# Forecasting China's regional energy demand by 2030: a Bayesian approach

Xiao-Chen Yuan [a, b, *], Xun Sun [c, d], Weigang Zhao [b, e], Zhifu Mi [f, *], Bing Wang [b, g], Yi-Ming Wei [b, e, *]

*a. Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China*

*b. Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, China*

*c. School of Geographic Sciences, East China Normal University, Shanghai 200241, China*

*d. Columbia Water Center, Columbia University, New York, NY 10027, USA*

*e. School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China*

*f. School of Construction and Project Management, The Bartlett School, University College London, 1-19 Torrington Place, WC1E 7HB, London, UK*

*g. School of Resources and Safety Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China*

[*] Corresponding author.

E-mail address: xc.yuan@outlook.com (X.-C. Yuan), Z.Mi@uea.ac.uk (Z. Mi), wei@bit.edu.cn (Y.-M. Wei)

***Abstract***

China has been the largest energy consumer in the world, and its future energy demand is of concern to policy makers. With the data from 30 provinces during 1995-2012, this study employs a hierarchical Bayesian approach to present the probabilistic forecasts of energy demand at the provincial and national levels. The results show that the hierarchical Bayesian approach is effective for energy forecasting by taking model uncertainty, regional heterogeneity, and cross-sectional dependence into account. The eastern and central areas would peak their energy demand in all the scenarios, while the western area would continue to increase its demand in the high growth scenario. For the country as a whole, the maximum energy demand could appear before 2030, reaching 4.97/5.25 billion tons of standard coal equivalent in the low/high growth scenario. However, rapid economic development would keep national energy demand growing. It also suggests that most western provinces still have great potential for energy intensity reduction. The energy-intensive industries should be cut down to improve energy efficiency, and the development of renewable energy is essential.

***Keywords:*** *energy demand; model uncertainty; Bayesian; forecast*

## 1 INTRODUCTION

China has been the largest energy consumer in the world, and its future energy demand is of concern to police makers due to the significance for strategic planning. In 2015, China's energy consumption totaled 4.30 billion tons of standard coal equivalent (SCE) of which coal accounted for 64.0%. The desire for strong economic growth as well as the ongoing processes of industrialization and urbanization will contribute to the increased energy use which eventually exerts pressure on the security and environmental issues (Chen et al., 2017; Hao et al., 2015; Jiang and Lin, 2012; Mi et al., 2016). Especially, in recent years, some ambitious carbon reduction targets have been explicitly proposed by China. This implies that more efforts may be needed to control the total amount of energy consumption so as to peak carbon dioxide emissions around 2030 (Mi et al., 2017). As a result, from a policy perspective it is imperative to investigate the potential ranges of energy demand in China (Brockway et al., 2015).

For medium- and long-term energy demand prediction, we argue that there is a need for informative estimates by integrating various information. This can be specified in the following ways. First, the analysis of energy use at the regional level is more useful. The regional pattern of energy demand would help make reasonable and specific policies since there are different situations across regions. Besides, it is suggested by You (2013) that the disaggregated information could improve the accuracy of energy demand forecasts. Second, it is of necessity to detect the uncertainty of energy demand predictions with regard to model estimation and possible adjustments of development polices. The possible range of energy demand could advance policy-making (Shao et al., 2015). Third, a combination of forecasts would make full use of the information carried by individual models, which is assumed to have a better predictive performance. When uncertainty is under consideration, incorporating probabilistic forecasts eventually presents a mixture distribution of energy demand that is supposed to be more reliable.

Previous studies employed various methods for energy forecasting (Suganthi and Samuel, 2012). Table 1 indicates that grey models and statistical models are more concise and less data/parameter-intensive. In particular, statistical approaches are easily applied to the analysis with multi-level information and provide an opportunity to estimate model uncertainty in a formal way. In practice, it is common to predict energy demand on the basis of the developed statistical relationship and the identified driving factors.

**Table 1** Comparisons of energy demand forecasting models

| Classification | Example | Model complexity | Data/parameter requirement |
|---|---|---|---|
| Bottom-up models | MARKAL (Tsai and Chang, 2015)<br>TIMES (Comodi et al., 2012)<br>LEAP (Kumar, 2016) | High level | High level |
| Intelligent models | ANN (Gunay, 2016)<br>PSO (Ünler, 2008)<br>GA (Li et al., 2015) | High level | High level |
| Grey models | GM(1,1) (Hamzacebi and Es, 2014) | Low level | Low level |
| Hybrid models | MPSO-RBF (Yu et al., 2012)<br>GP-GM (Lee and Tong, 2011) | High level | High level |
| Statistical models | ARIMA (Yuan et al., 2016)<br>Econometric model (You, 2013)<br>Semiparametric model (Shao et al., 2015) | Low level | Low level |

At present, traditional statistical techniques in the literature have not considered the uncertainties in the structural relations for energy estimation. Besides, they often use the common coefficient of regions for prediction at the sub-national level without fully accounting for heterogeneity. These problems make it difficult to obtain reasonable ranges of the estimated energy demand. Recent studies indicate that hierarchical Bayesian approach well addresses the uncertainties of model and parameter and provides for partial pooling of the common information from different regions while considering heterogeneity (Gelman and Hill, 2007). Moreover, it could flexibly model the dependence between variables to improve estimation. Therefore, this could help present the informative results of future energy demand so as to give useful insights for energy policies.

This paper aims to forecast China's energy demand and the associated uncertainties at the provincial and national levels. Our study contributes to the existing literature by formally modeling the uncertainties in the structural relations for energy estimation while considering regional heterogeneity and cross-sectional dependence, and offering a predictive distribution of energy demand.

# 2 METHODOLOGY

## 2.1 Influence factors of energy use

The possible influence factors of energy consumption has been extensively investigated in the literature. The major classifications are drawn as follows.

(1) *Economic level*. It shows that economic activity is a major contributor to energy consumption (Liao et al., 2016). Zhang and Xu (2012) examine the causal relationship between energy consumption and economic growth, and find that economic growth causes more energy consumption in China not only at the national level but also at the regional and sectoral levels. Furthermore, some studies indicate that there is a potentially nonlinear effect of economic development on energy consumption (Yoo and Lee, 2010; You, 2013).

(2) *Industrial structure*. There are significant differences in the energy consumed by industries. Especially, heavy industry is a primary consumer. It is commonly viewed that industrialization increases energy consumption (Sadorsky, 2014). However, Li and Lin (2015) find negative effects for both middle-/low-income and high-income groups. This suggests that the change in industrial structure caused by development would affect the pattern of energy use.

(3) *Demographic change*. The demographic factor (e.g. population and age structure) is an essential role considered for energy use in the literature (Liddle, 2014). Liu et al. (2015) find that the negative effect of

population density on energy consumption vary across regions of China. The given interpretation is the result of modernization.

(4) *Urbanization process*. The inconsistent findings exist in the historical studies (Al-mulali et al., 2012; York, 2007). Poumanyvong and Kaneko (2010) show that urbanization decreases energy use in the low-income group, while it increases energy use in the middle- and high-income groups. The reduction is interpreted as the effects of fuel switching from inefficient traditional fuels to efficient modern fuels. However, development raises the use of private and public infrastructure so that more energy resources are required to support urban population and urban economies.

(5) *Technological progress*. The advancement of technology has impacts on energy efficiency and energy structure. These are essential for energy consumption. To cope with climate change, there is a need of new technologies to change the pattern of energy use in the future.

Based on the identified influence factors, the research framework of this study for forecasting regional energy demand in China is shown as Figure 1. The causal effects of influence factors on energy consumption are constructed by hierarchical Bayesian approach which accounts for the uncertainties in the structural relations with regional heterogeneity and cross-sectional dependence. The estimated region-specific regression coefficients are used to obtain the energy demand predictions with uncertainty bounds. On the basis of individual models, the mixed probabilistic forecasts for energy demand with the specified development scenarios are made. We attempt to investigate the regional and national patterns of energy demand, the changes in energy intensity, and the impacts of energy structure adjustment.

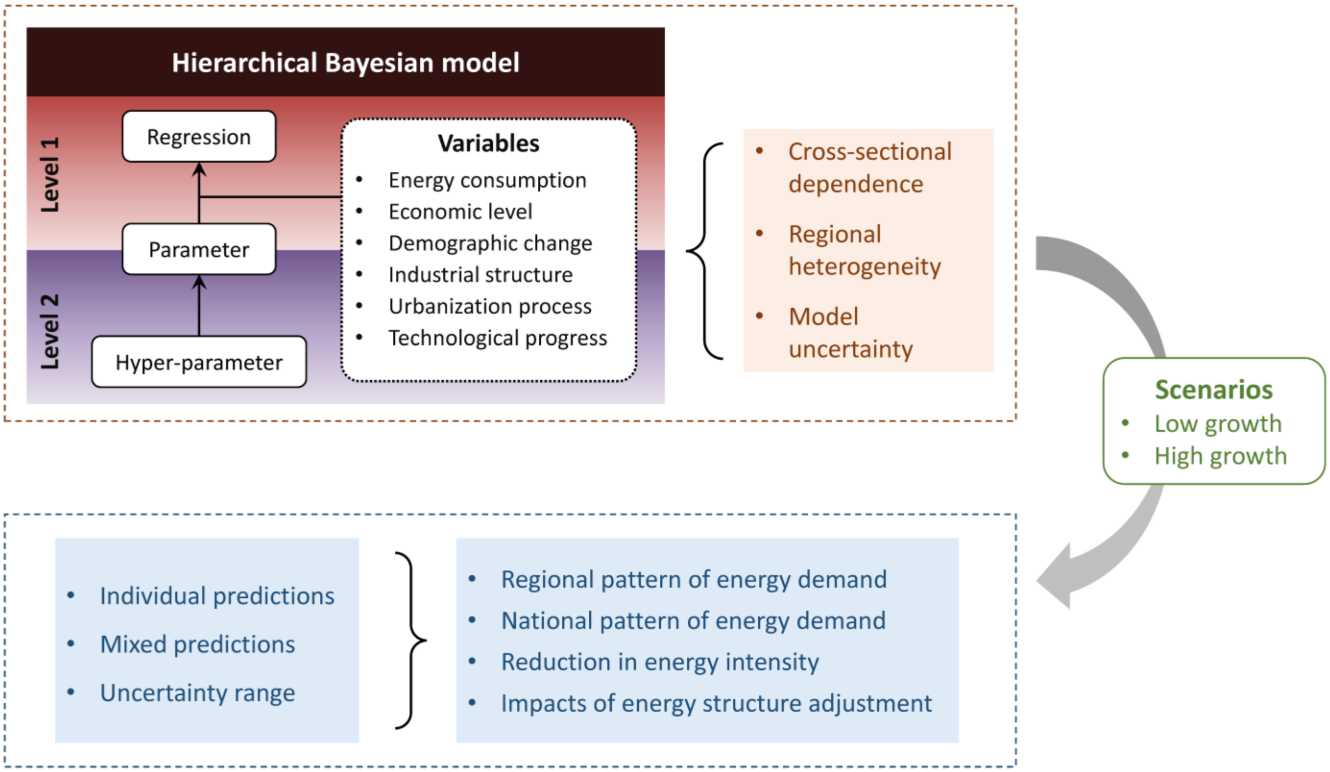


**Figure 1** Research framework for forecasting regional energy demand in China

## 2.2 Hierarchical Bayesian model

The empirical model for energy consumption per capita is shown as Eq. (1). For the $g$th group ($g$=1,2,…$G$) of $S^{(g)}$ provinces in year $t$, the energy consumption per capita $(y_{1t}, y_{2t}, ..., y_{S^{(g)}t})$ (log transformed) is modelled with a multivariate normal distribution which considers the dependence across provinces in group $g$.

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \cdots \\ y_{S^{(g)}t} \end{pmatrix} \sim \mathrm{MVN}\begin{pmatrix} \begin{matrix} \alpha_1 + x_{1,1t}\beta_{1,1} + x_{2,1t}\beta_{2,1} + \ldots + x_{J,1t}\beta_{J,1} \\ \alpha_2 + x_{1,2t}\beta_{1,2} + x_{2,2t}\beta_{2,2} + \ldots + x_{J,2t}\beta_{J,2} \\ \cdots \\ \alpha_{S^{(g)}} + x_{1,S^{(g)}t}\beta_{1,S^{(g)}} + x_{2,S^{(g)}t}\beta_{2,S^{(g)}} + \ldots + x_{J,S^{(g)}t}\beta_{J,S^{(g)}} \end{matrix}, \mathbf{\Sigma}_g \end{pmatrix} \quad (1)$$

where $\mathbf{x}_{st} = (x_{1,st}, x_{2,st}, ..., x_{J,st})$ is a set of $J$ explanatory variables associated with energy consumption per capita of province $s$ ($s$=1,2,…$S^{(g)}$) in year $t$. The regression coefficients $\boldsymbol{\beta}_s^{(g)} = (\beta_{1,s}, \beta_{2,s}, ..., \beta_{J,s})$, the intercepts $\boldsymbol{\alpha}_s^{(g)} = (\alpha_1, \alpha_2, ..., \alpha_s)$ and the covariance matrix $\mathbf{\Sigma}_g$ for group $g$ all need to be estimated. If there are $G$ groups in total, the coefficients for each province can be denoted by $\boldsymbol{\beta}_{s'} = (\beta_{1,s'}, \beta_{2,s'}, ..., \beta_{J,s'})$ with $s' = 1, 2, ..., (S^{(1)} + S^{(2)} + ... + S^{(G)})$. To describe the spread of covariate effects across all provinces, another multivariate normal distribution is applied to the regression coefficients. This indicates the second level of the hierarchical Bayesian model, and the equation is shown as follows (Chen et al., 2014; Devineni et al., 2013).

$$\boldsymbol{\beta}_{s'} \sim \mathrm{MVN}(\boldsymbol{\mu}_{\beta}, \boldsymbol{\Sigma}_{\beta}) \quad (2)$$

where $\boldsymbol{\mu}_{\beta}$ (a vector of length $J$+1) represents the common mean regression coefficients for all the provinces from $G$ groups; correspondingly, $\boldsymbol{\Sigma}_{\beta}$ is the covariance matrix. If the estimated variances of $\boldsymbol{\beta}_{s'}$ (diagonal of $\boldsymbol{\Sigma}_{\beta}$) are large, then it tends towards a no-pooling model where each province is regressed independently; by contrast, the small variances imply a full pooling model with homogeneous responses to the influencing factors (Gelman and Hill, 2007). We apply uninformative priors to the parameters $\boldsymbol{\Sigma}_{g}$, $\boldsymbol{\alpha}_{\beta}$, $\boldsymbol{\mu}_{\beta}$, and $\boldsymbol{\Sigma}_{\beta}$, and use Markov Chain Monte Carlo (MCMC) sampling to estimate posterior distributions. The convergence of the MCMC chain is evaluated by the potential scale reduction factor (Gelman and Rubin, 1992), and all the calculations are conducted by R and RStan (Stan Development Team, 2016).

Considering the effects of economic development, industrialization, and urbanization, the explanatory variables in Eq. (1) are selected as gross domestic product per capita, share of secondary industry, and urbanization rate (the share of urban population in the total population). In addition, the quadratic term of GDP per capita is introduced into the model to detect the nonlinear relationship between energy consumption and economic development. Also, the lagged energy consumption per capita is involved to establish dynamic models.

## 2.3 Model validation

Since our object is to extrapolate energy demand in the future, it is important to validate the models' out-of-sample forecast performance. Accordingly, motivated by the typical leave-one-out cross-validation (LOOCV) (James et al., 2013), we use the following root mean squared error (RMSE) criterion to measure the out-of-sample performance of energy demand forecasting models considered in this study:

$$\text{RMSE} = \frac{1}{T}\sum_{t=1}^{T}\sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{it} - y_{it})^2}{n}} \qquad (3)$$

where $T$ and $n$ are the numbers of years and provinces respectively. $y_{it}$ ($i$=1,2,…,$n$) is the actual energy consumption of the $i$th province in the $t$th year, while $\hat{y}_{it}$ is its corresponding forecast which is obtained from the following procedure: hold out the observations of all studied provinces in the $t$th year (i.e. $y_{1t}, y_{2t}, K, y_{nt}$) at first; then re-estimate the model on the remaining observations; finally, use this estimated model to obtain the required forecasts $\hat{y}_{1t}, \hat{y}_{2t}, K, \hat{y}_{nt}$. Clearly, this error can be reasonably used as a measure for the out-of-sample forecast performance. The smaller error indicates higher forecast accuracy.

## 2.4 Data and scenarios

### 2.4.1 Data description

This study takes 30 provinces (including municipalities and autonomous regions) of China as a study area, and they are divided into three groups (Table A1). The annual data of provinces during 1995-2012 are collected from China Statistical Yearbooks, provincial Statistical Yearbooks, and China Energy Statistical Yearbooks, including GDP, population, urbanization rate, energy consumption, and share of secondary industry. Note that GDP is converted into 2010 price (Chinese Yuan, CNY). Table 2 presents the variables used for models and their descriptive statistics, and all the observations are taken for our analysis.

**Table 2** Variables for models and descriptive statistics

| Variable | Definition | Observations | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|---|
| GDPPC | GDP per capita (2010 CNY, thousands) | 540 | 19.69 | 15.53 | 2.91 | 86.50 |
| SEC | Share of secondary industry (%) | 540 | 0.46 | 0.08 | 0.20 | 0.62 |
| URB | Urbanization rate (%) | 240 | 0.50 | 0.14 | 0.27 | 0.89 |
| ENGP | Energy consumption per capita (tons of SCE) | 540 | 2.16 | 1.32 | 0.42 | 7.95 |

### 2.4.2 Scenario assumptions

The projections of provincial GDP, population, urbanization rate, and share of secondary industry by 2030 are made based on the national projections from previous studies and some assumptions. Table 3 shows the national projections in some specific years, and basically there are two scenarios designed to describe the possible development. This paper intends to present the full range of energy demand, so the low and high growth scenarios are adopted. The details for developing provincial scenarios are introduced as follows.

**Table 3** Scenario assumptions for China

| | S1 (low growth) | | | S2 (high growth) | | |
|---|---|---|---|---|---|---|
| | 2020 | 2025 | 2030 | 2020 | 2025 | 2030 |
| GDP (trillions CNY) | 90.4 | 118.1 | 147.2 | 92.6 | 123.9 | 158.1 |
| Population [a] (millions) | 1390.5 | 1385.6 | 1367.3 | 1415.2 | 1444.2 | 1463.8 |
| Share of secondary industry (%) | 38.9 | 35.7 | 32.6 [b] | 39.8 | 37.4 | 35.0 |
| Urbanization rate (%) | 58.6 | 61.8 | 65.0 | 61.0 [c] | 65.4 [c] | 68.7 [c] |

Sources: a. World Population Prospects: The 2015 Revision (United Nations, 2015)

b. China 2030 (Hu et al., 2014)

c. World Urbanization Prospects: The 2014 Revision (United Nations, 2014)

The GDP growth rates for the whole country in these scenarios refer to those in the World Energy Outlook (2015). Specifically, the growth rate for scenario S1 (S2) is 6.0% (6.5%) in 2015-2020, 5.5% (6.0%) in 2021-2025, and 4.5% (5.0%) in 2026-2030. The structure of provinces' growth rate is assumed to be the same as that in 2014. Accordingly, the future annual GDP of provinces are obtained.

There are small changes in the share of provincial population in the national population over the past years, and thus the one in 2014 is taken to allocate the national population.

The smaller reduction in the share of secondary industry is made in scenario S2. On the basis of the assumed values in Table 3, the annual national projections are linearly interpolated. Furthermore, the structure of provinces' share of secondary industry in 2014 is taken to obtain annual provincial projections.

A comparatively small increase in urbanization rate is set in scenario S1. We linearly interpolate the values over the period to get annual urbanization rate of China. The share of provincial urban population in the national urban population was stable over the past years, and thus the one in 2014 is taken to calculate each province's projections of urbanization rate.

# 3 RESULTS

## 3.1 Empirical models for energy demand

The static and dynamic models for energy demand are established with various explanatory variables. The results estimated by fixed effects method and hierarchical Bayesian method are both made to reveal their differences. Note that the province fixed effects are only considered in the fixed effects estimation, and the common mean coefficients for all provinces in the hierarchical Bayesian model are taken for comparisons. The coefficient whose 90% interval of posterior distribution does not overlap with 0 is regarded to have significant effect.

The estimated regression coefficients of static models are shown in Table 4. The share of secondary industry is considered in all the models, and the significant positive impacts are found in model M1 by both fixed effects method and hierarchical Bayesian method. Urbanization rate is introduced into model M2 and M3, and the results suggest that energy consumption would increase with urbanization effect. However, the regression coefficients are not statistically significant. It also reveals that economic development would raise energy use, and particularly a significant nonlinear effect is indicated by two methods in model M3. Note that there are different situations for the provinces. Figure B1 displays the posterior distributions of the regression coefficients for each province in model M3. We notice that most provinces have significant nonlinear relationship between energy consumption and economic development. This implies that the energy demand is expected to decrease with further economic development, and the turning point varies from region to region. On the other hand, a time trend (TIME) is put into the models to represent technological effect. The significantly negative coefficients suggest that energy consumption would decrease with time.

The estimated regression coefficients of dynamic models are given in Table 5. In general, the current energy consumption is positively correlated to that in a former period. Yet, the insignificant coefficient is found by hierarchical Bayesian method in model M6. Also, it shows that industrialization and urbanization have positive impacts on energy consumption, though the regression coefficients estimated by two methods are not always statistically significant. In addition, model M6 attempts to investigate the nonlinear effect of economic development. But Figure B2 shows that there is no significant relationship. We notice that the regression coefficients of time trend estimated by two methods in model M4 are contrary to each other. Specifically, the positive rather than negative impact of time trend is indicated by fixed effect method.

**Table 4** The estimated regression coefficients of static models

| | M1 | | M2 | | M3 | |
|---|---|---|---|---|---|---|
| | FE | HB | FE | HB | FE | HB |
| Ln(GDPPC ) | 0.816***<br>(0.070) | 1.091<br>[0.928, 1.289] | 0.892***<br>(0.058) | 1.054<br>[0.984, 1.134] | 1.036***<br>(0.083) | 1.176<br>[1.046, 1.308] |
| SEC | 1.489***<br>(0.169) | 0.258<br>[0.030, 0.508] | 0.226**<br>(0.105) | 0.091<br>[-0.013, 0.193] | 0.166<br>(0.106) | 0.076<br>[-0.026, 0.186] |
| URB | | | 0.280<br>(0.211) | 0.179<br>[-0.015, 0.374] | 0.250<br>(0.209) | 0.194<br>[-0.004, 0.384] |
| $Ln(GDPPC)^2$ | | | | | -0.026**<br>(0.011) | -0.019<br>[-0.039, -0.002] |
| TIME | -0.016**<br>(0.007) | -0.040<br>[-0.062, -0.021] | -0.035***<br>(0.006) | -0.054<br>[-0.063, -0.046] | -0.032***<br>(0.006) | -0.054<br>[-0.062, -0.045] |
| RMSE | 0.117 | 0.080 | 0.030 | 0.026 | 0.030 | 0.024 |
| Observations | 540 | 540 | 240 | 240 | 240 | 240 |

Note: FE indicates the fixed effects method while HB indicates the hierarchical Bayesian method. The medians of common mean regression coefficients and the associated 5-95% uncertainty bounds (in square brackets) of HB model are presented. The standard errors are given in the parentheses for the regression coefficients of FE model.

* indicates significance at 10% level

** indicates significance at 5% level

*** indicates significance at 1% level

The RMSE of model forecasts is calculated in Table 4 and Table 5 for comparison. It can be found that all the models estimated by hierarchical Bayesian method have smaller RMSE which means better forecast performances. This is partially because the region-specific coefficients (Figure B1 and B2) are provided by hierarchical Bayesian method. These are essential for estimating future energy demand which needs to fully account for regional heterogeneity. Besides, the cross-sectional dependence is also introduced into the models which could improve estimation. Our study finally intends to give informative results of energy demand, accordingly, we retain the variables of insignificant effects since they can still provide some information for the posterior distribution of energy demand.

**Table 5** The estimated regression coefficients of dynamic models

| | M4 | | M5 | | M6 | |
|---|---|---|---|---|---|---|
| | FE | HB | FE | HB | FE | HB |
| Ln(GDPPC) | 0.085* | 0.521 | 0.596*** | 1.030 | 0.738*** | 1.121 |
| | (0.047) | [0.403, 0.642] | (0.066) | [0.932, 1.120] | (0.084) | [0.984, 1.275] |
| SEC | 0.671*** | 0.258 | 0.119 | 0.090 | 0.059 | 0.078 |
| | (0.100) | [0.061, 0.472] | (0.094) | [-0.025, 0.194] | (0.095) | [-0.029, 0.182] |
| URB | | | 0.102 | 0.156 | 0.073 | 0.175 |
| | | | (0.189) | [-0.053, 0.369] | (0.187) | [-0.035, 0.377] |
| Ln(ENGP(-1)) | 0.749*** | 0.523 | 0.358*** | 0.043 | 0.358*** | 0.035 |
| | (0.025) | [0.436, 0.606] | (0.049) | [0.000, 0.089] | (0.048) | [-0.008, 0.083] |
| $\text{Ln(GDPPC)}^2$ | | | | | -0.026*** | -0.016 |
| | | | | | (0.010) | [-0.036, 0.002] |
| TIME | 0.011*** | -0.017 | -0.026*** | -0.054 | -0.024*** | -0.052 |
| | (0.004) | [-0.026, -0.007] | (0.005) | [-0.063, -0.044] | (0.005) | [-0.061, -0.043] |
| RMSE | 0.063 | 0.061 | 0.029 | 0.026 | 0.028 | 0.025 |
| Observations | 510 | 510 | 210 | 210 | 210 | 210 |

Note: FE indicates fixed effects method while HB indicates hierarchical Bayesian method. The medians of common mean regression coefficients and the associated 5-95% uncertainty bounds (in square brackets) of HB model are presented. The standard errors are given in the parentheses for the regression coefficients of FE model.

* indicates significance at 10% level

** indicates significance at 5% level

*** indicates significance at 1% level

## 3.2 Forecasting regional energy demand by 2030

The posterior distributions of energy demand estimated by the models (M2, M3, M5, and M6) reflecting the integrated effects of economic development, industrialization, and urbanization are adopted. Then, these are mixed to present the probabilistic forecasts. The provincial energy demand with uncertainty bound during 2016-2030 is presented in Figure 2. There are different situations across provinces due to the various development stages.

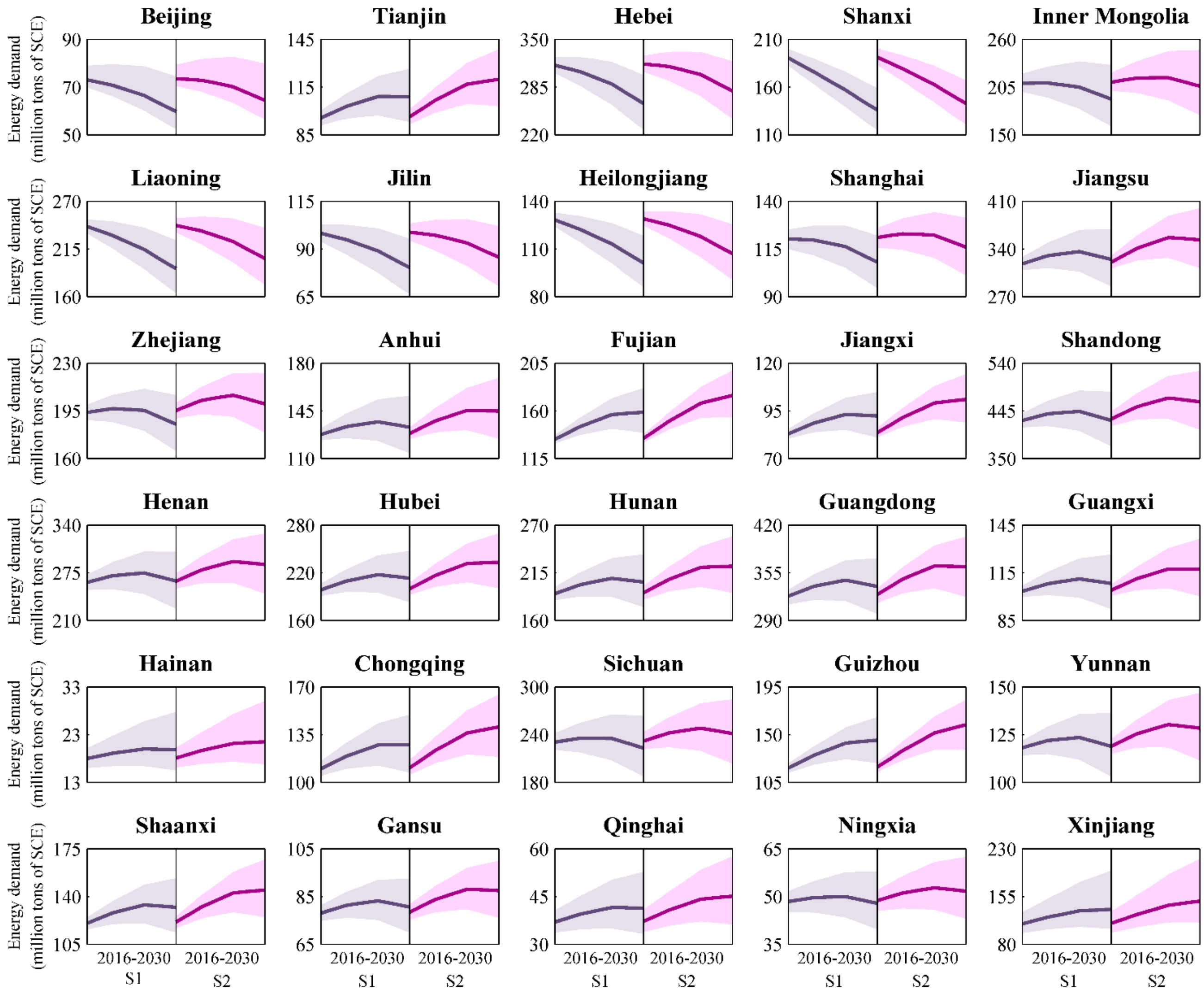


**Figure 2** Provincial energy demand during 2016-2030 in scenario S1 and S2. The lines indicate median value while the range indicates 2.5-97.5% uncertainty.

The ongoing economic transformation in China makes slower economic development in Hebei, Heilongjiang, Jilin, Liaoning and Shanxi, so that the desire for energy is expected to continuously decrease in the future. From 2016 levels by 2030, the amount of energy use (median value) in scenario S1/S2 would reduce by 16.6%/11.6% in Hebei, 21.1%/17.0% in Heilongjiang, 18.3%/13.2% in Jilin, 20.1%/15.7% in Liaoning, and 28.6%/25.4% in Shanxi. Also, there are significant reductions in Beijing, Inner Mongolia, and Shanghai.

In contrast, the energy demand in Fujian, Guizhou, and Xinjiang would keep rising in both the scenarios, and during 2016-2030 their increments (median value) in scenario S1/S2 are 19.2%/30.1%, 22.1%/33.4%, and 20.3%/31.0% respectively. Different from the continuous increase in scenario S2, there seems to be a flat after the increase in scenario S1 for Chongqing, Hainan, Qinghai, and Tianjin, or even a reduction for Jiangxi and Shaanxi.

The obvious turning points are found in Gansu, Henan, Jiangsu, Ningxia, Shandong, Sichuan, Yunnan, and Zhejiang in the two scenarios. Yet, the turning point would appear only in scenario S1 for Anhui, Hubei, Hunan, Guangdong, and Guangxi. Their energy demand seems stable after increment in scenario S2.

The regional energy demand is displayed in Figure 3. There are three groups in total, namely eastern, central, and western areas (Table A1). It can be seen that eastern and central areas could peak their energy demand in the scenarios. However, western area would keep energy demand (median value) growing to 1.27 billion tons of SCE by 2030 in scenario S2. The western provinces which are less developed generally have higher growth of economy in recent years, and they are also assumed to own faster economic development in the future. As a result, more energy is required.

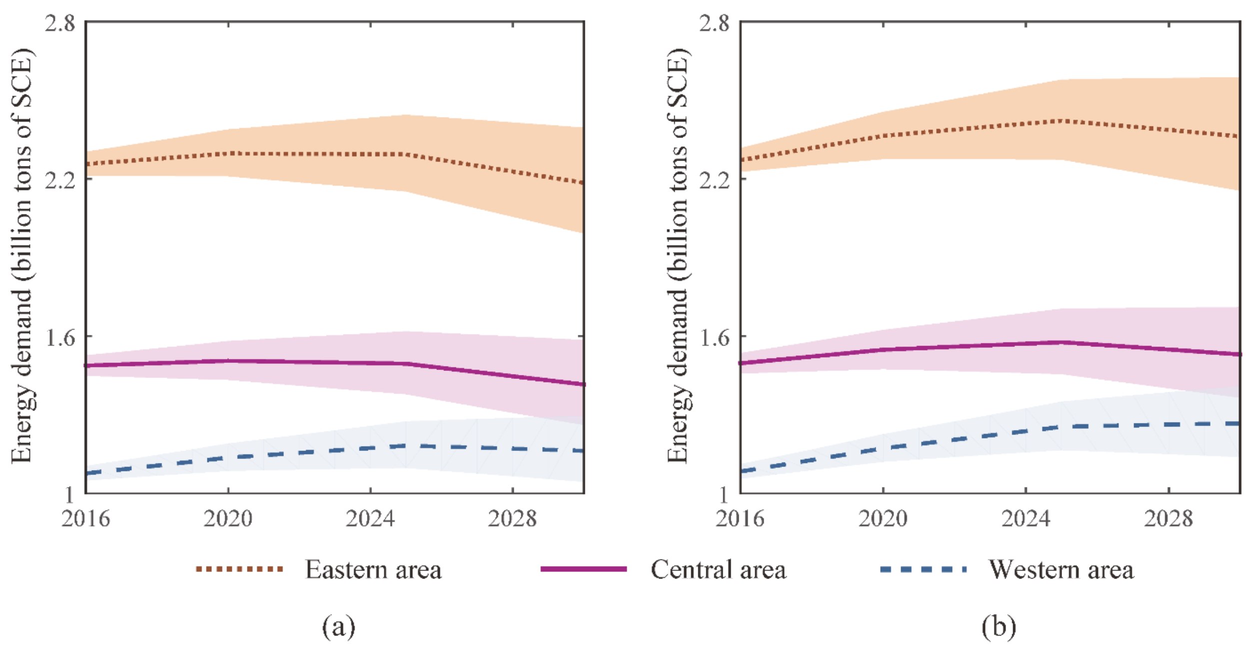


**Figure 3** Regional energy demand during 2016-2030 in scenario (a) S1 and (b) S2. The lines indicate median value while the range indicates 2.5-97.5% uncertainty.

## 3.3 Forecasting national energy demand by 2030

Based on the provincial estimates, the national energy demand and the associated uncertainty bound are shown in Figure 4. Here, the distributions predicted by the selected individual models and the mixture of distributions are all presented for comparison. Basically, these individual models indicate that the total energy demand (median value) could reach the peak in the scenarios. Specifically, model M2 suggests higher demand (median value) than other models, reaching 5.04/5.33 billion tons of SCE by 2025 in scenario S1/S2. Comparatively, the smaller peak (median value) of 4.94/5.21 billion tons of SCE in scenario S1/S2 is found by model M6. By aggregating the predicted distributions of all individual models, the mixed distribution shows that the maximum demand (median value) could rise to 4.97/5.25 billion tons of SCE in scenario S1/S2. However, the upper uncertainty bounds in scenario S2 suggest that the energy demand would keep growing. The mixed predictions show that it could be as much as 5.67 billion tons of SCE in 2030.

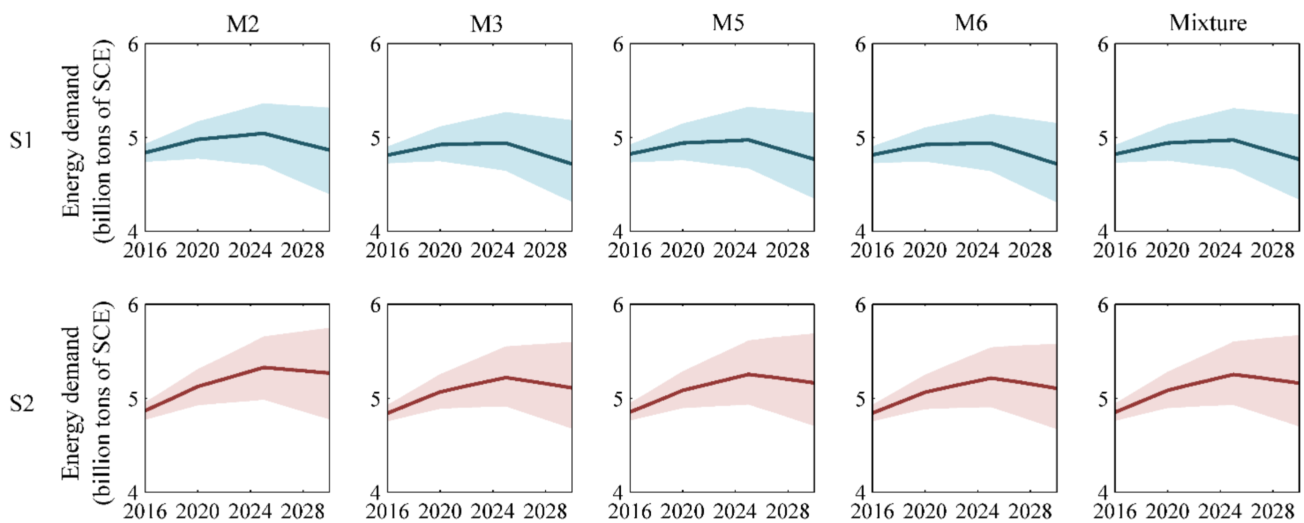


**Figure 4** National energy demand predicted in scenario S1 and S2 by individual models (M2, M3, M5 and M6) and the mixed predictions. The line indicates median value while the range indicates 2.5-97.5% uncertainty.

We also attempt to investigate the changes in energy demand with a focus on the effect of economic development. Accordingly, three economic development scenarios (Table 6) are established. For the prediction, other variables such as population, urbanization rate, and share of secondary industry are all consistent with those in scenario S1. The growth rates of GDP in the medium and low economic growth scenarios are the same as those in scenario S2 and S1. Since the government aims to make the average GDP growth higher than 6.5% during the 13th Five-Year Plan of China (2016-2020), we set the growth rate as 6.8% in the high economic growth scenario. Besides, the slower economic decline is also assumed during 2026-2030 for the high economic growth scenario. As shown in Figure 5, the energy demand (median value) would increase to 5.21 and 4.97 billion tons of SCE in the medium and low economic growth scenarios, respectively, and decrease to 5.12 and 4.76 billion tons of SCE by 2030. However, the energy demand (median value) in the high economic growth scenario is likely to increase continuously to 5.32 billion tons of SCE.

**Table 6** Scenario assumptions for GDP growth rate

| Scenario | 2016-2020 | 2021-2025 | 2026-2030 |
|---|---|---|---|
| High economic growth | 6.8% | 6.0% | 5.5% |
| Medium economic growth | 6.5% | 6.0% | 5.0% |
| Low economic growth | 6.0% | 5.5% | 4.5% |

Note: other variables in the three scenarios are the same as those in scenario S1.

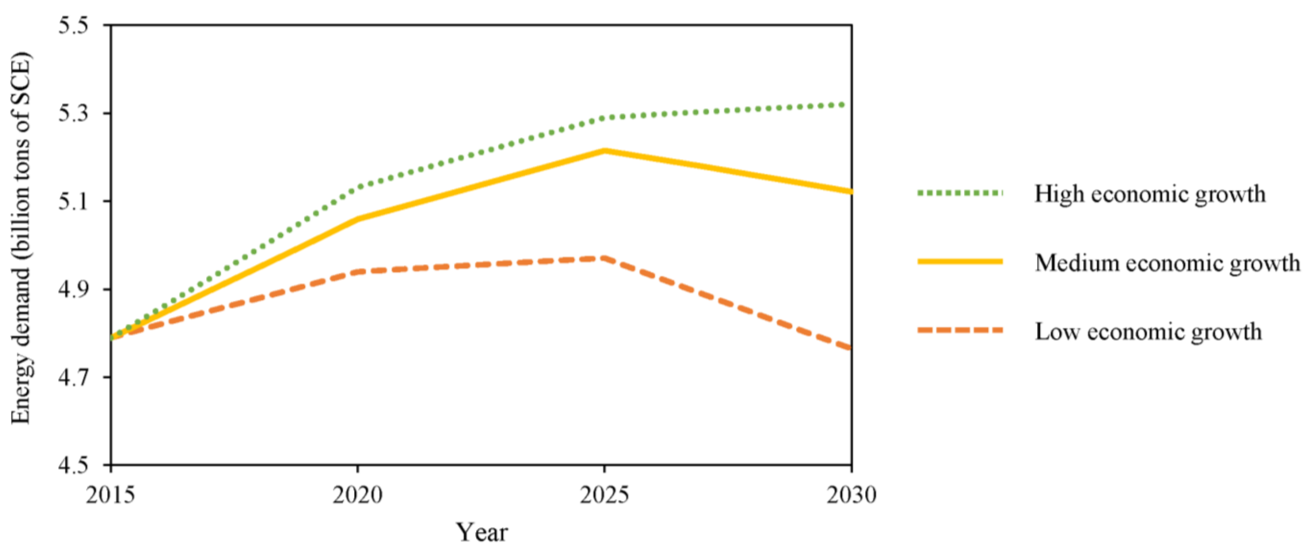


**Figure 5** Energy demand (median value) in three scenarios of different economic development

### 3.4 Energy intensity reduction

Energy intensity (the amount of energy consumed by per unit of GDP) is a key indicator in China's energy planning. The smaller value suggests better energy efficiency of economy. According to the mixed predictions, the changes in energy intensity (median value) are calculated. Figure 6 presents the regional energy intensity by 2030 and the associated reductions from 2010 levels in scenario S1.

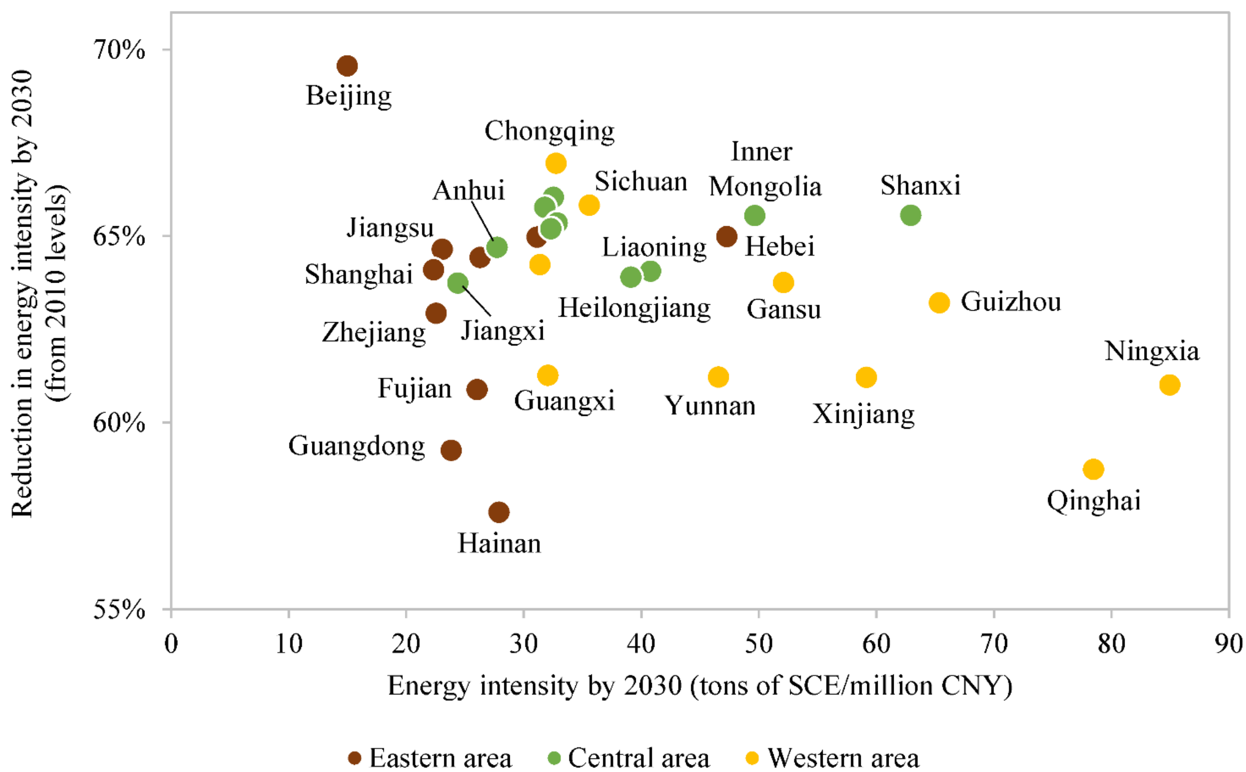


**Figure 6** Energy intensity (median value) by 2030 and the associated reductions from 2010 levels in scenario S1

The eastern provinces have lower energy intensity by 2030 that is 26.6 tons of SCE per million CNY on average. The amount of energy consumed by per unit GDP in Beijing by 2030 is 15.0 tons of SCE per million CNY, the smallest among all the provinces. By comparison, the western provinces have higher mean energy intensity of 51.8 tons of SCE per million CNY. In particular, the largest value of 85.0 tons of SCE per million CNY is found in Ningxia. Generally, economic development in the provinces of more energy-intensive industries such as Gansu, Guizhou, Inner Mongolia, Ningxia, Qinghai, Shanxi, and Xinjiang would consume more energy. Most of them are less developed in the central and western areas, and thus it is necessary to help them adjust industrial structure. Yet, the developed provinces in the eastern area such as Beijing, Guangdong, Jiangsu, Shanghai, and Zhejiang have advanced technology to make energy intensity lower.

On the other hand, there are different performances in energy intensity reduction across provinces. The central provinces decrease by 65.0% on average, and they seem to have similar changes. In contrast, the eastern provinces vary greatly in the reductions in energy intensity. Specifically, Beijing would decrease by 69.6%, larger than other provinces. Yet, the smallest reduction of 57.6% is found in Hainan. In addition, Chongqing and Sichuan would make bigger improvements in energy efficiency than other western provinces.

### 3.5 Impacts of energy structure adjustment

Energy structure adjustment becomes essential for carbon dioxide emissions reduction. On the basis of the mixed primary energy demand predictions, we attempt to investigate whether the proposed emissions reduction targets can be achieved. Accordingly, the energy structure in 2020 and 2030 (Table 7) provided by Hao et al. (2016) is used for our analysis.

The carbon emissions are calculated based on the coefficients given by Zhu et al. (2015). By 2020, the carbon intensity (carbon dioxide emissions per unit of GDP) would decrease by 55.6-58.9% in scenario S2 and 55.8-59.1% in scenario S1 from 2005 levels. Meanwhile, the reductions in carbon intensity would reach 74.5-78.9% in scenario S2 and 74.7-79.1% in scenario S1 by 2030. This implies that the current reduction targets could be realized with the energy structure adjustment. On the other hand, the carbon dioxide emissions peak would appear before 2030 since the decreases in energy demand.

The shares of non-fossil energy in Table 7 are close to the expected goals in China’s energy plans. As a result, it can be obtained that the non-fossil energy demand (median value) by 2020 and 2030 is as much as 0.74/0.76 and 1.00/1.08 billion tons of SCE in scenario S1/S2, respectively. We notice that the increment in non-fossil energy consumption during 2010-2015 was 0.19 billion tons of SCE, and the non-fossil energy consumption in 2015 was 0.59 billion tons of SCE. This means that future development of non-fossil energy can roughly achieve the goal by 2020.

**Table 7** China’s energy structure in 2020 and 2030

| | Coal | Oil | Natural gas | Non-fossil energy |
|---|---|---|---|---|
| 2020 | 57.3% | 17.6% | 10.1% | 15.0% |
| 2030 | 50.3% | 16.8% | 11.9% | 21.0% |

Source: Hao et al. (2016)

## 4 DISCUSSION

Recent studies have also investigated China’s future energy demand using different methods and data (Table 8). By comparison, the regional analysis can gather more information to give more specific insights for energy planning. As argued by the literature, the analysis with a focus on a panel of different regions needs to account for heterogeneity and cross-sectional dependence. It is improper to assume that the impacts on energy consumption are homogeneous across regions due to the varying development stages. Meanwhile, a relation of energy consumption between two regions may naturally exist, especially for those in similar geographical, economic, and political conditions. These issues are addressed in this study by hierarchical Bayesian approach. Furthermore, the distribution of the projected energy demand is presented to provide detailed information. It should also be noticed that the Bayesian approach can incorporate prior information to improve estimates. This requires the specific knowledge of the effect of influence factor.

**Table 8** China’s energy demand projected by the literature and this study (billion tons of SCE)

| Source | Data level | Energy demand in 2030 (low value) |
|---|---|---|
| Lin and Ouyang (2014) | Provincial | 5.59 |
| Wu and Peng (2017) | National | 4.60 |
| This study | Provincial | 4.76 (4.34-5.24) |

Note: this study gives the median and the 2.5-97.5% range in the parentheses.

Although various scenarios are assumed, our low estimate of China’s energy demand in 2030 is close to those in the literature (Table 8). Besides, this study shows that the energy demand peak would reach 5.25 (median value) billion tons of SCE in the high growth scenario. This is approximate to the amount of 5.30 billion tons of SCE given by Wu et al. (2017). As a result, the projections in this study are reasonable.

We find that there is no obvious energy demand peak in most western provinces. The unbalanced regional development in China makes the possible rapid economic growth in western area in the future, and thus the desire for energy might keep growing. It needs to be cautious to make plans for controlling energy use, since there might be a restriction on regional economy. Our analysis shows that most western provinces still have great potential for energy intensity reduction, and the advanced technology adoption and industrial structure adjustment are effective measures (Mi et al., 2015). Especially for Guizhou and Qinghai, their energy intensity is higher with the increased energy use. The energy-intensive industries should be cut down to improve energy efficiency.

Energy structure adjustment is still a critical issue in China's energy planning. The results suggest that China's new normal of economy might maintain the increase in national energy demand. Besides, current population policy may delay the decrease in national population, and further urbanization process could raise energy demand. In that case, energy structure adjustment is essential for reducing carbon dioxide emissions. Specifically, it needs to further decrease coal consumption but increase renewable energy (Li et al., 2017). The western provinces have abundant resources of renewable energy. Therefore, the development of renewable energy could promote industry in this region.

## 5 CONCLUSIONS

This study investigates the ranges of China's future energy demand accounting for sub-national heterogeneity, cross-sectional dependence, and quantification of uncertainty, and offers a predictive distribution of energy demand. The following primary conclusions are drawn from our analysis.

(1) The hierarchical Bayesian approach partially pools the common information from different regions and provides region-specific regression coefficients and associated uncertainty with flexibly modeling the dependence between variables. It indicates that the hierarchical Bayesian approach has better performance in model fitting than the fixed effect method. The probabilistic forecasts are informative and of great importance for energy policy making.

(2) The eastern and central areas could peak their energy demand in the scenarios designed by this study. However, the western area would keep energy demand growing to 1.27 billion tons of SCE by 2030 in the high growth scenario. As for the whole country, the mixed predictions show that the maximum demand could rise to 4.97/5.25 billion tons of SCE in the low/high growth scenario.

(3) The economic development in the provinces of more energy-intensive industries such as Gansu, Guizhou, Inner Mongolia, Ningxia, Qinghai, Shanxi, and Xinjiang would consume more energy. Yet, the developed provinces in the eastern area such as Beijing, Guangdong, Jiangsu, Shanghai, and Zhejiang would have lower energy intensity. On the other hand, the central provinces could decrease energy intensity by 65.0% on average, and they seem to have similar changes. In contrast, the eastern provinces vary greatly in the reductions in energy intensity. In addition, Chongqing and Sichuan would make bigger improvements in energy efficiency than other western provinces.

(4) By 2020, the carbon intensity would decrease by 55.6-58.9% in scenario S2 and 55.8-59.1% in scenario S1 from 2005 levels. Meanwhile, the reductions in carbon intensity would reach 74.5-78.9% in scenario S2 and 74.7-79.1% in scenario S1 by 2030. Moreover, the non-fossil energy demand (median value) by 2020 and 2030 is as much as 0.74/0.76 and 1.00/1.08 billion tons of SCE in scenario S1/S2, respectively.

There are still large uncertainties in future socioeconomic development. Further work should give more information to improve projection, such as scenario design and parameter setting. The Bayesian framework could actually integrate all the uncertainties and provide a more informative estimate of energy demand.

## APPENDIX A

**Table A1** Abbreviations for China's provinces in three areas

| Eastern area | | Central area | | Western area | |
|---|---|---|---|---|---|
| Beijing | BJ | Anhui | AH | Chongqing | CQ |
| Fujian | FJ | Henan | HEN | Gansu | GS |
| Guangdong | GD | Heilongjiang | HLJ | Guangxi | GX |
| Hebei | HEB | Hubei | HUB | Guizhou | GZ |
| Hainan | HN | Hunan | HUN | Ningxia | NX |
| Liaoning | LN | Inner Mongolia | IM | Qinghai | QH |
| Jiangsu | JS | Jilin | JL | Sichuan | SC |
| Shandong | SD | Jiangxi | JX | Shaanxi | SHX |
| Shanghai | SH | Shanxi | SX | Xinjiang | XJ |
| Tianjin | TJ | | | Yunnan | YN |
| Zhejiang | ZJ | | | | |

## APPENDIX B

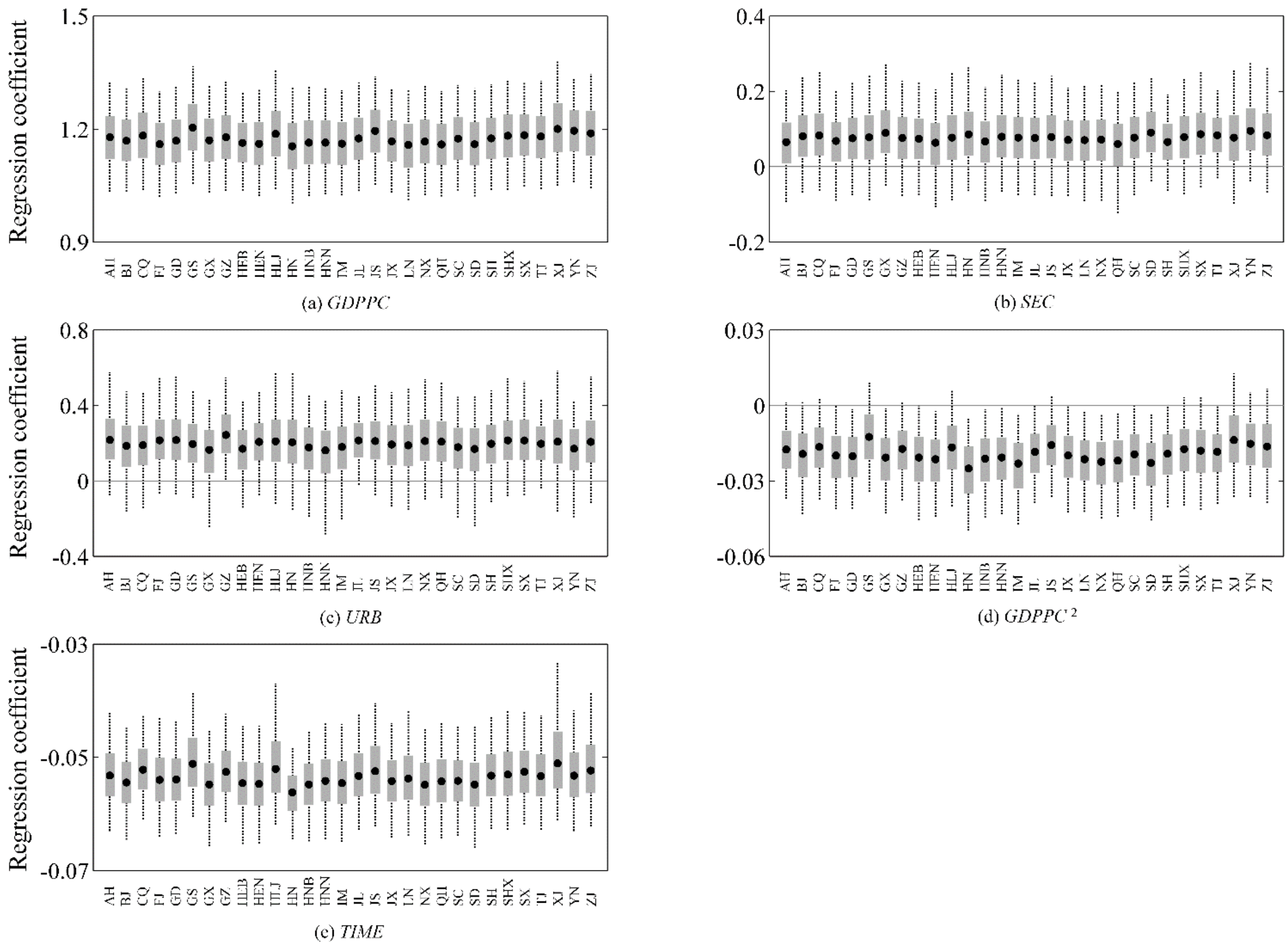


**Figure B1** Regression coefficients and the associated uncertainty bounds for model M3. Each box shows the posterior distribution of the regression coefficient of a province with the 25th, median, and 75th percentile, and whiskers extend to the 5th and 95th percentile.

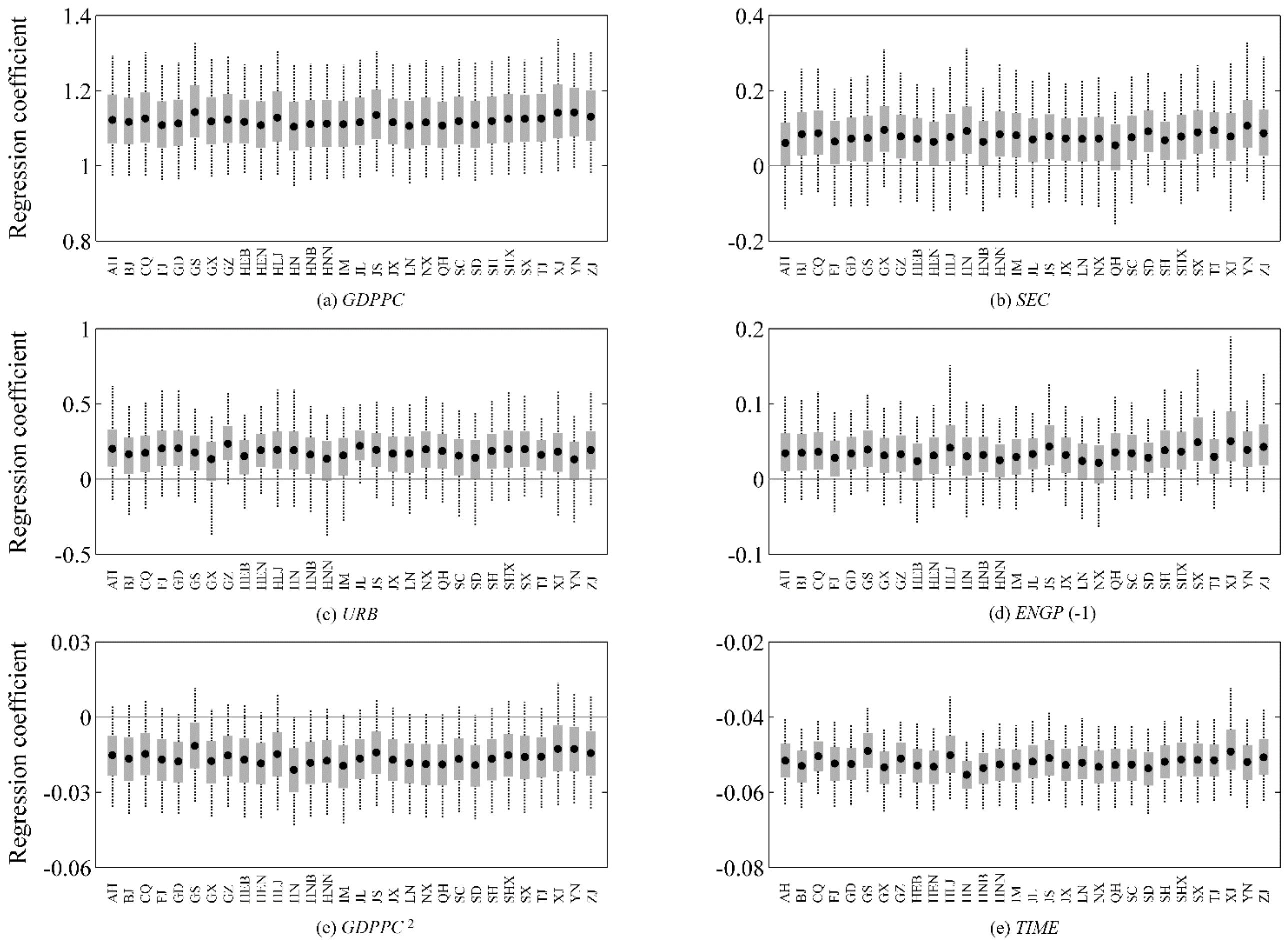


**Figure B2** Regression coefficients and the associated uncertainty bounds for model M6. Each box shows the posterior distribution of the regression coefficient of a province with the 25th, median, and 75th percentile, and whiskers extend to the 5th and 95th percentile.


## ACKNOWLEDGMENTS

The authors are grateful for the financial support from the National Key R&D Program (2016YFA0602603), the National Natural Science Foundation of China (NSFC) (No. 71521002), the project of China Postdoctoral Science Foundation (2016M600046), and the Fundamental Research Funds for the Central Universities (FRF-TP-16-053A1).

## REFERENCES


Al-mulali, U., Binti Che Sab, C.N., Fereidouni, H.G., 2012. Exploring the bi-directional long run relationship between urbanization, energy consumption, and carbon dioxide emission. Energy 46, 156-167.

Brockway, P.E., Steinberger, J.K., Barrett, J.R., Foxon, T.J., 2015. Understanding China's past and future energy demand: An exergy efficiency and decomposition analysis. Appl. Energy 155, 892-903.

Chen, W.D., Wu, F.Y., Geng, W.X., Yu, G.Y., 2017. Carbon emissions in China's industrial sectors. Resour. Conserv. Recycl. 117, 264-273.

Chen, X., Hao, Z., Devineni, N., Lall, U., 2014. Climate information based streamflow and rainfall forecasts for Huai River basin using hierarchical Bayesian modeling. Hydrology and Earth System Sciences 18, 1539-1548.

Comodi, G., Cioccolanti, L., Gargiulo, M., 2012. Municipal scale scenario: Analysis of an Italian seaside town with MarkAL-TIMES. Energy Policy 41, 303-315.

Devineni, N., Lall, U., Pederson, N., Cook, E., 2013. A tree-ring-based reconstruction of Delaware river basin streamflow using hierarchical Bayesian regression. Journal of Climate 26, 4357-4374.

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Statistical science 7, 457-472.

Gunay, M.E., 2016. Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicators and climatic conditions: Case of Turkey. Energy Policy 90, 92-101.

Hamzacebi, C., Es, H.A., 2014. Forecasting the annual electricity consumption of Turkey using an optimized grey model. Energy 70, 165-171.

Hao, Y., Zhang, Z.-Y., Liao, H., Wei, Y.-M., 2015. China’s farewell to coal: A forecast of coal consumption through 2020. Energy Policy 86, 444-455.

Hao, Y., Zhang, Z., Liao, H., 2016. China's energy 'New Normal': the forecasts for energy economy during the '13th Five-Year Plan' and by 2030. Journal of Beijing Institute of Technology (Social Sciences Edition) 18, 1-7 (in Chinese).

Hu, A., Yan, Y., Wei, X., 2014. China 2030. Springer Berlin Heidelberg.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. Springer, New York.

Jiang, Z.J., Lin, B.Q., 2012. China's energy demand and its characteristics in the industrialization and urbanization process. Energy Policy 49, 608-615.

Kumar, S., 2016. Assessment of renewables for energy security and carbon mitigation in Southeast Asia: The case of Indonesia and Thailand. Appl. Energy 163, 63-70.

Lee, Y.S., Tong, L.I., 2011. Forecasting energy consumption using a grey model improved by incorporating genetic programming. Energy Conv. Manag. 52, 147-152.

Li, B.B., Liang, Q.M., Wang, J.C., 2015. A comparative study on prediction methods for China's medium- and long-term coal demand. Energy 93, 1671-1683.

Li, K., Lin, B.Q., 2015. Impacts of urbanization and industrialization on energy consumption/CO2 emissions: Does the level of development matter? Renewable & Sustainable Energy Reviews 52, 1107-1122.

Li, N., Zhang, X.L., Shi, M.J., Zhou, S.L., 2017. The prospects of China's long-term economic development and CO2 emissions under fossil fuel supply constraints. Resour. Conserv. Recycl. 121, 11-22.

Liao, H., Cai, J.-W., Yang, D.-W., Wei, Y.-M., 2016. Why did the historical energy forecasting succeed or fail? A case study on IEA's projection. Technol. Forecast. Soc. Chang. 107, 90-96.

Liddle, B., 2014. Impact of population, age structure, and urbanization on carbon emissions/energy consumption: evidence from macro-level, cross-country analyses. Population and Environment 35, 286-304.

Lin, B., Ouyang, X., 2014. Energy demand in China: Comparison of characteristics between the US and China in rapid urbanization stage. Energy Conv. Manag. 79, 128-139.

Liu, Y., Zhou, Y., Wu, W., 2015. Assessing the impact of population, income and technology on energy consumption and industrial pollutant emissions in China. Appl. Energy 155, 904-917.

Mi, Z., Wei, Y.-M., Wang, B., Meng, J., Liu, Z., Shan, Y., Liu, J., Guan, D., 2017. Socioeconomic impact assessment of China's CO2 emissions peak prior to 2030. Journal of Cleaner Production 142, 2227-2236.

Mi, Z., Zhang, Y., Guan, D., Shan, Y., Liu, Z., Cong, R., Yuan, X.-C., Wei, Y.-M., 2016. Consumption-based emission accounting for Chinese cities. Appl. Energy 184, 1073-1081.

Mi, Z.F., Pan, S.Y., Yu, H., Wei, Y.M., 2015. Potential impacts of industrial structure on energy consumption and CO2 emission: a case study of Beijing. Journal of Cleaner Production 103, 455-462.

Poumanyvong, P., Kaneko, S., 2010. Does urbanization lead to less energy use and lower CO2 emissions? A cross-country analysis. Ecological Economics 70, 434-444.

Sadorsky, P., 2014. The Effect of Urbanization and Industrialization on Energy Use in Emerging Economies: Implications for Sustainable Development. American Journal of Economics and Sociology 73, 392-409.

Shao, Z., Gao, F., Zhang, Q., Yang, S.L., 2015. 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 156, 502-518.

Stan Development Team, 2016. Modeling Language User's Guide and Reference Manual, v2.12.0.

Suganthi, L., Samuel, A.A., 2012. Energy models for demand forecasting-A review. Renew. Sust. Energ. Rev. 16, 1223-1240.

Tsai, M.S., Chang, S.L., 2015. Taiwan's 2050 low carbon development roadmap: An evaluation with the MARKAL model. Renew. Sust. Energ. Rev. 49, 178-191.

United Nations, 2014. World Urbanization Prospects: The 2014 Revision, CD-ROM Edition.

United Nations, 2015. World Population Prospects: The 2015 Revision, DVD Edition.

Ünler, A., 2008. Improvement of energy demand forecasts using swarm intelligence: The case of Turkey with projections to 2025. Energy Policy 36, 1937-1944.

Wu, J., Mohamed, R., Wang, Z., 2017. An Agent-Based Model to Project China’s Energy Consumption and Carbon Emission Peaks at Multiple Levels. Sustainability 9, 893.

Wu, Q.L., Peng, C.Y., 2017. A hybrid BAG-SA optimal approach to estimate energy demand of China. Energy 120, 985-995.

Yoo, S.H., Lee, J.S., 2010. Electricity consumption and economic growth: A cross-country analysis. Energy Policy 38, 622-625.

York, R., 2007. Demographic trends and energy consumption in European Union Nations, 1960-2025. Soc. Sci. Res. 36, 855-872.

You, J., 2013. China's challenge for decarbonized growth: Forecasts from energy demand models. J. Policy Model. 35, 652-668.

Yu, S.W., Zhu, K.J., Zhang, X., 2012. Energy demand projection of China using a path-coefficient analysis and PSO-GA approach. Energy Conv. Manag. 53, 142-153.

Yuan, C.Q., Liu, S.F., Fang, Z.G., 2016. Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model. Energy 100, 384-390.

Zhang, C., Xu, J., 2012. Retesting the causality between energy consumption and GDP in China: Evidence from sectoral and regional analyses using dynamic panel data. Energy Econ. 34, 1782-1789.

Zhu, B., Wang, K., Chevallier, J., Wang, P., Wei, Y.-M., 2015. Can China achieve its carbon intensity target by 2020 while sustaining economic growth? Ecological Economics 119, 209-216.