Analysis of The Impact of Urban Social Electricity Consumption Based on GAM Model: Empirical Evidence from China From 2010 to 2020

Zichun Liu[1, #], Jinghang Zou[2, *, #], Hongbiao Zeng[3, #]

[1]School of Science, Sichuan Agricultural University, Ya'an, China

[2]School of Mathematics, South China University of Technology, Guangzhou, China

[3]College of Computer Science, Guangdong University of Science and Technology, Dongguan, China

*Corresponding author: 1092654852@qq.com

[#]These authors contributed equally.

Abstract. As China continues to be the world's largest country in terms of electricity generation, evaluating the factors affecting its urban social electricity consumption can provide a scientific basis for its energy conservation and sustainable development. Based on the data from Chinese cities from 2009 to 2019, this study uses the fixed effect model (FEM) and the generalized additive model (GAM) to establish the relationship between urban social electricity consumption and its influencing factors and the linear regression equation to compare and evaluate the prediction and fitting effects of the models. The results show that FEM and GAM can effectively evaluate various factors affecting urban social electricity consumption. However, the input variables in FEM do not fit well with the output variables, and the interpretability of GAM is better than the FEM model. The number of hospitals, Total Natural Gas, and Per Capita GRP have the highest contribution to Social electricity consumption. Per Capita GRP, the Number of Hospital and Total Natural Gas are the highest contributors to Industrial electricity consumption, Number of GAM has the highest contribution to Social electricity consumption, Industrial electricity consumption and Household electricity consumption. The fitting and prediction results of GAM for Social electricity consumption, Industrial electricity consumption and Industrial electricity consumption are better than those of FEM.

Keywords: Urban Social Electricity Consumption, Fixed Effect Model, Generalized Additive Model, Comparative Analysis.

1. Introduction

With the continuous improvement of people's living standards and the easing of the domestic epidemic, China's total social electricity consumption in 2021 was 8,312.8 billion kilowatt-hours, an increase of 10.7% year-on-year, ranking first in the world with nearly 29% of power generation. The influencing factors in various aspects need to be further studied and explored for the rapid growth of social electricity consumption.

The previous judgmental analysis of the social electricity consumption structure is too old and lacking in foresight, and the influence of each social factor on the fluctuation of electricity consumption has not been studied in depth. This study explores the correlation between each factor and social electricity consumption from influencing factors in the city, and the Hausman test confirms the feasibility and operability of the idea.

Theoretical significance reveals the intrinsic correlation between the two, comprehensively analyzes the degree of influence, opens up new research ideas for the development trend of social electricity consumption, and provides reliable theoretical support for urban electricity planning. In terms of practical significance, it serves as a reference for urban planners to guide the overall development of society and serves as an inspiration to improve the revision of the social electricity structure.

This paper focuses on the influence of various essential factors of cities on social electricity consumption. It proposes a GAM model adapted to different cities, which is the first time the GAM model has been used to determine the correlation and degree of influence between the two. The main research objectives are twofold：

(1) To construct a reasonable mathematical model by selecting independent and dependent variables from various aspects of different cities in different periods.

(2) The combination of independent and dependent variables are connected through linear and non-linear relationships of the GAM model to reduce the model risk caused by a single linear setting and comprehensively and reasonably judge the degree of connection and influence between factors related to social electricity consumption and their correlation.

2. Literature review

2.1 Status of foreign research

Most of the foreign research results are studying the influence of a few specific factors on the change of social electricity consumption. Nowadays, research on the impact of electricity consumption expands the scope of studies and diverse research methods. In 2020, Nguyen Can Trong et al. (2020) [7] applied ground-based remote sensing data to analytically explore the relationship between surface environmental factors (SEF) and electricity consumption (EC) in the Bangkok metropolitan area, concluding with a combined analysis that temperature extremes will significantly increase the total electricity consumption in urban areas.

Following this, Sarkodie Samuel Asumadu et al. (2021) [6] used a machine learning estimator to study climate and energy consumption and the impact on the household, industrial sector, and commercial electricity consumption in Norway. It was concluded that climate and weather patterns change equally to influence electricity consumption. Mathpal Megha et al. (2022) [8] used the Johansen cointegration test to analyze the long-run electricity consumption (PCEC), per capita gross domestic product (PCGDP), and urban population (UP) in India from 1991 to 2018 based on association, while the Granger causality test indicates a two-way causality between PCEC and PCGDP and a one-way causality between EMP, UP and PCGDP.

2.2 Status of domestic research

There are many factors related to social electricity consumption and a large data surface, so many studies on the factors influencing electricity consumption based on extensive data analysis have appeared. For example, Li Yongyi et al. (2019) [5] made a detailed analysis and study in terms of the growth of residential electricity consumption and the proportion of residential electricity consumption through statistical calculations of socioeconomic as well as big data of electricity consumption of residents in Shaanxi Province since 2000.

Similarly, correlation and regression analysis can be studied to correct the model and improve the accuracy. Luo Wenyun et al. (2021) [4] conducted a linear regression analysis on the data of urbanization rate and social electricity consumption in Guizhou province using SPSSAU. It was analyzed the calculated predicted values against the actual values; Zhang Youguo et al. (2021) [9] conducted a qualitative analysis of the data of each variable from 2004 to 2018 in Shanghai by unit root and Granger causality test with gray correlation quantitative analysis, and PCA was used to correct the model-related factors. From the comprehensive research results. It can be seen that domestic and foreign scholars have done much work on the impact of electricity consumption and have achieved many valuable research results. The research results of domestic and foreign scholars on the impact factors of electricity consumption are shown in Table 1.

Table 1. Summary of research on factors influencing electricity consumption by domestic and foreign scholars

No.	Year	Author	Title of the literature	Processing
1	2019	Li Yong Yi Shi Rong	Research on residential electricity consumption behavior and influencing factors in Shaanxi Province based on big data analysis	Big Data Analytics
2	2020	Nguyen Can Trong Phan Diem Kieu	Factors affecting urban electricity consumption: a case study in the Bangkok Metropolitan Area uses an integrated earth observation data and data analysis approach.	Data Combination Analysis
3	2021	Ahmed Maruf Yakubu	Ambient air pollution and meteorological factors escalate electricity consumption	Machine learning estimator
4	2021	Luo Wenyun Wang Hao	Analysis of the application of urbanization rate in electricity demand forecasting	SPSSAU linear regression analysis
5	2021	Zhang You Guo Gao Yan	Forecasting electricity demand in Shanghai-based on gray correlation analysis and multiple regression	Granger causality test and grey correlation
6	2022	Mathpal Megha	Economic Growth and Electricity Consumption in India: An Econometric Analysis	Johansen cointegration test and Granger causality test

3. Methodology

3.1 GAM model principles and step-by-step implementation

Hastie first proposed Generalized Addition Model (GAM) and Tibshirani (1986) [10], which can contain both parametric and non-parametric components. It has a flexible setting and requires fewer samples, which has broad applicability. It can objectively express the linear and non-linear relationships between the explanatory and explained variables. It can objectively express the linear and non-linear relationships between the explanatory variables and the explanatory variables and reduces the model risk caused by linear settings. Its general form is:

$$g(E(Y \mid X)) = \sum{}_i \beta_i X_i + \sum{}_j f_j(X_j) + \varepsilon \tag{1}$$

Among them, $g(\cdots)$ is the connection function whose form depends on the specific form of the distribution of the explanatory variable Y, ε is the random error term, the name of the connection function where the explanatory variable is usually distributed is Identity, and the connection function is of the form $g(\mu) = \mu$, where $\mu = E(Y|X)$ and the model requires only additivity and $E(\varepsilon|X) = 0$. X_i is the explanatory variable that strictly obeys the parametric form, β_i is the corresponding parameter, and $f_j(\cdots)$ is the smoothing function corresponding to the explanatory variable X_j that obeys a non-parametric form. The generalized addable model approach can be extended to handle different types of data.

For the data used in this paper, the data were first organized and preprocessed using Excel to form a panel data set, and basic descriptive statistics were done using STATA, and then the Hausman test under the three independent variables was done separately.

The implementation of the generalized additive model in this paper can be divided into the following 3 steps: variable pre-analysis, model construction, and model prediction. The generalized additive model is a combination of a generalized linear model and an additive model according to the specific situation, where a generalized linear model is a functional form that changes the conditional expectation of the response variable, which is written as $g(muY)$, where $muY = E(Y/X_1, \cdots, X_n)$, which can be expressed as:

$$g(muY) = \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n + \varepsilon \tag{2}$$

It can also be used in a non-parametric form to describe the correspondence between the conditional expectation of the response variable and the explanatory variables, denoted by $f(x)$, which can be expressed as $E(Y/X) = f(x)$. When it is generalized to multiple explanatory variables, it is noted as the additive model:

$$E(Y / X_1, \cdots, X_n) = f(X_1) + \cdots + f(X_n) \tag{3}$$

Combining the two approaches, the generalized additive model can be expressed as:

$$g(muY) = \alpha_0 + f_1(X_1) + \cdots + f_n(X_n) + \varepsilon \tag{4}$$

Among them, in the $g(muY)$ term, muY is the expectation of Y, $g(\cdots)$ is the connection function, α_0 is the intercept moment, and in the $f_n(X_n)$ term, $f_n(\cdots)$ is the univariate function of the explanatory variable X_n and ε is the random variable.

3.2 Indicator Selection

The number of influencing factors related to social electricity consumption is significant. It is supposed that all of them are introduced into the prediction model. In that case, the prediction model will often be disturbed by non-essential factors, resulting in little realistic and economic significance of the parameter estimates, which may also lead to the loss of significance tests for some explanatory variables, etc. Therefore, it is necessary to screen the variables and then determine the input variables of the prediction model, and in this paper, the independent variables are selected by referring to the literature and comparing the effects.

There are external relationships and internal logic between each indicator and dependent variable for the basis of indicator selection.

(1) Analysis from a macro perspective. Population: electric power products are an essential part of residents' daily life, and at the same time, human social activities are accompanied by the consumption of electric power, and as the size of the population continues to expand, its electricity consumption will also continue to increase. GDP: the development of the economy cannot be supported by the development of electric power, and the development of electric power also requires economic development at the same time. When there is a shortage of electricity, it will seriously affect the operation of all industries in the national economy, thus affecting the growth of the national economy [12]. The GDP of a region reflects the overall electricity demand and the size of electricity consumption in a city. GDP per capita of the municipality: considering only the total GDP of the city without considering the population size can cause errors, and for places with a small population size but high GDP per capita, the consumption of electricity remains high and is also an important indicator to consider affecting electricity consumption.

(2) Analysis from the perspective of building facilities. The number of schools: Schools build many classrooms and meeting rooms, and each indoor space requires a large amount of electricity to support the teaching equipment, and there is an evident demand for lighting. The number of hospitals: Lighting and the use of many medical devices consume a large amount of electricity and are an enormous source of electricity consumption for society. The number of industrial enterprises: Enterprises related to light and heavy industries will use large machinery, lighting, lighting, office electricity, etc. Such machinery and equipment are sources of social electricity consumption. Area of built-up areas in municipalities: Wu Wei (2021) [11] and others showed that increased density of

built-up areas was significantly negatively correlated with residential electricity consumption. Urban roads: The number and length of roads determine several electricity consumptions factors such as street lighting electricity consumption, the number of traffic lights, and camera electricity consumption, which can reflect the prosperity of the city to a certain extent and also influence the size of social electricity consumption to a certain extent.

(3) Analysis from other aspects. Natural gas: As the main source of energy for all types of activities, natural gas can reduce the economic burden of households through substitution, so natural gas and electricity are, to a certain extent, substitution relations. Total annual bus (electric) passenger trips: The start-up of a bus will consume electricity. For the current large number of buses, the total number of trips can directly reflect the number of bus trips, which can also affect the electricity consumption to a certain extent.

All data are derived from the China Urban Statistical Yearbook from 2009 to 2019, and the main editor is the Department of Urban Socio-Economic Survey of the National Bureau of Statistics. Based on the above basis reasons, we selected a total of 10 indicators above as the independent variables of the model. Therefore, the impact indicator system constructed in this paper is shown in Table 2.

Table.2. Introduction to the Impact Indicator System

Dimensionality	Indicators	Unit of quantity
Demographic factors	Population	10000 persons
Economic Development Level	GDP	10000 yuan
	Per Capita GRP	Yuan
Building Facilities	School	Number
	Hospital	Number
	Industrial enterprises	Number
	Area of built-up area	Square kilometers
	Urban road area	10000 square meters
Alternative Energy	Total Natural Gas	10000 cubic meters
Traveling Mobility	Total number of public bus (electric) passenger transport in the whole year	10000 person-times

4. Result And Analysis

4.1 Sample profile and statistical analysis

(1) Variable selection and data sources

This paper focuses on the research on the effects of various factors on the social power consumption city, which aims to build the influence of different indicators and the connection of social electricity consumption. Therefore, we have collected all the cities in the Chinese provinces in 2009-2019 power consumption(Social electricity consumption, Industrial electricity consumption, and Household electricity consumption) as the dependent variable after analyzing population, Population, GDP, Per Capita GRP, Number of Schools, Number of Hospitals, Number of industrial enterprises, built-up area, Total Natural Gas, Total number of the public bus(electric) passenger transport in the whole year were selected as independent variables. The relevant data are from the *Statistical Yearbook of Chinese Cities* from 2009 to 2019.

Before starting the empirical analysis, we minorize all variables at the regional level by year at less than 1% and more than 99% to mitigate the possible influence of outliers on the regression results. Secondly, to eliminate the possible clustering characteristics of the sample data, the paper also adjusted the cluster at the regional level for the standard error of the regression coefficient. Due to the large gap between the data, logarithmic transformation was performed on all independent variables. Because the panel data is extensive, the above adjustments to the data are not described in detail. After a series of adjustments, cities corresponding to variables with insignificant data and low variance interpretation (For example, Lhasa City in Tibet, Sansha City in Hainan Province, Chaohu City in Anhui Province, etc.). Data analysis and processing are mainly accomplished through Stata 16.0.

(2) Model selection

Generally speaking, there are two analysis models for panel data: the fixed effect model and the random effect model. Therefore, it is necessary to determine which model to adopt before starting empirical regression. The Hausmann test was performed using the Stata, and the results are shown in Table 3.

Table 3 Hausman test results

Social electricity	Industrial electricity	Household electricity
H_0：difference in coefficients not systematic chi2(11) =108.81 Prob>chi2=0.0000	H_0：difference in coefficients not systematic chi2(11) =69.64 Prob>chi2=0.0000	H_0：difference in coefficients not systematic chi2(11) =245.02 Prob>chi2=0.0000

It can be found from the Test results that the P values of the Hausman Test of the three dependent variables are all 0.0000. It is suggested that the fixed effect model should be used for regression analysis instead of the random effect model.

(3) Statistical analysis

In order to **preliminarily** understand the characteristics of variables, this section conducts statistical analysis on the data selected in the study, and the result is shown in Table 4.

Table 4 Descriptive statistics of major variables

	N	MEAN	St. D	MIN	MEDIAN	MAX
lnSocial electricity	3113	13.3401	1.1764	9.9120	13.3450	16.3552
lnIndustrial electricity	3113	12.8520	1.3347	7.6089	12.9550	15.8927
lnHousehold electricity	3113	11.1434	1.0462	6.7968	11.0387	14.3476
lnPopulation	3113	4.6745	0.7691	2.8332	4.6052	7.2428
lnNumber of industrial enterprises	3113	6.5879	1.0820	3.3322	6.5709	9.4328
lnSchool	3113	6.5353	0.8075	3.8918	6.5653	9.7510
lnPer Capita GRP	3113	10.8505	0.5924	8.9948	10.8613	12.4918
lnbuilt-up area	3113	4.4742	0.8236	2.6391	4.3438	7.1709
lnGDP	3113	16.3806	0.9722	13.6096	16.3123	19.5299
lnUrban road area	3113	7.0326	0.9558	4.1744	6.9460	9.6998
lnNumber of Hospital	3113	4.9602	0.7651	2.1972	5.0239	7.5088
lnTotal Natural Gas	3113	8.7771	1.9886	1.0986	9.0033	13.5214
lnTotal number of public bus (electric) passenger transport in the whole year	3113	9.0480	1.3300	4.0943	8.9693	12.5028

It can be seen from the descriptive statistical results that all the statistical characteristics of each variable are within a reasonable range, so it can be considered that the data we selected are rigorous and reasonable to a certain extent.

(4) Model checking

Based on the above processing, we examine the fixed effect panel regression and then adjust the number in parentheses (i.e., significance) below each number by clustering adjustment, which would reduce the individual interaction within the group. The debugging results of the FEM model are shown in Table 5.

Table 5 FEM inspection results

	(1)	(2)	(3)
	lnSocial	lnIndustrial	lnHousehold
lnPopulation	0.3413***	0.4037***	0.2198***
	(4.22)	(3.36)	(3.08)
lnenterprises	0.1969***	0.2034**	0.1162***
	(2.62)	(2.21)	(2.78)
lnSchool	-0.0118	0.0105	-0.0106
	(-0.32)	(0.21)	(-0.33)
lnGRP	0.2759***	0.3380***	0.1486***
	(4.20)	(3.85)	(2.80)
lnArea	0.2195***	0.2602***	0.0581
	(3.45)	(3.20)	(1.08)
lnGDP	-0.3498***	-0.3870***	-0.2379***
	(-6.74)	(-5.75)	(-6.35)
lnurbanroadarea	0.0536	0.0393	0.0920**
	(1.02)	(0.56)	(2.18)
lnHospital	-0.2055***	-0.2611***	-0.1107***
	(-4.48)	(-4.63)	(-3.49)
lnGas	0.0280*	0.0383*	0.0152
	(1.86)	(1.89)	(1.04)
lnbus	0.0410	0.0581	0.0452**
	(1.41)	(1.42)	(2.22)
_cons	11.8051***	10.7841***	10.3596***
	(9.30)	(6.90)	(12.18)
Firm	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	3113	3113	3113
R^2	0.742	0.623	0.580
adj. R^2	0.740	0.621	0.577

Note: *、**、*** means significant at the significance level of 10%, 5%, and 1%, respectively (two-tailed test); In parentheses are the t-test values adjusted by the cluster at the regional level.

4.2 Comparative analysis of three dependent variable models of GAM

Note: For the convenience of statistics, we record School, Per Capita GRP, GDP, Urban Road Area, Number of Hospital, Total Natural Gas, and Total Number of the public are included in the following analysis Bus (Electric) Passenger transport in the whole year, Population, Number of industrial enterprises and built-up area as School, GRP, GDP, URA, NH, TNG, PB, POP, NIE, and BA.

(1) Social electricity consumption

The relationship between social electricity consumption and ten dependent variables is analyzed by the GAM model, as shown in Table 6.

Table 6 GAM model test results of Social electricity consumption

	edf	ref.df	f	p-value
s(School)	7.945	8.691	10.510	< 2e-16 ***
s(GRP)	5.228	6.444	31.811	< 2e-16 ***
s(GDP)	5.490	6.729	13.229	< 2e-16 ***
s(URA)	4.813	6.035	3.876	0.000728 ***
s(NH)	6.944	7.981	61.165	< 2e-16 ***
s(TNG)	4.045	5.123	35.410	< 2e-16 ***
s(PB)	5.948	7.167	4.525	5.21e-05 ***
	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.95290	0.21966	45.310	< 2e-16 ***
POP	0.22113	0.02814	7.859	5.30e-15 ***
NIE	0.08509	0.01787	4.761	2.02e-06 ***
BA	0.40062	0.03809	10.518	< 2e-16 ***

R-sq.(adj) = 0.76 Deviance explained = 76.3%

According to the test results of the model, 10 independent variables were compared and analyzed. School, Per Capita GRP, GDP, Number of Hospital, Total Natural Gas, and Built-up Area made outstanding contributions to the model. The three biggest contributors are the Number of Hospital, Total Natural Gas, and Per Capita GRP, whose F test values are 61.165, 35.410, and 31.811, respectively. Urban Road Area, Total number of public bus (electric) Passenger transport in the whole year, Population, and Number of industrial enterprises contributed little to the model of industrial enterprises. The f-value of the F-test for the three methods were 3.876, 4.525, 7.859, and 4.761, respectively. Their p-values were all less than 0.1, and Deviance explained 76.3%, indicating that these factors significantly influence at the $p<0.01$ level.

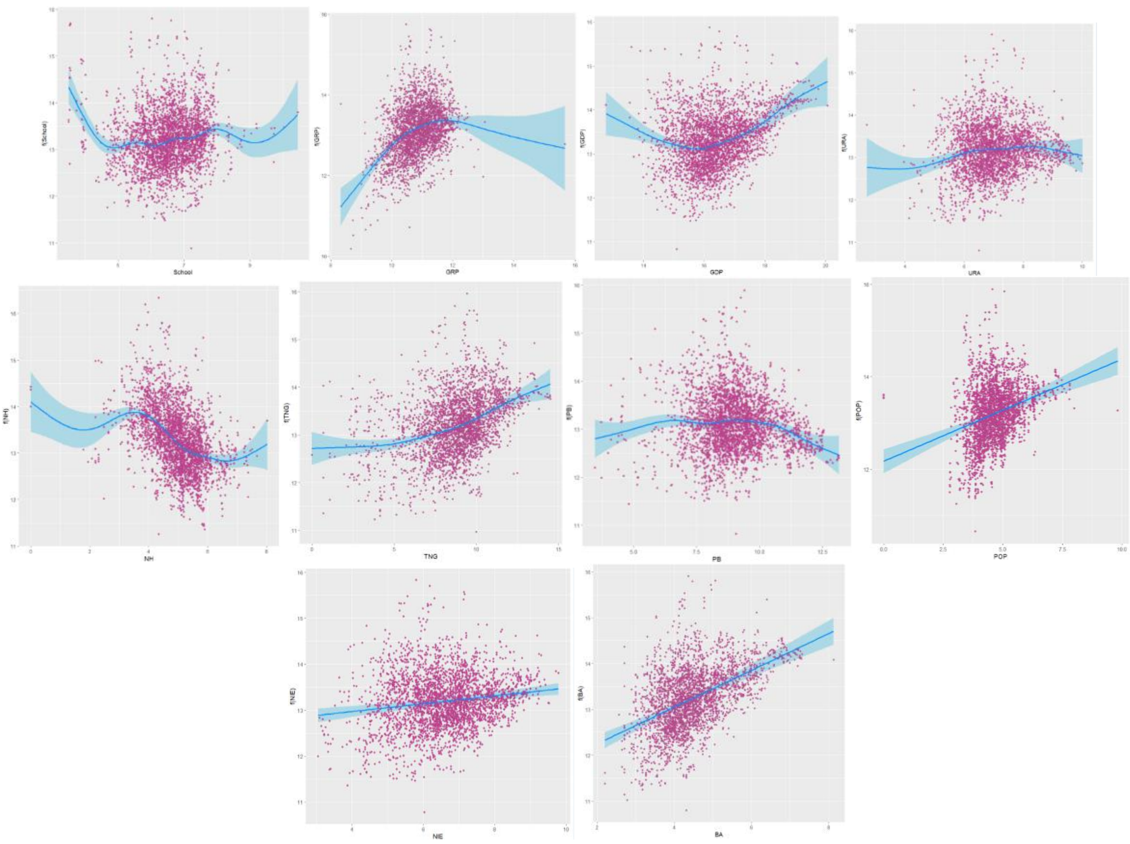

Fig.1 GAM diagram of the influence of independent variables on social electricity consumption

(2) Industrial electricity consumption

The relationship between Industrial electricity consumption and ten dependent variables is analyzed by the GAM model, as shown in Table 7.

Table 7 GAM model test results of Industrial electricity consumption

	edf	ref.df	f	p-value
s(School)	6.800	7.889	8.798	< 2e-16 ***
s(GRP)	5.161	6.372	28.265	< 2e-16 ***
s(GDP)	4.766	5.961	5.666	8.94e-06 ***
s(URA)	5.031	6.262	3.731	0.000864 ***
s(NH)	5.760	6.904	40.050	< 2e-16 ***
s(TNG)	4.527	5.675	29.941	< 2e-16 ***
s(PB)	4.944	6.126	6.595	7.78e-08 ***
	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	8.40767	0.29166	28.827	< 2e-16 ***
Population	0.21801	0.03728	5.848	5.50e-09 ***
NIE	0.18959	0.02366	8.012	1.58e-15 ***
BA	0.48582	0.05056	9.608	< 2e-16 ***

R-sq.(adj) = 0.673   Deviance explained = 67.8%

According to the test results of the model, 10 independent variables were compared and analyzed. Per Capita GRP, a Number of Hospital, Total Natural Gas made outstanding contributions to the model. Their F test values were 28.265, 40.050, and 29.941, respectively. Schools, GDP, and population contribute very little to the industrial enterprise model. Urban Road Area, GDP, and Population have the least influence among them. The f-values of the F test for the three methods were 3.731, 5.666, and 5.848, respectively. Their p-values were all less than 0.01, and Deviance explained equals 67.8%, indicating that these factors significantly influence at the p<0.01 level.

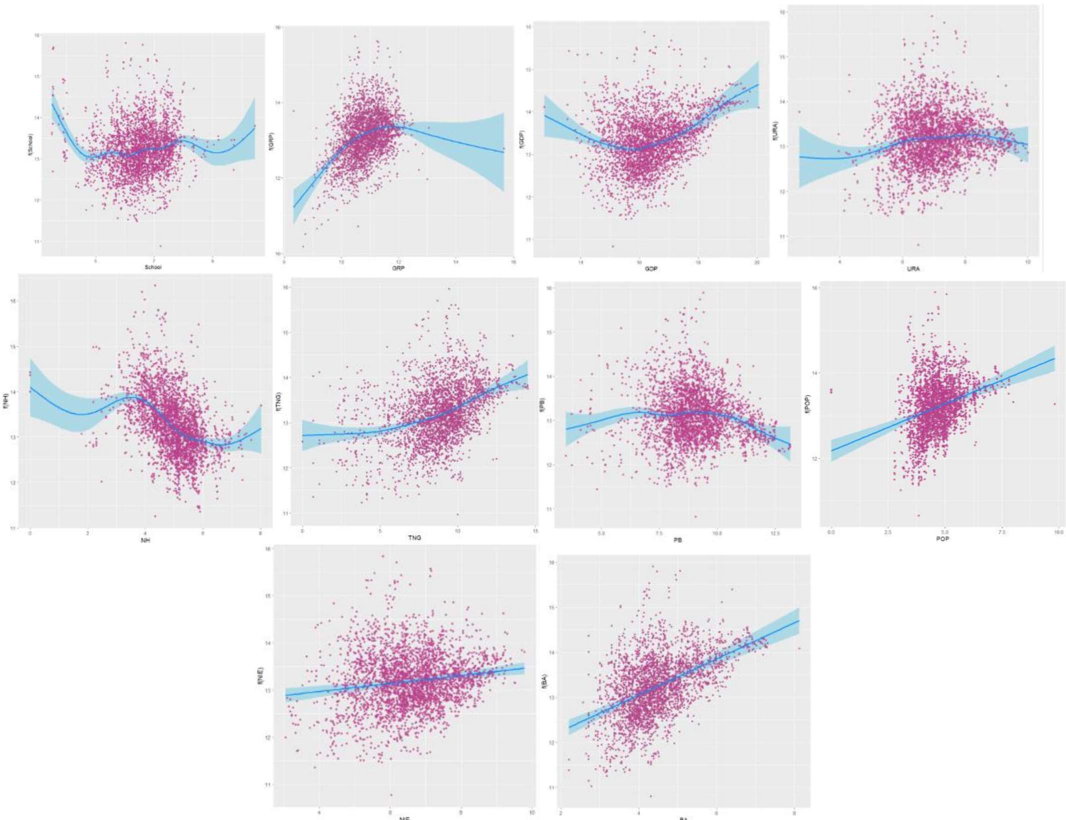

Fig.2 GAM diagram of the influence of independent variables on Industrial electricity consumption

(3) Household electricity consumption

The relationship between Household electricity consumption and ten dependent variables is analyzed by the GAM model, as shown in Table 8.

Table 8 GAM model test results of Household electricity consumption

	edf	ref.df	f	p-value
s(School)	4.245	5.310	13.221	< 2e-16 ***
s(GRP)	5.771	6.995	7.611	< 2e-16 ***
s(GDP)	5.547	6.781	16.610	< 2e-16 ***
s(URA)	1.957	2.567	3.369	0.02674 *
s(NH)	5.785	6.923	42.244	< 2e-16 ***
s(TNG)	4.600	5.761	3.199	0.00493 **
s(PB)	6.848	7.974	14.660	< 2e-16 ***
	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	7.92023	0.17381	45.568	< 2e-16 ***
POP	0.33438	0.02231	14.990	< 2e-16 ***
NIE	0.11953	0.01417	8.433	< 2e-16 ***
BA	0.19473	0.03020	6.449	1.31e-10 ***

R-sq.(adj) = 0.808 Deviance explained = 81.1%

According to the test results of the model, 10 independent variables were compared and analyzed. The number of Hospital made the biggest contributor to the model, and its f-value of the F test was 42.244. School, GDP, Total number of public buses (electric), Passenger transport in the whole year, and population make an outstanding contribution to the household model. Total Natural Gas and Urban Road Area have the least influence. The f-values of the F test for the three methods were 3.199 and 3.369, respectively. Their p-values were all less than 0.1, and Deviance explained equal to 81.1%, indicating that these factors have a significant influence at $p<0.01$ level.

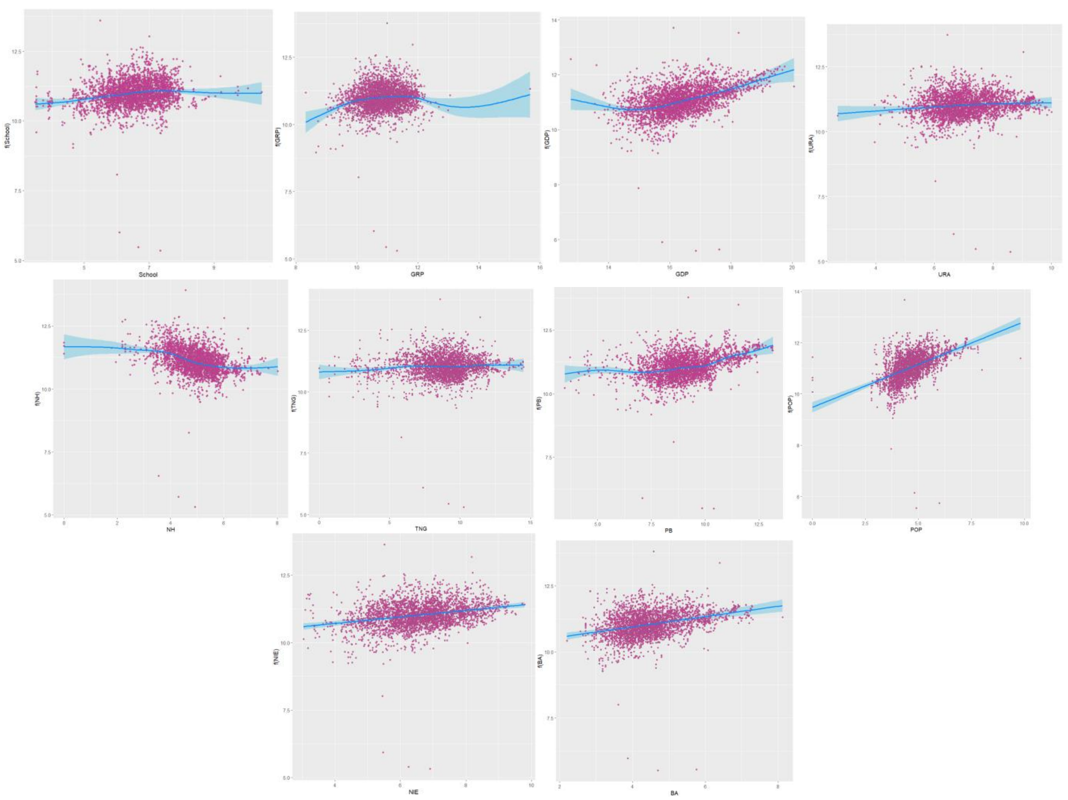

Fig.3 GAM diagram of the influence of independent variables on Household electricity consumption

4.3 Comparative analysis of FEM and GAM

Through the debugging test results of the two models, it can be seen that both models can obtain significant prediction results, but by contrast, the correlation of prediction results obtained by the FEM is 10%, 5%, and 1%, respectively. In contrast, these factors in the prediction results of GAM are significantly affected at the level of $P < 1\%$, which indicates that the results obtained by GAM have vital significance. Secondly, the coefficient of determination (R2) of FEM for the verification of social electricity consumption, Industrial electricity consumption, and Household electricity consumption are 0.742, 0.623, and 0.580, respectively, while the coefficient of determination (R2) of GAM for the verification of three dependent variables are 0.760, 0.673 and 0.808 respectively. Therefore, the coefficient of determination (R2) measured by GAM is greater than that of the FEM model. The above results indicate that GAM has a better fitting effect than FEM.

GAM is used to analyze the relationship between electricity consumption and its influencing factors in this paper. The relationship between them has non-linear characteristics. Compared with the traditional regression method (fixed effect mode), GAM is considered a more informative tool. [13]

5. Conclusion

Based on the GAM, this paper selects all cities in each province of China from 2009 to 2019 as the study area to explore the factors affecting urban social electricity consumption. This paper analyzes the relationship between urban social electricity consumption and each factor that affects it based on GAM, where Population, Number of industrial enterprises, Urban road area, and the Total number of public bus (electric) passenger transportation in the whole year are less significant. GAM produced more significant results and stronger correlations than the fixed effects panel regression model. However, the results of both methods have similarities in that they both verify that Per Capita GRP, GDP, and Number of Hospital are highly significantly correlated with urban social electricity consumption. In addition, this paper lacks an assessment of the interaction term introduced by GAM, such as the interaction effect between Total Natural Gas and the Number of industrial enterprises and the interaction effect between School and Population.

The results of the empirical analysis also show that GAM has a better fit and prediction accuracy than FEM, and the ten independent variables verified by GAM have significant effects on all three dependent variables. The significant advantage of GAM over FEM is that it can explain the relationship between the input and output variables. Therefore, the use of GAM can effectively study the degree of influence of each factor on urban social electricity consumption, which can provide reliable theoretical support and a decision-making basis for the reworking of the revised social electricity consumption structure and the government's formulation of energy-saving and environmental protection measures to reduce electricity consumption.

References

[1] Miao Yijing. (2021). Research on logistics demand forecasting model for residential consumer goods in different types of cities (Master's thesis, Beijing Jiaotong University). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD202201&filename=1021875662.nh

[2] Liu, Chun, Ma, Chao, Feng, Yongchun & Wang, Zhixiu. (2022). The relationship between electricity consumption, industrial structure and economic growth - A study based on panel vector autoregression (PVAR) model for Gansu Province prefecture-level city data. Systems Science and Mathematics (03), 599-613.

[3] Zeng, R. Y. (2018). Research on the application of weighted plain Bayesian algorithm based on LDWPSO in logistics demand forecasting (Master's thesis, Nanjing University of Posts and Telecommunications). https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201901&filename=1018898906. nh

[4] Luo W.-Y., Wang H., He X.-G., Liu W.-X. & Zhong E.-L. (2021). Analysis of the application of urbanization rate in electricity demand forecasting. Power Survey and Design (05), 62-66. doi: 10.13500/j.dlkcsj.issn1671-9913.2021.05.012.

[5] Chen X, Zhao YP, Li X, Xiao CH, Li ZY & Zhang LM. (2021). Analysis of the effect of temperature and humidity factors on the population dynamics of Drosophila melanogaster in the field based on generalized additive model. Journal of Environmental Entomology (04), 858-866.

[6] Sarkodie Samuel Asumadu, Ahmed Maruf Yakubu, Owusu Phebe Asantewaa. Ambient air pollution and meteorological factors escalate electricity consumption[J]. Science of the Total Environment,2021,795.

[7] Nguyen Can Trong, Nguyen Diep Thi Hong, Phan Diem Kieu. Factors affecting urban electricity consumption: a case study in the Bangkok Metropolitan Area using an integrated approach of earth observation data and data analysis. [J]. Environmental science and pollution research international,2020,28(prepublish).

[8] Aneja Ranjan, Mathpal Megha. Economic Growth and Electricity Consumption in India: An Econometric Analysis [J]. The Indian Economic Journal,2022,70(1).

[9] Zhang Youguo & Gao Yan. (2021). Forecasting electricity demand in Shanghai based on grey correlation analysis and multiple regression. Logistics Science and Technology (12), 26-29+35. doi: 10.13714/j.cnki.1002-3100.2021.12.009.

[10] Hastie T, Tibshirani R. Generalized additive models [J]. Statistical Science, 1986, 1(3): 297- 318.

[11] Wu Wei, Xueshan Bai & Shuai Ni. (2021). Study on the influence of built environment on residential electricity consumption in residential communities--Tianjin Eco-city as an example. Journal of Hebei Engineering University (Social Science Edition) (04), 23-29.

[12] Sun Yan & Luo Xiaoqing. (2016). An empirical consideration of the relationship between GDP and electricity consumption in China. Statistics and Decision Making (05), 98-100. doi: 10.13546/j.cnki.tjyjc.2016.05.026.

[13] Bellido J M，Pierce C I J，Wang J．Modelling in tra-annual variation in a bundance of squid，Loligo forbesi in Scottish waters using generalized additive models[J]．Fish Res,2001,（52）：23－39