Scientific Data

https://doi.org/10.1038/s41597-026-07327-8

Article in Press

Hourly electricity load curve dataset for Chinese provinces derived from meteorological variables

Received: 4 January 2026

Accepted: 22 April 2026

Cite this article as: Yi, B., Luo, Q., Zhang, S. *et al.* Hourly electricity load curve dataset for Chinese provinces derived from meteorological variables. *Sci Data* (2026). https://doi.org/10.1038/s41597-026-07327-8

Bowen Yi, Qian Luo, Shaohui Zhang, Yulu Ji, Songmin Yu & Ying Fan

Hourly electricity load curve dataset for Chinese provinces derived from meteorological variables

Bowen Yi [a,b*], Qian Luo [a,b], Shaohui Zhang [c], Yulu Ji [a,b], Songmin Yu [d], Ying Fan [a,b*]

[a] School of Economics & Management, Beihang University, Beijing, China

[b] MOE Laboratory for Low-carbon Intelligent Governance (LLIG), Beihang University, Beijing, China

[c] International Institute for Applied Systems Analysis, Laxenburg, Austria

[d] Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany

* Corresponding author, E-mail: ybw2018@buaa.edu.cn; yfan1123@buaa.edu.cn

Abstract: As the world's largest electricity consumer, China has long faced a shortage of publicly available electricity load data at a high temporal resolution. To address this limitation, this study leverages the strong correlation between electricity demand and meteorological conditions, as well as the broad accessibility of hourly weather data. By combining the limited publicly released load statistics from the National Development and Reform Commission with detailed hourly observations of temperature, wind speed, solar irradiance, and relative humidity, we estimated province-level power coefficients and threshold temperatures for both cooling and heating. These coefficients were then used to reconstruct hourly electricity load profiles for all provinces. The resulting dataset provides hourly electricity demand for 31 provincial-level regions across China for the period 2015–2024. Importantly, the proposed methodology is highly scalable and can be extended to any target year using annual electricity consumption data, air conditioner ownership per household, and corresponding hourly meteorological inputs. This dataset offers a valuable empirical foundation for research on electricity demand dynamics and long-term energy system planning in China.

Background & Summary

As countries around the world continue to announce carbon neutrality goals, the power sector has been receiving increasing attention in the context of low-carbon transitions, and related research has been growing steadily[1]. Due to the real-time matching requirement of electricity supply and demand, this field typically requires high temporal resolution, at least at the hourly level, which in turn demands access to hourly electricity load data[2]. Such data are readily available in regions such as Europe and the United States. However, in China, complete hourly demand data are not publicly available[3]. As the world's largest electricity consumer, this lack of data presents a significant obstacle to research on China's electricity transition.

Currently, most macro-level studies on provincial electricity demand in China rely on reconstructed data[4-7]. The primary data source is the set of typical daily demand curves for weekdays and weekends in 2018 published by the National Development and Reform Commission (NDRC), along with the maximum and minimum daily demand values. Researchers typically assume that the daily load profile is the same as that of a typical day and adjust the values at each time point based on the daily maximum and minimum demand levels[8]. This approach has two main limitations. First, it assumes a uniform daily load curve throughout the year, which clearly does not reflect reality. For example, electricity consumption patterns differ significantly between winter and summer. Second, the method lacks generalizability because the absolute daily values are based entirely on the characteristics of 2018. These values may reflect one-off events or be highly correlated with a year's specific weather conditions. Applying the same pattern to other years is therefore not a reasonable practice.

For a given region, when the proportions of residential and industrial electricity consumption remain relatively stable, electricity demand tends to be highly correlated with weather conditions[9]. For example, in many areas, the annual peak electricity demand is most likely to occur during the coldest or hottest days. Weather conditions are playing an increasingly significant role in the power system, influencing both the supply side such as wind and solar power generation, and the demand

side. While there has been extensive research on the effects of weather on electricity supply[10-13], studies focusing on its impact on electricity demand are still relatively limited. The impact on demand is mainly reflected in heating and cooling within buildings[14-16]. According to a report by the International Energy Agency (IEA), nearly 10% of global electricity consumption comes from air conditioners and electric fans[17], while approximately 15% of final energy consumption is attributed to heating, a portion of which is powered by electricity[18]. In the literature, estimates of China's electricity demand are often spatially detailed but temporally coarse[19,20], and they rarely take into account the impact of weather on demand.

Using meteorological data to reconstruct electricity load curves is a clever approach. On the one hand, there is a strong correlation between weather conditions and electricity demand; on the other hand, the high temporal resolution and public availability of meteorological data align well with the characteristics of load curves. Based on this, the paper builds upon the building-adjusted internal temperature (BAIT) framework proposed by Staffell et al. (2023)[21] and extends it to a regionalized analysis at the provincial level in China. Such an extension is particularly meaningful in the Chinese context, as it helps address key limitations in existing studies, including the challenges in estimating temperature sensitivity coefficients and the scarcity of high-resolution electricity load data.

The main contributions of this paper can be summarized in three aspects. First, this study provides the first systematic assessment of cooling- and heating-related electricity demand coefficients and associated temperature thresholds at the provincial level in China, offering new empirical evidence on regional heterogeneity in temperature–load responses. Second, this study is the first to estimate hourly provincial-level electricity load curves based on meteorological features and to extend these estimates across multiple years, covering the period from 2015 to 2024. The resulting dataset offers valuable demand-side data support for research on the transition of China's power system. Third, unlike Staffell et al. (2023) [21], who estimate daily BAIT and subsequently derive hourly profiles using typical-day cooling and heating patterns, this study directly estimates

BAIT at the hourly scale through a series of methodological steps. This approach effectively mitigates the scarcity of typical-day cooling and heating data in China. Consequently, the estimated cooling and heating demand differs substantially from that reported by the Renewables Ninja platform and provides a more accurate representation tailored to provincial characteristics in China. Overall, against the backdrop of substantial data constraints related to electricity load, cooling demand, and heating demand in China, this study reconstructs provincial hourly load curves to the greatest extent possible using an integrated methodological framework, representing a relatively systematic and comprehensive attempt in this line of research.

Methods

Overall framework

The method of this study was to use the limited load data for 2018 released by the NDRC, combined with detailed meteorological data, to first estimate the power coefficients and threshold temperatures for heating and cooling. It was then assumed that the non-weather-related load curves for weekdays and weekends remained consistent. By combining hourly meteorological data, this method captured the temporal variation of loads and thus reconstructed hourly load curves for provinces in China. Based on the estimated typical daily load curves, power coefficients and threshold temperatures for heating and cooling in 2018, this method could be extended to any other year, as long as hourly meteorological data, air conditioner ownership per household and the total annual load for that year were provided. The research framework is shown in Fig. 1.

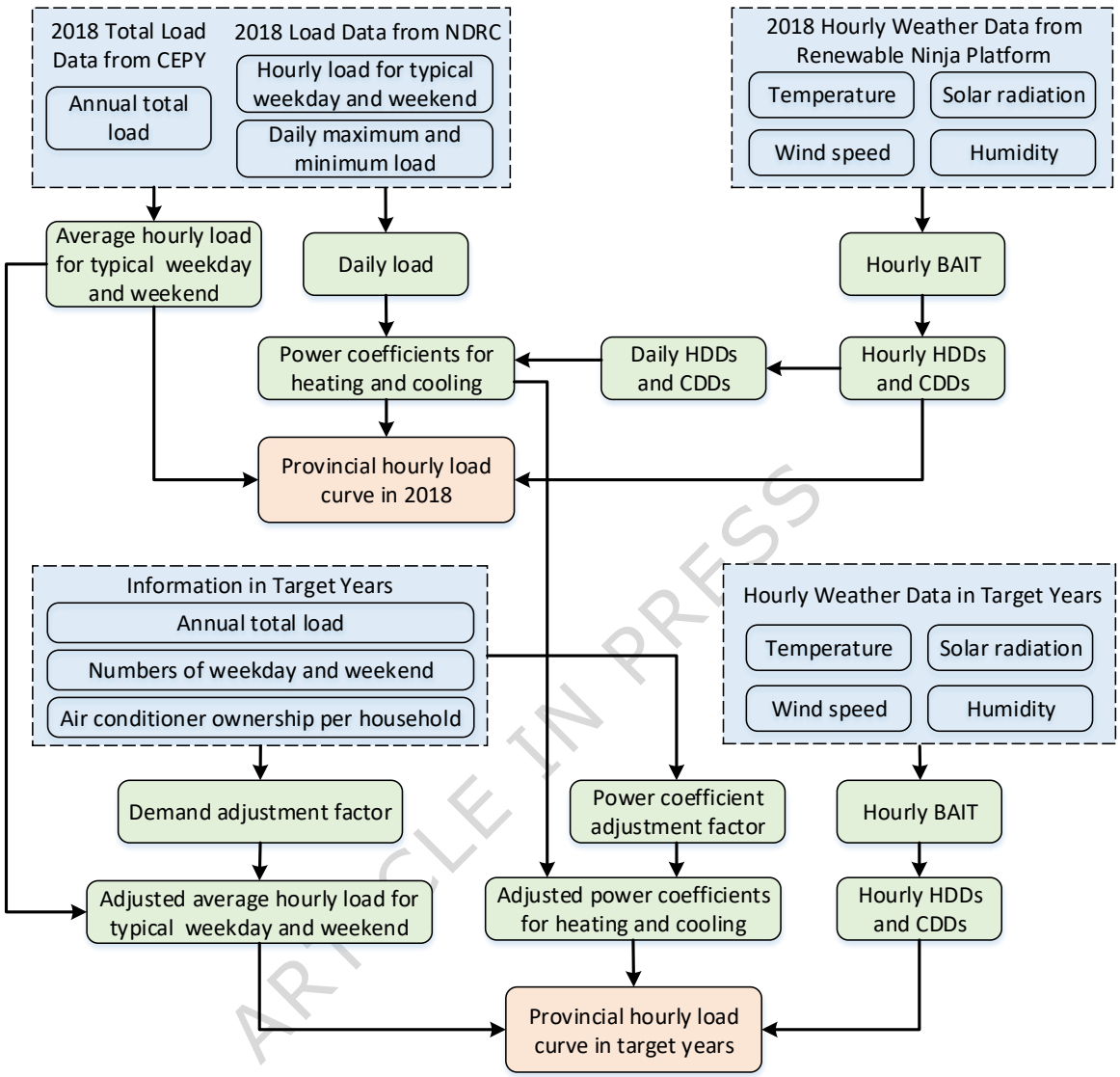

Fig. 1. Research framework.

Data input

The data input required for this study is shown in Table 1, which mainly includes hourly meteorological data over the years, annual total electricity demand, annual air conditioner ownership per household, and typical weekday and weekend load curves as well as daily maximum and minimum loads released by NDRC in 2018. All of these data are publicly available. Among them, the load information from the NDRC is provided in image format, and we digitized it using the GetData software. Meteorological data include temperature, wind speed, solar radiation, and

humidity. The original data came from the MERRA-2 dataset with a spatial resolution of 0.625° × 0.5°, and were aggregated to the provincial level using the population-weighted average method.

Table 1. Summary of the input data.

Data description	Temporal resolution	Temporal range	Spatial resolution	Data source
Temperature, wind speed, solar radiation, and humidity	Hourly	2015-2024	Provincial	Renewable Ninja Platform[22]
Total electricity demand	Annual	2015-2024	Provincial	China Electric Power Yearbook[23]
Typical load curves for weekdays and weekends	Hourly	2018	Provincial	National Development and Reform Commission[24]
Maximum and minimum loads	Daily	2018	Provincial	National Development and Reform Commission[24]
Air conditioner ownership per household	Annual	2015-2024	Provincial	China Statistical Yearbook[25]

Calculating daily electricity demand

The 2018 provincial electricity demand data released by the NDRC contain 24-hour typical daily load curves for weekdays and weekends, as well as daily maximum and minimum load values for 365 days. Based on this dataset, we aimed to estimate the daily total electricity demand (D_d^i) for each province.

First, we used the typical daily load curve to calculate the ratio of the daily total demand to the sum of the daily maximum and minimum load values ($\beta_{v(d)}^i$), as shown in Equation (1). Assuming that the ratio is consistent for each typical day (weekday or weekend), we then used the corresponding daily maximum and minimum values to estimate the original daily electricity demand (oD_d^i), as shown in Equation (2).

$$\beta_{v(d)}^i = \frac{\sum_{h=1}^{24}\left(oD_{v(d),h}^i\right)}{max\left(oD_{v(d),h}^i\right)+min\left(oD_{v(d),h}^i\right)} \quad (1)$$

$$oD_d^i = \left(oUD_d^i + oLD_d^i\right) \cdot \beta_{v(d)}^i \quad (2)$$

where the superscript *i* denotes the province, and the subscripts *h* and *d* represent the hour and the day, respectively. *v(d)* identifies the type of day (weekday or weekend). $oD_{v(d),h}^i$ represents the hourly electricity demand from the typical day profiles provided by the NDRC. The functions $max(\cdot)$ and $min(\cdot)$ denote the maximum and minimum, respectively. oUD_d^i and oLD_d^i represent the daily maximum and minimum load values reported in the original NDRC dataset.

Given that the aggregated load data from the NDRC shows some discrepancies compared to the annual total electricity demand (D^i) reported in the China Electric Power Yearbook (CEPY) for certain provinces, we further adjusted the estimated daily total electricity demand to align with the official annual totals. This calibration process is described in Equation (3).

$$D_d^i = \frac{D^i}{\sum_{d=1}^{365}\left(oD_d^i\right)} \cdot oD_d^i \quad (3)$$

Calculating degree days in hourly resolution

When estimating the heating and cooling needs of each province, it was found that these needs are closely related to weather conditions. This study used hourly population-weighted temperature, solar radiation, wind speed, and relative humidity data for each province in China in 2018 from the Renewable Ninja platform. First, we calculated the BAIT for each province, which is an indicator that reflects the perceived indoor temperature in the absence of active heating or cooling. We follow the day-level BAIT algorithm of Staffell et al. (2023)[21], but differ from their approach in that we compute BAIT directly at an hourly resolution. In their study, the hourly resolution was obtained using representative 24-hour diurnal profiles, whereas at the provincial level in China no sufficiently representative data of this kind are available. Computing BAIT directly at the hourly scale leads to differences in the smoothing process as well as in the subsequent selection of threshold temperatures.

The BAIT is affected by outdoor temperature, solar radiation, wind speed, humidity, and building thermal inertia. The calculation method of the BAIT is shown in Equation (4). Among these factors, outdoor temperature plays the most important role. Solar radiation helps to increase indoor temperature, while increased wind speed promotes ventilation, thereby reducing indoor temperature. The effect of humidity on perceived temperature is more subtle and changes with changes in ambient temperature. Under hot conditions, increased humidity hinders the body's ability to cool through sweating, resulting in an increase in perceived temperature. In cold conditions, humid air causes cold moisture to condense on warm skin, increasing thermal conductivity and lowering perceived temperature.

$$oBAIT_t^i = T_t^i + x \cdot (S_t^i - kS_t^i) - y \cdot (W_t^i - kW_t^i) + z \cdot (H_t^i - kH_t^i) \cdot sign(T_t^i - T^*) \quad (4)$$

where subscript t represents the consecutive hours throughout the year. $oBAIT_t^i$ represents the baseline BAIT. T_t^i, S_t^i, W_t^i, and H_t^i denote the hourly values of outdoor temperature, solar radiation, wind speed, and relative humidity for each province, respectively. $sign(\mathrm{g})$ function returns 1 for positive inputs, 0 for zero, and −1 for negative inputs. The coefficients *x*, *y*, and *z* correspond to solar radiation, wind speed, and relative humidity, with values set at 0.012 °C $W^{-1}m^2$, 0.2 °C $m^{-1}s$, and 0.05 °C $g^{-1}kg$, respectively. kS_t^i, kW_t^i, kH_t^i, and T^* represent counterfactual values for solar radiation, wind speed, humidity, and temperature, and are calculated as shown in Equation (5).

$$kS_t^i = 100 + 7T_t^i, \quad kW_t^i = 4.5 - 0.025T_t^i, \quad kH_t^i = e^{1.1+0.06T_t^i}, \quad T^* = 16 \quad (5)$$

Building structures possess the ability to store heat, which provides an inertial buffering effect against fluctuations in outdoor temperatures. This thermal inertia is incorporated into the analysis using a temperature smoothing method, as described in Equation (6). For the smoothing process, we adopted a 48-hour (2-day) moving window consistent with the daily calculation in Staffell et al. (2023)[21]. By converting the daily-level assumption that the weight of the previous day is half

of the current day's weight into an hourly weight ratio of 0.9027 between adjacent hours, we assigned weights to each hourly BAIT using an exponential weighting curve. In this curve, the weight at the current hour is 1, and the weight for each subsequent hour is 0.9027 times that of the previous hour, extending over a 48-hour period.

$$sBAIT_t^i = \frac{\sum_{t'=t-48}^{t} \left(\sigma_{t'} \cdot oBAIT_t^i\right)}{\sum_{t'=t-48}^{t} \left(\sigma_{t'}\right)} \tag{6}$$

where the subscript *t'* also represents consecutive hours throughout the year. $sBAIT_t^i$ represents the smoothed BAIT. σ is the smoothing coefficient. The weight curve follows the function $\sigma_q = e^{-0.10232q}$, where q represents the number of hours from the current time point (ranging from 0 to 48).

However, due to the influence of behavioral factors, the application of the smoothing method may reduce the accuracy of cooling demand estimation to a certain extent. In hot weather, people are more likely to open windows, resulting in the dissipation of building thermal inertia. In contrast, this behavior is less common in cold weather. To address this issue, this study constructed a weighted comprehensive index combining BAIT and raw outdoor temperature, as shown in Equations (7) and (8).

$$B_t^i = \frac{0.5}{1+e^{-kB_t^i}} \tag{7}$$

$$kB_t^i = \begin{cases} -5, & T_t^i < B_L \\ \left(T_t^i - 0.5 \cdot \left(B_U + B_L\right)\right) \cdot \dfrac{10}{B_U - B_L}, & T_t^i \in \left[B_L, B_U\right] \\ 5, & T_t^i > B_U \end{cases} \tag{8}$$

where B_t^i represents the blending factor. kB_t^i is the input to a sigmoid function. B_U and B_L are the upper and lower thresholds for blending, with values of 23 °C and 15 °C, respectively. The

final calculation formula for BAIT was obtained through this blending process, as shown in Equation (9).

$$BAIT_t^i = sBAIT_t^i \cdot (1 - B_t^i) + T_t^i \cdot B_t^i \quad (9)$$

Subsequently, the calculated hourly BAIT was converted into heating degree days (HDDs) and cooling degree days (CDDs) to quantify the heating and cooling demand, which was necessary to divide by 24 to align with daily scales. These transformations are defined in Equations (10) and (11), respectively.

$$HDD_t^i = \frac{(T_{heat}^i - BAIT_t^i)^+}{24} \quad (10)$$

$$CDD_t^i = \frac{(BAIT_t^i - T_{cool}^i)^+}{24} \quad (11)$$

where T_{heat}^i and T_{cool}^i represent the threshold temperatures for heating and cooling in each province, which are obtained through subsequent regression analysis. The + symbol indicates that only positive values are considered.

Estimating electricity demand for heating and cooling

Based on the calculations above, we aggregated the hourly HDD and CDD values for each day to obtain day-level HDD_d^i and CDD_d^i. Combined with the daily electricity demand data for 2018, we then estimated the power coefficients associated with heating and cooling through regression analysis conducted separately for each province, as shown in Equation (12).

$$D_d^i = D_{base}^i + P_{heat}^i \cdot HDD_d^i + P_{cool}^i \cdot CDD_d^i + \alpha^i \cdot M_{v(d)} + \varepsilon_d^i \quad (12)$$

where $M_{v(d)}$ is a binary variable that equals 1 on weekdays and 0 on weekends. ε_d^i is the residual term. D_{base}^i, P_{heat}^i, P_{cool}^i, and α^i are estimated parameters. Among them, D_{base}^i represents the base load, which is the intercept term obtained from the regression. P_{heat}^i and P_{cool}^i

are the power coefficients for heating and cooling, respectively. α^i represents the additional demand on weekdays compared to weekends.

The parameters P^i_{heat} and P^i_{cool} in Equation (12) are of particular importance in this study. They quantify the increase in electricity demand associated with heating and cooling needs, respectively, and vary across provinces. These values tend to be positively correlated with the absolute level of electricity demand in each province and are also influenced by the structure of electricity consumption. Provinces with a larger proportion of residential and commercial electricity use typically exhibit higher power coefficients for heating and cooling.

Estimating threshold temperatures for heating and cooling

The choice of threshold temperatures is crucial when calculating HDD and CDD. Staffell et al. (2023) used 14°C and 20°C in their global model, but these values are relatively low compared with living habits in China[21]. In practice, however, there is currently no robust China-specific dataset to support alternative choices. Moreover, given China's vast geographic span, threshold temperatures may vary across provinces. Therefore, this study adopts endogenous threshold temperatures, identifying the values that yield the best provincial demand fit.

To this end, we proposed three schemes. The first assigns each province its own threshold temperatures; the second applies a uniform threshold nationwide; and the third divides provinces into northern and southern regions based on the presence of district heating, assigning a common set of thresholds within each region. Although the first scheme can theoretically achieve the best fit, the limited sample size often causes the estimated thresholds for some provinces to reach the preset bounds, which is inconsistent with reality. The second scheme, on the other hand, overlooks interprovincial heterogeneity and yields inferior fitting performance. Overall, we adopted the third scheme, and a comparison of the three approaches is presented in the Technical Validation section.

First, we defined the 16 provinces with district heating as the northern region (*N*), including Beijing, Gansu, Hebei, Henan, Heilongjiang, Jilin, Liaoning, Inner Mongolia, Ningxia, Qinghai,

Shaanxi, Shandong, Shanxi, Tianjin, Xinjiang, and Tibet. The remaining 15 provinces were defined as the southern region (*S*), with $i \in \{N, S\}$. The classification of regions with district heating is based on the urban district heating area reported by the Ministry of Housing and Urban-Rural Development. In some provinces, district heating is only available in a limited number of districts, with Anhui and Jiangsu being typical examples. However, given the very low share of heated area, Anhui and Jiangsu are classified as part of the non-heating southern region. For each regional set, we constructed an objective function that minimizes the weighted mean squared error (WMSE) across all provinces within that region, as shown in Equation (13). The decision variables were the temperature thresholds for HDD and CDD.

$$\min \quad WMSE = \sum_i \sum_d \left(\frac{MSE_d^i}{D_{aver}^i} \right) \tag{13}$$

where D_{aver}^i represents the daily mean load, and MSE_d^i represents the mean squared error from the regression. Using this as the objective function not only avoids the problem of low-load regions being neglected due to differences in load magnitude across provinces when using mean squared error (MSE), but also avoids the issue of assigning equal weight to all regions when using normalized mean squared error (NMSE), which would overlook the importance of high-load regions in the national-level assessment. The current approach tends to achieve a better fit for high-load regions while still taking low-load regions into account, although with relatively lower weight.

The constraints include that each province within a set obtains its own four optimal coefficients in Equation (12) through regression, as shown in Equation (14). λ_i represents the coefficient matrix, composed of D_{base}^i, P_{heat}^i, P_{cool}^i, and α^i. Y_i is the dependent variable matrix, composed of D_d^i. X_i is the independent variable matrix, composed of 1, HDD_d^i, CDD_d^i, and $M_{v(d)}$, as shown in Equation (15). Therefore, when deciding on the unified threshold temperatures, we simultaneously determine the parameters in Equation (12) from Section 2.3, meaning they are

decided together. The threshold temperatures affect HDD_d^i and CDD_d^i, which in turn influence the coefficient matrix and the overall MSE.

$$X_i^T X_i \lambda_i = X_i^T Y_i \tag{14}$$

$$X_i = \begin{bmatrix} 1 & HDD_1^i & CDD_1^i & M_{v(1)} \\ 1 & HDD_2^i & CDD_2^i & M_{v(2)} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ 1 & HDD_d^i & CDD_d^i & M_{v(d)} \end{bmatrix}, \quad \lambda_i = \begin{bmatrix} D_{base}^i \\ P_{heat}^i \\ P_{cool}^i \\ \alpha^i \end{bmatrix}, \quad Y_i = \begin{bmatrix} D_1^i \\ D_2^i \\ \mathrm{M} \\ D_d^i \end{bmatrix} \tag{15}$$

In addition, to ensure that the threshold temperatures remain within a reasonable range, we imposed bounds as shown in Equation (16). The heating threshold is constrained between 12°C and 18°C, while the cooling threshold is constrained between 20°C and 26°C. In practice, when thresholds are determined separately for the northern and southern regions or uniformly for the whole country, the values never reach these bounds. However, when thresholds are estimated individually for each province, they frequently hit the limit values.

$$T_{heat}^L \le T_{heat}^i \le T_{heat}^M, \quad T_{cool}^L \le T_{cool}^i \le T_{cool}^M \tag{16}$$

Generating reconstructed hourly load curves

Using the power coefficients for heating and cooling estimated in Section 2.3, we then calculated the total weather-related hourly electricity demand (WD_t^i) for each province in 2018, as defined in Equation (17).

$$WD_t^i = P_{heat}^i \cdot HDD_t^i + P_{cool}^i \cdot CDD_t^i \tag{17}$$

The variable WD_t^i can be indexed by both day and hour (denoted $WD_{d,h}^i$), and categorized according to whether it falls on a weekday or weekend. The average 24-hour weather-related electricity demand was then computed separately for all weekdays and weekends, yielding two representative daily demand profiles, as shown in Equation (18).

$$WD^{i}_{v(d),h} = \frac{\sum_{d \in v(d)} \left(WD^{i}_{d,h} \right)}{n_{v(d)}} \tag{18}$$

where $n_{v(d)}$ represents the total number of days of that type in the year.

It should be noted that the typical day curve published by NDRC does not represent the annual average of the corresponding day type. Therefore, using the daily demand results from Section 2.1, we categorized and computed the average electricity demand for each type of typical day. By comparing this with the corresponding total demand from the NDRC's typical day curves, we derived an adjustment factor. This factor was then applied to the original hourly demand profiles to obtain the average hourly demand profiles ($D^{i}_{v(d),h}$) for weekdays and weekends, as described in Equation (19).

$$D^{i}_{v(d),h} = oD^{i}_{v(d),h} \cdot \frac{\sum_{d \in v(d)} \left(D^{i}_{d} \right)}{n_{v(d)}} \cdot \frac{1}{\sum_{h=1}^{24} \left(oD^{i}_{v(d),h} \right)} \tag{19}$$

Finally, by subtracting the weather-related portion from the total demand of each typical day, we isolated the non-weather-related demand. We assumed that this component maintains a fixed 24-hour profile for each day type. The variation arises from hourly weather-related demand, which is dynamically calculated using Equation (17). By combining this variable component with the fixed non-weather-related profile, we obtained the final reconstructed hourly load curve ($D^{i}_{d,h}$), as expressed in Equation (20). This approach ensures that the aggregated hourly load exactly matches the annual total load.

$$D^{i}_{d,h} = WD^{i}_{d,h} + D^{i}_{v(d),h} - WD^{i}_{v(d),h} \tag{20}$$

Extending to other years

The load curve reconstruction method described above was mainly based on data from 2018,

but can be extended to any other year. To achieve this extension, the required inputs include hourly weather variables such as temperature, wind speed, solar radiation, and relative humidity for the target year, as well as annual electricity consumption data and the distribution of weekdays and weekends. In addition, another important factor is how the future power coefficients for heating and cooling may change. This determines whether, at the same temperature, more or less electricity will be used for heating or cooling in the future. This trend is difficult to estimate. A feasible approach is to use the air conditioner ownership per household to characterize this change, as suggested by Jiang et al. (2025)[26], as shown in Equation (21). This extension assumes that the basic electricity consumption structure and user behavior patterns in each province remain relatively stable over time. Therefore, the method is more accurate when applied to recent years. However, if there is a major electrification transition or structural changes in electricity consumption, the estimation accuracy may decrease.

$$P_{heat}^{i,y} = P_{heat}^{i} \cdot \left(\frac{AIR_i^y}{AIR_i^{2018}} \right)^{\eta}, \quad P_{cool}^{i,y} = P_{cool}^{i} \cdot \left(\frac{AIR_i^y}{AIR_i^{2018}} \right)^{\eta} \tag{21}$$

where the superscript y denotes the target year. $P_{heat}^{i,y}$ and $P_{cool}^{i,y}$ represent the power coefficients for heating and cooling in the target year, respectively. *AIR* denotes the air conditioner ownership per household in each province. η is the elasticity coefficient, which is set to 0.8 in this study.

The calculation of the demand adjustment factor (e^y), as shown in Equation (22), is primarily influenced by the difference in total electricity demand between the target year and the baseline year (2018). In addition, to account for the difference in the number of weekdays between the target year and the baseline year, a corresponding adjustment was made in the denominator of Equation (22).

$$e^y = \frac{D^y}{\sum_{v(d)} \left(n_{v(d)}^y \cdot D_{v(d)}^{2018} \right)} \tag{22}$$

where $D_{v(d)}^{2018}$ refers to the daily demand for each type of typical day in the baseline year. For clarity,

regional index *i* was omitted here, implying that the year extension method is applicable to any province.

The absolute electricity demand levels of the two typical day types (weekdays and weekends) were then adjusted to the target year using the calculated demand adjustment factor, as shown in Equation (23).

$$D^{y}_{v(d),h} = e^{y} \cdot D^{2018}_{v(d),h} \tag{23}$$

With the availability of hourly weather data for the target year, HDDs and CDDs can be calculated accordingly. Using the adjusted power coefficients, the weather-related electricity demand for the target year was estimated. Finally, by integrating this demand with the adjusted hourly demand profiles for typical days and applying Equation (20), the hourly load curve for the target year can be reconstructed.

Data Records

The dataset generated in this study is publicly available on Figshare[27] and covers 31 provincial-level regions in China, excluding Hong Kong, Macao, and Taiwan. The data are provided as a single Microsoft Excel file (Data output.xlsx) containing 11 worksheets. One worksheet, titled Power coefficients, presents the estimated heating and cooling power coefficients for each province over the period 2015–2024. The remaining ten worksheets correspond to individual years from 2015 to 2024, each containing hourly electricity load data for all provinces over the full year. In these worksheets, provinces are arranged as columns and hourly timestamps as rows.

Technical Validation

Coefficient estimates and significance

Table 2 lists the regression coefficients and temperature thresholds for heating and cooling. Since many production activities were suspended during the Spring Festival, resulting in a

significant reduction in power load, and this effect lasted for a long time, centered on the Spring Festival, we excluded this effect from the regression analysis. Specifically, we deleted 21 days of data from the Spring Festival holiday and the week before and after (i.e., February 8 to February 28). The remaining 344 days of data in 2018 were used for the regression analysis.

Based on our estimates, the cooling and heating threshold temperatures in northern China are 22.253 °C and 14.713 °C, respectively, while in southern China they are 22.631 °C and 16.818 °C. The heating threshold in the south is about 2 °C higher, which is consistent with the fact that residents in northern regions are generally more cold-tolerant. In contrast, the difference in cooling thresholds is small, indicating that people in the south are only slightly more heat-tolerant than those in the north. These results reflect a clear pattern: regional differences in tolerance to cold are more pronounced, whereas differences in tolerance to heat are relatively minor. Overall, the estimated thresholds for China are higher than those found in European studies, which aligns with lifestyle characteristics commonly observed among Chinese residents. In addition, the provincial cooling threshold is usually lower than individual perceptions because it captures the statistical point where overall electricity load begins to rise. Public buildings and some households start cooling at lower temperatures, which lowers the aggregate threshold relative to typical personal comfort levels.

Noticeably, the vast majority of the estimates were highly statistically significant. For the cooling coefficient, except for a few provinces where the CDD is close to zero (Tibet, Yunnan, Guizhou, and Ningxia), the estimates of all provinces were significant. For heating, only Hainan, Guangdong, and Fujian had insignificant coefficients. However, due to the continued high temperatures in these three regions and the low heating demand, their insignificance had little impact on the overall results. In general, the regression results demonstrate a high level of significance, with insignificant values appearing only in provinces where the corresponding demand is extremely low.

The coefficients reflect the differences in the additional power load caused by cooling and

heating in different provinces. The results of our estimation are basically consistent with the actual situation. For example, these coefficients are related to total electricity demand as well as industrial structure. In provinces with large electricity production, such as Guangdong, the coefficients were significantly higher. In regions such as Beijing and Shanghai, where residential and commercial electricity consumption account for a large proportion, the coefficients are also significantly higher than in other regions with similar overall demand. In contrast, industrial provinces such as Shandong have relatively low coefficients. In addition to these two influencing factors, hotter regions often have more widespread access to cooling equipment, resulting in higher electricity demand per unit of CDD. However, the reverse is not necessarily true. This is mainly because cooling is usually completely dependent on electricity, while heating is not. The most typical example is that a large part of the winter centralized heating in northern China relies directly on coal. Therefore, the heating coefficients in these regions did not show the same trend as cooling. For example, even after taking into account differences in scale and structure, many provinces in the northeast and northwest still have lower heating coefficients.

Table 2. Regression coefficient results.

Province	Power coefficient for heating (GWh/℃/Day)	Power coefficient for cooling (GWh/℃/Day)	Weekday increment (GWh)	Intercept (GWh)	Threshold temperature for heating (℃)	Threshold temperature for cooling (℃)
Anhui	8.843***	38.231***	14.341	485.522	16.818	22.631
Beijing	8.192***	21.695***	20.092	220.325	14.713	22.253
Chongqing	6.066***	31.569***	8.529	245.322	16.818	22.631
Fujian	0.912	29.176***	5.371	597.598	16.818	22.631
Guangdong	1.193	102.582***	106.126	1485.549	16.818	22.631
Gansu	1.855***	11.644***	4.066	337.077	14.713	22.253
Guangxi	8.674***	12.184***	24.191	423.969	16.818	22.631
Guizhou	9.391***	1.029	5.739	370.849	16.818	22.631
Hainan	0.559	3.364***	1.857	80.699	16.818	22.631
Hebei	9.755***	31.768***	37.789	889.187	14.713	22.253
Henan	13.750***	52.669***	17.206	789.308	14.713	22.253
Heilongjiang	1.604***	2.623***	6.468	243.461	14.713	22.253
Hunan	12.063***	34.832***	25.442	373.668	16.818	22.631

Hubei	9.268***	39.147***	2.997	472.680	16.818	22.631
Jilin	2.001***	4.352***	12.878	174.123	14.713	22.253
Jiangsu	19.996***	80.452***	30.038	1480.806	16.818	22.631
Jiangxi	6.010***	20.092***	9.843	342.183	16.818	22.631
Liaoning	4.030***	10.500***	32.572	571.383	14.713	22.253
Inner Mongolia	1.686***	5.382*	16.777	891.956	14.713	22.253
Ningxia	0.591***	1.296	0.729	287.972	14.713	22.253
Qinghai	0.643***	18.341**	0.190	195.652	14.713	22.253
Shaanxi	7.466***	27.720***	20.698	363.826	14.713	22.253
Sichuan	12.469***	44.829***	17.740	587.528	16.818	22.631
Shandong	13.782***	53.237***	44.305	1470.980	14.713	22.253
Shanghai	9.503***	33.087***	14.512	344.333	16.818	22.631
Shanxi	4.645***	12.297***	35.252	532.734	14.713	22.253
Tianjin	3.116***	10.983***	2.868	202.243	14.713	22.253
Xinjiang	3.134***	14.759***	6.091	698.889	14.713	22.253
Tibet	0.307***	0.000	0.696	15.182	14.713	22.253
Yunnan	4.070***	0.000	4.720	451.942	16.818	22.631
Zhejiang	9.118***	65.804***	0.342	1161.497	16.818	22.631

Notes: *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Comparison with actual values

Fig. 2 shows the difference between the actual and estimated daily electricity load values for each province. We can see that the fit of the two curves is quite good in the vast majority of provinces. The normalized root mean square error (NRMSE) relative to average demand is within 8% in almost all provinces. Of course, since our model only considered the impact of weather factors on electricity load, there are certainly many other factors that were not considered, so it is impossible to achieve a perfect fit. In particular, short-term and sudden events are difficult to capture. The most notable example is the sharp drop in electricity load in many provinces in February, which coincided with the Spring Festival when most production activities were suspended. These factors cannot be estimated based on weather data alone.

At the regional level, provinces dominated by residential and commercial electricity demand, such as Beijing, Shanghai, Chongqing, and Henan, have the highest fit. For example, Chongqing experienced a sharp drop in electricity load in early September, which was caused by a sharp drop in temperature. Historical data show that similar sharp changes occur annually in late August or

early September in this region, and such variations were well captured by our method. In contrast, the few provinces with less accurate fitting results are mainly those where the electricity load is less affected by weather, such as Ningxia, Inner Mongolia, and Yunnan. These provinces have less seasonal variation in electricity loads and a very low cooling demand in summer. In addition, Ningxia and Inner Mongolia have limited electricity heating demands in winter. Therefore, the load fluctuations in these provinces are mostly unrelated to the weather, making it difficult for our weather-based model to achieve a good fit.

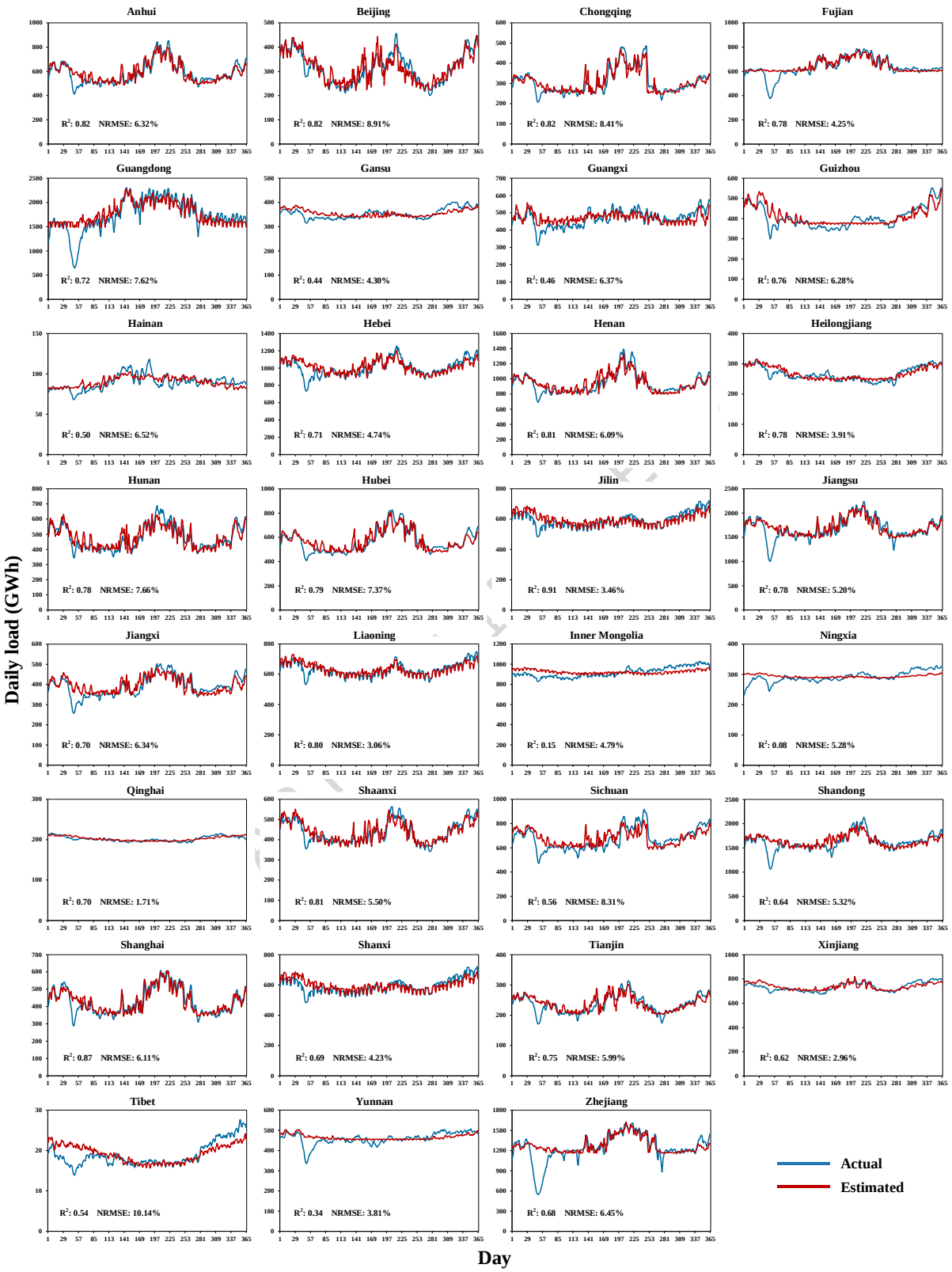

Fig. 2. Comparison between actual and estimated daily loads.

Although Fig. 2 compares the estimates with daily averages, daily averages themselves are not raw observations and do not capture intraday dynamics. Therefore, we further compare the estimated daily hourly maximum and minimum loads with the original data at the hourly level, as shown in Fig. 3. Our method generally performs well in simulating summer peak electricity loads, which, for most provinces, also correspond to the annual maximum and thus play a critical role in determining overall capacity requirements, highlighting the practical relevance of our approach.

Nevertheless, two aspects exhibit relatively weaker alignment. First, for some provinces, the estimated summer minimum load is lower than observed, leading to a slightly overestimated intraday load variability in summer. Second, for several southern provinces, the estimated intraday load variation in winter is narrower than in the observed data, whereas the winter estimates for northern provinces are generally more accurate. These intraday discrepancies can largely be attributed to electricity-use behavior rather than weather-driven factors alone.

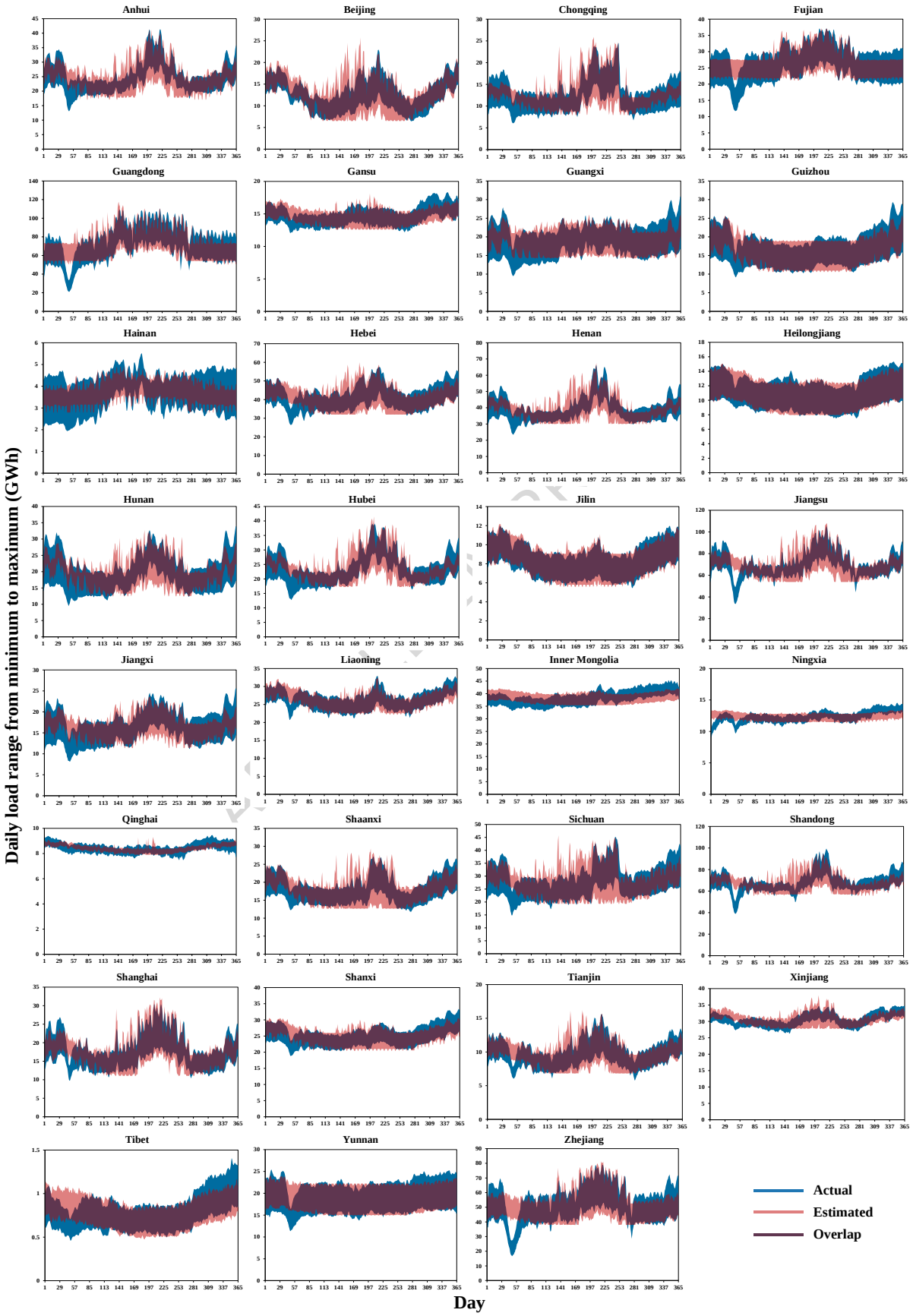
Anhui
Beijing
Chongqing
Fujian
Guangdong
Gansu
Guangxi
Guizhou
Hainan
Hebei
Henan
Heilongjiang
Hunan
Hubei
Jilin
Jiangsu
Jiangxi
Liaoning
Inner Mongolia
Ningxia
Qinghai
Shaanxi
Sichuan
Shandong
Shanghai
Shanxi
Tianjin
Xinjiang
Tibet
Yunnan
Zhejiang
Actual
Estimated
Overlap
Daily load range from minimum to maximum (GWh)
Day

Fig. 3. Comparison between actual and estimated daily maximun and minimun loads.

To provide validation at a more aggregated scale, we compare the reconstructed results with officially reported data at the national monthly level. Monthly aggregation is adopted because it represents the finest temporal resolution currently available from official statistics, while still capturing variations driven by meteorological conditions. As shown in Fig. 4, the reconstructed series closely follows the observed national electricity consumption in terms of overall trends. Noticeable deviations are observed during exceptional events, such as the COVID-19 pandemic and the Spring Festival, which introduce abrupt fluctuations not directly related to meteorological factors. Aside from these anomalies, the model achieves a good fit, indicating that the proposed approach effectively captures the weather-driven component of electricity demand at the aggregate level.

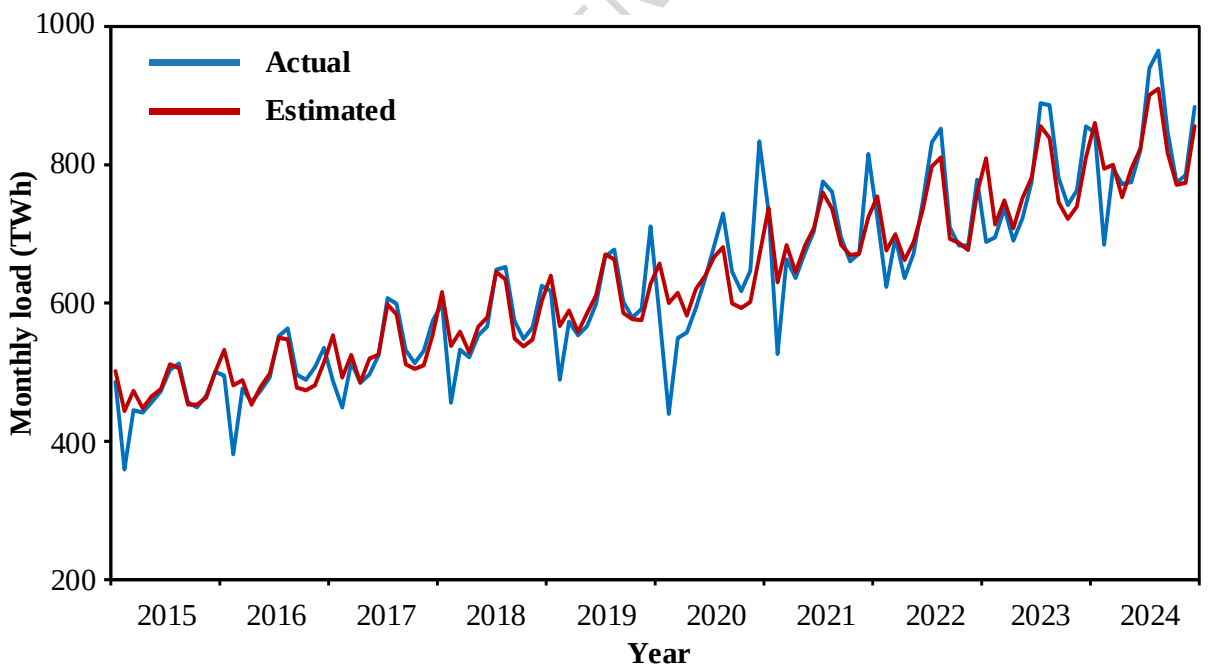

Fig. 4. Comparison between actual and estimated national monthly electricity load.

Hourly patterns of electricity load

Fig. 5 presents our estimated hourly heating and cooling demand across Chinese provinces in

2018. In approximately 64% of provinces, the peak cooling demand is significantly higher than the peak heating demand. These provinces are mostly located in the economically developed southeastern region, where electricity consumption tends to be higher. As a result, the national average shows that summer electricity demand peaks remain much higher than those in winter. In contrast, about 23% of the provinces exhibit the opposite pattern, with higher peak heating demand, mainly located in the northeastern, northwestern, and southwestern regions. The remaining 13% of the provinces, including Shanxi, Liaoning, Qinghai, and Guangxi, show relatively similar peak levels for both heating and cooling.

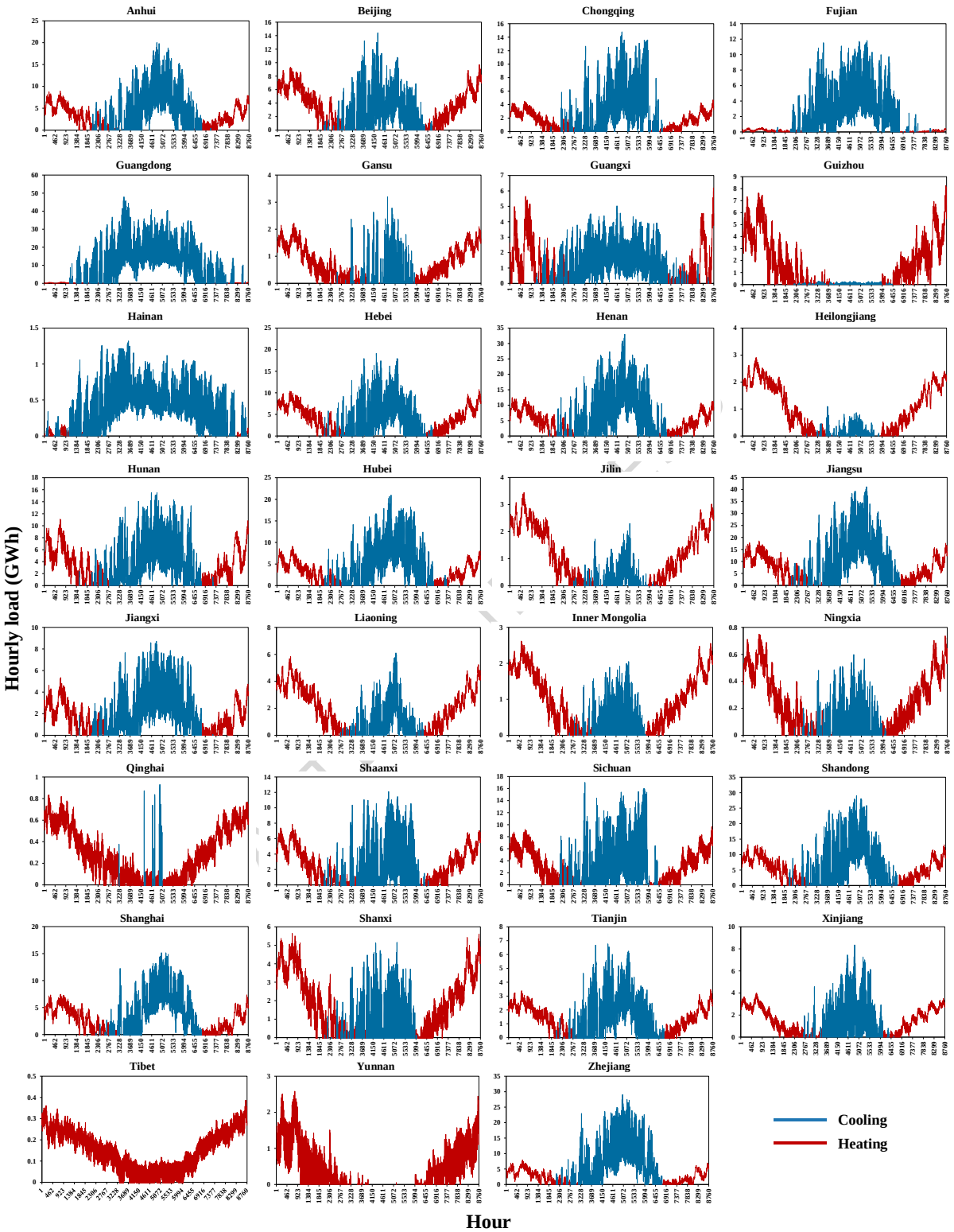

Fig. 5. Estimated hourly heating and cooling demand.

Fig. 6 presents the range and characteristics of intraday fluctuations in cooling and heating

demand across provinces. The two demand profiles generally exhibit opposite patterns: cooling demand typically peaks at 15:00 and reaches a minimum around 5:00, whereas heating demand peaks at 7:00 and bottoms out near 16:00. In terms of fluctuation amplitude, cooling demand shows significantly higher variability than heating demand. Regarding median fluctuation values, cooling ranges from 0.24 to 2.09, while heating falls within the range of 0.71 to 1.35. This is largely because, under the current heating supply structure, the marginal coefficient of CDD on electricity demand is generally much larger than that of HDD. In addition, during winter, closed windows improve thermal insulation and indoor thermal inertia, further dampening the responsiveness of electricity demand to temperature variations. This also helps explain the greater variability in cooling demand observed in Fig. 5, in contrast to the relatively stable pattern of heating demand. Although differences in diurnal temperature variation among provinces lead to some variation in fluctuation amplitude, the overall differences remain limited.

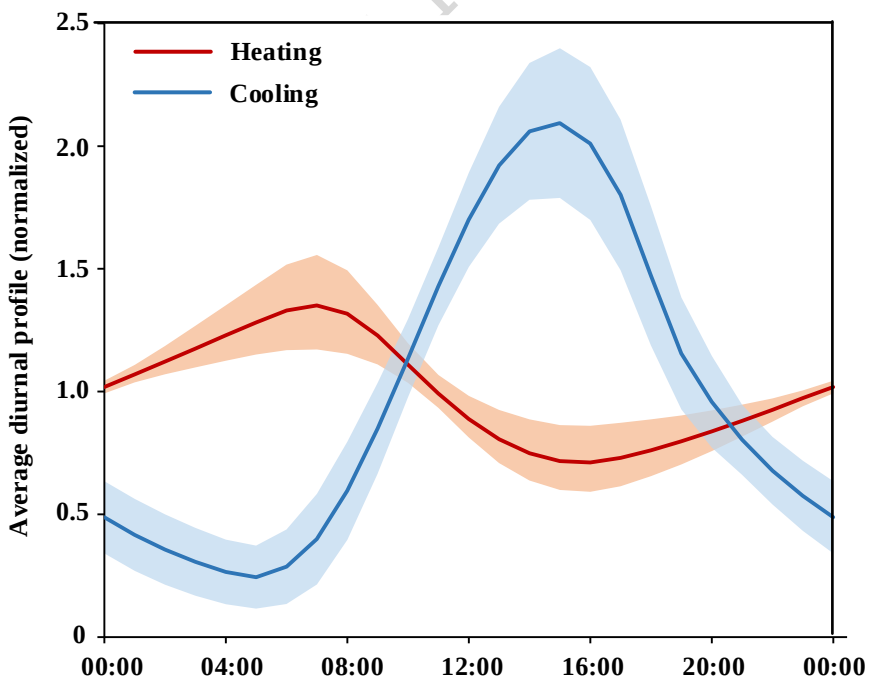

Fig. 6. Average diurnal profiles for heating and cooling demand.

Notes: The line represents the weighted mean across provinces of the daily mean values at each time point. The shaded area indicates the weighted standard deviation.

Since cooling and heating demands primarily occur in summer and winter, respectively, and their intraday fluctuation patterns are nearly opposite, this results in pronounced seasonal differences in the shapes of typical daily load curves. This also explains why applying the same typical day profile across the entire year is insufficient. In general, a typical load curve may exhibit a peak around noon or in the evening, with a trough in the early morning hours. As shown in Fig. 6, a higher midday cooling demand in summer raises the load during that period while reducing the evening demand and further lowering the early morning trough. In contrast, winter heating demand tends to suppress the midday peak, increase the evening load, and elevate the early morning baseline. Consequently, the peak-to-trough variation is usually larger in summer, with a single midday peak, whereas in winter, there may be both midday and evening peaks.

The influence of cooling and heating on the total load curve, however, varies across provinces due to significant differences in the relative proportions of these demands, making the overall impact more complex. Although our dataset captures daily variations, for clarity, Fig. 7 presents only the average typical daily load curves for summer and winter in each province. Most provinces follow the seasonal patterns described above, with the contrasts being especially prominent in residential and commercial centers, such as Beijing, Shanghai, and Chongqing.

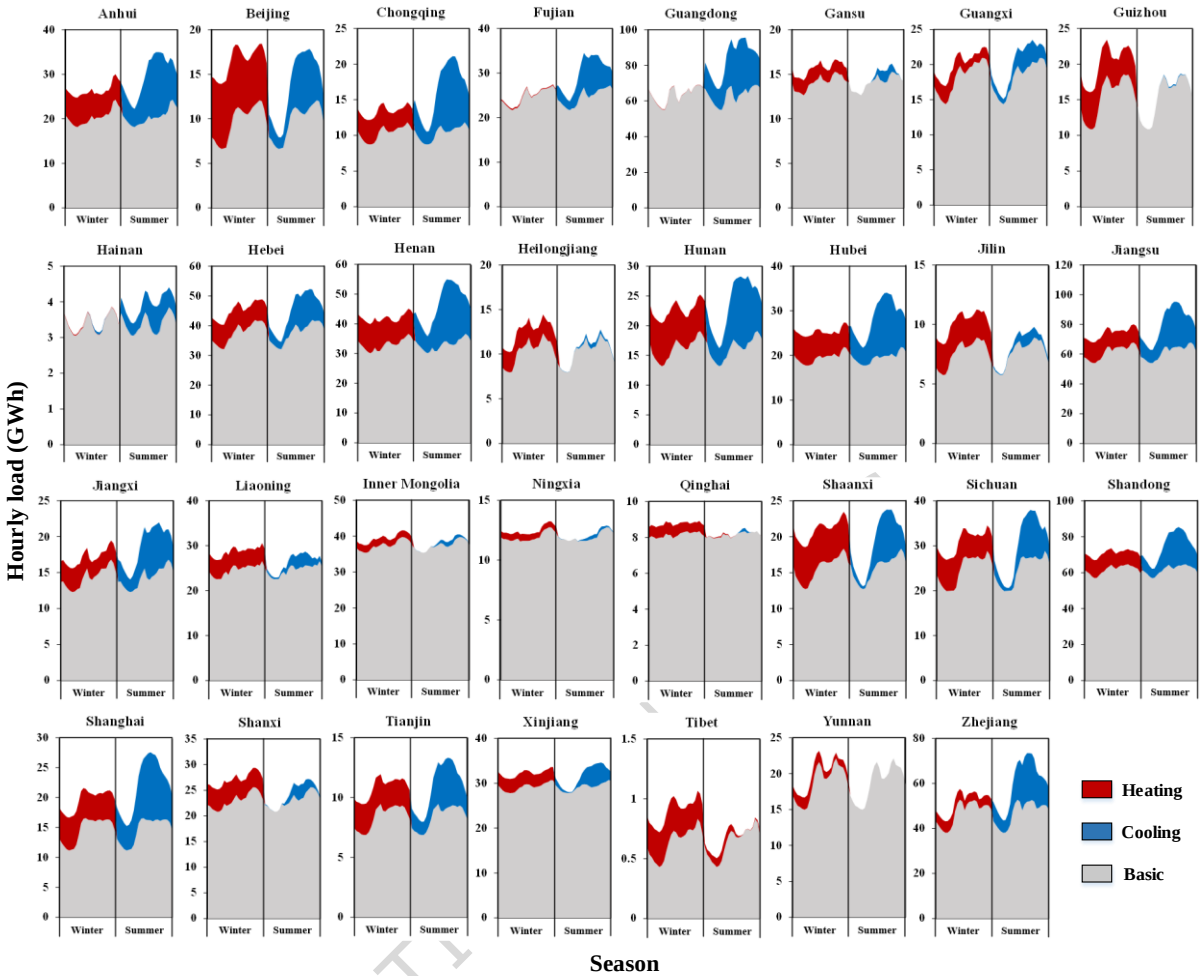

Fig. 7. Typical daily load curves for winter and summer.

Notes: The winter data is based on the average values for January, while the summer data is based on the average values for July.

Consistency checks

Industrial electricity consumption accounts for a dominant share of total electricity demand in China. Compared with weather-driven loads, this component is relatively stable and exhibits limited short-term variability associated with meteorological conditions. In this study, industrial load is therefore treated as a relatively smooth baseline component, controlled for by province-specific, year-specific, and day-type (weekday or weekend) effects. Fig. 8 illustrates the relationship between the annual industrial electricity share and the estimated share of electricity consumption for cooling and heating across 31 provinces over the period 2015–2022. A clear

negative association is observed: provinces with higher industrial electricity shares tend to exhibit lower shares of weather-driven electricity demand. This pattern is consistent with expectations, as a larger industrial base reduces the relative contribution of weather-driven loads. It also provides supporting evidence that the estimated cooling and heating components capture meaningful meteorological variation, rather than reflecting spurious correlations.

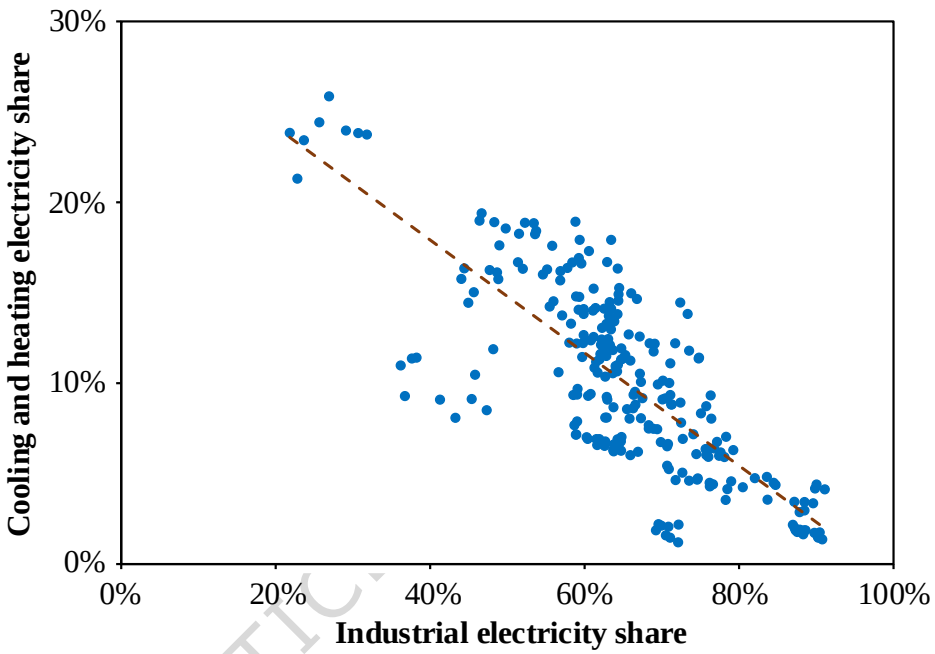

Fig. 8. Relationship between cooling and heating electricity share and industrial electricity share.

Another aspect that merits further validation concerns whether the air conditioner ownership per household used in the temporal extension of our analysis can adequately capture the evolving trends in the power coefficients associated with heating and cooling. In practice, such validation is challenging because China lacks direct data on actual cooling and heating demand. To address this limitation, we adopt an indirect approach by comparing trends in residential electricity consumption with the estimated cooling and heating electricity demand obtained in this study, as shown in Fig. 9. Residential electricity consumption is based on observed data, whereas cooling and heating electricity demand is model based. We acknowledge that these two indicators are highly correlated but do not fully overlap in scope. For example, residential electricity

consumption includes non cooling and non heating uses, while cooling and heating demand also partly arises from the commercial sector. Nevertheless, comparing their trends allows us to provide indirect evidence supporting the reliability of our results. As illustrated in Fig. 9, the two series display broadly similar trends with closely aligned growth rates. Over the period 2015–2024, the average growth rates of the two indicators are nearly identical, further confirming their consistency. Cooling and heating demand is more strongly influenced by weather conditions, which leads to weaker linearity and the occurrence of occasional single year extremes. Notably, 2018, which serves as the primary benchmark year in this study, was characterized by relatively higher temperatures in southern provinces, whereas 2020 represents a relatively cooler year. Overall, these findings suggest that the extension indicator employed in this study exhibits strong adaptability.

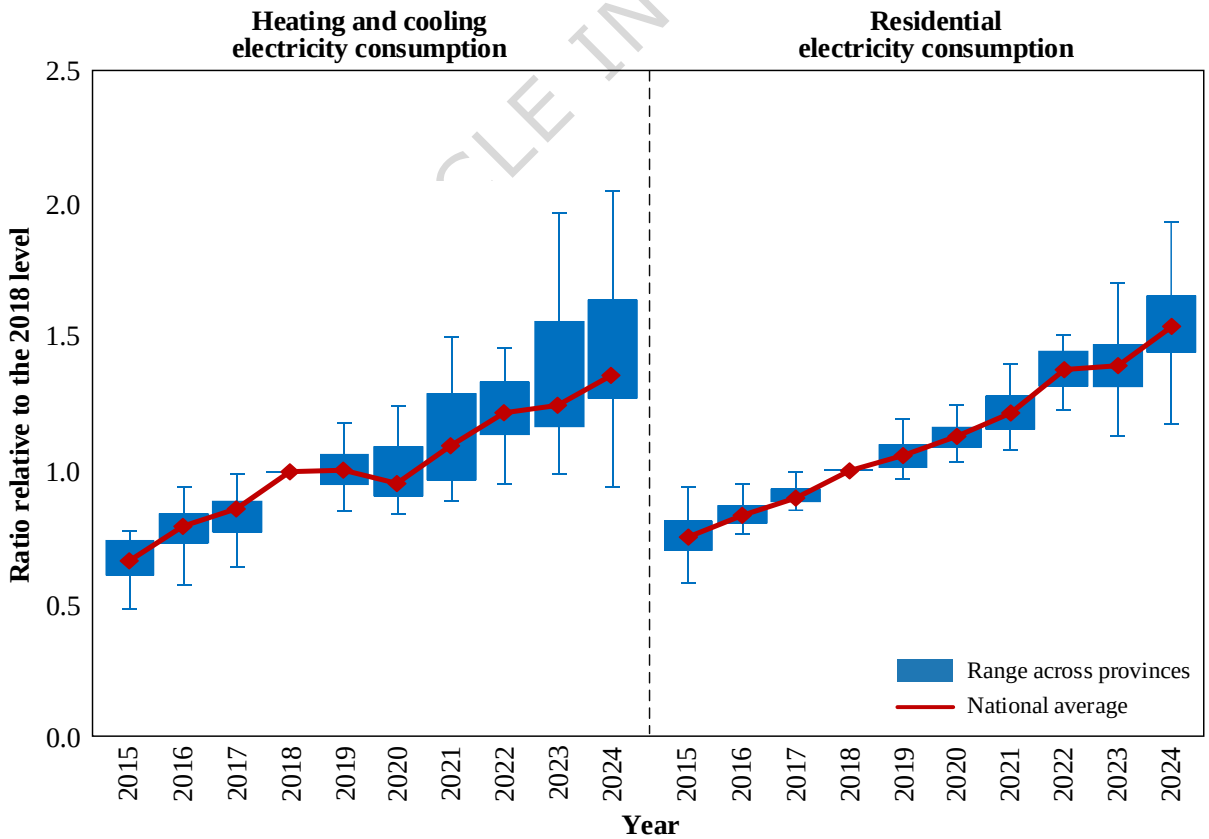

Fig. 9. Comparison of cooling and heating electricity demand with residential electricity consumption.

Notes: Taking the 2018 level as 1, values for other years are expressed as ratios relative to it. The line plot

represents the national average, while the box plots indicate the range across provinces.

Robustness checks

As part of the robustness checks for threshold temperatures, we examine the sensitivity of the results to an alternative specification by comparing region-specific estimates based on a north–south division with those obtained under a nationally uniform specification. Under the unified framework, the estimated cooling and heating thresholds are 23.471°C and 15.041°C, respectively, which are broadly consistent with the region-specific estimates for the north (22.253°C and 14.713°C) and the south (22.631°C and 16.818°C), indicating limited variation across different estimation strategies. As reported in Table 3, allowing for separate estimation by region improves model fit in approximately 68% of provinces, with more pronounced gains observed in Beijing, Tianjin, Yunnan, and Guizhou, where the R^2 increases by more than 0.02. At the same time, the overall differences remain modest, as most provinces exhibit only minor changes in goodness of fit, and decreases exceeding 0.01 are observed only in Hainan and Shandong. These results suggest that the main findings are not very sensitive to the specific choice of threshold temperature specification, supporting the robustness of the model.

As part of the robustness checks for smoothing parameters, we examine the sensitivity of the results to an alternative specification using a shorter smoothing window. In addition to the baseline 48-hour (2-day) moving window, we implement a 24-hour (1-day) window, which corresponds to an hourly decay factor of 0.8224. The results, reported in Table 3, show that the shorter window generally leads to a deterioration in model performance relative to the baseline specification. Improvements are observed only in a small number of provinces, notably Heilongjiang and Yunnan, while most provinces experience declines in goodness of fit. The reduction in R^2 is typically in the range of 0.01 to 0.04, indicating that although the shorter smoothing window performs slightly worse, the magnitude of the change remains moderate. These findings suggest that the main results are not highly sensitive to the choice of smoothing window, while also supporting the use of the 48-hour specification as the preferred baseline.

Table 3. Changes in goodness of fit under alternative threshold and smoothing specifications.

Province	Nationally unified threshold temperature		24-hour (1-day) moving window	
	R^2 change	NRMSE change	R^2 change	NRMSE change
Anhui	-0.0065	0.115%	-0.0230	0.394%
Beijing	0.0356	-0.825%	-0.0148	0.352%
Chongqing	-0.0029	0.068%	-0.0308	0.693%
Fujian	0.0075	-0.073%	-0.0218	0.209%
Guangdong	0.0039	-0.052%	-0.0227	0.302%
Gansu	0.0113	-0.043%	-0.0022	0.008%
Guangxi	0.0080	-0.047%	-0.0135	0.079%
Guizhou	0.0277	-0.351%	-0.0073	0.095%
Hainan	-0.0166	0.108%	-0.0311	0.198%
Hebei	0.0095	-0.077%	-0.0175	0.141%
Henan	-0.0070	0.114%	-0.0342	0.528%
Heilongjiang	-0.0012	0.011%	0.0031	-0.028%
Hunan	0.0030	-0.052%	-0.0255	0.442%
Hubei	-0.0022	0.040%	-0.0238	0.410%
Jilin	0.0015	-0.028%	-0.0057	0.107%
Jiangsu	0.0017	-0.021%	-0.0196	0.230%
Jiangxi	0.0041	-0.043%	-0.0235	0.246%
Liaoning	-0.0020	0.015%	-0.0087	0.065%
Inner Mongolia	-0.0014	0.004%	-0.0029	0.008%
Ningxia	-0.0005	0.001%	-0.0033	0.009%
Qinghai	0.0044	-0.013%	-0.0007	0.002%
Shaanxi	0.0109	-0.158%	-0.0247	0.351%
Sichuan	0.0022	-0.021%	-0.0418	0.382%
Shandong	-0.0137	0.102%	-0.0357	0.258%
Shanghai	0.0044	-0.105%	-0.0107	0.253%
Shanxi	0.0033	-0.022%	-0.0136	0.091%
Tianjin	0.0286	-0.337%	-0.0186	0.221%

Xinjiang	0.0133	-0.052%	-0.0099	0.039%
Tibet	0.0005	-0.006%	-0.0007	0.008%
Yunnan	0.0461	-0.131%	0.0034	-0.010%
Zhejiang	0.0064	-0.063%	-0.0147	0.145%

As part of the robustness checks for coefficient adjustment procedures, we assess the sensitivity of the results to the assumed elasticity coefficient used to adjust cooling and heating electricity demand when extending from the benchmark year 2018 to other years. In the baseline specification, an elasticity coefficient of 0.8 is adopted, while a plausible range of 0.5–1.0 is considered for sensitivity analysis. Fig. 10 presents the resulting growth rates of cooling and heating electricity demand under different elasticity assumptions, together with the observed growth in residential electricity consumption. The results indicate that variations in the elasticity coefficient do affect the estimated intertemporal evolution of cooling and heating demand, although the overall range remains moderate. Specifically, the increase in 2024 relative to 2018 ranges from 1.23 to 1.46 across the considered interval, compared to 1.36 under the baseline assumption. When benchmarked against residential electricity consumption, the baseline specification appears more consistent with observed trends: residential electricity consumption in 2024 is approximately 2.05 times that in 2015, while the corresponding multiples implied by elasticity coefficients of 0.5, 0.8, and 1.0 are 1.68, 2.04, and 2.33, respectively. These findings suggest that, although the elasticity assumption influences the magnitude of estimated growth, the main conclusions remain stable and the baseline parameter choice is well supported by observed data.

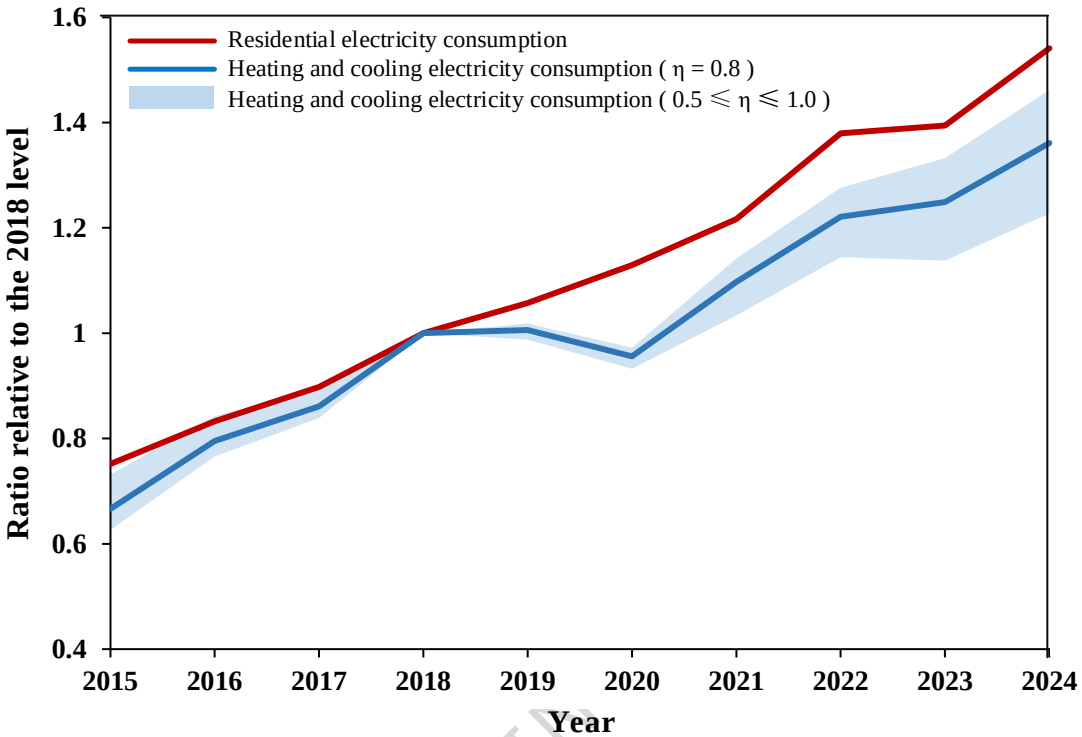

Fig.10. Sensitivity of elasticity coefficients in temporal extrapolation.

Limitations and extension

This study utilized the limited electricity load data released by the NDRC, combined with high-resolution meteorological data, to estimate the power coefficients and threshold temperatures for space cooling and heating across Chinese provinces. Using these coefficients, we reconstructed hourly electricity load profiles at the provincial level, filling a notable gap in existing empirical research. The proposed methodology is highly scalable and requires only the total annual electricity demand, air conditioner ownership per household for the target year, along with hourly meteorological variables such as temperature, wind speed, solar radiation, and relative humidity. This framework can be applied not only to the reconstruction of historical load profiles but also to projections of future demand. In particular, climate simulation outputs and electricity demand forecasts under various SSP scenarios from CMIP6 offer valuable inputs for modeling future electricity load curves.

This study has two primary limitations. First, it considered only weather-related drivers of

electricity demand and did not account for other influencing factors, which may result in discrepancies between the reconstructed load profiles and actual observations. Second, the reconstructed profiles are more applicable to the near term, since evolving patterns of electricity consumption, especially those driven by increasing electrification, including the adoption of electric heating, may significantly alter the empirical relationships derived from historical data. To improve the accuracy of future load projections, further disaggregation of electricity demand is required. This includes making assumptions about the penetration and technological advancement of cooling and heating systems. When combined with future meteorological scenarios, such enhancements can provide a more robust foundation for estimating long-term electricity demand and supporting planning efforts for power system transformation over extended horizons.

Code availability

The code can be obtained by contacting the corresponding author.

Data availability statement

The dataset generated by this study is openly accessible on Figshare[27]. The website is https://doi.org/10.6084/m9.figshare.29832701.

References

[1]. Markard J. The next phase of the energy transition and its implications for research and policy. Nature Energy, 2018, 3(8): 628-633.

[2]. Wang Y, Wang R, Tanaka K, et al. Accelerating the energy transition towards photovoltaic and wind in China. Nature, 2023, 619(7971): 761-767.

[3]. Zhang X, Patino-Echeverri D, Li M, et al. A review of publicly available data sources for models to study renewables integration in China's power system. Renewable and Sustainable Energy Reviews, 2022, 159: 112215.

[4]. Xu J H, Yi B W, Fan Y. Economic viability and regulation effects of infrastructure investments for inter-regional electricity transmission and trade in China. Energy Economics, 2020, 91: 104890.

[5]. Zhuo Z, Du E, Zhang N, et al. Cost increase in the electricity supply to achieve carbon neutrality in China. Nature Communications, 2022, 13: 3172.

[6]. Chen X, Liu Y, Wang Q, et al. Pathway toward carbon-neutral electrical systems in China by mid-century

with negative CO_2 abatement costs informed by high-resolution modeling. Joule, 2021, 5(10): 2715-2741.

[7]. Li M, Shan R, Abdulla A, et al. The role of dispatchability in China's power system decarbonization. Energy & Environmental Science, 2024, 17(6): 2193-2205.

[8]. Yi B, Zhang S, Fan Y. Economics of planning electricity transmission considering environmental and health externalities. iScience, 2022, 25(8): 104815.

[9]. Liu G, Liu J, Bai Y, et al. EWELD: A large-scale industrial and commercial load dataset in extreme weather events. Scientific Data, 2023, 10(1): 615.

[10]. Feron S, Cordero R R, Damiani A, et al. Climate change extremes and photovoltaic power output. Nature Sustainability, 2021, 4(3): 270-276.

[11]. Gernaat D E H J, de Boer H S, Daioglou V, et al. Climate change impacts on renewable energy supply. Nature Climate Change, 2021, 11(2): 119-125.

[12]. Pryor S C, Barthelmie R J. A global assessment of extreme wind speeds for wind energy applications. Nature Energy, 2021, 6(3): 268-276.

[13]. Perera A T D, Nik V M, Chen D, et al. Quantifying the impacts of climate change and extreme climate events on energy systems. Nature Energy, 2020, 5(2): 150-159.

[14]. Ruhnau O, Hirth L, Praktiknjo A. Time series of heat demand and heat pump efficiency for energy system modeling. Scientific Data, 2019, 6(1): 189.

[15]. Ramon D, Allacker K, De Troyer F, et al. Future heating and cooling degree days for Belgium under a high-end climate change scenario. Energy and Buildings, 2020, 216: 109935.

[16]. Antonini E G A, Di Bella A, Savelli I, et al. Weather-and climate-driven power supply and demand time series for power and energy system analyses[J]. Scientific Data, 2024, 11(1): 1324.

[17]. International Energy Agency (2018). The Future of Cooling. IEA, Paris, https://www.iea.org/reports/the-future-of-cooling.

[18]. International Energy Agency (2020). Is cooling the future of heating. IEA, Paris, https://www.iea.org/commentaries/is-cooling-the-future-of-heating.

[19]. Yan X, Huang Z, Ren S, et al. Monthly electricity consumption data at 1 km × 1 km grid for 280 cities in China from 2012 to 2019. Scientific Data, 2024, 11(1): 877.

[20]. Chen J, Gao M, Cheng S, et al. Global 1 km × 1 km gridded revised real gross domestic product and electricity consumption during 1992–2019 based on calibrated nighttime light data. Scientific Data, 2022, 9(1): 202.

[21]. Staffell I, Pfenninger S, Johnson N. A global model of hourly space heating and cooling demand at multiple spatial scales. Nature Energy, 2023, 8(12): 1328-1344.

[22]. Renewable Ninja Platform. https://www.renewables.ninja/.

[23]. China Electric Power Yearbook Editorial Committee (various years). China Electric Power Yearbook. China Electric Power Press, Beijing, China. (in Chinese)

[24]. National Development and Reform Commission (2019). Notice on the Signing of Medium and Long-term Electric Power Contracts in 2020. (in Chinese)

[25]. National Bureau of Statistics (various years). China Statistical Yearbook. China Statistics Press, Beijing, China. (in Chinese)

[26]. Jiang K, Liu N, Wang K, et al. Spatiotemporal assessment of renewable adequacy during diverse extreme weather events in China[J]. Nature Communications, 2025, 16(1): 5198.
[27]. Yi B (2026). Electricity load curve. Figshare. Dataset. https://doi.org/10.6084/m9.figshare.29832701.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant No.W2412161, No.72374018, and Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China under Grant No. JYB2025XDXM904.

Author contributions

B.Y. and Y.F. designed the research and prepared the manuscript. B.Y., Q.L. and Y.J. contributed to data collection and data processing. S. Z. and S. Y. contributed to data analysis.

Competing interests

The authors declare no competing interests.