Analyzing Residential Energy Consumption in the Greater Bay Area: A STIRPAT Model Approach

Yi Yan*

*2482516799@qq.com

Tsinghua University, Beijing, China

Abstract. This paper utilizes the STIRPAT model to analyze residential energy consumption in the Greater Bay Area, focusing on the interplay of population growth, economic development, and urban residential infrastructure. Incorporating factors such as per capita population, GDP, residential building area, and climatic variables like cooling degree days and geographical coordinates, the study reveals significant correlations between these factors and per capita electricity consumption in residential buildings. Empirical analysis indicates that increases in both per capita population and GDP significantly elevate residential electricity consumption, whereas changes in residential building area show minimal impact. The research highlights the importance of considering socio-economic and environmental factors in urban energy planning, providing insights for developing sustainable energy strategies in urban settings. The findings particularly emphasize the need for targeted energy efficiency measures in rapidly urbanizing areas, taking into account the diverse socio-economic and climatic conditions prevalent in such regions.

Keywords. STIRPAT Model, Residential Energy Consumption, Greater Bay Area, Socio-Economic Factors, Urban Energy Efficiency

1 Introduction

Residential energy consumption is a pressing concern, particularly in urban areas. This study delves into the unique context of the Greater Bay Area, characterized by stable climatic conditions but varying levels of socio-economic development across its cities. To explore the intricate factors influencing energy use in urban households, we adopt the STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology) model, building upon the foundational IPAT (Impact, Population, Affluence, Technology) framework. The Greater Bay Area serves as an intriguing backdrop for our investigation due to the diversity among its cities and their distinct developmental trajectories. The study's primary aim is to unravel the interplay of population growth, economic prosperity, and residential infrastructure in shaping energy consumption patterns. Given the rapid urbanization in this region, gaining a nuanced understanding of these factors is imperative for the formulation of sustainable energy policies.

Our research goes beyond mere quantification of energy consumption; it seeks to reveal the underlying drivers responsible for variations in energy usage among different cities within the Greater Bay Area. Our analysis encompasses diverse data, including per capita population, GDP, residential building area, and climate variables. These insights are poised to inform

BDEDM 2024, January 12-14, Ningbo, People's Republic of China

DOI 10.4108/eai.12-1-2024.2347177

policymakers and urban planners, underscoring the necessity of tailored strategies in the context of swiftly urbanizing regions.

2 STIRPAT Model Theory

2.1. The IPAT Model

To quantitatively reveal the impact of population growth, economic development, and technological progress on environmental pressure, American demographers Ehrlich and Holden proposed the IPAT model in 1971:

$$I = PAT \tag{1}$$

Where I (Impact) represents environmental impact, P (Population) refers to the population, A (Affluence) represents affluence, and T (Technology) stands for technological progress. The environmental impact I on the left side of the equation can be expressed using different indicators such as carbon emissions and energy consumption [1]. On the right side of the equation, the population P is typically represented by the size of the population, affluence A is often measured using indicators like per capita GDP and per capita disposable income, and technological progress T is determined based on the type of environmental impact I, such as carbon emissions per unit of GDP or energy consumption per unit of building area. Many applications of this model exist, including the well-known Kaya Identity:

$$CO_2\ emissions = population \times \left(\frac{GDP}{population}\right) \times \left(\frac{CO_2 emission}{GDP}\right)$$
$$= population \times \left(\frac{GDP}{population}\right) \times \left(energy\frac{comsumption}{GDP}\right) \times \left(\frac{CO_2 emissions}{energy} comsumption\right)) \tag{2}$$

In most cases in China, the population P in the IPAT equation is continuously increasing, while affluence A continues to grow with economic development. Under the influence of these two factors, environmental impact I also continues to increase. Therefore, only by improving technological levels, i.e., reducing emission intensity and energy consumption intensity, can environmental pressure be effectively mitigated. The IPAT model clearly illustrates that population size, economic development, and technological level are the main drivers of changes in environmental impact. Furthermore, the model points out that each factor does not act independently of the others in affecting environmental impact [2]. For example, in a certain country over a period of time, its population P and technological level T remain relatively constant, but economic level A continues to grow. In this case, it cannot be assumed that the growth in environmental impact I is solely due to the increase in A but is rather influenced by both P and T, which are reflected in the changes in economic level A.

2.2. STIRPAT Model

Due to the IPAT model, which examines the impact of a particular factor on the environment while keeping other factors constant, resulting in proportional effects on the dependent variable, this is also the main limitation of the model. To address this limitation, Dietz and Rosa (1994) built upon the IPAT framework by establishing a stochastic model to study the non-proportional impacts of factors such as population on environmental pressures, known as the STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology) model:

$$I = aP^b A^c T^d e \tag{3}$$

In the equation above, the newly added "a" is the model exponent, "b," "c," and "d" represent the exponents for factors such as population size, affluence, and technology level, and "e" represents the model residuals [3]. The STIRPAT model retains the three key elements of population, affluence, and technology level from the original IPAT model, while transforming the simple linear model into a multivariate nonlinear model. It can be observed that the IPAT model is a special form of the STIRPAT model when $a = b = c = d = e = 1$. Taking the natural logarithm of both sides of the equation, we get:

$$\ln I = \ln a + b \ln P + c \ln A + d \ln T + \ln e \tag{4}$$

Since both the dependent and independent variables are in natural logarithmic form, the coefficients in the model represent percentage changes. Additionally, the three factors P, A, and T can be further decomposed, meaning that other related indicators can be introduced for more detailed analysis and research.

Furthermore, York et al. (2003) introduced the concept of Ecological Elasticity (EE) for a deeper analysis. Ecological Elasticity refers to the proportionate change in environmental impact I when a particular influencing factor changes. For example, the population elasticity EEP of environmental impact I refers to the extent of change caused by variations in population size, while the affluence elasticity EEA is the magnitude of change in environmental impact due to changes in economic levels (such as per capita GDP or GNP). In the equation above, "b" and "c" represent population elasticity and affluence elasticity, respectively. However, as of now, there is no defined concept of technology elasticity, and there is still some controversy surrounding the measurement indicators for technology level [4].

3 Analysis Model of Urban Residential Building Energy Consumption Factors Based on STIRPAT

3.1. Factor Selection

Building energy use is influenced by various factors, including not only physical characteristics like climate, heating methods, housing size, and building age, but also socio-economic factors and resident behaviors. In the subtropical Greater Bay Area, climate conditions are relatively uniform, but significant socio-economic disparities among its cities contribute to varying levels of residential energy consumption. For instance, a notable positive correlation exists between residential energy consumption and population size across the cities of the Greater Bay Area. This is because as populations grow, the demand for energy naturally increases, aligning residential energy use with city populations.

In addition to population, there's a link between residential energy consumption and real GDP, except in the Hong Kong and Macao Special Administrative Regions. Here, economic growth boosts residents' incomes, leading to increased energy usage for activities like using more home appliances and extended air conditioning. However, in Hong Kong and Macao, residential energy consumption hasn’t shown much sensitivity to GDP growth over the past 18 years, likely due to stable living standards and consistent household appliance ownership [5].

Moreover, total residential energy consumption is positively related to the amount of residential stock. While some new housing units may remain unoccupied initially, an overall increase in housing stock tends to drive up energy consumption. In cities like Guangzhou, Shenzhen, and Foshan, residential energy consumption has shown an exponential relationship with the increase in housing stock, suggesting that new residential developments lead to significant rises in energy use. Consequently, there is a pressing need to enhance building energy efficiency in these areas like figure 1.

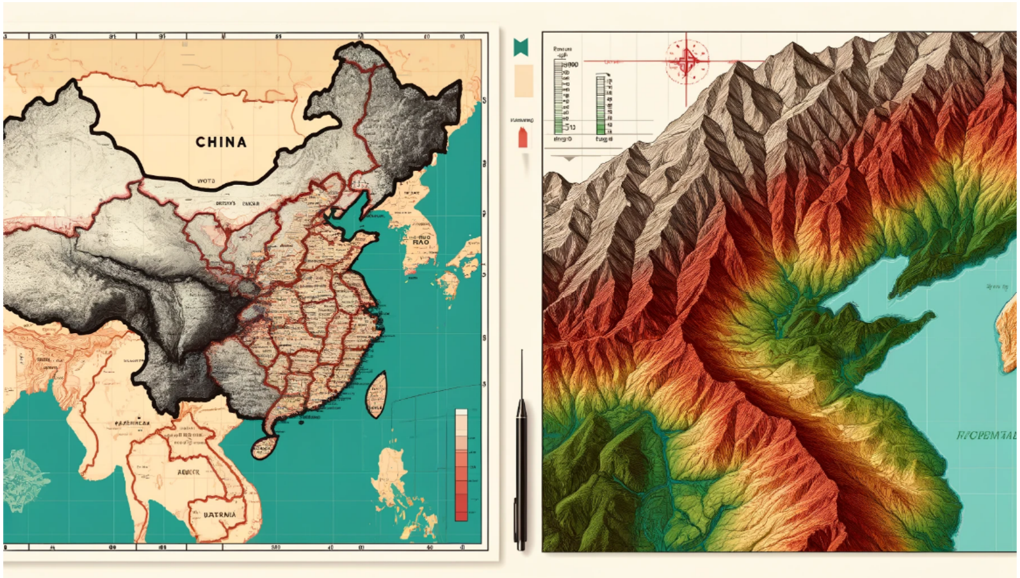

Figure 1. Establishing coordinated development index of urbanization based on multi-source data: A case study of Guangdong-Hong Kong-Macao Greater Bay Area, China.

3.2. Model Construction

The differential analysis presented in the previous chapter revealed that the variations in residential building energy consumption among cities in the Greater Bay Area are primarily concentrated in electricity usage. Therefore, this paper takes the per capita electricity consumption in residential buildings in each city as the focus of analysis. Based on the STIRPAT model theory, this paper regards the per capita electricity consumption in residential buildings in cities of the Greater Bay Area as the environmental impact "I." According to the qualitative analysis of factors influencing building energy consumption in the previous section, this paper selects per capita population, per capita residential building area, and per capita GDP of urban residential buildings to reflect the influence of population and affluence [6]. Additionally, it introduces the cooling degree days (CDD_26), the longitude, and latitude of the cities as control variables representing climatic conditions. Based on this, this paper establishes a regression analysis model for the factors affecting residential building energy consumption:

$$\ln(EI_{it}) = \alpha + \beta_1 \cdot \ln(HS_{it}) + \beta_2 \cdot \ln(pFA_{it}) + \beta_3 \cdot \ln(pGDP_{it}) + \beta_4 \cdot CDD\_26_{it} + \beta_5 \cdot longitude_i + \beta_6 \cdot latitude_i + \varepsilon_{it} \quad (5)$$

In the equation above, "EI_{it}" represents the per capita electricity consumption in residential buildings in city "i" in year "t," "HS_{it}" represents per capita population, " pFA_{it}" represents per capita residential building area, " $pGDP_{it}$" represents per capita GDP, " CDD_26_{it}" represents cooling degree days, "$longitude_i$" and "$latitude_i$" represent the longitude and latitude of the city, respectively. The remaining parameters include the intercept term " " and the model error " ε." Since both the independent and dependent variables are in logarithmic form, the regression coefficient " β " in the model indicates that when the independent variable changes by 1%, the dependent variable changes by " β %." When only the dependent variable is log-transformed, it indicates that a one-unit change in the independent variable results in a " %" change in the dependent variable.

4 Results Analysis

4.1. Descriptive Statistical Analysis

The descriptive statistical results of the variables related to the regression model in this paper are shown in Table 1. below.

Table 1. Descriptive Statistics of Regression Model Variables.

Model Variables	Sample Size	Mean	Standard Deviation	Minimum	Maximum
EI (units: kgce(hh·a))	198	603.0	290.1	182.8	1626.7
HS (units: person)	198	3.109	0.329	2.284	4.189
pFA (units: m2)	198	28.06	8.344	12.13	59.63
pGDP (units: 104 RMB)	198	7.900	9.872	0.742	5.595
CDD_26 (units: °C·d)	198	393.98	81.300	105.20	660.00
Longitude (units: °)	198	143.53	0.582	102.47	118.42
Latitude (units: °)	198	12.61	0.480	21.05	28.13

4.2. Model Specification

For panel data, there are two extreme estimation strategies: one treats it as mixed cross-sectional data, requiring that each individual in the sample has the same regression equation, and the other estimates a separate regression equation for each individual. Both of these approaches have their limitations; the former ignores unobservable or omitted heterogeneity, while the latter neglects commonality among individuals and may suffer from inadequate sample size. Therefore, previous research often adopted a compromise estimation strategy, assuming that individual regression equations have the same slopes but can have different intercept terms to capture heterogeneity. The model corresponding to this estimation strategy is the "individual effects model":

$$Y_{it} = X'_{it}\beta + u_i + \gamma_t + \varepsilon_{it} \tag{6}$$

Where 'X_{it}'' represents explanatory variables that vary with individuals and time, and the disturbance term is composed of three parts, i.e., the "composite disturbance term," denoted as $(u_i + \gamma_t + \varepsilon_{it})$. If '$u_i$' is correlated with an explanatory variable, the model is a "fixed effects model." If 'γ_t' varies with time, the model is a "time fixed effects model." If both of these conditions are present, the model is an "individual-time fixed effects model." Furthermore, if 'u_i' is not correlated with any explanatory variable, the model is a "random effects model."

Since most of the individual effects models in previous studies were fixed effects models, this paper first verifies whether the constructed model is a mixed regression model or an individual effects model from the perspective of fixed effects. The paper also employs both F-tests (Chow tests) and LR tests to avoid bias resulting from using a single testing method.

The null and alternative hypotheses for the two tests are as follows:

(1) H_0: Under the assumption that γ_t varies with time,$u_1 = u_2 = \cdots = u_n$, , i.e., the model is a mixed model. H_1: Under the assumption that γ_t varies with time, not all 'u_i' are equal, i.e., the model is an individual-time fixed effects model.

(2) H_0: $u_1 = u_2 = \cdots = u_n$, i.e., the true model is a mixed model. H_2: not all 'u_i' are equal, i.e., the true model is an individual fixed effects model.

(3) H_0:$\gamma_1 = \gamma_2 = \gamma_3 = \cdots \gamma_T$, i.e., the true model is a mixed model. H_3: λt is not all equal, i.e., the true model is a time fixed effects model.

Based on the results of the F-test and LR-test, it can be preliminarily inferred that the model constructed in this paper is not a mixed model. Then, the Hausman test (H_0: 'u_i' is uncorrelated with 'X_{it}'') is used to determine whether to use fixed effects or random effects. The calculated test statistic is 5.53 (P = 0.2367). Since the P-value is greater than 0.1, the null hypothesis is not rejected, and the regression model constructed in this paper is a random effects model, not a fixed effects model.

4.3. Model Estimation Results

The estimated results of the regression model constructed in this paper are shown in Table 2. In this table, the Wald test statistic is 912.63, and the corresponding associated probability is less than 0.001, indicating that the choice of a panel data model with individual random effects is reasonable, and the model has a high coefficient of determination (R-squared of 0.911), suggesting a good fit. Regarding model coefficients, per capita population, per capita GDP, and city latitude in the cities of the Greater Bay Area significantly affect the per capita electricity consumption of residential buildings, while per capita residential building area, air conditioning degree days, and city longitude do not have a significant impact.

Table2. The estimation results of the regression model.

Independent Variables	Model Coefficients
ln HS	0.917***（>0.001）
ln pFA	0.124（0.244）
ln pGDP	0.500***（>0.001）
CDD_26	>0.001（0.806）
Longitude	0.414（0.596）

Latitude	-0.299**（0.021）
Constant Term	5.068（0.554）
Wald	272.37***（>0.001）
Sample Size	178
R^2	0.811

Through empirical analysis, it is evident that there is a significant positive relationship between per capita electricity consumption in residential buildings and per capita population in households:

When the per capita population in households increases by 1%, per capita electricity consumption on average increases by 0.919%. The economic level of residents also has a significant positive impact on per capita energy consumption in residential buildings. When per capita GDP increases by 1%, per capita electricity consumption on average increases by 0.600%. This is in line with previous research findings, indicating that both population growth and an improvement in residents' economic status contribute to the growth of residential energy consumption. Population growth drives the essential demand for energy, while an increase in economic level corresponds to an improved standard of living, prompting residents to purchase more household appliances and use them for longer periods.

5 Conclusion

Based on the STIRPAT model theory, three influencing factors—per capita population, per capita residential building area, and per capita GDP—along with three control factors—air conditioning degree days, urban longitude, and latitude—were selected for regression analysis of per capita electricity consumption in residential buildings across cities in the Greater Bay Area. The results indicate that per capita population and per capita GDP have significant promoting effects on per capita electricity consumption in residential buildings: for every 1% increase in per capita population, there is a 0.919% increase in per capita electricity consumption, and for every 1% increase in per capita GDP, there is a 0.600% increase in per capita electricity consumption. Per capita residential building area does not have a significant impact on per capita electricity consumption in cities across the Greater Bay Area, mainly because the increase in per capita residential building area in some cities in the Greater Bay Area over the past eighteen years has been minimal. Among the climate conditions, only urban latitude has a significant impact on residential energy consumption in cities in the Greater Bay Area: as latitude increases, the annual average temperature gradually decreases, reducing air conditioning usage by household residents during the summer to some extent, thereby lowering electricity consumption in residential buildings.

References

[1] Kim, Tae-Young, and Sung-Bae Cho. "Predicting residential energy consumption using CNN-LSTM neural networks." Energy 182 (2019): 72-81.

[2] Tsemekidi Tzeiranaki, Sofia, et al. "Analysis of the EU residential energy consumption: Trends and determinants." Energies 12.6 (2019): 1065.

[3] Belaïd, Fateh, David Roubaud, and Emilios Galariotis. "Features of residential energy consumptio

n: Evidence from France using an innovative multilevel modelling approach." Energy policy 125 (2019): 277-285.
[4] Pinjari, Abdul Rawoof, and Chandra Bhat. "Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: Application to residential energy consumption analysis." Journal of choice modelling 39 (2021): 100283.
[5] Usman, Ahmed, et al. "Revealing the nexus between nuclear energy and ecological footprint in STIRPAT model of advanced economies: Fresh evidence from novel CS-ARDL model." Progress in Nuclear Energy 148 (2022): 104220.
[6] Usman, Muhammad, and Nesrine Hammar. "Dynamic relationship between technological innovations, financial development, renewable energy, and ecological footprint: fresh insights based on the STIRPAT model for Asia Pacific Economic Cooperation countries." Environmental Science and Pollution Research 28.12 (2021): 15519-15536.