	145.6941	1.5695	0.7833	−0.0040	0.0020
	AR	13.3924	55.6238	79.3940	1.2403	0.8054	−0.0027	0.0020
Quarter III	**ARIMA**	11.7223	51.7582	69.0258	1.0770	0.8686	−0.0025	0.0012
	ELMAN	15.7889	70.1882	116.2407	1.4432	0.7257	−0.0033	0.0017
	BPNN	17.0724	67.8703	102.2224	1.4015	0.7384	−0.0036	0.0018
	PSO-BP	17.0063	65.8411	100.3045	1.3830	0.7633	−0.0031	0.0011
	MODA-BP	16.3213	65.0572	98.5833	1.3345	0.7701	−0.0033	0.0009
	EEMD-FTS-MODA-BP	2.0165	43.9885	66.0198	1.0958	0.8725	−0.0008	0.0003
	ICE-FTS-MODA-BP	**1.0438**	**41.5602**	**57.1816**	**0.8557**	**0.8958**	**−0.0002**	**0.0001**
	Naïve	5.2401	80.0778	125.9585	1.8716	0.6927	−0.0009	0.0008
	AR	6.6482	65.4412	70.7058	1.3212	0.8207	−0.0009	0.0016
Quarter IV	**ARIMA**	5.5848	51.4666	65.5011	1.1111	0.8559	−0.0012	0.0006
	ELMAN	−3.5854	70.6211	101.1660	1.4903	0.7215	0.0008	0.0004
	BPNN	2.7865	64.5021	89.4172	1.3773	0.7468	−0.0006	0.0003
	PSO-BP	2.7752	62.7497	87.1616	1.2894	0.7748	−0.0004	0.0003
	MODA-BP	2.7127	59.7035	87.7709	1.2200	0.7915	−0.0002	0.0003
	EEMD-FTS-MODA-BP	0.6242	44.4048	60.8851	0.9295	0.8514	−0.0004	0.0001
	ICE-FTS-MODA-BP	**0.3693**	**39.4680**	**53.6836**	**0.8505**	**0.8792**	**−0.0001**	**0.0000**

non-linear models are difficult to parametrize, while nearly all methods are adaptive if the model parameters are updated on a regular basis. The update of hybrid models requires the solution of large optimization problems and thus is computationally costly. In practical application, the model parameters in the proposed model could be continuously updated by using historical data [49]. For example, parameters can be redetermined for the respective type of the day (working days vs. weekend days) by using historical data [50]. On the other hand, numbers of training days and the correction parameters can also be redetermined dynamically by online implementation.

The proposed model enables to obtain accurate and reliable short-term forecasting results but unfortunately, it cannot adapt to future changing conditions automatically and dynamically. For example, certain parameters such as the improvement of transmission technology and the change of consumption habits will affect load demand but they are not included in the proposed model. Further, the proposed model may encounter problems when catastrophic events happen in the future. However, the time resolution of short-term forecasting models is always from several minutes-ahead to several hours-ahead, but not for a long period. Short-term forecasting models focus more on instability, intermittency and complex fluctuation of short-term load demand. While the proposed model is applicable for long-term (e.g., day-ahead, week-ahead or even longer) power prediction, the forecasting accuracy will be largely reduced. The longer the time interval of prediction, the greater the fluctuation and the worse the regularity of the electricity. In terms of long-term load forecasting, there is no state-of-art method to address it. Most of long-term load forecasting models are based on certain scenarios and rough estimation, which is out of the scope of this paper. However, short-term forecasting with long-term scenarios analysis could provide practical solutions in parallel with the changing and uncertain future.

Despite the limitations, the proposed short-term forecasting model could be used, for instance, by policy makers and industrial stakeholders to support and orient the development of power systems toward the use of the most stable and efficient type of renewable energies. Moreover, the proposed method may provide statistics to support renewable energy-oriented policy scenarios and accurate forecasting information to project developers. This novel approach designed for large territories, here applied to Australian cases, is generic and can be applied to other territories. Furthermore, the proposed model could be implemented to various real-time tasks, such as RE generation and distribution, real-time maintenance and security check [51] in power systems, and the household energy consumption reducing and management.

5. Conclusions

In this paper, a hybrid weekday-weekend short-term load forecasting model was proposed that consists of data cleaning, forecasting, optimization and evaluation. The raw short-term load time series is decomposed and reconstructed by the ICEEMDAN. Then, the FTS further mines the characteristics of the preprocessed data and the fuzzified data is imported to ANNs. A multi-objective optimization algorithm is employed to boost forecasting accuracy and stability simultaneously. Finally, the forecasting results are

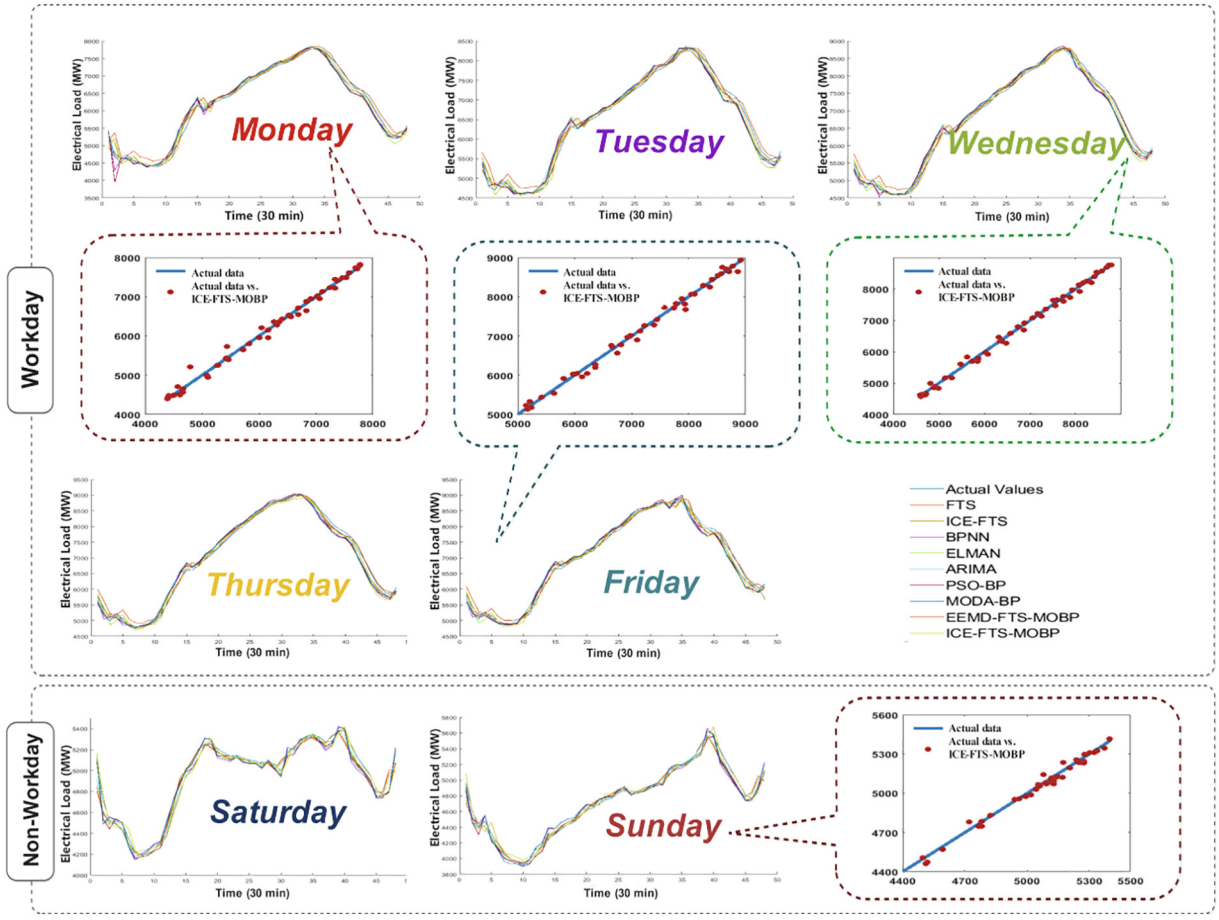

Fig 2. The forecasting result on workdays and non-workdays.

Table 4
The interval forecasting results under different significance levels.

Quarter	α	Workday		Non-workday	
		IFCP	IFAW	IFCP	IFAW
Quarter I	0.10	0.3575	0.0511	0.4133	0.1593
	0.20	0.6335	0.1022	0.7143	0.3186
	0.30	**0.7679**	**0.1533**	**0.9184**	**0.4778**
	0.40	0.8575	0.2044	0.9796	0.6371
	0.50	0.9185	0.2556	0.9949	0.7964
Quarter II	0.10	3.6276	0.1175	0.2296	0.1351
	0.20	0.4830	0.2088	0.4847	0.2701
	0.30	**0.7007**	**0.3131**	**0.7296**	**0.4052**
	0.40	0.8414	0.4175	0.9133	0.5403
	0.50	0.9167	0.5219	0.9745	0.6753
Quarter III	0.10	0.4185	0.1080	0.3316	0.1423
	0.20	0.5726	0.2160	0.5918	0.2845
	0.30	**0.7177**	**0.3240**	**0.8061**	**0.4268**
	0.40	0.8495	0.4319	0.8980	0.5690
	0.50	0.9086	0.5399	0.9694	0.7113
Quarter IV	0.10	0.4203	0.0552	0.3622	0.1141
	0.20	0.6362	0.1104	0.6122	0.2282
	0.30	**0.7733**	**0.1657**	**0.8571**	**0.3423**
	0.40	0.8602	0.2209	0.9490	0.4564
	0.50	0.9238	0.2761	0.9694	0.5705

obtained by defuzzying. Both point and interval forecasting were conducted to the short-term load forecasting. To consider the seasonal and workday-weekend patterns, separate forecasting results regarding those patterns are provided. Comprehensive measurement and evaluation were used to verify the effectiveness of the proposed forecasting model. The results and discussion show that

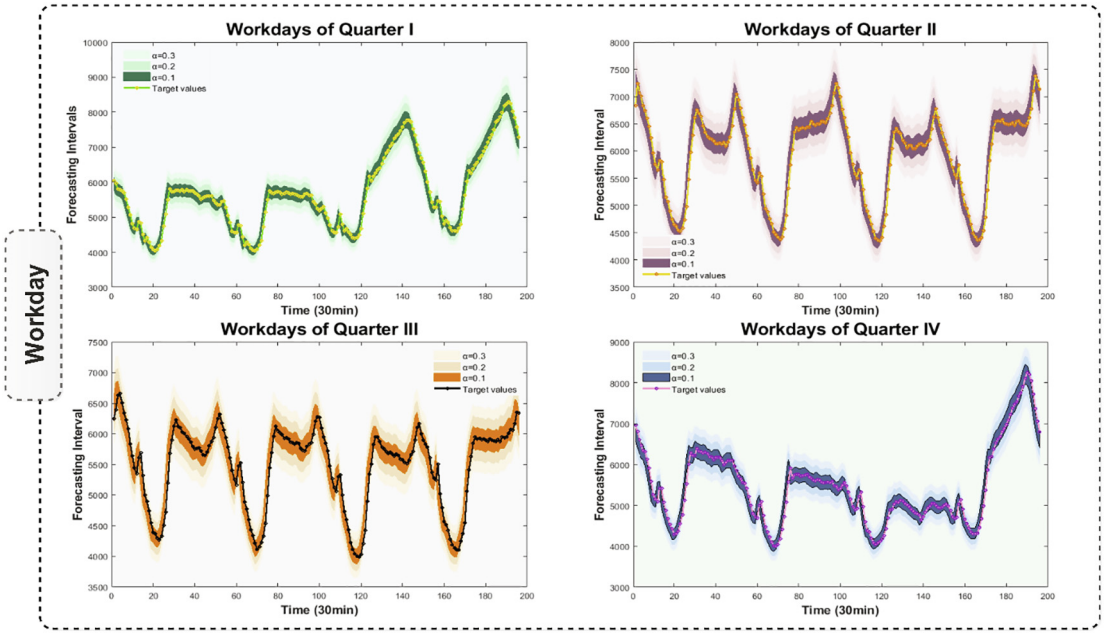

Fig 3. Electrical power interval forecasting performance on workdays.

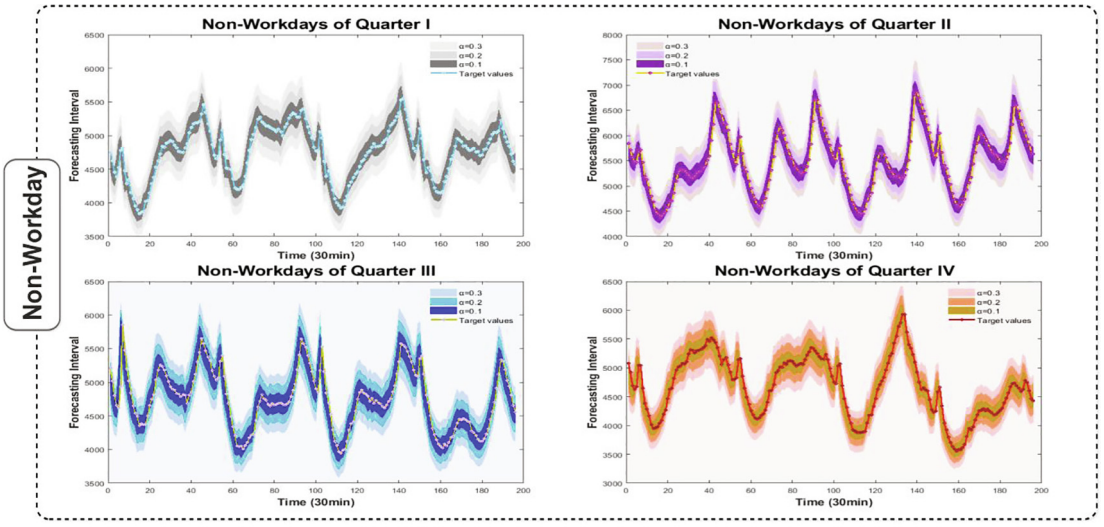

Fig 4. Electrical power interval forecasting performance on non-workdays.

Table 5
The DM test and the FE results.

Test	Datasets	Proposed Model	BPNN	GRNN	ELMAN	ARIMA
DM test	Quarter I	–	10.0736***	12.0238***	11.4516***	12.2955***
	Quarter II	–	4.5084***	10.015***	9.3209***	13.2921***
	Quarter III	–	6.3854***	10.4656***	4.7530***	21.9588***
	Quarter IV	–	1.8411*	10.5858***	2.3242**	13.8611***
FE[a]	Quarter I	0.9897	0.9852	0.9821	0.9839	0.9537
	Quarter II	0.9896	0.9857	0.9815	0.9846	0.9485
	Quarter III	0.9876	0.9818	0.9772	0.9846	0.9530
	Quarter IV	0.9857	0.9832	0.9756	0.9814	0.9552
FE[b]	Quarter I	0.9798	0.9704	0.9650	0.9686	0.9145
	Quarter II	0.9797	0.9733	0.9627	0.9714	0.9087
	Quarter III	0.9748	0.9642	0.9540	0.9669	0.9142
	Quarter IV	0.9658	0.9679	0.9478	0.9587	0.9168

The values in bold indicate the best values.
* Indicates the 10% significance level, ** indicates the 5% significance level, and *** indicates the 1% significance level.
[a] Indicates the 1st-order forecasting effectiveness.
[b] Indicates the 2nd-order forecasting effectiveness.

Table 6
The REE results of the proposed model and comparison models.

Dataset	Model	Workdays		Non-workdays	
		MAPE	RRE	MAPE	REE
Quarter I	FTS	1.2745	24.8038	1.1453	45.5895
	ICE-FTS	1.0843	11.6236	0.9544	34.6960
	BPNN	1.4427	33.5645	1.3380	53.4380
	FTS-MODA-BP	1.1080	13.5379	0.8380	25.6563
	ICE-FTS-BP	1.0022	4.3912	0.7095	12.1298
	ICE-FTS-MODA-BP	**0.9580**	/	**0.6231**	/
Quarter II	FTS	2.1599	53.0801	2.4378	59.0480
	ICE-FTS	1.8156	44.1873	2.0313	50.8616
	BPNN	1.4376	29.5059	1.2878	22.4553
	FTS-MODA-BP	1.4434	29.7990	1.7525	43.0365
	ICE-FTS-BP	1.3837	26.7534	1.5344	34.9413
	ICE-FTS-MODA-BP	**1.0130**	/	**0.9982**	/
Quarter III	FTS	1.7902	31.5642	1.9404	55.8763
	ICE-FTS	1.4948	18.0054	1.6175	47.0625
	BPNN	1.5566	21.2725	1.4020	38.9444
	FTS-MODA-BP	1.7286	29.1088	1.5939	46.2649
	ICE-FTS-BP	1.4917	17.8404	1.4213	39.7607
	ICE-FTS-MODA-BP	**1.2251**	/	**0.8557**	/
Quarter IV	FTS	1.1886	9.2593	2.1571	60.5471
	ICE-FTS	**0.9974**	−8.1244	1.7976	52.6433
	BPNN	1.5844	31.9444	1.3773	38.1990
	FTS-MODA-BP	1.4435	25.2945	1.7423	51.1481
	ICE-FTS-BP	1.3261	18.7029	1.5576	45.3436
	ICE-FTS-MODA-BP	1.0780	/	**0.8505**	/

the proposed hybrid model ICE-FTS-MODA-BP takes advantage of each component and achieves high forecasting stability and accuracy. The proposed model can provide an effective and efficient weekday-weekend and seasonal forecasting, which assists the decision-makers for strategy making in the field of energy and electricity system development.

CRediT authorship contribution statement

Chen Li: Conceptualization, Methodology, Software, Data curation, Writing - original draft, Visualization, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author would like to thank Jingjing Ding (former master student at school of statistics in Dongbei University of Finance and Economics, China) for her assistance in writing and visualizing. China Scholarship Council (CSC) is gratefully acknowledged for its support to Chen Li (file No. 201908210319).

Appendix A. . List of abbreviations

RE	**Renewable energy**	***CEEMDAN***	**Complete ensemble empirical mode decomposition with adaptive noise**
ANN	Artificial neural network	*FLR*	Fuzzy logical relationship
SVM	Support vector machines	*LHS*	Left-hand side
ARMA	Autoregressive moving average	*RHS*	Right-hand side
SOM	Self-organizing map	*DA*	Dragonfly algorithm
FTS	Fuzzy time series	*MODA*	Multi-objective dragonfly algorithm
BPNN	Back propagation neural network	*IF*	Interval forecasting
DM	Diebold–Mariano	*FVD*	Forecasting validity degree
AI	Artificial intelligence	*PSO*	Particle swarm optimization
FE	Forecasting effectiveness	*MOPSO*	Multi-objective particle swarm optimization
ICEEMDAN	Improved complete ensemble empirical mode decomposition with adaptive noise	*MOALO*	Multi-objective ant lion optimization
EMD	Empirical mode decomposition	*IGD*	Inverted generational distance
IMF	Intrinsic mode functions	*SP*	Spacing
EEMD	Ensemble empirical mode decomposition	*RRE*	educed relative error

Appendix B. List of nomenclature

$\boldsymbol{Y}$	Set of continuous values	$\boldsymbol{H}$	Output of the output layer
$\boldsymbol{f}$	Fuzzy set	K	Input signal
$\boldsymbol{u}$	Universe of discourse	δ	Error signals
$\boldsymbol{F}$	Fuzzy time series (set of f)	θ	Thresholds of the BPNN
$\boldsymbol{R}$	Fuzzy logical relationship	α, β	Learning parameters
$\boldsymbol{A_i}$	Left-hand side of the FLR	T	The target output of the output layer
$\boldsymbol{A_j}$	Right-hand side of the FLR	o	Number of objectives
$\boldsymbol{R_{min}}$	Upward range	m	Number of unequal constraints
$\boldsymbol{R_{max}}$	Downward range	p	Number of equal constraints
$\boldsymbol{a}$	A constant	L_i	The lower frontiers of the i_{th} variables
$\boldsymbol{W\_s}$	Standardized weighting matrix	U_i	The upper frontiers of the i_{th} variables
$\boldsymbol{D}$	Defuzzied matrix	$\vec{x}, \vec{y}$	Vectors
$\boldsymbol{W}$	Unstandardized weights matrix factors	E	Mathematical expectation
$\hat{W}_i$	Standardized weights matrix factors	I_{min}, I_{max}	Interval limits
$\boldsymbol{F\_u}$	Ultimate forecasting value	α	Confidence level
ζ	A coefficient	$\hat{\Theta}$	A Function
$\boldsymbol{W}$	Weights between layers of the BPNN	ε	Forecasting error
$\boldsymbol{k}$	Number of input nodes	S	The variance
$\boldsymbol{l}$	Number of hidden nodes	L	The loss function
$\boldsymbol{h}$	Number of output nodes	Z	The critical z-value
$\boldsymbol{L}$	Output of the hidden layer	Q	Discrete probability distribution

Appendix C. The basic theory of the FTS

Definition 1: $\boldsymbol{Y}_t$ is denoted as a set of continuous values. Fuzzy sets f_j and the universe of discourse $\boldsymbol{u}$ are obtained on the basis of $\boldsymbol{Y}_t$. After that, $\boldsymbol{F}_t$, a set of $f_{1,t}$, $f_{1,t}$, ⋯, refers to the fuzzy time series [52].

Definition 2: $\boldsymbol{F}_t$ is supposed to be only related to $\boldsymbol{F}_{t-1}$. A forecasting model is defined as $\boldsymbol{F}_t = \boldsymbol{F}_{t-1} * \boldsymbol{R}_{t-1}$, in which $\boldsymbol{F}_{t-1}$ and $\boldsymbol{F}_t$ are two fuzzy sets and $\boldsymbol{R}_{t-1,t}$ is the fuzzy logical relationship (FLR) between this two fuzzy sets.

Definition 3: Given $\boldsymbol{F}_{t-1} = \boldsymbol{A}_i$ and $\boldsymbol{F}_t = \boldsymbol{A}_j$. The FLR between two fuzzy values can be denoted as $\boldsymbol{A}_i \rightarrow \boldsymbol{A}_j$, where $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ reflect the left-hand side (LHS) and right-hand side (RHS) of the FLR, respectively.

Definition 4: Based on the same LHS of the FLR, every single FLR can be assembled into several groups. Then, the calculation steps of the weighted fuzzy time series can be expressed as in [53]:

Step 1: Define universe of discourse $\boldsymbol{u}_i = [R_{min}\text{-}a, R_{max} + a]$, where R_{min} and R_{max} are the upward and downward of Y_i, respectively. a is a constant and $\boldsymbol{u}$ is afterwards split into several intervals based on the equidistant interval partitioning methods.

Step 2: Set a fuzzy membership function and then obtain the fuzzy sets. The fuzzy sets $\boldsymbol{A}_i$ is divided based on the above-mentioned intervals. In this paper, five subsets are employed as in [54].

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 \\ A_5 &= 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 \end{aligned} \quad (1)$$

Step 3: Fuzzifier the observation values. For instance, the fuzzified result of one single data is $\boldsymbol{A}_j$, and the maximum degree of membership of this value is in $\boldsymbol{A}_j$.

Step 4: Determine the fuzzy logical relationships and assemble them. For instance, if $\boldsymbol{A}_i \rightarrow \boldsymbol{A}_j$, $\boldsymbol{A}_i \rightarrow \boldsymbol{A}_k$, and $\boldsymbol{A}_i \rightarrow \boldsymbol{A}_l$ can be assembled to $\boldsymbol{A}_i \rightarrow \boldsymbol{A}_j$, $\boldsymbol{A}_k$, $\boldsymbol{A}_l$.

Step 5: Construct weights matrix. The fuzzified matrix can be expressed by using the centroid defuzzification method and the weights matrix also can be further standardized.

Step 6: Obtain the prediction results. Prediction results can be obtained by multiplication of the fuzzified and standardized weights matrix that determined as follows:

$$\mathrm{W\_s}_t = (\hat{\mathbf{W}}_1, \hat{\mathbf{W}}_2, \cdots, \hat{\mathbf{W}}_k),\ \hat{\mathbf{W}}_i = \mathrm{W}_i / \sum_{i=1}^{k} \mathrm{W}_i \quad (2)$$

$$\mathrm{F}_t = \mathrm{D}_{t-1} * \mathrm{W\_s}_{t-1} \quad (3)$$

where $\boldsymbol{W\_s}$ is the standardized weighting matrix, and $\boldsymbol{D}$ is the defuzzied matrix. $\boldsymbol{W}$ is defined as the unstandardized weights matrix factors; $\hat{W}$is the standardized factor, and $\boldsymbol{F}$ represents the forecasting result.

Step 7: Finally, the forecasting results are improved by applying Eq. (3) to achieve the ultimate forecasting results. Where $\mathbf{y}_{t-1}$ indicates the actual value on time $t-1$, and F_u_t represents the ultimate forecasting value.

$$F\_u_t = y_{t-1} + \zeta * (F_t - y_{t-1}) \quad (4)$$

where ζ is a coefficient.

Appendix D. . Test on the MODA

To testify the efficiency and superiority of the MODA, four test functions (shown in Appendix I) are used for algorithm test in this paper. Two multi-objective optimization algorithms, e.g., multi-objective particle swarm optimization (MOPSO) and multi-objective ant lion optimization algorithm (MOALO) are employed for comparison. In order to achieve effective and robust simulation results, each algorithm has been repeatedly simulated for 20 times, and the final result is obtained by averaging those 20 results. All experiments were performed with 100 iterations, 50 search agents, and 100 number of archives. Inverted generational distance (IGD) [55,56] and spacing (SP) [57,58] are implemented to quantitatively evaluate the performance of three algorithms. The statistic values of IGD and SP are shown in Table 5. The obtained Pareto optimal solutions by each algorithm are shown in Appendix J. Two main conclusions are made from the algorithm test:

The MODA algorithm obtains the optimal IGD and SP values in most statistic magnitudes, which indicates its better optimization capacity than two comparison algorithms.

Table D1
Statistical values of the IGD and SP for four test functions.

Statistic Magnitude	Algorithm	ZDT1		ZDT2		ZDT3		ZDT1 in linear front	
		ICP	SP	ICP	SP	ICP	SP	ICP	SP
Ave.	MODA	0.0023	0.0200	0.0030	0.1653	0.0247	0.0211	0.0026	0.1310
	MOPSO	0.0024	0.0251	0.0127	0.0250	0.0254	0.0290	0.0027	0.0262
	MOALO	0.0078	0.0145	0.0102	0.0145	0.0266	0.0260	0.0070	0.0165
Std.	MODA	0.0014	0.0121	0.0007	0.3040	0.0003	0.0054	0.0013	0.2411
	MOPSO	0.0005	0.0033	0.0001	0.0031	0.0009	0.0042	0.0005	0.0043
	MOALO	0.0047	0.0047	0.0067	0.0069	0.0020	0.0154	0.0070	0.0061
Median	MODA	0.0015	0.0179	0.0032	0.0202	0.0246	0.0203	0.0028	0.0309
	MOPSO	0.0023	0.0253	0.0127	0.0243	0.0252	0.0285	0.0026	0.0258
	MOALO	0.0068	0.0160	0.0081	0.0136	0.0260	0.0219	0.0046	0.0147
Worst	MODA	0.0055	0.0634	0.0038	0.7501	0.0253	0.0308	0.0045	0.7900
	MOPSO	0.0040	0.0305	0.0129	0.0322	0.0276	0.0376	0.0040	0.0334
	MOALO	0.0198	0.0209	0.0212	0.0292	0.0316	0.0566	0.0328	0.0291
Best	MODA	0.0010	0.0087	0.0021	0.0166	0.0244	0.0149	0.0010	0.0109
	MOPSO	0.0018	0.0155	0.0126	0.0207	0.0244	0.0221	0.0020	0.0203
	MOALO	0.0022	0.0033	0.0031	0.0040	0.0243	0.0082	0.0024	0.0083

The MODA algorithm covers more Pareto optimal solutions than the MOPSO and MOALO, which verifies its better universality.

***Algorithm* 1: Pseudo code of the MODA**

Fitness function:

$$\min\begin{cases} fitness_1 = |Bias(\hat{x})| \\ fitness_2 = Std(x - \hat{x}) \end{cases}$$

***Output*:**

$\hat{X} - X$with the best fitness

Parameters:

$Iter_{Max}$ — the maximum iterations
n — the dragonflies' number
F_i —the fitness of *i-th* dragonfly
$[L_i,U_i]$ — the boundaries of the variable
X_i — the position of *i-th* variable
ΔX_i — the step vactor
t — the current iterations
d — the dimension of the optimize problem

1. /*Set the basic parameters of the MODA.*/
2. /*Initialize the dragonflies population $\boldsymbol{X}_i$ (i = 1, 2,. . ., n) randomly.*/
3. /*Initialize the step vectors $\Delta\boldsymbol{X}_i$ (i = 1, 2,. . ., n) .*/
4. /*Define the maximum number of hyper
5. spheres (segments).*/
6. /*Define the archive size.*/
7. **FOR EACH** i: $1 \le i \le n$ **DO**
8. Calculate the corresponding $\boldsymbol{F_i}$ using ranking process
9. **END FOR**
10. /*Determine the best dragonflies and suppose it as the elite.*/
11. **WHILE** (t < $iter_{\mathrm{Max}}$) **DO**
12. /*Calculate the objective values of all dragonflies.*/
13. /*Find the non-dominated solutions.*/
14. /*Update the archive in regard to the obtained non-dominated solutions.*/
15. **IF** the archive is full **DO**
16. /*Omit some solutions from the archive to add the new solutions.*/
17. **END IF**
18. **IF** any new added solutions to the archive is outside hyper spheres **DO**
19. /*Update and re-position all of the hyper to cover the new solutions.*/
20. **END IF**

(*continued*)

***Algorithm* 1: Pseudo code of the MODA**

21. Select a food source from archive: $\boldsymbol{X}^{+}$ = SelectFood (*archive*)
22. Select an enemy from archive: $\boldsymbol{X}^{-}$= SelectEnemy (*archive*)
23. /*Update the step vectors.*/
24. $\Delta\boldsymbol{X}_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta\boldsymbol{X}_t$
25. /*Update the position vectors according to different conditions.*/
26. $\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \Delta\boldsymbol{X}_{t+1}$
27. $\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + Le'vy(d) \times \boldsymbol{X}_t$
28. Check and correct the new positions based on the boundaries [$\boldsymbol{L}_i$,$\boldsymbol{U}_i$]
29. $t = t + 1$
30. END WHILE
31. RETURN x_b

Appendix E. . Pareto optimal solutions

Minimize:

$$\mathrm{F}(\vec{\mathbf{x}}) = \{f_1(\vec{\mathbf{x}}), f_2(\vec{\mathbf{x}}), ..., f_o(\vec{\mathbf{x}})\} \tag{1}$$

Subject to:

$$\mathrm{g}_i(\vec{\mathbf{x}}) \geqslant 0, \quad i = 1, 2, ..., m \tag{2}$$

$$\mathrm{h}_i(\vec{\mathbf{x}}) \geqslant 0, \quad i = 1, 2, ..., p \tag{3}$$

$$\mathrm{L}_i \leqslant \mathrm{x}_i \leqslant \mathrm{U}_i, \quad i = 1, 2, ..., n \tag{4}$$

Suppose that there are o objectives, m unequal constraints, and p equal constraints. And $\mathbf{L}_i$ and $\mathbf{U}_i$ denote the lower and upper frontiers of the i_{th} variables, respectively.

Definition 6. Pareto dominance [59]:

Assume that there are two vectors $\vec{\mathbf{x}} = (\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_l)$, $\vec{\mathbf{y}} = (\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_l)$.

Vector $\vec{x}$ dominates $\vec{y}$, defined as $\vec{\mathbf{x}} \succ \vec{\mathbf{y}}$, if:

$$\forall i\{1, 2, ..., l\}, [f(\mathrm{x}_i) \geqslant f(\mathrm{y}_i)] \wedge [i \in 1, 2, ..., l : f(\mathrm{x}_i)] \tag{5}$$

Definition 7. Pareto optimality [60]:

A Pareto optimal can be through $\vec{\mathbf{x}} \in \mathrm{X}$, if:

$$\exists \vec{\mathbf{y}} \in \mathrm{X} | \mathrm{F}(\vec{\mathbf{y}}) \succ \mathrm{F}(\vec{\mathbf{x}}) \tag{6}$$

If neither of the solutions governs the other, and then they are non-dominated.

Definition 8. Pareto optimal set:

Pareto set are the set that contains all non-dominated solutions:

$$P_s = \{\vec{\mathbf{x}}, \vec{\mathbf{y}} \in X | \exists F(\vec{\mathbf{y}}) \succ F(\vec{\mathbf{x}})\} \quad (7)$$

Definition 9. Pareto optimal front:

A set including the corresponding values of Pareto optimal solutions in a Pareto optimal set is named as Pareto optimal front:

$$P_f = \{F(\vec{\mathbf{x}}) | \vec{\mathbf{x}} \in P_s\} \quad (8)$$

To handle the multi-objective optimization problems through the MODA approach, it is essential to construct an archive which is utilized to store and then retrieve the optimal approximations of the true Pareto optimal solutions. The food resources are selected from the archive and the updating position of each search agent is identical to the DA algorithm. Finally, through the least populated field of the Pareto optimal front, a food resource is chosen to attain a well-spread front.

The key to finding the least populated region of the Pareto optimal front is to divide the search space into several segments. Next, the selection applies the roulette wheel mechanism. The worst hyper-sphere enemies from the archive are to be chosen according to the roulette-wheel results, which prevents the dragonflies from searching through non-promising crowded regions. During the whole process, the archive is supposed to update regularly as it can become gradually complete. In this paper, the research of Coello et al. [61] is used to manage the archive.

Appendix F. The evaluation metrics

It is essential to make a comprehensive evaluation whereas there is no unified standard for model measurement [62]. Therefore, this paper employs multiple error criteria. AE roughly measures the difference between the forecasted values and actual values; MAE shows the degree of the difference between the forecasted values and actual values; RMSE is another relative error estimator that pays more attention to the impact of extreme values based on the MAE; MAPE is a common index in statistics for evaluating the accuracy of forecasting models; DA indicates the forecasting direction of each model; FB measures of mean bias and indicate only systematic errors which lead to always underestimate or overestimate the measured values; TIC provides a measure of how well a time series of estimated values compares to a corresponding time series of observed values. IFCP is a quantitative measure that shows the probability of actual values covered by the lower and upper bounds; and IFNAW is a significant characteristic of IFs. A general description of nine criteria is shown below, where F_i and A_i indicate the forecasted and actual values at time i, respectively.

Metric	Definition	Equation
AE	The average error of N forecasting results	**AE** $=\frac{1}{N}\sum_{i=1}^{N}(F_i - A_i)$
MAE	The mean absolute error of N forecasting results	**MAE** $=\frac{1}{N}\sum_{i=1}^{N}\lvert F_i - A_i\rvert$
RMSE	The square root of the mean square error	**RMSE** $=\sqrt{\frac{1}{N}\times\sum_{i=1}^{N}(F_i - A_i)^2}$
MAPE	The mean absolute percent error of N forecasting results	**MAPE** $=\frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{A_i - F_i}{A_i}\right\rvert \times 100\%$
DA	The direction accuracy of forecasting results	**DA** $=\begin{cases}1, & if\ (A_{i+1}-A_i)(F_{i+1}-A_i)>0\\ 0, & otherwise\end{cases}$
FB	The fractional bias of forecasting results	**FB** $=2(\bar{A}-\bar{F})/(\bar{A}+\bar{F})$
TIC	Theil's inequality coefficient of forecasting results	**TIC** $=\frac{\sqrt{\frac{1}{N}\times\sum_{i=1}^{N}(F_i-A_i)^2}}{\sqrt{\frac{1}{N}\times\sum_{i=1}^{N}A_i^2}+\sqrt{\frac{1}{N}\times\sum_{i=1}^{N}F_i^2}}$
IFCP	The internal forecasting coverage probability	**IFCP** $=\frac{1}{N}\sum_{i=1}^{N}y_i, y_i=\begin{cases}1, & if\ y_i[L_i,U_i]\\ 0, & otherwise\end{cases}$
IFAW	The internal forecasting average width	**IFAW** $=\frac{1}{N}\sum_{i=1}^{N}(U_i - L_i)$

Appendix G. Test methods

G-1: Diebold-Mariano (DM)

The DM test is a statistical test to evaluate whether two models have a significant difference in regard to forecasting performance [63]. The null and alternative hypothesis of the DM test are shown below:

$$H_0:\ E(d_h)=0,\ \forall n;\ H_1:\ E(d_h)\neq 0,\ \exists n \quad (1)$$

The statistics of DM test is

$$DM = \frac{\sum_{h=1}^{k}(L(\varepsilon_{t+h}^{(A)}) - L(\varepsilon_{t+h}^{(B)}))/k}{\sqrt{S^2/k}} s^2 \quad (2)$$

where ε_{t+h}, S^2 and L represent the forecasting error, the variance and the loss function, respectively. In this paper, the square error loss function is used. The statistics of DM is convergent to the standard normal distribution, where the null hypothesis cannot be accepted if

$$|DM| > \boldsymbol{z}_{\alpha/2} \quad (3)$$

where $z_{\alpha/2}$ is the critical z-value and α is the significance level. The statistics of DM is subject to a normal distribution, so it will be rejected the null hypothesis if **|DM|** > 1.96, given the 5% significant level [64].

G-2: Forecasting effectiveness (FE)

This paper uses forecasting effectiveness (FE) to evaluate average accuracy along forecasting periods [65]. Assume the forecasted short-term load series x, $(x_t, t = 1, 2, \ldots, N)$ and m forecasting methods are to be compared. x_{it} represents the forecasting value at time t with the i_{th} method, where i = 1,2,...,m.

Definition 1. The value of e_{it} is the relative error of the i_{th} method at time i = 1,2,. . .,m, t = 1,2,. . .,N. $E = (e_{it})_{m\times N}$ is the matrix of relative errors.

$$e_{it} = \begin{cases} -1, & \frac{x_t - x_{it}}{x_t} < -1 \\ \frac{x_t - x_{it}}{x_t}, & -1 < \frac{x_t - x_{it}}{x_t} < 1 \\ 1, & \frac{x_t - x_{it}}{x_t} > 1 \end{cases} \quad (4)$$

Definition 2. The forecasting accuracy of the i_{th} method at time t is calculated by $A_{it} = 1 - |e_{it}|$ (i = 1,2,. . .,m, t = 1,2,. . .,N). Naturally, $0 \le A_{it} \le 1$ and $A_{it} = 0$ when $(x_t - x_{it})/x_t > 1$.

Definition 3. The element of the k-order forecasting validity degree (FVD) with the i_{th} method:

$$m_i^k = \sum_{t=1}^{N} Q_t A_{it}^k \quad (5)$$

where k is the positive integer and (Q_t, t = 1,2,. . .,N) is the discrete probability distribution of the m_{th} forecasting method at time t:

$$\sum_{t=1}^{N} Q_t = 1, (Q_t > 0) \quad (6)$$

Definition 4. $H\,(m_i{}^1\ m_i{}^2,\ldots,\ m_i^k)$ is a k-element continuous function that illustrates the k-order FVD.

Definition 5. When $H(x) = x$ is a *one*-element continuous function, $H(m_i{}^1) = m_i{}^1$

is the *one*-order forecasting validity of the i_{th} forecasting method; when $H(x_i) = x_i\,(1 - \sqrt{y - x_i^2})$ is a *two*-element continuous function, the *two*-order forecasting validity of the i_{th} forecasting method is represented as:

$$H(m_i^1, m_i^2) = m_i^1(1 - \sqrt{m_i^2 - (m_i^1)^2}) \quad (7)$$

Definition 6. The *1st*-order forecasting effectiveness and the *2nd*-order forecasting effectiveness are the expectation forecasting accuracy sequence, and the difference between the expectation and standard deviation of the forecasting accuracy sequence, respectively.

Appendix H. . Parameter settings

Model	Experimental Parameter	Value
BPNN	Maximum iteration times	1000
	Learning rate	0.01
	Training accuracy goal	0.00001
	Neuron number of the input layer	3
	Neuron number of the hidden layer	7
	Neuron number of the output layer	1
ELMAN	Neuron number of the input layer	3
	Neuron number of the hidden layer	7
	Neuron number of the output layer	1
	Iteration number of displays once in an image	20
	Maximum iteration times	1000
ARIMA (p, d, q)	Autoregressive term (p)	1
	Moving average number (q)	1
	Difference times (d)	1

Appendix I. Test functions

ZDT1	ZDT1 with linear PF
$Minimize: f_1(\mathrm{x}) = \mathrm{x}_1$ $Minimize: f_2(\mathrm{x}) = g(\mathrm{x}) \times h(f_1(\mathrm{x}), g(\mathrm{x}))$ $Where: G(x) = 1 + \frac{9}{N-1}\sum_{i=2}^{N}\mathrm{x}_i$ $h(f_1(\mathrm{x}), g(\mathrm{x})) = 1 - \left(\sqrt{\frac{f_1(\mathrm{x})}{g(\mathrm{x})}}\right)^2$ $0 \leqslant \mathrm{x}_i \leqslant 1, 1 \leqslant i \leqslant n$	$Minimize: f_1(\mathrm{x}) = \mathrm{x}_1$ $Minimize: f_2(\mathrm{x}) = g(\mathrm{x}) \times h(f_1(\mathrm{x}), g(\mathrm{x}))$ $Where: G(\mathrm{x}) = 1 + \frac{9}{N-1}\sum_{i=2}^{N}\mathrm{x}_i$ $h(f_1(\mathrm{x}), g(\mathrm{x})) = 1 - \frac{f_1(\mathrm{x})}{g(\mathrm{x})}$ $0 \leqslant \mathrm{x}_i \leqslant 1, 1 \leqslant i \leqslant n$
ZDT2	**ZDT3**
$Minimize: f_1(\mathrm{x}) = \mathrm{x}_1$ $Minimize f_2(\mathrm{x}) = g(\mathrm{x}) \times h(f_1(\mathrm{x}), g(\mathrm{x}))$ $Where: G(\mathrm{x}) = 1 + \frac{9}{N-1}\sum_{i=2}^{N}\mathrm{x}_\mathrm{i}$ $h(f_1(\mathrm{x}), g(\mathrm{x})) = 1 - \left(\sqrt{\frac{f_1(\mathrm{x})}{g(\mathrm{x})}}\right)^2$ $0 \leqslant \mathrm{x}_i \leqslant 1, 1 \leqslant i \leqslant n$	$Minimize: f_1(\mathrm{x}) = \mathrm{x}_1$ $Minimize: f_2(\mathrm{x}) = g(\mathrm{x}) \times h(f_1(\mathrm{x}), g(\mathrm{x}))$ $Where: G(\mathrm{x}) = 1 + \frac{9}{29}\sum_{i=2}^{N}\mathrm{x}_\mathrm{i}$ $h(f_1(\mathrm{x}), g(\mathrm{x})) = 1 - \sqrt{\frac{f_1(\mathrm{x})}{g(\mathrm{x})}} - \left(\frac{f_1(\mathrm{x})}{g(\mathrm{x})}\right)\sin(10\pi f_1(\mathrm{x}))$ $0 \leqslant \mathrm{x}_i \leqslant 1, 1 \leqslant i \leqslant n$

Appendix J. Pareto optimal results for MODA, MOPSO and MOALO (Figure 5)

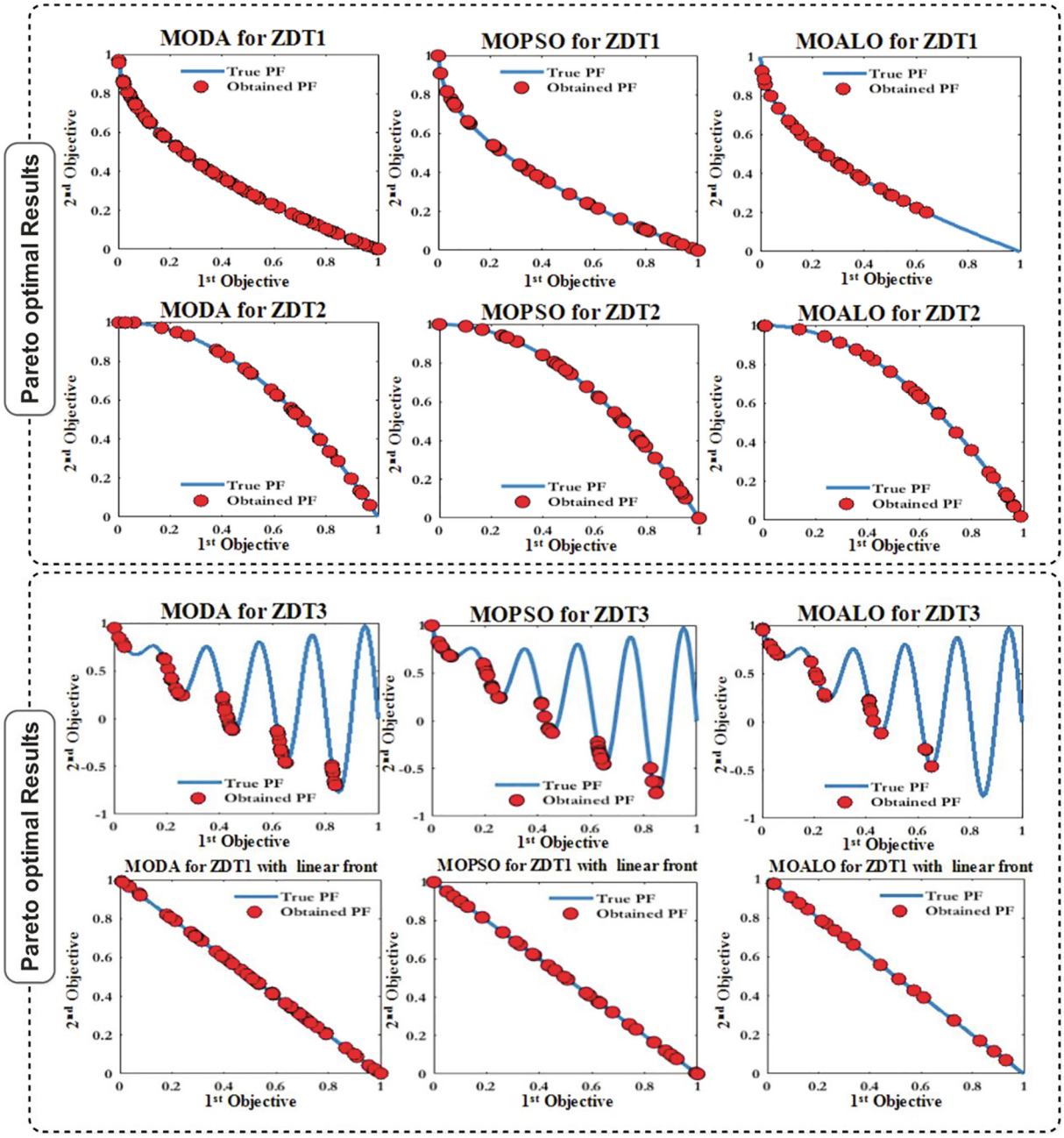

Fig 5.

References

[1] K.-B. Song, Y.-S. Baek, D.H. Hong, G. Jang, Short-term load forecasting for the holidays using fuzzy linear regression method, IEEE Trans. Power Syst. 20 (1) (2005) 96–101.

[2] N.J. Johannesen, M. Kolhe, M. Goodwin, Relative evaluation of regression tools for urban area electrical energy demand forecasting, J. Cleaner Prod. 218 (2019) 555–564.

[3] H.J. Sadaei, P.C. de Lima e Silva, F.G. Guimarães, M.H. Lee, Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series, Energy 175 (2019) 365–377.

[4] K. Lang, M. Zhang, Y. Yuan, X. Yue, Short-term load forecasting based on multivariate time series prediction and weighted neural network with random weights and kernels, Cluster Comput. 22 (S5) (2019) 12589–12597.

[5] M. Gilanifar, H. Wang, L.M.K. Sriram, E.E. Ozguven, R. Arghandeh, Multi-task Bayesian spatiotemporal Gaussian processes for short-term load forecasting, IEEE Trans. Ind. Electron. (2019).

[6] A. Azadeh, S.F. Ghaderi, S. Sohrabkhani, A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran, Energy Policy 36 (2008) 37–44.

[7] P.C. Chang, C.Y. Fan, J.J. Lin, Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach, Electr. Power Syst. Res. 33 (2011) 17–27.

[8] D.C. Park, M.A. Sharkawi, R.J. Marks, Electric load forecasting using a neural network, IEEE Trans. Power Syst. 6 (1991) 442–449.
[9] J.Z. Wang, F. Liu, Y.L. Song, J. Zhao, A novel model: Dynamic choice artificial neural network (DCANN) for an electricity price forecasting system, Appl. Soft Comput. 48 (2016) 281–297.
[10] F. Yu, X. Xu, A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network, Appl. Energy 134 (2014) 102–113.
[11] L. Hernandez, C. Baladron, J.M. Aguiar, Short-term load forecasting for microgrids based on artificial neural networks, Energies 6 (2013) 1385–1408.
[12] B. Shah, B.H. Trivedi, Artificial neural network-based intrusion detection system: a survey, Int. J. Comput. Appl. 39 (6) (2012) 13–18.
[13] J. Morshed, J.J. Kaluarachchi, Application of artificial neural network and genetic algorithm in flow and transport simulations, Adv. Water Resour. 22 (2) (1998) 145–158.
[14] L.W. Liu, H.J. Zong, E.D. Zhao, C.X. Chen, J.Z. Wang, Can China realize its carbon emission reduction goal in 2020: From the perspective of thermal power development, Appl. Energy 124 (2014) 199–212.
[15] J.J. Wang, J.Z. Wang, Y.N. Li, S.L. Zhu, J. Zhao, Techniques of applying wavelet de-noising into a combined model for short-term load forecasting, Int. J. Electr. Power 62 (2014) 816–824.
[16] S. Goudarzi, M.H. Anisi, N. Kama, F. Doctor, S.A. Soleymani, A.K. Sangaiah, Predictive modelling of building energy consumption based on a hybrid nature-inspired optimization algorithm, Energy Build. 196 (2019) 83–93.
[17] F. Zhang, C. Deb, S.E. Lee, J. Yang, K.W. Shah, Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique, Energy Build. 126 (2016) 94–103.
[18] S.C. Pandian, K.D. Duraiswamy, C.C. Rajan, N. Kanagaraj, Fuzzy approach for short-term load forecasting, Electr. Power Syst. Res. 76 (2006) 541–548.
[19] P.F. Pai, Hybrid ellipsoidal fuzzy systems in forecasting regional electricity loads, Energy Convers. Manag. 47 (2006) 2283–2289.
[20] E.M. Burger, S.J. Moura, Gated ensemble learning method for demand-side electricity load forecasting, Energy Build. 109 (2015) 23–34.
[21] S. Itaba, H. Mori, A fuzzy-preconditioned GRBFN model for electricity price forecasting, Procedia Comput. Sci. 114 (2017) 441–448.
[22] R. Azimi, M. Ghofrani, M. Ghayekhloo, A hybrid wind power forecasting model based on data mining and wavelets analysis, Energy Convers. Manag. 127 (2016) 208–225.
[23] C.J. Lu, Y.W. Wang, Combining independent component analysis and growing hierarchical self-organizing maps with support vector regression in product demand forecasting, Int. J. Prod. Econ. 128 (2010) 603–613.
[24] I. Okumus, A. Dinler, Current status of wind energy forecasting and a hybrid method for hourly predictions, Energy Convers. Manag. 123 (2016) 362–371.
[25] L.N. Elvira, Annual Electrical Peak Load Forecasting Methods with Measures of Prediction Error 2002 Ph.D. Thesis, Arizona State University, Tempe, AZ, USA, 2002.
[26] M.A. Colominas, G. Schlotthauer, M.E. Torres, Improved complete ensemble EMD: a suitable tool for biomedical signal processing, Biomed. Signal Process Contr. 14 (2014) 19–29.
[27] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, N. Yen, et al., The empirical model decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, R. Soc. London Proc. Ser. A 454 (1996).
[28] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition, Adv. Adapt. Data Anal. (2009) (1):1e41.
[29] M. Ali, R. Prasad, Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition, Renewable Sustainable Energy Rev. 104 (2019) 281–295.
[30] C. Li, Z. Zhu, Research and application of a novel hybrid air quality early-warning system: a case study in China, Sci. Total Environ. 626 (2018) 1421–1438.
[31] M.S. AL-Musaylh, R.C. Deo, Y. Li, J.F. Adamowski, Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting, Appl. Energy 217 (2018) 422–439.
[32] J. Wang, S. Xiong, A hybrid forecasting model based on outlier detection and fuzzy time series—a case study on Hainan wind farm of China, Energy 76 (2014) 526–541.
[33] H.J. Sadaei, F.G. Guimarães, C.J. Da Silva, M.H. Lee, T. Eslami, Short-term load forecasting method based on fuzzy time series, seasonality and long memory process, Int. J. Approx. Reason. 83 (2017) 196–217.
[34] Y.S. Chen, C.H. Cheng, W.L. Tsai, Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting, Appl. Intell. 41 (2014) 327–347.
[35] S.T. Li, Y.C. Cheng, Deterministic fuzzy time series model for forecasting enrollments, Comput. Math. Appl. 53 (2007) 1904–1920.
[36] Y.C. Lee, C.H. Wu, S.B. Tsai, Grey system theory and fuzzy time series forecasting for the growth of green electronic materials, Int. J. Prod. Res. 52 (2014) 2931–2945.
[37] Abdullah, L.,0 Taib, I., High order fuzzy time series for exchange rates forecasting. In: Proceedings of the 2011 3rd Conference on Data Mining and Optimization (DMO), Putrajaya, Malaysia, 28–29 June 2011, pp. 1–5.
[38] J.Z. Wang, H. Jiang, Y.J. Wu, Y. Dong, Forecasting solar radiation using an optimized hybrid model by cuckoo search algorithm, Energy 81 (2015) 627–644.
[39] C. Ren, N. An, J.Z. Wang, L. Li, B. Hu, D. Shang, Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting, Knowl.-Based Syst. 56 (2014) 226–239.
[40] S. Mirjalili, Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems, Neural Comput. Appl. 27 (4) (2016) 1053–1073.
[41] S.R. K.S., S. Murugan, Memory based hybrid dragonfly algorithm for numerical optimization problems, Expert Syst. Appl. 83 (2017) 63–78.
[42] S. Mirjalili, A. Lewis, Novel performance metrics for robust multi-objective optimization algorithms, Swarm Evol. Comput. 21 (2015) 1–23.
[43] J. Wang, J. Heng, L. Xiao, C. Wang, Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting, Energy 125 (2017) 591–613.
[44] S. Qin, F. Liu, J. Wang, B. Sun, Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models, Atmos. Environ. 98 (2014) 665–675.
[45] Y. Song, S. Qin, J. Qu, F. Liu, The forecasting research of early warning systems for atmospheric pollutants: a case in Yangtze River Delta region, Atmos. Environ. 118 (2015) 58–69.
[46] Jie Zhang, A.J. Morri., A sequential learning approach for single hidden layer neural networks, Neural Networks 11 (1) (1998) 65–80.
[47] P.J. Gemperline, J.R. Long, V.G. Gregoriou, Nonlinear multivariate calibration using principal components regression and artificial neural networks, Anal. Chem. 63 (20) (1991) 2313–2323.
[48] J. Wang, J. Hu, A robust combination approach for short-term wind speed forecasting and analysis - Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model, Energy 93 (2015) 41–56.
[49] J.J. Liao, R. Liu, Re-parameterization of five-parameter logistic function, J. Chemometr.: J. Chemometr. Soc. 23 (5) (2009) 248–253.
[50] Thomas Nigitz, Markus Gölles, A generally applicable, simple and adaptive forecasting method for the short-term heat load of consumers, Appl. Energy 241 (2019) 73–81.
[51] G. Sudheer, A. Suseelatha, Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models, Int. J. Electr. Power Energy Syst. 64 (2015) 340–346.
[52] Q. Song, B.S. Chissom, Fuzzy Time Series and Its Models, Elsevier North-Holland, Inc., Amsterdam, The Netherlands, 1993.
[53] H.K. Yu, Weighted fuzzy time series models for TAIEX forecasting, Phys. A Stat. Mech. Appl. 349 (2012) 609–624.
[54] Van Veldhuizen Da, Lamont GB. Evolutionary computation and convergence to a pareto front. Late Break Pap Genet Program 1998 Conf 1998, pp. 221–228.
[55] S. Mirjalili, S. Saremi, S.M. Mirjalili, L.D.S. Coelho, Multi-objective grey wolf optimizer: a novel algorithm for multi-criterion optimization, Expert Syst. Appl. 47 (2016) 106–119.
[56] J.R. Schott, OH AIRFIOFTW-PAFB. Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization 1995; 37(1):1–13.
[57] S. Conti, R. Nicolosi, S.A. Rizzo, H.H. Zeineldin, Optimal dispatching of distributed generators and storage systems for MV islanded microgrids, IEEE Trans. Power Delivery 27 (3) (2012) 1243–1251.
[58] C.A.C. Coello, Evolutionary multi-objective optimization: some current research trends and topics that remain to be explored, Front. Comput. Sci. China 3 (1) (2009) 18–30.
[59] P. Ngatchou, A. Zarei, A. El-Sharkaw, Pareto multi objective optimization. In: Proc 13th Int Conf on, Intell Syst Appl to power Syst; 2005. p. 84e91.
[60] C.A.C. Coello, M.S. Lechuga, MOPSO: a proposal for multiple objective particle warm optimization. In: Proc. 2002 Congr. Evol. Comput. Cec 2002, vol. 2; 2002. p. 1051e6.
[61] C.A.C. Coello, G.T. Pulido, M.S. Lechuga, Handling multiple objectives with particle swarm optimization, Evol. Comput. IEEE Trans. 8 (3) (2004) 256–279.
[62] Y. Xu, W. Yang, J. Wang, Air quality early-warning system for cities in China, Atmos. Environ. 148 (2017) 239–257, https://doi.org/10.1016/j.atmosenv.2016.10.046.
[63] F.X. Diebold, R.S. Mariano, Comparing predictive accuracy, J. Bus. Econ. Stat. 13 (3) (1995) 253, https://doi.org/10.2307/1392185.
[64] H. Chen, Q.L. Wan, Y.R. Wang, Refined Diebold-Mariano test methods for the evaluation of wind power forecasting models, Energies 7 (2014) 4185–4198.
[65] F. Herrera, V.E. Herrera, F. Chiclana, Multiperson decision-making based on multiplicative preference relations, Eur. J. Oper. Res. 129 (2001) 372–385.