Energy 167 (2019) 1144–1154

Contents lists available at ScienceDirect

Energy

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

Long-term electricity consumption forecasting based on expert prediction and fuzzy Bayesian theory

Lei Tang*, Xifan Wang, Xiuli Wang, Chengcheng Shao, Shiyu Liu, Shijun Tian

The State Key Laboratory on Electrical Insulation and Power Equipment, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China

ARTICLE INFO

Article history:
Received 7 June 2018
Received in revised form
18 September 2018
Accepted 13 October 2018
Available online 1 November 2018

Keywords:
Long-term
Load forecasting
Annual electricity consumption
Fuzzy Bayes
Expert prediction

ABSTRACT

Long-term electricity consumption (EC) forecasting is a very important part for the expansion planning of power system. Instead of point forecasting, based on fuzzy Bayesian theory and expert prediction, a novel long-term probability forecasting model is proposed to predict the Chinese per-capita electricity consumption (PEC) and its variation interval over the period 2010–2030. The special model structure can improve the reliability and accuracy of expert prediction through econometric methodology. It contains three components: fuzzy relation matrix, prior prediction, and fuzzy Bayesian formula. To contend with the long-term uncertainty, the prior prediction is implemented to combine the advantages of expert's experience with other time-based methods from the perspective of probability. With the utilization of fuzzy technique, the multiple effects of influencing factors (IFs) on PEC can be expressed as a fuzzy relation matrix. It can rule the results of prior prediction to obey the long-run equilibrium relationship of natural evolution thorough probability calibration. To demonstrate its efficiency and applicability, the result of this method is compared with that of other 6 approaches and 4 agencies. The case study shows that the proposed methodology has higher accuracy and adaptability.

1. Introduction

Long-term annual electricity consumption (EC) forecasting plays a crucial role in country's power mix planning and energy development. Accurate and reliable results can effectively improve the performance of planning scheme. However, they are always influenced by many influencing factors (IFs), like economy growth, technology advancement and policy adjustment, which make long-term forecasting a complex task [1]. During the past decade, many efforts at various levels have been made. Some examples are regression analysis, econometrics, fuzzy logic, artificial neural network (ANN) and grey model (GM) [2,3]. Generally, these forecasting techniques can be sorted into three categories: time series models, artificial intelligence models, and uncertainty models.

Time series models are the widely used methods in electricity consumption forecasting. These methods are based on a common assumption that the historical data already discount all correlation factors and their trends in an internal structure [4]. The nature of this internal structure has a practical function which allows us to forecast the future values from current and past values. One of the most popular time series models is auto-regressive moving average (ARMA) [5]. Based on the multi-model partitioning theory, the study in Ref. [6] discussed the application conditions and order selection criteria of ARMA model under the presence of noise. In order to achieve more accurate results, a bootstrap aggregating technique was introduced to combine autoregressive integrated moving average (ARIMA) with exponential smoothing (ES) methods [7]. With improved grey model and ARMA (GM–ARMA) [8], introduced a method to forecast the energy consumption in China which was based on Hodrick-Prescott (HP) Filter. Additionally, time series techniques can be divided into parametric and non-parametric methods. In parameters models, the parameter identification mainly relies on distributional assumptions and predetermined functional form. This predictor can be implemented easily, but the flexibility and fitting results of this model may not be satisfactory in practice. To address these issues, the semi-parametric additive models were proposed to estimate the relationships between demand and the driver factors [9]. Combining the semi-parametric model with bootstrap resampling based estimates and similarity measure algorithm, an improved probability forecasting methodology was presented in Ref. [10]. An intrinsic feature of the time series is that it is highly dependent on past

* Corresponding author.
E-mail address: leitang@stu.xjtu.edu.cn (L. Tang).

https://doi.org/10.1016/j.energy.2018.10.073
0360-5442/

observations, so if there is a lot of variability in the historical data, it may not be the right choice [4].

Artificial intelligent models like ANN can deal with the non-linear problem and give more flexible relations between load and impacting variables [11]. In Ref. [12], to facilitate future predictions of EC, a separate SVR model was created for each of the input variables by using their current and past values, and then these models were combined to yield consumption prediction values. A multilayer perceptron ANN model was addressed in Ref. [13] to predict the problem of Greek long-term energy consumption [14] proposed a generalized regression neural network (GRNN) to deal with the non-linear characteristic of annual power load. Based on improved particle swarm optimization (PSO) and shuffled frog-leaping (SFL) algorithms, the optimal ANN models were developed to investigate the effects of demand side management on accuracy of electrical energy consumption and long-term forecasting in US [15]. The electricity energy demand in Turkey was forecasted by using the ant colony optimization (ACO) technique with independent variables such as gross domestic product (GDP), population, and import and export [16]. Nevertheless, similar to time-series methods, intelligent models have a strong dependence on the training data [17]. Therefore, intelligent models can match the past pattern very well, but it has no idea of the new changes in future.

The uncertainty prediction models that mainly include GM and fuzzy theory are committed to describing the stochastic factors in the prediction process by means of probability or interval-valued [3]. Comparing with the point forecasts, probability (interval) forecasting results can provide an important risk management reference for policymakers when making important decisions on power mix planning [18]. Since GM models have better performance on describing the characteristics of an uncertain system even in the presence of sparse data, many researchers optimized the long-term forecasting methods by using grey theory. A relevant study reported an optimized grey modeling which is applied to predict total electric energy demand of Turkey for 2013–2025 [19]. Combining a new initial condition with rolling mechanism, a novel optimized grey prediction model was designed in Ref. [20] to accurately forecast China's overall and industrial electricity consumption. In Fuzzy theory, the solution to the problem can be expressed as terms which human can understand, so the expert experience prediction can be used for algorithm design [21]. However, the simple fuzzy prediction often fails to meet the requirement of precision, so it needs to be combined with other methods. To deal with the uncertainty of the long-term load, a collaborative principal component analysis and fuzzy feed-forward neural network (PCA-FFNN) approach was reported [22]. In addition, some of the latest techniques such as Bayesian vector auto-regression (BVAR) and Bayesian network (BN) were proposed to improve the performance of probability density forecast methodology [23].

In summary, the primary task for prediction is to capture the inherent characteristics and interactions between per-capita electricity consumption (PEC) and numerous IFs from historical data. However, the fixed parameters of traditional method make the long-term prediction inflexible and difficult to take the future uncertain factors into account. To deal with this inherent drawback, a novel forecasting approach for long-term electricity consumption based on fuzzy Bayes (FB) and expert prediction (EP) is proposed to combine the statistical intrinsic pattern with knowledge-based architecture. Firstly, the EP is implemented to describe the possible future changes with different scenarios by reasoning through bodies of knowledge. Next, a fuzzy long-run equilibrium relationship is developed to modify the EP results from multiple information dimensions through the fuzzy Bayesian formula. Distinct from other Bayes forecasting methods which are mainly adopted to optimize the parameters of forecasting method [24], in this paper, the Bayesian formula is utilized as a framework. It can improve the rationality and flexibility of forecasting result through expert's experience and statistical information. In fact, the key idea of this algorithm can be expressed as fuzzy probability calibration. At last, the effectiveness of the proposed methodology is illustrated with a real case study of China in 2010–2030. To confirm the performance, this method is compared with other well-established forecasting approaches, i.e. LR, ARMA, ES, ANN, GM, and EP. Furthermore, the forecasting results of 4 international agencies are also analyzed in this paper. The comprehensive testing results demonstrate that this algorithm has better accuracy, flexibility, and reliability. The sensitive analysis for Chinese new economic mode can provide more intuitive results to help decision-makers adopt reasonable policies.

2. Methodology

As some factors are not stationary random, but rather nonstationary random like policy, the accurate forecasting becomes a challenging task. Fortunately, interval forecasting based on fuzzy theory has the advantage to take the variability and uncertainty into account. Thus it can reduce the amount of random variation that is relative to classic point series. Fig. 1 shows the appearance of interval forecasting result.

As shown in Fig. 2, the proposed methodology consists of 4 parts, econometric analysis (A), fuzzy relation matrix (B), prior prediction (C) and fuzzy Bayesian formula (D). Firstly, the fuzzy relation matrix is developed to describe the long-run equilibrium relationship between PEC and IFs form the perspective of probability by using multi-regression and fuzzification. Secondly, with fuzzy technique, prior prediction can make the full use of residual series and synthesize the multi-scenarios of expert prediction. Therefore, this model is more flexible to deal with the random variation in future. Finally, the fuzzy Bayesian formula is mainly used to modify the prior probability through the fuzzy relation matrix. The concrete steps of this proposed method can be described in the following sections:

In part A:

Step 1: After unit root test, the cointegration and causal relationships between per capita electricity consumption and IFs are examined.

Step 2: Delete the influencing factor that has no cointegration and causal relationship.

Step 3: Repeat Step 1 and Step 2 until all statistics are significant to obtain the set of influencing factors.

In part B:

Step 4: Based on factors set, the panel model is employed to describe the long-run equilibrium relationships.

Step 5: Fuzzify the equilibrium equation to obtain the fuzzy relation matrix.

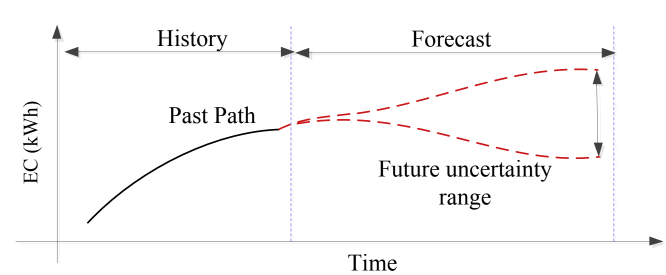


Fig. 1. Interval forecasting and action.

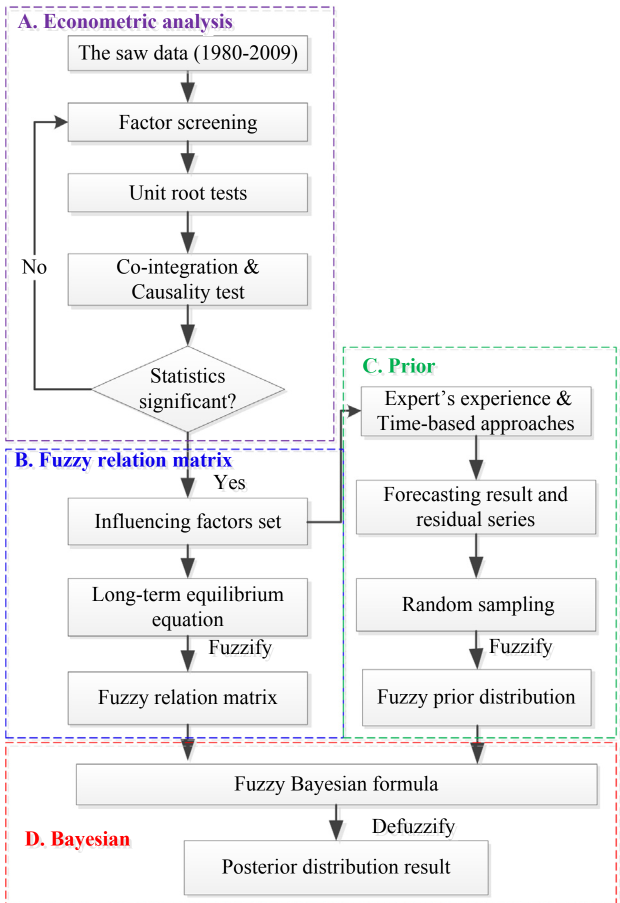


Fig. 2. The flowchart of the proposed methodology.

In part C:

Step 6: The expert's experience and other time-based approaches (to be discussed in Section 5) are adopted to predict the deterministic results of PEC and IFs, respectively.

Step 7: With randomly sampling in these forecasting results and the corresponding residual series, the sample space can be built.

Step 8: After all the samples have been fuzzified and clustered, the Bayesian prior distribution can be constructed to describe the timing-based character of PEC and IFs.

In part D:

Step 9: At last, the posterior distribution can be calculated by combining Bayesian formula with the prior distribution and the fuzzy relation matrix.

3. Econometric analysis and results

The primary objective of this empirical analysis is to study the long-run cointegration and causality relationship between PEC and IFs. On this basis, a range of factors is screened to form the IFs set. Through extensive literature research and analysis, this paper draws up four IFs as the explanatory variables of the multiple regression equation.

1) As reported in Ref. [25], there is a cointegration relationship between electricity consumption and economic growth. It is the main reason to promote electricity consumption. Per capita gross domestic product (PGDP) is adopted to represent economic growth.
2) The electricity intensity (EI) is related to the technological progress and economic structure, which directly reflects the utilization efficiency of electricity. It is calculated as units of electricity per unit of GDP. The decrease of EI will reduce the energy consumption significantly [26].
3) The PEC of urban area is far greater than that in rural area, so the process of urbanization must be accompanied by the change of electricity consumption [27]. Urbanization rate (UR) is equal to the proportion of urban population to total population.
4) The adjustment of the industrial structure (IS) can affect the electricity consumption by changing the way of its usage. In this paper, the proportion of industrial output to GDP is used to mirror the IS.

Compared with time-series, the plane model is more powerful. It allows us to raise the degrees of freedom and study the cross-sectional interdependence with different individual effects. Hence, this paper applies the panel unit root, panel cointegration, and vector error correction model to investigate the relationship between PEC and IFs for 8 countries (China, France, Indian, Japan, Korea, England, USA and Thailand) from 1980 to 2009. In this paper, the data are sourced from World Bank (2018) [28] and all variables and data are expressed in nature logarithm.

3.1. Panel unit root tests

Before conducting the panel cointegration analysis, the panel unit root test must be taken to verify the stationary properties of all variables. In order to enhance the reliability of unit root test result, four widely-used methods, namely Levin-Lin-Chu (LLC) test, Im-Pesaran-Shin (IPS) test, ADF-Fisher test and PP-Fisher test, are adopted to identify the order for each variable in this paper. The null hypothesis of the above four tests is that the variables are non-stationary. If all series are subject to the same differenced order d, it is said to be integrated of order d, i.e. $I(d)$. Table 1 summarizes the results of panel unit root test, where the number inside the bracket is the corresponding probability of t-statistics. It can be seen from Table 1 that all the variables in level form are no stationary under all these four tests. However, after the first-order difference, except LLC test of UR (just a little bit bigger than 0.05), all other statistics significantly reject the null hypothesis of no cointegration at the 0.05 level. Thus, it is reasonable to believe that these five series (*PEC*, *PGDP*, *EI*, *UR* and *IS*) are integrated of order one I(1).

3.2. Panel cointegration tests

After unit root test, the second step is to test whether there is a long-run cointegration relationship among these variables. The equation for Pedroni cointegration tests which allow for cross-section interdependence with different individual effects can be expressed as follows:

$$PEC_{g,t} = \alpha_{g,t} + \delta_g \cdot t + \beta_{PGDP,g} \cdot PGDP_{g,t} + \beta_{EI,g} \cdot EI_{g,t} + \beta_{UR,g} \cdot UR_{g,t} + \beta_{IS,g} \cdot IS_{g,t} + \varepsilon_{g,t} \tag{1}$$

where $g = 1,2,\cdots,N$ represents each country in the pool and $t = 1,2,.\cdots,T$ refers to the time period. The parameters $\alpha_{g,t}$ and δ_g are the fixed effects for each country and deterministic trend effects. $\varepsilon_{g,t}$ is the estimated residual from the long-run relationship. *PEC* represents the per capita electricity consumption, *PGDP* denotes per capita gross domestic product. *EI*, *UR* and *IS* represent the electricity intensity, urbanization rate, and industrial structure, respectively. Since all variables are expressed in natural logarithms, the slope coefficient β can be interpreted as the elasticity of corresponding variables.

$$\varepsilon_{g,t} = \rho_{g,t} \cdot \varepsilon_{g,t-1} + \varsigma_{g,t} \tag{2}$$

If the residual $\varepsilon_{g,t}$ is stationary, the null hypothesis of no cointegration ($\rho_{g,t} = 1$), can be rejected and these variables will be co-integrated. Table 2 shows the results of panel cointegration tests which are composed of panel cointegration tests (4 statistics) and group cointegration tests (3 statistics). In panel cointegration tests, three panel statistics admit that there is a cointegration relationship between PEC and IFs (*PGDP*, *EI*, *UR*, *IS*) at 5% level. In group cointegration tests, two panel statistics, PP and ADF, are quite significant at a 5% level of significance, only rho-stat cannot reject the null hypothesis. As most of the statistics support the existence of cointegration relationship and the rho-statistics has lower power than the PP-statistics [29], so it is reasonable to accept the existence of cointegration relationship among these variables.

3.3. Panel Granger causality tests

Having established the long-run cointegration, the next step is to implement the Granger causality test with an error correction model. The main idea of Granger causality testing is whether the unconstrained equation is more powerful than the constraint one in explaining variables. Thus, the direction of causality between PEC and IFs can be examined in a panel context which is based on the following regressions

$$\Delta PEC_{g,t} = \sum_{\tau} \theta_{11,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{12,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{13,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{14,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{15,g,\tau} \Delta IS_{g,t-\tau} + \phi_{1,g} + \psi_{1,g} ECT_{g,t-1} + \omega_{1,g,t} \tag{3}$$

$$\Delta PEC_{g,t} = \sum_{\tau} \theta_{11,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{12,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{13,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{14,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{15,g,\tau} \Delta IS_{g,t-\tau} + \phi_{1,g} + \psi_{1,g} ECT_{g,t-1} + \omega_{1,g,t} \tag{4}$$

$$\Delta PGDP_{g,t} = \sum_{\tau} \theta_{21,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{22,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{23,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{24,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{25,g,\tau} \Delta IS_{g,t-\tau} + \phi_{2,g} + \psi_{2,g} ECT_{g,t-1} + \omega_{2,g,t} \tag{5}$$

$$\Delta EI_{g,t} = \sum_{\tau} \theta_{31,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{32,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{33,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{34,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{35,g,\tau} \Delta IS_{g,t-\tau} + \phi_{3,g} + \psi_{3,g} ECT_{g,t-1} + \omega_{3,g,t} \tag{6}$$

$$\Delta UR_{g,t} = \sum_{\tau} \theta_{41,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{42,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{43,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{44,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{45,g,\tau} \Delta IS_{g,t-\tau} + \phi_{4,g} + \psi_{4,g} ECT_{g,t-1} + \omega_{4,g,t} \tag{7}$$

$$\Delta IS_{g,t} = \sum_{\tau} \theta_{51,g,\tau} \Delta PEC_{g,t-\tau} + \sum_{\tau} \theta_{52,g,\tau} \Delta PGDP_{g,t-\tau} + \sum_{\tau} \theta_{53,g,\tau} \Delta EI_{g,t-\tau} + \sum_{\tau} \theta_{54,g,\tau} \Delta UR_{g,t-\tau} + \sum_{\tau} \theta_{55,g,\tau} \Delta IS_{g,t-\tau} + \phi_{5,g} + \psi_{5,g} ECT_{g,t-1} + \omega_{5,g,t} \tag{8}$$

where φ is the serially uncorrelated error term with mean zero, ψ is the speed of adjustment. Let Δ denote the first difference operator. θ and τ are the coefficients set and lag length set for difference variables, respectively. The optimal lag length $\tau = 2$ is selected by using the Schawrtz−Bayesian criteria (SBC). The Error correction term ECT could be derived from (1).

$$ECT_{g,t} = PEC_{g,t} - \alpha_{g,t} - \beta_{PGDP,g} \cdot PGDP_{g,t} - \beta_{EI,g} \cdot EI_{g,t} - \beta_{UR,g} \cdot UR_{g,t} - \beta_{IS,g} \cdot IS_{g,t} \tag{9}$$

The source of short-run and long-run causality can be identified by testing the significance of coefficients from (3)–(8). For short-run causality, if the results reject the null hypothesis $\theta_{12,g,r} = 0$ for all g and τ in (4), it indicates that there is Granger causality running from PGDP to PEC. Similarly, if null hypothesis $\theta_{21,g,\tau} = 0$ for (5) is rejected, it means that there is Granger causality running from *PEC* to *PGDP*. Table 3 reports the results of Granger causality tests which incorporate Wald F-test statistics and the corresponding probability. In the short-run, it suggests that there are three bilateral Granger causal relationships between $\Delta PGDP$ and ΔPEC, ΔEI and ΔPEC, ΔIS and ΔPEC at 1%, 5% and 1% levels, respectively. On the contrary, only unidirectional Granger causality running from ΔUR to ΔPEC has been found. The empirical exercise reveals that the increase in PEC can be driven by PGDP, EI, UR and IS in short-run. For long-run causality, we can test the significance of the speed of adjustment, i.e. H0: $\psi_{1,g} = 0$ for all g in (3) or H0: $\psi_{2,g} = 0$ in (4). The causal relationship exists if the null hypothesis is rejected. In Table 3, error correction terms are statistically significant, suggesting that PGDP and UR are the long-run causalities to PEC. Furthermore, it also appears that EI and IS do not have a statistically significant impact on PEC in the long-run. In conclusion, the above empirical results can provide some effective advices on what factors should be taken into account to construct the fuzzy relation

Table 1
The results of panel unit root tests.

Variables		LLC test	IPS test	Fisher-ADF	Fisher-PP
EC	Level	−3.88776 (0.0001)	−1.90395 (0.0285)	32.7896 (0.0079)	27.9511 (0.0320)
	First difference	−4.11549 (0.0000)	−5.71979 (0.0000)	71.7974 (0.0000)	101.134 (0.0000)
PGDP	Level	−1.37590 (0.0844)	1.92217 (0.9727)	14.2750 (0.5782)	31.9635 (0.0101)
	First difference	−9.72928 (0.0000)	−8.01144 (0.0000)	90.4989 (0.0000)	91.2638 (0.0000)
EI	Level	−2.33601 (0.0097)	0.58449 (0.7206)	12.7102 (0.6938)	11.9186 (0.7496)
	First difference	−11.1660 (0.0000)	−9.89350 (0.0000)	114.996 (0.0000)	113.470 (0.0000)
UR	Level	0.96915 (0.8338)	3.82854 (0.9999)	9.31226 (0.9000)	36.5059 (0.0025)
	First difference	−1.54861 (0.0607)	−1.83639 (0.0332)	35.5693 (0.0080)	38.8184 (0.0030)
IS	Level	−3.03798 (0.0012)	−1.52066 (0.0642)	22.6068 (0.1247)	23.2068 (0.1083)
	First difference	−12.2178 (0.0000)	−10.9130 (0.0000)	134.074 (0.0000)	134.849 (0.0000)

Table 2
The results of panel cointegration tests.

	Method	T Statistics
Panel	v-Statistic	3.233382 (0.0006)
	rho-stat	−1.759781 (0.0392)
	PP-stat	−4.356610 (0.0000)
	ADF-stat	−2.627227 (0.0043)
Group	rho-stat	−0.04287 (0.4829)
	PP-stat	−4.482981 (0.0000)
	ADF-stat	−2.004724 (0.0225)

matrix. Specifically, on the basis of co-integration, if there is a long-run Granger causality between PEC and one of IFs, it is reasonable to take this factor into account. If there is no long-run causality, but the regression coefficient has a significance level, it is also acceptable.

4. Fuzzy Bayes algorithm

4.1. Fuzzy relation matrix

To facilitate the presentation of proposed fuzzy Bayesian algorithm later, we only describe the modeling process between PEC and EI, and the others are similar. The introduction of basic concepts is given as follows.

Definition 1. Set $U_{EI}=[l_{EI,\min}, l_{EI,\max}]$ as the universe of discourse for the differential of electricity intensity (ΔEI) where $l_{EI,\ \min}$ and $l_{EI,\max}$ are the minimum and maximum of ΔEI in 8 countries that depend on their own historical data and the multiple regression function.

Fuzzy approaches always employ the designed intervals to construct the fuzzy sets. For example, in Fig. 3 (A1), set $n_{EI}=4$, U_{EI} can be divided into ($2n_{EI}$-1) pieces of equal length based on the size of the universe ($L_{EI}=l_{EI,\max}$-$l_{EI,\min}$), and let $l_{EI,i}$ denote the corresponding section point, $i\in[1,\ 2n_{EI}]$. $u_{EI,1}\cdots\ u_{EI,j}\cdots$ represent the all possible classifications of the universe of discourse, $j\in[1,\ n_{EI}]$, which are drawn with solid lines of different colors. Then it gets

$$U_{EI}=\bigcup_{j=1}^{n_{EI}} u_{EI,j}=\{u_{EI,1},\ u_{EI,2},\cdots,u_{EI,n_{EI}}\} \tag{10}$$

Definition 2. Assume that *f* is the fuzzy membership function of a fuzzy set. It is used to describe the grade of a number belonging to a fuzzy set by mapping it to a probability value in the range from 0 to 1. As shown in Fig. 3 (A1), a trapezoid-shaped membership function is introduced in this paper.

$$f_{EI,j}(\Delta EI)=\begin{cases} 0 & \Delta EI\le l_{EI,2j-2} \\ \dfrac{\Delta EI-l_{EI,2j-2}}{l_{EI,2j-1}-l_{EI,2j-2}}+1 & l_{EI,2j-2}\le \Delta EI\le l_{EI,2j-1} \\ 1 & l_{EI,2j-1}\le \Delta EI\le l_{EI,2j} \\ \dfrac{l_{EI,2j+1}-\Delta EI}{l_{EI,2j+1}-l_{EI,2j}} & l_{EI,2j}\le \Delta EI\le l_{EI,2j+1} \\ 0 & l_{EI,2j+1}\le \Delta EI \end{cases} \tag{11}$$

Similarly, the trapezoid-shaped membership function is also used in Fig. 3 (A2) and U_{PEC} is set as the universe of discourse for the differential of per-capita electricity consumption (ΔPEC), $n_{PEC}=5$.

Definition 3. Let Λ_{EI} denote the fuzzy set of variable ΔEI in U_{EI}. A fuzzy set is a series of pairs ($u_{EI,j}$, $f_{EI,j}$) and it is often denoted by

$$\varLambda_{EI}(\Delta EI)=\frac{f_{EI,1}(\Delta EI)}{u_{EI,1}}+\frac{f_{EI,2}(\Delta EI)}{u_{EI,2}}+\cdots+\frac{f_{EI,n_{EI}}(\Delta EI)}{u_{EI,n_{EI}}} \tag{12}$$

where $f(u_{\mathrm{EI},j})$ denotes the membership degree of the interval $u_{\mathrm{EI},j}$ to the fuzzy set Λ_{EI}, and the symbol "+" indicates which elements are in Λ_{EI} and "−" denotes their corresponding membership degree.

Definition 4. Assume that *G* is the transformation function which can map the membership degree of ΔEI to the universe of discourse for $\varDelta PEC$ (U_{PEC})

Table 3
The results of Granger causality tests.

Independent variables	Dependent variables				
	ΔPEC	*ΔGDP*	*ΔEI*	*ΔUR*	*ΔIS*
ΔPEC	–	14.84853 (0.0000)	3.337619 (0.0374)	1.906325 (0.1512)	17.78410 (0.0000)
ΔGDP	17.75465 (0.0000)	–	36.46152 (0.0000)	3.111908 (0.0466)	0.480832 (0.6189)
ΔEI	3.588616 (0.0293)	37.05309 (0.0000)	–	0.197268 (0.8211)	2.039544 0.1326)
ΔUR	4.577973 (0.0113)	0.545791 (0.5802)	0.348797 (0.7059)	–	0.532447 (0.5879)
ΔIS	15.33786 (0.0000)	0.529884 (0.5895)	1.806938 (0.1667)	0.765724 (0.4663)	–
ECT	12.38207 (0.0005)	5.411060 (0.0210)	0.136161 (0.7125)	8.565237 (0.0038)	0.041052 (0.8396)

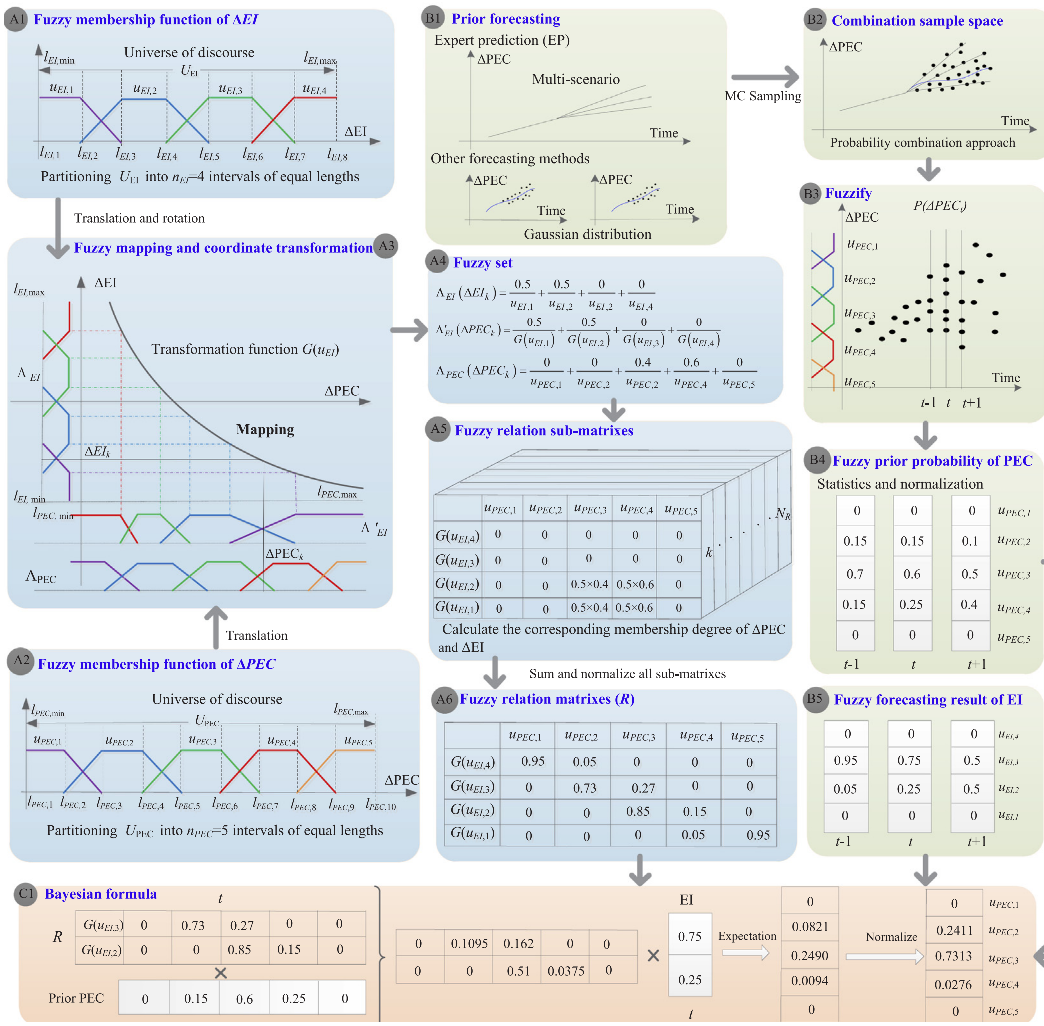


Fig. 3. The processes of fuzzy Bayes algorithm and a simple case study.

$$\Lambda'_{EI}(\Delta PEC) = \frac{f_{EI,1}(\Delta EI)}{G(u_{EI,1})} + \frac{f_{EI,2}(\Delta EI)}{G(u_{EI,2})} + \cdots + \frac{f_{EI,n_{EI}}(\Delta EI)}{G(u_{EI,n_{EI}})} \tag{13}$$

$$G(u_{EI,j}) = \beta_{EI} \cdot u_{EI,j} \tag{14}$$

where Λ'_{EI} is the fuzzy set of ΔEI which is mapped to U_{PEC}. As shown in Fig. 3 (A3), firstly, the fuzzy set Λ_{EI} in U_{EI} is projected on U_{PEC} through transformation function G. Secondly, N_R points are taken evenly from U_{PEC} and the corresponding membership degree is calculated in both Λ'_{EI} and Λ_{PEC}, respectively. For example, the mapping results of k-th point are depicted in Fig. 3 (A4). Thirdly, multiply these two membership degrees (vectors) to form the fuzzy relation sub-matrices which are shown in Fig. 3 (A5). Finally, sum and normalize all the sub-matrices by Eq. (15), and then we can get the fuzzy relation matrix (R).

$$R(u_{EI}/u_{PEC}) = \sum_{k=1}^{N_R} \Lambda'_{EI}(\Delta PEC_k) \cdot \Lambda^T_{PEC}(\Delta PEC_k) \Big/ N_R \tag{15}$$

where k indicates the k-th point in U_{PEC}, Λ'_{EI} and Λ_{PEC} are the column vectors of n_{EI} and n_{PEC}. Normally, a classical relation matrix can only be used to represent a binary relation between a pair of variables like ΔPEC and ΔEI. However, in the fuzzy relation matrix, the element can vary continuously, just as shown in Fig. 3 (A6). This design of fuzzy relation matrix can make this model more efficient to consider the correction effects of IFs on prior distribution in the uncertain environment. Specifically, fuzzy relationship matric reflects the long-run intrinsic pattern of historical data which is constant during the period of study.

4.2. Prior forecasting

The prior forecasting can be either a combination method including EP or just an EP method. Assume that the residual sequence of each method used in prior forecasting obeys a Gaussian distribution. Let μ_t and σ_t denote the expectation and standard deviation of Gaussian distribution in t period. Taking EP as an example, assume that the three multi scenarios (high, middle, low) of EP correspond to 3σ, μ and -3σ, respectively, and then we can calculate the probability density function of EP. On this basis, the Monte Carlo (MC) method is used to sample all the results of various prior forecasting methods, and then the combination sample space of prior distribution $P(\Delta PEC_t)$ in year t can be easily obtained, just as shown in Fig. 3 (B1) and (B2). Next, count the membership degree of all sampling points in different intervals ($u_{PEC,j}$) and normalize these degrees in every period, we can finally get the fuzzy prior distribution (shown in B4 and B5). The vertical line in Fig. 3 (B3) represents the t-th period.

4.3. Bayesian validation

Fig. 3 (C1) provides a simple case study. The fuzzy relation matrix is used to modify the prior forecasting result through the Bayes formula. In the period t, firstly, we multiply the corresponding elements in fuzzy relation matrix (R) and prior distribution of PEC, and then calculate the expectation with all possible states of EI. Finally, normalize the expectation, obtaining Bayesian posterior distribution. Similar to other algorithms, the complexity and computation time increases considerably with the number of IFs. Fortunately, since (1) allows us to analyze the effects of IFs on PEC separately, the fuzzy relation matrix for each influencing factor can be constructed to overcome the curse of dimensionality by using fuzzy addition operation, and then the Bayes formula of multi-IFs can be formulated as (16).

$$P(\Delta PEC_t | u_{PGDP} \cdot u_{EI} \cdots u_{IS}) = \frac{P_s(u_{PGDP} \cdot u_{EI} \cdots u_{IS}) \cdot \left[R_s(u_{PGDP}/u_{PEC_t}) + R_s(u_{EI}/u_{PEC_t}) \cdots R_s(u_{IS}/u_{PEC_t})\right] \cdot P(\Delta PEC_t)}{\sum_{s=1}^{Ns} \left[P_s(u_{PGDP} \cdot u_{EI} \cdots u_{IS}) \cdot R_s(u_{PGDP} \cdot u_{EI} \cdots u_{IS}/u_{PEC_t})\right] \cdot P(\Delta PEC_t)} \tag{16}$$

where Ns is the number of combination states, s denotes one of the combination states of IFs, for example, ($u_{GDP,2}$, $u_{EI,1}$,$u_{UR,3}$,$u_{IS,1}$). For simplicity, in Fig. 3, we consider only one influence factor (EI), and the Ns is equal to 2. The symbol '+' here represents the fuzzy addition operator, $P_s(u_{PGDP} \cdot u_{EI} \cdots u_{IS})$ and R_s are the joint probability and conditional probability of combination state. The conditional probability $P(\Delta PEC_t | u_{PGDP} \cdot u_{EI} \cdots u_{IS})$ denotes the Bayesian posterior distribution of period t. If all elements in fuzzy relation matrix are equal to 1, the proposed method will degrade to the expert prediction. On the other hand, if the Bayesian prior distribution $P(\Delta PEC_t)$ is equal to 1, this approach will behave as a fuzzy multiple regression of the ΔPEC_t. In particular, the Bayesian formula will improve the probability of predicted value which meets the long-run cointegration relationship, while suppressing the others. Finally, fuzzy forecasting result of PEC can be defuzzified with the center-of-gravity (COG) method.

5. Case study

To verify the performance and applicability of proposed FB method, the practical data of the annual PEC, GDP, EI, UR and IS in China from 1980 to 2009 were employed to forecast China's PEC from 2010 to 2030. After all samples that extract from residual distribution are fuzzified and clustered, the prior distribution of IFs is shown in Fig. 4. For ease of presentation, we divide Fig. 4 into four groups, (a), (b), (c) and (d). In these graphs, red scatters indicate the distribution range of the predicted values and dotted lines represent the bounds of 95% confidence intervals.

This paper compares the forecast value in 2010–2016 with the actual data and the percentage errors (PE) of IFs are illustrated in the sub-graph of Fig. 4. It reveals that, in 2030 of China, the expectation of PGDP will reach 18,392 US dollars. With the technical advance, the expectation of EI will drop from 0.645 in 2010 to 0.311 kWh/dollar in 2030. Meanwhile, the UR will increase from 49.2% to 72.7% rapidly. In Fig. 4 (d), as there are great differences in several approaches, the confidence interval of IS is also larger than other IFs. The expectation will slowly decline to 42.5%. Comparatively, except for IS, all actual values of IFs are included in the confidence interval. Besides, the time-based methods that employed in combination forecasting are listed in Table 4.

After Hausman test and F test, an entity fixed effects panel data varying coefficient model is built by using pooled least squares. The coefficient, t-statics and corresponding probability of China are given as follows

$$PEC_t = \underset{7.196(0.0000)}{0.583} \cdot PGDP_t + \underset{3.888(0.000)}{0.431} \cdot EI_t + \underset{4.346(0.000)}{0.91} \cdot UR_t + \underset{0.624(0.533)}{0.16} \cdot IS_t + 13.16 \tag{17}$$

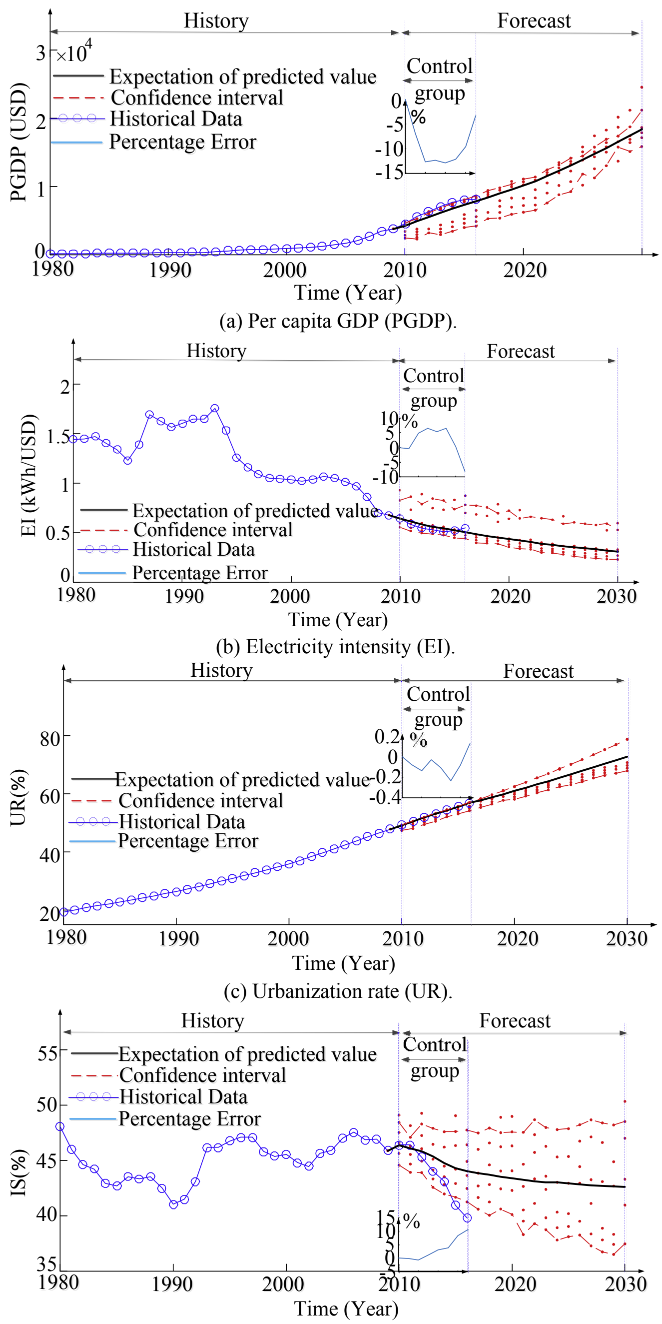
Given the non-significant IS coefficient and a poor forecasting result of IS, it is reasonable to remove IS from this model. Rewriting (17), we have

$$\begin{aligned} PEC_t &= \underset{6.476(0.000)}{0.567} \cdot PGDP_t + \underset{3.420(0.001)}{0.409} \cdot EI_t + \underset{4.051(0.000)}{0.940} \cdot UR_t + 3.304 \\ &\Rightarrow \begin{cases} \Delta PEC_t = 0.567 \cdot \Delta PGDP_t, \ \Delta PEC_t = 0.409 \cdot \Delta EI_t \\ \Delta PEC_t = 0.940 \cdot \Delta UR_t \end{cases} \end{aligned} \tag{18}$$

Eq. (18) shows that the long-run elasticity of PGDP, EI and UR with respect to PEC are 0.567, 0.409 and 0.940, respectively. For example, an increase in PEC by 1% is associated with an increase in PGDP by 0.56%. In addition, the t-statistics are quite significant at a 1% level and R-squared is equal to 99.79%. Next, the long-run co-integration equation (18) can be used to construct the fuzzy relation matrix $R(u/u_{PEC})$ reasonably. In this study, the universe of discourse of PEC is divided into 13 pieces ($n_{PEC} = 13$), while the others are 11, i.e. $n_{PGDP} = 11$, $n_{EI} = 11$, $n_{UR} = 11$, $n_{IS} = 11$. With the increase of n, the algorithm will gradually convert to the deterministic algorithm and the elasticity will become worse.

To demonstrate the advantages of this proposed method, it is compared with most well-known approaches in this field, namely LR, ARMA, ES, ANN, GM and EP. The expert forecasting results of PEC are mainly extracted from a report of Chinese Academy of Sciences (CAS). Since the difference in statistics between National Bureau of Statistics of China (NBSC) and World Bank, the forecasting results of EP based on NBSC are converted into growth rates which are tabulated in Table 5. Assume that growth rates in these three scenarios fellow Gaussian distribution, and then these samples can be included in sample space to form the prior probability of PEC. The expectation of expert prediction (EEP) is indicated in Fig. 5.

In order to facilitate the analysis of forecasting results, Fig. 5 (a) shows the forecasting results of above approaches. The PEC in 2030 is 13120.38, 14892.96, 32021.51, 14083.74, 21661.44, 7000.00 kWh, respectively. As expected, even though these methods can achieve a good forecasting performance for the training data, their performances are not satisfied in long-term forecasting, especially for ES. In Fig. 5 (b), EP, ES and ANN are selected to compose prior prediction. In order to highlight the correction effect of fuzzy relation matrix on prior distribution, the result of ES is also included as a distractor. Comparing the performance of Fig. 5 (b) with Fig. 5 (c), it



(a) Per capita GDP (PGDP).

(b) Electricity intensity (EI).

(c) Urbanization rate (UR).

(d) Industrial structure (IS).

Fig. 4. The combination forecasting results of IFs.

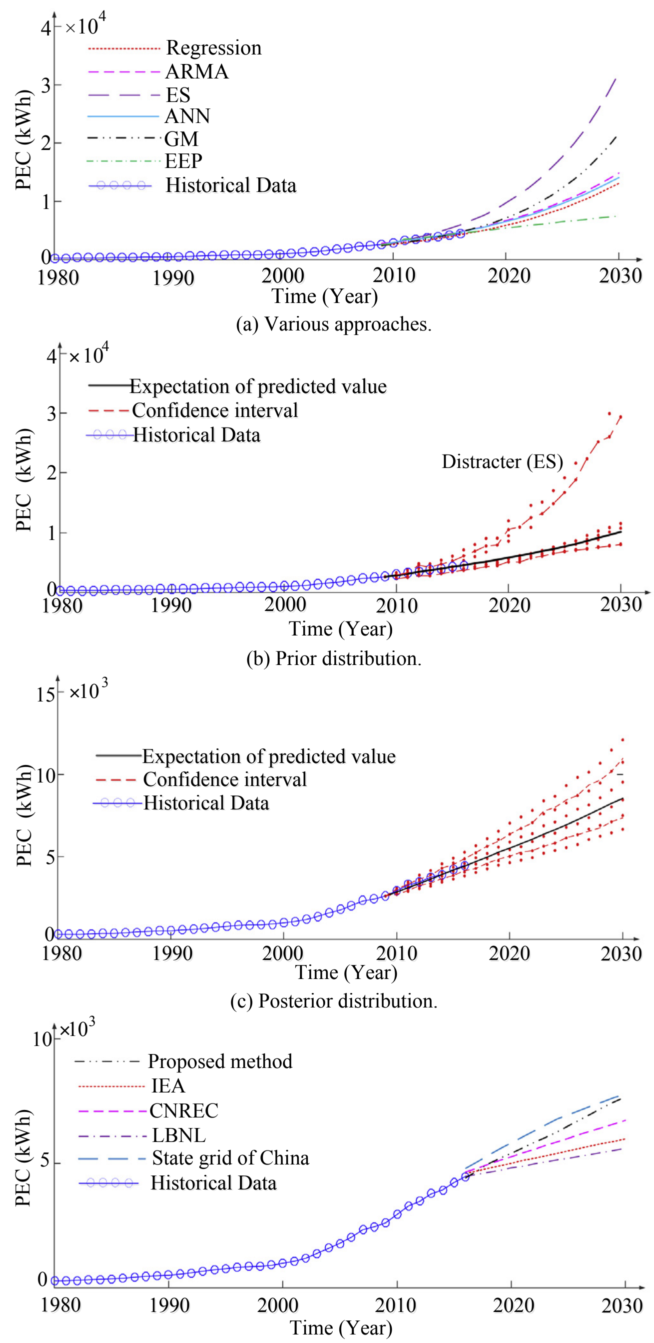
Table 4
The components of prior combination method.

IFs	combination method
(a)PGDP	ARMA and ANN
(b)EI	Nonlinear regression, ES and ANN
(c)UR	Logistic regression and GM
(d)IS	ARMA, GM and ANN

is easy to see that the forecasting results of ES which are observed in prior distribution disappear in posterior distribution. On the contrary, the forecasting results that conform to the long-run cointegration relation will be strengthened. Given the uncertainty surrounding, it is not advisable to follow one single forecast, but rather interval forecasting. Compare with the traditional method in Fig. 5 (a), this method can make a full use of the multi-dimensions information from the perspective of probability.

Table 5
The multi-scenarios forecasting results of experts.

Growth rate	2011	2020	2030
High	9.60%	3.90%	1.89%
Medium	9.60%	3.53%	1.58%
Low	9.60%	3.12%	1.42%



(a) Various approaches.

(b) Prior distribution.

(c) Posterior distribution.

(d) International agencies.

Fig. 5. The forecasting results of PEC.

Table 6
Forecasting results of the proposed methodology for PEC (kWh).

Year	Expectation	Lower bound	Upper bound
2010	2612.46	2612.46	2612.46
2011	2857.24	2698.15	2921.34
2012	3122.14	2957.80	3202.47
2013	3371.88	3178.66	3581.11
2014	3623.59	3416.01	3848.50
2015	3892.49	3671.07	4135.87
2016	4165.70	3867.58	4533.88
2017	4424.55	4156.36	4872.42
2018	4696.81	4378.84	5133.23
2019	4974.90	4613.23	5627.22
2020	5194.92	4860.16	5928.43
2021	5394.78	4920.83	6245.76
2022	5613.20	5082.24	6712.12
2023	5832.42	5248.95	7213.31
2024	6044.66	5421.12	7449.92
2025	6268.49	5598.94	7694.28
2026	6506.21	5782.59	7946.67
2027	6741.38	5972.27	8207.33
2028	6983.59	6168.17	8820.16
2029	7238.32	6245.17	9292.28
2030	7459.97	6450.02	9597.08

Fig. 5 (d) shows the PEC forecasting result of various international agencies, i.e. international energy agency (IEA), China national renewable energy center (CNREC), Lawrence Berkeley national laboratory (LBNL) and State grid of China. The PEC in 2030 is 5979.49, 6724.41, 5592.85, and 7762.41 kWh, respectively [30–32]. Obviously, the forecast results of domestic agencies are higher than that of abroad. It may be caused by the differences in statistical data. As some prediction results are given every 5 and 10 years, the intermediate values in Fig. 5 (d) are filled by linear equation. Hence we can only compare the values in a specific time, like 2020 or 2030. It is not reasonable to compare the prediction performances. Table 6 shows the forecasting results of the proposed method and the 95% confidence interval from 2010 to 2030. The PEC of proposed method is 7459.97 kWh which is close to the forecasting result of state grid.

In this study, different performance measures have been adopted to evaluate the accuracy of the prediction scheme, such as percentage error, mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE).

$$PE(t) = \frac{P_t - A_t}{T} \tag{19}$$

$$MAE = \frac{\sum_{t=2010}^{2016} |P_t - A_t|}{T} \tag{20}$$

$$MAPE = \frac{\sum_{t=2010}^{2016} |P_t - A_t|/A_t}{T} \tag{21}$$

Table 7
The percentage error (2010–2016) (%).

Approach	2010	2011	2012	2013	2014	2015	2016
Proposed	−2.93	−5.33	−2.97	−3.97	−0.88	−1.54	−1.15
Regression	−9.84	−12.8	−10.4	−10.6	−6.97	−6.47	−4.24
ARMA	1.63	−1.7	1.09	0.88	5.04	5.64	8.2
ES	−5.22	0.57	7.54	11.6	20.8	26.4	34.6
ANN	0	0.28	0.28	0	5.29	4.9	6.87
GM	−8.53	−10.4	−6.56	−5.35	0.15	2.48	6.9
EEP	−6.77	5.61	6.73	4.28	5.96	3.69	3.05

Table 8
The forecasting performances (2010–2016).

Approach	MAE (kWh)	MAPE (%)	RMSE (kWh)
Proposed	83.36	2.35	99.75
Regression	316.05	8.77	325.16
ARMA	139.89	3.46	184.76
ES	621.62	15.24	809.64
ANN	105.87	2.51	160.70
GM	206.20	5.77	232.92
EEP	186.64	5.17	189.95

Table 9
The criteria of MAPE.

Threshold	MAPE
1st level (perfect)	<1%
2nd level (well)	<5%
3rd level (acceptable)	<10%
4th level (incapable)	>10%

$$RMPE = \sqrt{\frac{\sum_{t=2010}^{2016} (P_t - A_t)^2}{T}} \tag{22}$$

The percentage error of proposed methodology and other 6 approaches are present in Table 7 and MAE, MAPE and RMSE are listed in Table 8. It is reported that the performance of the proposed method with respect to percentage error is better than those of other approaches, especially for LR, SD and ES. The proposed method can provide a very good fit for the collected data where the MAPE is less than 2.35%, while other methods can reach 8–15%. If MAPE is less than 10%, it is accepted as a successful prediction [19], the criteria are presented in Table 9. Although ANN can better fit training data for short-term prediction, it fails to perform well when forecasting several steps ahead. For example, the result of ANN in 2030 is 14083.74 kWh which is quite different from the expert's experience (about 7000 kWh). It is confirmed that the proposed method is more suitable for long-term forecasting. As expected, the ARMA also has a very good forecasting accuracy and it is quite close to the proposed methodology.

As depicted in Fig. 5 (c) and Table 6, during 2010–2016, it is clear that all the values of actual PEC are in the confidence interval. The results of case study fully demonstrate that the proposed fuzzy Bayes method is more feasible and reasonable in long-term forecasting. Another advantage of adopting panel model is that it can help us to analyze different scenarios of many countries. It is further assumed that the economic development mode of China will transit to that of Japan, France, USA and England gradually from 2010 to 2030. Consequently, a sensitivity analysis for Chinese new

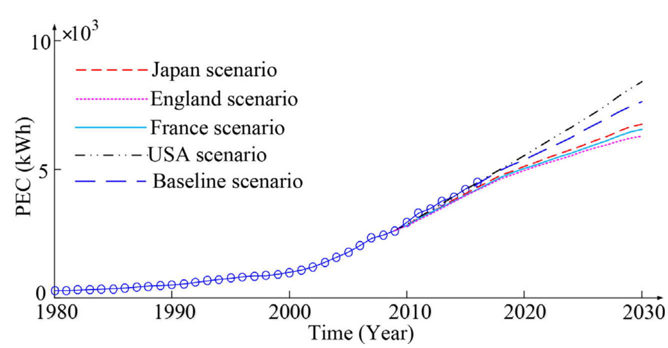


Fig. 6. The multiple scenarios analysis.

Table 10
The forecasting results of multiple scenarios in 2030 (kWh).

Scenario	Expectation	Lower bound	Upper bound
Japan	6757.81	6031.61	8288.88
France	6549.00	5796.63	7655.62
USA	8396.38	6661.59	11619.49
England	6282.13	5682.59	6931.63

Table 11
The percentage error of multi-scenarios (2010–2016) (%).

Scenario	2010	2011	2012	2013	2014	2015	2016
Baseline	−2.93	−5.33	−2.97	−3.97	−0.09	−1.54	−1.15
Japan	−3.62	−6.00	−4.17	−5.58	−3.08	−4.29	−3.89
France	−3.97	−6.66	−4.76	−6.20	−4.22	−5.53	−5.66
USA	−2.69	−4.73	−2.13	−2.67	0.19	−0.37	0.12
England	−4.8	−7.76	−6.10	−7.31	−4.97	−6.23	−6.45

Table 12
The forecasting performances of multi-scenarios (2010–2016).

Scenario	MAE (kWh)	MAPE (%)	RMSE (kWh)
Baseline	83.36	2.35	99.76
Japan	142.12	3.83	155.68
France	173.59	4.62	190.50
USA	54.78	1.61	76.284
England	204.23	5.46	222.96

economic mode can be performed through a dynamically changing fuzzy relation matrix. For example, in the scenario of Japan, the long-run elasticity of PGDP, EI and UR with respect to PEC of China will change linearly from 0.567, 0.409 and 0.940 to 0.445, 0.126 and 0.205. Meanwhile, the fuzzy relation matrixes $R(u/u_{\mathrm{PEC}})$ will also change correspondingly to get different posterior results. The results of four scenarios are shown in Fig. 6 and Table 10. To facilitate comparison of forecasting performances in multi-scenarios later, the baseline scenario is also illustrated in Table 11 and Table 12.

Compared with other scenarios, the PEC in USA is relatively high. Even though it has excellent performance in terms of MAPE, the high growth rate in electricity is unsustainable, especially with economic expansion having moderated to a “new normal” pace. Thus the scenario of Japan may be more reasonable for China. Considering the decoupling between economy and electricity, the growth rate of PEC will be further reduced. This is very important for expansion planning of power system.

6. Conclusions

In this paper, a Bayesian framework is developed to improve the reliability and accuracy of expert prediction through econometric methodology. Given the uncertainty surrounding, it cannot be done by simply deterministic forecast method, but rather probability forecasting which has the advantage of taking into account the uncertainty and variability. The salient features of this proposed approach lie in two aspects:

- The prior distribution can improve the algorithm's robustness by synthesizing the advantages of expert's experience and various traditional methods, such as ANN and GM. Furthermore, the fuzzy processing allows us to incorporate the residual series of these methods into the Bayesian probability model to get interval forecasting results that can reduce the random variation and increase the result's flexibility.
- The fuzzy relation matrix is constructed to modify the prior distribution from the perspective of probability through Bayes formula. In brief terms, we can verify, modify and optimize the prior probabilities with the long-run cointegration relationship between IFs and PEC. This technique, based on the regular pattern of natural evolution and development, can make the expert's result more reasonable from multi-dimensions (various factors).

The case study demonstrates that the proposed method has higher accuracy and better stability for long-term electricity consumption forecasting. The sensitive analysis provides a variety of future scenarios for decision makers. Furthermore, the key idea of this algorithm is probability calibration, it can be implemented in various forecasting scenarios which have long-run equilibrium relationship (linear or nonlinear) between independent variable and dependent variable, such as GDP forecasting, CO2 emission, energy/electricity consumption forecasting.

Acknowledgments

This work is financially supported by the State Grid Science (B3440818K010) & Technology Project and National Natural Science Foundation of China (U1766205).

References

[1] Xia C, Wang J, McMenemy K. Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks. Int J Electr Power Energy Syst 2010;32(7):743–50.
[2] Suganthi L, Samuel AA. Energy models for demand forecasting—a review. Renew Sustain Energy Rev 2012;16:1223–40.
[3] Shao Z, Chao F, Yanga SL, Zhou KL. A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting. Renew Sustain Energy Rev 2017;75:123–36.
[4] Khuntia SR, Rueda1 JL, Meijden M. Forecasting the load of electrical power systems in mid- and long-term horizons: a review. IET Gener, Transm Distrib 2016;10(16):3971–7.
[5] Filik ÜB, Gerek ÖN, Kurban M. A novel modeling approach for hourly forecasting of long-term electric energy demand. Energy Convers Manag 2011;52:199–211.
[6] Pappasa SS, Ekonomoub L, Karamousantasc DC, Chatzarakisb GE, Katsikasd SK, Liatsis P. Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. Energy 2008;33(9):1353–60.
[7] Oliveira EM, Oliveira FLC. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. Energy 2018;144:776–88.
[8] Xu W, Gu R, Liu Y, Dai Y. Forecasting energy consumption using a new GMARMA model based on HP filter: the case of Guangdong Province of China. Econ Modell 2015;45:127–35.
[9] Hyndman RJ, Fan S. Density forecasting for long-term peak electricity demand. IEEE Trans Power Syst 2010;25:1142–53.
[10] Shao Z, Gao F, Zhang Q, Yang SL. Multivariate statistical and similarity measure based semiparametric modeling of the probability distribution: a novel approach to the case study of mid-long term electricity consumption forecasting in China. Appl Energy 2015;156:502–18.
[11] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, Saidur R. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9.
[12] Kavaklioglu K. Modeling and prediction of Turkey's electricity consumption using support vector regression. Appl Energy 2011;88:368–75.
[13] Ekonomou L. Greek long-term energy consumption prediction using artificial neural networks. Energy 2010;35(2):512–7.
[14] Li H, Guo S, Li C, Sun J. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. Knowl Base Syst 2013;37:378–87.
[15] Ardakani FJ, Ardehali MM. Novel effects of demand side management data on accuracy of electrical energy consumption modeling and long-term forecasting. Energy Convers Manag 2014;78:745–52.
[16] Toksari MD. Estimating the net electricity energy generation and demand using the ant colony optimization approach: case of Turkey. Energy Pol 2009;37(3):1181–7.
[17] Azadeh A, Ghaderi SF, Tarverdian S, Saberi M. Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. Appl Math Comput 2007;186:1731–41.
[18] Hu Z, Bao Y, Chiong R. Mid-term interval load forecasting using multi-output

support vector regression with a memetic algorithm for feature selection. Energy 2015;84:419–31.

[19] Coskun H, Huseyin AEs. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 2014;70:165–71.

[20] Ding S, Hipel KW, Dang Y. Forecasting China's electricity consumption using a new grey prediction model. Energy 2018;149:314–28.

[21] Askari S, Montazerin N. A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering. Expert Syst Appl 2015;42:2121–35.

[22] Chen T, Wang YC. Long-term load forecasting by a collaborative fuzzy-neural approach. Electr Power Energy Syst 2012;43:454–64.

[23] Morris P, Vine D, Buys L. Application of a Bayesian Network complex system model to a successful community electricity demand reduction program. Energy 2015;84:63–74.

[24] Henze GP, Pavlak GS, Florita AR, Dodier RH, Hirsch AI. An energy signal tool for decision support in building energy systems. Appl Energy 2015;138:51–70.

[25] Ozturk I, Aslan A, Kalyoncu H. Energy consumption and economic growth relationship: evidence from panel data for low and middle income countries. Energy Pol 2010;38:4422–8.

[26] Liddle B. Revisiting world electricity intensity convergence for regional differences. Appl Energy 2010;87:3218–25.

[27] Elliott RJR, Sun P, Zhu T. The direct and indirect effect of urbanization on electricity intensity: a province-level study for China. Energy 2017;123:677–92.

[28] Worldbank.http://data.worldbank.org/indicator/EG.USE.ELEC.KH. [accessed on 20.03.18].

[29] Wang SS, Zhou DQ, Zhou P, Wang QW. CO2 emissions, energy consumption and economic growth in China: a panel data analysis. Energy Pol 2011;39:4870–5.

[30] International Energy Agency (IEA). World energy outlook. 2014. http://www.iea.org/weo.

[31] Zhou N, Fridley D, Khanna NZ, Ke J, McNeil M, Levine M. China's energy and emissions outlook to 2050: perspectives from bottom-up energy end-use model. Energy Pol 2013;53:51–62.

[32] China National renewable energy center. The energy system for Beautiful China 2050. 2016. http://boostre.cnrec.org.cn.