Team Control Number

For office use only
T1 ________________
T2 ________________
T3 ________________
T4 ________________

86837

Problem Chosen

C

For office use only
F1 ________________
F2 ________________
F3 ________________
F4 ________________

2018

MCM/ICM

Summary Sheet

Energy profile and targets for the four U.S. states

summary

First, we select relevant variables from four aspects: total energy consumption, per capita energy consumption, different sectors and energy structure to give a summary of the energy profiles of each state. By studying the total energy consumption and renewable energy consumption, we get the evolution of the energy profile. Their changing relationships over time are given. We define the rate of change and the contribution rate of change to describe the specific change. We analyze the similarities and difference between the four states and discuss their influencing factors, including geography, industry, population and climate. We analyze the use of clean energy in four states, the evolution process and its causes.

Second, using the previous conclusions, we set up an evaluation model for clean renewable energy. Criteria are selected from four aspects: technology, society, environment and economy. We use AHP to determine the weights of the indicators and calculate the final score to confirm that the CA have the “best” profile for use of cleaner, renewable energy in 2009. We chose six variables as parameters to predict the energy profile, use ARMA time series method to predict the parameters. We establish the GBDT prediction model to predict the total energy consumption and the renewable energy consumption in 2025 and 2050 in the four states. The results are shown in Table 7.

Third, based on the criteria and predictions, we establish a renewable energy target decomposition model and set targets for renewable energy use in all 2025 and 2050 states, production targets for various clean energy sources, and industry consumption targets. To achieve the energy compact goals, we propose the following measures: increase the production of various types of renewable energy, enhance cooperation between the states and create a favorable external environment.

Next, we conduct a sensitivity analysis of the evaluation model, the ranking doesn’t change when the weights fluctuated. Cross validation of the prediction model reduces the root mean square error. In predicting the energy profile, the GBDT model uses the idea of model fusion to make the forecast more accurate.

Finally, we give a memo.

Keywords: Energy Profile、AHP、GBDT、Cross-validation、Target decomposition

Content

1 Memo

TO: governor of a state
FROM: team86837
DATE: February 12, 2018
SUBJECT: energy profile and energy compact goals.

We will outline the energy profiles of the four states. In terms of the overall energy profile, the total energy consumption has been on an upward trend and $TX > CA > NM > AZ$. From the per capita energy profile, the overall per capita consumption is a downward trend, total per capita energy expenditure is generally an upward trend. In terms of energy consumption in different sectors, the transport sector of CA, AZ and NM is the largest, and TX is the industrial sector, which partly reflects TX's reliance on non-renewable energy. From an energy structure overview, non-renewable energy sources in each state continue to be the major consumers, especially oil and gas resources, but renewable energy sources in all states are developing.

By building an energy evolution model, we got the energy profile of each state from the four continents from 1960 to 2009. It is concluded from the analysis that the CA's clean energy occupies a large proportion of total energy and has a diversified structure. AZ has the largest clean resources, water and nuclear energy rich. NM clean energy development tends to diversify after 2000. Because of the vast area, biomass can be rich. TX clean energy accounted for a smaller proportion of total energy, but in recent years there has been a change wind energy output is the largest.

Considering the evaluation model from the four aspects of technology, society, environment and economy, we have confirmed that the better use of clean energy in four states ranks $CA > AZ > NM > TX$.

Through the previous analysis, we established the GBDT forecasting model to predict the total energy consumption and renewable energy consumption of the four states in 2025 and 2050 to reflect the energy profile of each state. The predicted results are as follows:

Table 1 Predictors for *TETCB* and *RETCB*

value / year	CA		AZ		NM		TX	
	TETCB	RETCB	TETCB	RETCB	TETCB	RETCB	TETCB	RETCB
2025	8149793	738884	1529268	95801	689743	317951	11312929	312083
2050	8158872	761945	1548130	98742	692942	341313	11478427	356306

By combining the "best" profiles and forecasting predictions, we set the model to determine the renewable energy targets of the states in 2025 and 2050. Then we plan the production and use of clean energy.

Table 2 Energy allocation targets in the States

Year / State	CA	AZ	NM	TX
2025	653365.5	149325.3	155681.6	343459
2050	686924.8	197577.7	182632.2	365891

Finally, we put forward the following 3 suggestions for the result of the model and the analysis of the whole problem：increase the production of renewable energy sources, enhance cooperation between states and create a good external environment.

2 Introduction

2.1 Background

Energy is the material basis for the survival and development of human society, which occupy a particularly important place in the national economy. But now countries are facing two problems in using energy. Environmental pollution and fossil energy depletion. Thus, it affects the development of the national economy. In this case, the use of clean energy and renewable energy is getting more and more attention. There are four states on the border between the United States and Mexico, CA, AZ, NM, TX. It is hoped that through the establishment of a new energy compact, the use of new energy sources in four states and cooperation will be reached.

Therefore, it is essential to have an in-depth understanding of energy profiles in various states, in particular, the total consumption of energy and the consumption of renewable energies. A study of the similarities and differences between the four states and the influencing factors let us have an objective evaluation of them. Research into the evolution of energy profiles can help us predict future energy conditions. Based on evaluation and forecasting, we can set up scientific energy targets for four states and make relevant suggestions.

2.2 Our work

Part I A: We analyze energy profiles in each state from four aspects of overall, per capita, different sectors and energy structure. We filter the relevant variables from the table to describe them, and at the end give a summary of each state's energy profile.

Part I B: Our research on the energy profile starts with a study of total energy consumption and renewable energy consumption. Describing changes in energy profiles using curve fitting, defined change rate and change contribution rate. We have reached the similarities and differences between the four states and analyzed the influencing factors. Finally, we analyze the use of clean energy in four states, the evolution process and its causes.

Part I C: We have established a clean renewable energy assessment model. First of all, from the technical, social, environmental, economic four aspects to determine the selection criteria, and give the data of the indicators. Then, the weights of each index are determined by the analytic hierarchy process. Finally, the evaluation results are given, the CA has the best clean renewable energy profile.

Part I D: We established the GBDT model to predict the total energy consumption and the total amount of renewables consumed in each state in 2025 and 2050. We choose 6 variables as parameters and use ARMA time series to predict the parameters. Then the GBDT model is established, which is trained according to the known data and the accuracy is improved by cross-validation. We substitute the predicted variables into the model to get the predicted results.

Part II A：Based on the "best" profile and forecast results, we set up a target decomposition model for renewable energy planning. We identify renewable energy use targets for the 2025 and 2050 states. Then, we plan for each state's production targets for clean energy and consumption targets in various industries.

Part II B：To achieve the aforementioned energy contract objectives, we propose the following initiatives: increase the production of renewable energy sources, enhance cooperation between States and create a good external environment.

Part III: We give a memo.

3 Assumptions

The accuracy of our models rely on certain key, simplifying assumption. These assumptions are listed below:

1) We assume that the creation of GDP only comes from the consumption of energy.
2) The environment is good or bad only consider the consumption of energy.
3) Everyone's consumption of energy is same.
4) We assume that CO2 emissions are only related to the consumption of non-renewable energy.
5) We assume that the climate in one state does not change significantly over the study period.
6) We suppose that all the data is authentic.

4 Symbol definition and description

We will begin by defining a list of the nomenclature used in this report:

Table 3 Notation

Symbols	Descriptions
α_k	The rate of change
β_k	The change contribution rate
x_1、x_2、x_3、x_4	Technical、social、environmental、economic indicator
ω_j	The weight vector
y_i	The final score of each state
$R_{i,total}$	The total target share
$R_{i,fix}$	The fixed share
$R_{i,float}$	The floating share

5 Establishment and Solution of Model

5.0 Data preprocessing

By observing the data, we found that there are 22 variables without data, such as *CCEXB*, *CCEXD*, *CCEXP*, *CCEXV* and so on. Some variables lack of 10 or 17 consecutive data, most of which are average energy prices and total expenditures. Moreover, if there lacks of 10 data, it must be 10 consecutive data from 1960 to 1969; if there lacks of 17 data, it must be 17 consecutive data from 1960 to 1977. At the same time, the variables that containing the missing data are same in each state.

We think the lack of these data is due to objective reasons rather than statistical errors. Therefore, we will deal with these missing data in the concrete process of solving the problem. Next, we summarize the data from different states separately.

5.1 The Solution of Part Ⅰ

5.1.1 The Energy Profile of Each of Four States

First, we count the energy profile of the four states from the overall energy profile, the per-capita energy profile, the energy profiles of different sector and the energy structure profile. Due to space limitations, we summarize its energy profile for each state at the end of 4.1.1.

1) The Overall Energy Profile

We describe the overall energy profile from four perspectives: total energy production (*TEPRB*), total energy consumption (*TETCB*), total energy average prices (*TETCD*), and total energy expenditures (*TETCV*). We use the data of the above variables from 1960 to 2009 and make statistics. The results are as follows:

Table 4 The overall energy profile

Parameter / State	TEPRB/ (Billion Btu)	TETCB/ (Billion Btu)	TETCD/ (Dollars/ million Btu)	TETCV/ (Million Dollars)
CA	3192795.938	6556299.813	9.401	50327.040
AZ	397380.184	880081.365	10.330	7477.492
NM	2196199.552	526236.508	8.867	3097.570
TX	13067545.204	8930451.709	7.234	51960.895

From the data in the table above, we can intuitively see the value size of the four parameters, then, determine the relationship between supply and demand.

2) The Per-capita Energy Profile

We describe the per-capita energy profile from two perspectives: total energy consumption per capita (*TETPB*) and total energy expenditure per capita (*TETPV*). In the figure, the blue line which represents *TETPV* only have the data after 1970.

Figure 1 The per-capita energy profile

From the above table, we can see that the trends of the total energy consumption per capita and the total energy expenditure per capita are basically the same. *TETPB* first increase and then decrease, made the near maximum in 1970. *TETPV* first increased, then stabilized around 1980 and then increased again around 2000, and finally dropped after reaching its maximum in

2008.

3) The Energy Profiles of Different Sector

There are four departments involved in the data: the transportation sector, the commercial sector, the industrial sector and the residential sector. The total energy consumption of these four sectors are *TNACB*, *TNCCB*, *TNICB* and *TNRCB* respectively. We give the share of the total energy consumption of each sector in the summing of the total energy consumption of the four sectors over the past 50 years.

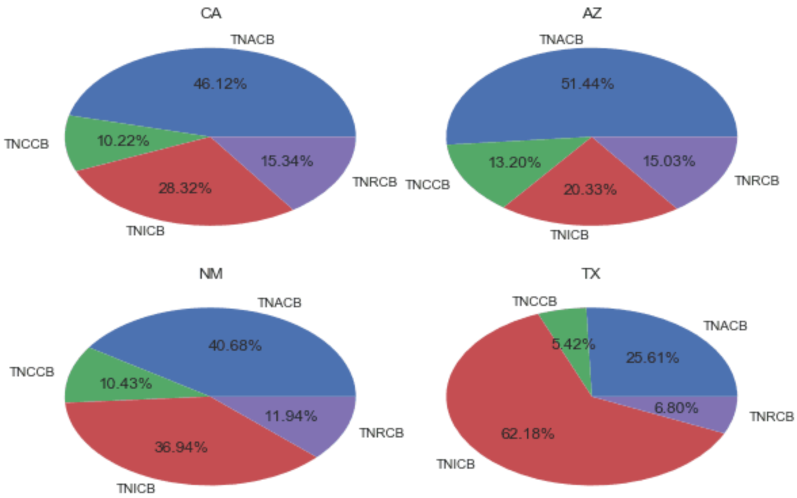

Figure 2 The energy profiles of different sector

From the figure above, we can see that in addition to the TX state, the proportions of *TNACB*, *TNICB*, *TNRCB*, and *TNCCB* in the other three states decreased in turn. However, TX state has the highest *TNICB*, followed by *TNACB*, *TNRCB* and *TNCCB*.

4) The Energy structure profile

We classify the resources which is given in the worksheet ("msncodes"), as shown in the following table:

Table 5 Energy classification

Energy	Branch	Example
Non-renewable energy	coal	coal, coke
	oil	asphalt, road oil, aviation gasoline, kerosene, lubricating oil, distillate oil, Petroleum coke liquefied petroleum gas, aviation gasoline, automobile gasoline, naphtha and so on
	natural gas	natural gas, refueling gas
	nuclear energy	nuclear power, nuclear fuel
Renewable Energy	Wind energy	wind power
	Water energy	hydropower, condensate
	Solar energy	photovoltaic and solar thermal power generated
	Geothermal Energy	geothermal energy and heat pumps, geothermal energy
	Biomass	fuel ethanol, wood, waste
Imported energy	import	net import electricity into the United States

We show coal total end-use consumption (*CLTXP*), all petroleum products total consumption (*PATCB*), Natural gas total consumption (excluding supplemental gaseous fuels) (*NNTCB*), electricity produced from nuclear power (*NUETB*), electricity produced from wind energy (*WYTCB*), photovoltaic and solar thermal energy total consumption (*SOTCB*), geothermal energy total consumption (*GETCB*), hydropower Consumption (*HYTCB*), and Biomass Energy Consumption (*BMTCB*) below. Energy consumption units are Billion Btu.

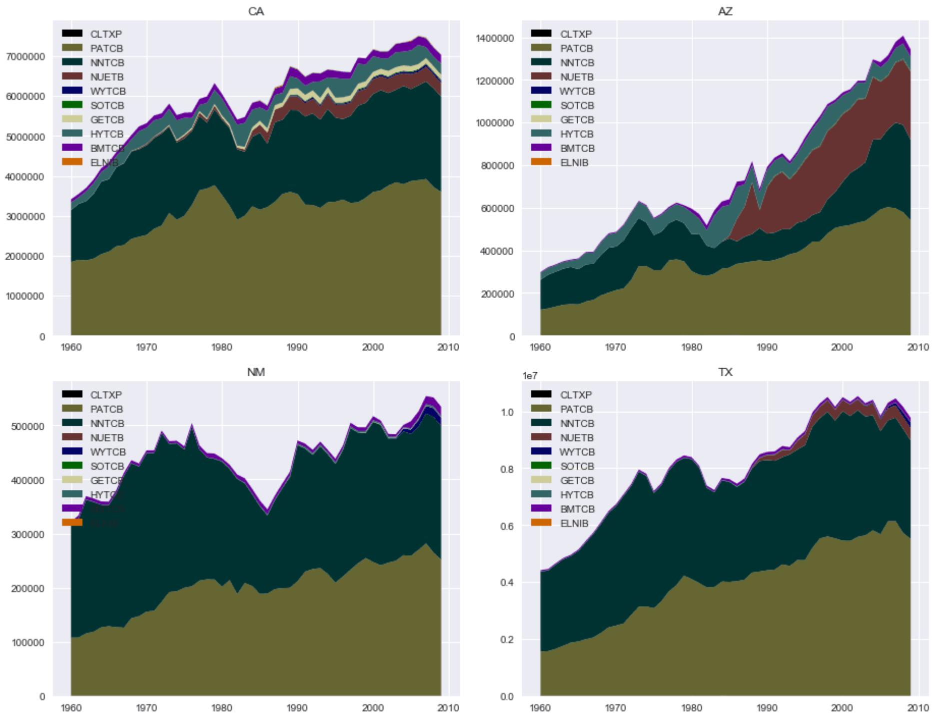

Figure 3 State energy structure chart

5) Summary of Energy Profile of Each State

(1) California---CA

Overall, total energy production in California ranks second among the four states, with total energy consumption at the third highest level well below AZ and TX, but total energy consumption is much larger than total energy production, so CA belongs to the energy input area. In per capita terms, per capita energy consumption and per capita total energy expenditure are the lowest among the four states, partly because of the mild climate, which makes air-conditioning and heating less dependent on the electricity generated by energy sources. For the four departments, the transportation sector accounted for the largest proportion of energy consumption 46.12%. Renewable energy is best developed among the four states, and the proportion of renewable energy in the state's total energy consumption is the largest, the type of renewable energy is also very complete. Water energy is one of the major renewable energy for CA.

(2) Arizona ---AZ

Overall, AZ's total energy production is the lowest among the four states. At the same time, total energy consumption is also less, supply is less than demand, so it belongs to the energy input area. In particular, the total energy consumption is rapidly increasing after 1980. However, due to the energy shortage in the area, the total energy average price is the highest. On a per capita basis, although total energy average prices are higher, AZ’s total energy consumption per capita is lower, so total energy expenditures per capita is lower. For the four departments, the transport sector accounted for the largest share of energy consumption, which exceeded 50%. AZ is also more diversified in its energy mix. In terms of non-renewable energy, coal consumption is low, oil is always the main source of energy, and nuclear energy consumption accounts for the largest share of the four states. At the same time, the proportion of renewable energy consumption is large, the type of renewable energy is complete, water energy is always a major source of renewable energy.

(3) New Mexico---NM

Overall, NM total energy production ranked the third in four states with the lowest total energy consumption and oversupply, so NM belongs to the energy export zone. In terms of per capita, total energy consumption per capita of NM is higher, and total energy expenditure per capita is also higher. For the four sectors, the transportation sector accounted for the largest share of total energy consumption, 40.68%. New Mexico's energy structure is more single, non-renewable energy is the largest share of total energy. Biomass is the state's major renewable energy source because of its larger acreage. Compared to other states, wind energy consumption is more.

(4) Texas---TX

Texas is rich in energy, they have abundant oil, lignite and natural gas resources. Overall, total energy production and consumption is the highest among the four states, and supply is much greater than demand, so they belong to the energy output zone, and the total energy average price is the lowest. In terms of per capita, Texas has the highest energy consumption per capita and the highest total energy expenditure per capita. Among the four departments, the industrial sector accounted for the largest proportion of resource consumption, 62.18%. Texas has the single most energy structure with nearly all energy consumption coming from non-renewable energy sources. Utilization of renewable energy started late, and it is mainly biomass energy and wind energy.

5.1.2 Analysis of the Evolution of the Energy Profile

First, we study the evolution of energy profile. Using curve fitting, we present the total energy consumption over time for four states. Next, we define the change rate and the change contribution rate to describe the specific changes. Then, we explain the similarities and differences and influencing factors of the four states. Finally, we analyze the profile of energy use, evolution and reasons in four states.

1) Model for the Evolution of Energy Profile

We select total energy consumption (*TETCB*) and total renewable energy consumption (*RETCB*) as the main variables to describe the resource profile. As can be seen from Figure 3, the total energy consumption in each state increases over time, but at a different rate and magnitude of growth. From Figure 4 below, we can see that the total renewable energy consumption in most states increases over time.

To characterize the evolution of the energy profile, we make *TETCB*, *RETCB* respectively fitting time and tested the fit using the goodness-of-fit R^2.

$$\begin{cases} TETCB = TETCB(T) \\ RETCB = RETCB(T) \end{cases}$$

The relationship between the total energy consumption (*TETCB*) and time (1960-2009) is as follows:

CA: $TETCB_{CA} = -1454t^2 + 5865000t - 590400000$, $R^2 = 0.9639$.
AZ: $TETCB_{AZ} = 25540t - 49800000$ $R^2 = 0.9790$.
NM: $TETCB_{NM} = 7265t - 13890000$ $R^2 = 0.9200$.
TX: $TETCB_{TX} = -1731t^2 + 7031000t - 7128000000$ $R^2 = 0.9647$.

To explain the magnitude of the change and the reasons for the change each year, we constructed two indicators.

① The rate of change α_k: Describe the change of total energy consumption between two years. When $\alpha_k > 0$,it shows that the amount of energy consumption is increasing; when $\alpha_k < 0$, it shows that energy consumption is declining. The size of $|\alpha_k|$ reflects the increase or decrease.

$$\alpha_k = \frac{Z_{k+1} - Z_k}{Z_k}, k = 1, 2, \ldots, 49. \quad (1)$$

②The change contribution rate β_k: Describe the contribution rate of a certain kind of energy consumption z change to the total energy consumption Z change in a year. When $\beta_k > 0$, it shows that the direction of change of such resources is in the same direction as the change of the overall resources, and it has a promoting effect. When $\beta_k < 0$, it shows that the change of this kind of resources has a hindering effect on the change of the overall resources. The size of $|\beta_k|$ reflects the size of the certain energy's role in overall.

$$\beta_k = \frac{z_{k+1} - z_k}{Z_{k+1} - Z_k}, k = 1, 2, \ldots, 49. \quad (2)$$

2) Model Analysis and Interpretation

(1) Similarities and Differences Between the Four States

According to the fitted function, we find that the total energy consumption of the four states increased roughly from 1960 to 2009. Among them, AZ and NM showed linear growth, that is, the growth rate was constant, but the growth rate of NM was far less than the growth rate of AZ. CA and TX showed a quadratic curve growth, the growth rate gradually slowed down, TX growth rate slowed down faster. The increase in energy consumption is mainly due to the increase of population and the development of various industries.

Calculating α_k of the four states and we can get that α_k is concentrated in the interval (-0.07,0.12), it shows that the energy consumption in the four states is relatively stable. The overall trend of the four states is growing, with several fluctuations, and the declining years are similar, for example, the energy consumption of the four states is dropped resulting from the 2008-2009 economic crisis.

For non-renewable energy, calculate the value of β_k. As we can see from Figure 3, for each state, the contribution of non-renewable energy consumption in total energy consumption is larger, which shows that non-renewable energy is the main source of energy. The contribution of renewable energy consumption to the growth of total energy consumption is also gradually increasing. At the same time, each state also has its own characteristics. The growth of CA states mainly comes from the growth of oil, which is related to the state's geographical location, rich oil resources and developed transportation industry. The change in total energy consumption in AZ state is consistent with changes in oil and gas resources, but the slow growth in oil and gas resources is due to the low resource of fossil fuels in the state. NM natural gas consumption changes in tandem with total energy consumption, and oil and gas energy consumption have always played a dominant role in total energy consumption, which is related to the state's abundant fossil fuel and mineral resources. TX total energy consumption is high and growing at a high rate, mainly due to the abundant oil and gas resources in the state. In the meantime, there are many energy-intensive industries in TX, which consume a large amount of resources.

(2) Use of Clean Energy

Clean energy includes nuclear energy and renewable energy. We plot *NUETB*, *BMTCB*, *HYTCB*, *GETCB*, *SOTCB*, *WYEGB* in the stacked plot below. The development of clean energy in different states is different, it mainly expresses in the following aspects: energy consumption, the proportion of clean energy in total energy consumption, energy structure and type, development time and speed. The main factors leading to these differences are: population, geography, industry, climate, policy, technology and so on.

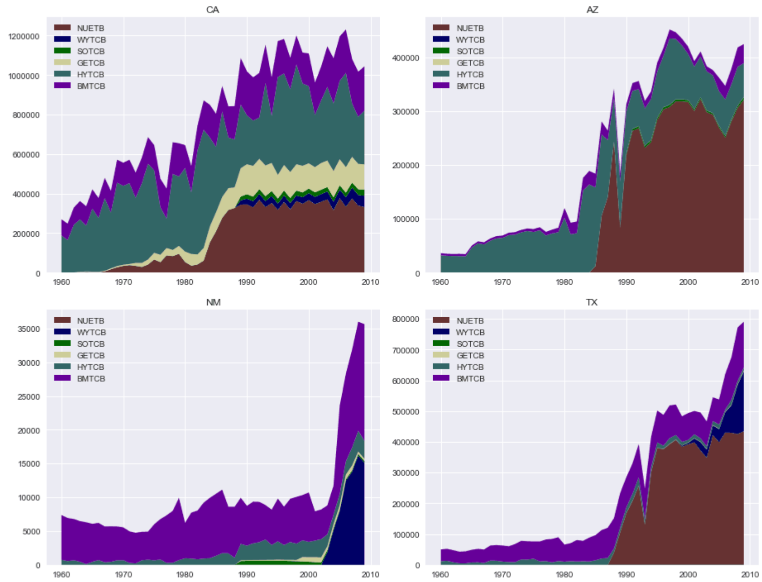

Figure 4 Clean energy stacked plot

All along, CA's clean energy consumption and the proportion of total resources consumption are larger, the structure of clean energy is diversified. As the state is located near the sea, the total amount of water energy is large and the proportion of it is stable. The development of nuclear and geothermal energy is significantly earlier than other states, due to geographical and technological advantages, they developed rapidly and reached a stable.

AZ has the largest share of clean energy, the structure of clean energy is diversified, which is related to the lack of renewable energy in the state. The total amount of water energy is large and the proportion of it is stable, which is related to the greater precipitation in some areas of the state. The proportion of nuclear energy is large between 1985 and 2009, with Palo Verde, the state's largest nuclear power plant is the second largest power plant in USA.

NM Clean energy consumption is less, after 2000, there is a greater increase, the development tends to be diversified, which is related to the state's policy. Due to the larger area planted, biomass energy has been the state's main clean energy source. In recent years, wind and hydropower energy have been developed.

TX clean energy in the proportion of total energy is relatively small, the development of energy structure is relatively simple, but in the past 20 years there has been a greater development. Biomass energy has always been an important source of clean energy, with a steady share. Like other states, wind develops rapidly and the produce of wind energy is the largest among the four states.

5.1.3 The Evaluation for the Use of Clean Renewable Energy

Using the previous conclusions, the selection of indicators for our clean renewable energy assessment model mainly from the technical, social, environmental and economic considerations. We select the relevant data, then, we use analytic hierarchy process to determine the weight of the indicators to calculate the final score to confirm which state has the "best" profile of using clean renewable energy.

1) Determination of Indicators

The selection of clean and renewable energy evaluation indicators mainly from the technical, social, environmental and economic considerations.

(1) Technical Indicator x_1

Technical indicator reflects the technical level of clean energy using. The use of biomass requires a higher level of technology, so we select the total consumption of biomass per capita as a quantitative indicator of the technical. The larger the biomass consumption per capita is, the better the clean renewable energy technologies develop.

$$x_1 = \frac{BMTCB}{TPOPP} \qquad (3)$$

Where: *BMTCB* is Biomass total consumption, *TPOPP* is resident population including Armed Forces.

(2) Social Indicator x_2

Social indicator refers to the contribution made by clean energy to society. Here, we use the proportion of clean, renewable energy production (*REPRB*) in total energy production (*TEPRB*) to describe the social indicators. The larger the social indicator is, the greater the proportion of the clean, renewable energy production in total energy production is. We express the relationship as:

$$x_2 = \frac{REPRB}{TEPRB} \qquad (4)$$

(3) Environmental Indicator x_3

Environmental indicator mainly refers to the amount of waste generated during the clean energy treatment. Here, we measure the emissions of waste with the per capita emissions of CO_2, and we assume that the emissions of CO_2 only related to the consumption of non-renewable energy. So, developing renewable energy helps mitigate the greenhouse effect. We have collected the CO_2 emissions of each state in 2009 from the website given in the problem,[1] as shown in the following Table 4. The larger the CO_2 emissions per capita is, the greater the positive impact of clean energy is. Here, we express the definition of Environmental Indicator as below:

$$x_3 = \frac{TTCDE}{TPOPP} \qquad (5)$$

Where: *TTCDE* is the total carbon dioxide emissions; *TPOPP* is resident population including Armed Forces.

Table 6 CO2 emissions from the four states in 2009

States	CA	AZ	NM	TX
The Total Carbon Dioxide Emissions （million metric tons of CO2）	372.0	37.8	57.3	550.1

（4） Economic Indicator x_4

Economic indicator shows the economy of clean, renewable energy. We assume that the creation of GNP only come from the consumption of energy, and everyone has the same energy use conditions. We use the ratio of income and cost of using clean energy for each person to characterize economic indicators.

$$x_4 = \frac{\frac{GDPRV \cdot x_2}{TPOPP}}{\frac{RETCB \cdot TETCD}{TPOPP}} = \frac{GDPRV \cdot x_2}{RETCB \cdot TETCD} \qquad (6)$$

Where: *GDPRV* is current-dollar gross domestic product; *TPOPP* is Resident population including Armed Forces; *RETCB* is renewable energy total consumption; *TETCD* is total energy average price.

2) Evaluation of the Four States

(1) Data Preprocessing

The values of the four evaluation indicators we obtained are shown in the following table, we mark the data which is in the i-th row, the j-th column in the table as a_{ij}, $i = 1,2,3,4$, $j = 1,2,3,4$.

states	x_1 （Billion Btu /Thousand）	x_2	x_3 （million Metric tons/Thousand）	x_4
AZ	5.375634249	0.155117877	0.005738007	0.019367004
CA	6.090459424	0.243756757	0.010084686	0.035017879
NM	8.616098181	0.014005848	0.028545594	0.001701899
TX	5.985453285	0.02548864	0.022207733	0.005303745

The data is normalized and the processed variables are recorded as b_{ij}. Taking the different types of indicators into account, treatment is also different.

$$b_{ij} = \begin{cases} \dfrac{a_{ij} - a_j^{\min}}{a_j^{\max} - a_j^{\min}}, & j = 1,2,4, \\ \dfrac{a_j^{\max} - a_{ij}}{a_j^{\max} - a_j^{\min}}, & j = 3. \end{cases}$$

(2) Determination of Weight

Using analytic hierarchy process to determine the weight of each indicator μ_j, $j = 1,2,3,4$.

First, create a hierarchy as shown in the following table.

Table 7 Hierarchy table

Target level	The state that has the "best" profile of using clean renewables
Guidelines level	Technical Indicator、Social Indicator、Environmental Indicator、Economic Indicator
Program level	AZ、CA、NM、TX

Then, we construct a paired comparison matrix.

$$\begin{bmatrix} 1 & 1/2 & 1/3 & 1/3 \\ 2 & 1 & 2/3 & 2/3 \\ 3 & 3/2 & 1 & 1 \\ 3 & 3/2 & 1 & 1 \end{bmatrix}$$

The maximum eigenvalue is 4, and the corresponding eigenvector is $u = \begin{bmatrix} 0.3592 & 0.4170 & 0.3592 & 0.3592 \end{bmatrix}$. Conformance index CI=0, conformance ratio $CR = 0 < 0.1$, so, we pass the consistency test. Thus, the weight vector:

$$\omega_j = \frac{\mu_j}{\sum_{j=1}^{4} \mu_j} \tag{7}$$

Then, we get the weight $\omega_j = [0.240331861, 0.279004416, 0.240331861, 0.240331861]$.

(3) Calculate the Final Score y_i

$$y_i = \sum_{j=1}^{4} b_{ij} w_j,\ j = 1,2,3,4. \tag{8}$$

We get the state's rating score as follows:

Table 8 The state's rating score

states	AZ	CA	NM	TX
Evaluation score	0.539126152	0.766881236	0.240331861	0.151939244

We can obtain from the above table, the highest score belongs to CA, AZ second, NM third, TX lowest. Therefore, CA has the best profile of clean, renewable energy. Among the four states, CA has the highest social benefits because renewables make up the largest share of total energy. CA's renewable energy creates the highest unit value and therefore has the highest economic benefits. At the same time, CA plays the superiority of environmental protection of clean energy.

5.1.4 Forecast of Energy Profile

In the analysis of 4.1.2, we obtained the evolution of the energy profile by curve fitting. The fitting results have better characterization ability to the year from 1960 to 2009, but not suitable for the long-term and medium-term prediction. Taking the energy profile to be affected by many factors into account. First, we conduct an analysis. Then we set up the GBDT model to predict the total energy consumption and the total renewable energy consumption in each state in 2025 and 2050.

1) Analysis of Influencing Factors

From the previous model results analysis, we can obtain that the consumption of energy is influenced by population, geography, technology, industry and climate and other factors. We assume that the climate in one state does not change significantly over the period of time studied. We search relevant data of the population, GDP, energy production, energy intensity, and industrial structure to study the relationship between them and energy consumption.

(1) Population: Due to population grows, consumption of energy by society increase. We use *TPOPP* to represent the number of population.

(2) Energy production: For the impact of geographical factors, we mainly consider the impact of mineral ownership. In order to facilitate the study, we use energy production reflect the amount of energy ownership. We select the total energy production *TEPRB* and renewable energy production *REPRB* to conduct research.

(3) Energy intensity: For technical factors, we think that technology promotes the realization of energy value, which is reflected by energy intensity. Energy intensity is the ratio of total primary energy use to gross domestic product, at the same time, it is one of the most commonly used indicators for comparing the efficiency of energy utilization in different countries and regions, and reflects the economic benefits of energy utilization. [2]Therefore, we use the total energy consumed per dollar of real gross domestic product (*TETGR*) to represent energy intensity.

(4) Industrial Structure: Different industrial structures have different impacts on energy consumption. According to the analysis in 4.1.1, we know that the proportion of the industrial sector has a great impact on the total energy consumption. We represent the impact of industry factors with the share of *TEICB* in *TETCB*.

(5) Gross Domestic Product: Economic development is at the expense of energy consumption. We choose the real gross domestic product (*GDPRX*) to indicate the impact of economic development on energy consumption.

2) Filling of Missing Quantity

Among the above indicators, the four states' real gross domestic product (*GDPRX*) and total energy consumed per dollar of real gross domestic product (*TETGR*) are missing. They all lost data from 1960 to 1976. For *GDPRX*, directly based on the data after the forward forecast there will be deviations, here we find the data of the gross national product of the United States over the years. Then, we calculate the Pearson correlation coefficient for the four states, and we find the value of it is close to 1. So, we populated the missing *GDPRX* data based on the U.S. gross national product. Based on the definition of energy intensity mentioned above, we look for the total primary energy consumption for the first 13 years, which is denoted as *TEOTE*. [3]*TETGR* can be obtained from the following equation:

$$TETGR = \frac{TEOTE}{GDPRX}$$

3) Use GBDT Model for Training Prediction

(1) Introduction to GBDT [4][4]

CART Decision Tree Algorithm: The CART Decision Tree Algorithm can be used for classification as well as for regression prediction. When it is used for regression, it uses the strategy of square error minimization. CART is a dichotomous recursive segmentation technique. By choosing the optimal segmentation feature A value, and then data is divided according to the value is greater than or less than the value of the tree split eventually generate regression tree

Given a set of training data $D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_N, y_N)\}, y_i \in R$. The input space is divided into M units $R_1, R_2, \cdots, R_M$ and the output value of the unit R_m is $c_m, m = 1, 2, \cdots, M$. Then the regression tree model is as follows. Where, $I(\cdot)$ is a characteristic function.

$$f(\mathbf{x}) = \sum_{m=1}^{M} c_m I(x \in R_m)$$

GBDT algorithm: GBDT is an integrated model, which consists of multiple decision trees. The GBDT model can be represented as the addition model of the basic classifier.

$$f(\mathbf{x}) = f_M(\mathbf{x}) = \sum_{m=1}^{M} h_m(\mathbf{x}, \Theta_m)$$

Where, $h_m(\mathbf{x}, \Theta_m)$ is decision tree, Θ_m is decision tree parameters, M is the number of the decision tree.

(2) Model Training and Prediction

First, train the model. We make the above five indicators as a feature and make the total energy consumption and renewable energy consumption as a label to train model. We take the AZ renewable energy consumption as an example, we get the root mean square error is 5540.4. It should be pointed out that this error appears to be large, but it is small compared with the magnitude of the forecast value. So, the model fits well. And then we will further reduce the error.

Then, we used the corresponding methods to predict four factors in 2025 and 2050. According to reference[5], we use the improved Logistic model to predict the population in 2025 and 2050. We use the ARMA model to predict energy intensity[6]and GDP[7].

Although the production of renewable energy and total energy is almost steady, we use the moving average method to predict. For the proportion of industry, we carry out regression analysis to get the predicted value. So, we get the predicted value of all the factors.

Last, we put the predictive value of the factor characteristics into the trained model, then, we get the prediction value of the energy profile, as shown in the following figure.

Table 9 Predictive values of *TETCB* and *RETCB*

value / year	CA		AZ		NM		TX	
	TETCB	*RETCB*	*TETCB*	*RETCB*	*TETCB*	*RETCB*	*TETCB*	*RETCB*
2025	8149793	738884	1529268	95801	689743	317951	11312929	312083
2050	8158872	761945	1548130	98742	692942	341313	11478427	356306

5.2 The Solution of Part Ⅱ

The purpose of establishing energy compact between the four states is to reach agreement at specific policy issues in the use of clean, renewable energy and adopt a set of standards to cooperate with regional or national affairs. In this section, first, we establish a decomposition model for the planning goal of renewable energy that allocates energy goals to states from the top down. Then, we give three actions that the four states may take to meet their energy contract goals.

5.2.1 Determination of Energy Goals

Based on that we get the state of energy profiles, similarities and differences, and best energy profile standards in each state, we establish a decomposition model for the planning goal of renewable energy and give the energy goals of renewable energy use in all states。

1）The Establishment of Decomposition Model for the Planning Goal of Renewable Energy

Based on the energy profiles and similarities and differences of states, we divide the energy target into three parts: fixed share and floating share, and allocate the total energy goal to each state. Here we step by step to establish the model.

(1) Determine the overall target share. The overall target share refers to the percentage of the total energy consumption of the target amount of renewable energies in each state in the predicted years. It consists of a fixed share and a floating share, which represents four states. Energy targets in four states are an integral part of the overall objective. The determination of the overall objective strengthens energy cooperation and distribution among States, and can achieve energy mutual assistance. Formally, we express the relationship as:

$$R_{i,total} = R_{i,float} + R_{i,fix} \tag{9}$$

(2) Determine the fixed shares of each state. The fixed share R_{fix} is the share that must be reached. It is the sum of the existing share R_{start} and the basic share R_{base}, it has nothing to do with other factors.

$$R_{i,fix} = R_{i,start} + R_{i,base} \tag{10}$$

Where: The basic growth share is the growth goal that each state must meet;

The existing share corresponds to the social indicators in the energy source profile, which is total renewable energy consumption in 2009 divided by total energy consumption, and it's 5.5% in developed country.[8]

$$R_{i,start} = \frac{RETCB_{i,2009}}{TETCB_{i,2009}} \tag{11}$$

(3) Determine the floating share of the states. [9]In determining this indicator, we take into account the economic indicators in the best energy profile, that is, the energy consumption produced per unit of gross domestic product. The floating share R_{float} is mainly related to the economic development status $GDPX_{i,average}$ and population quantity $TPOPP_i$ of each country.

The per capita floating renewable energy consumption in four states:

$$E_{float,average} = (20\% - R_{fix}) \cdot \sum_{i=1}^{4} TPOPP_i . \tag{12}$$

Per Capita Floating Renewable Resources by Province:

$$E_{i,float,average} = \frac{GDPX_{i,average} \cdot E_{float,average}}{GDPX_{average}} . \tag{13}$$

The floating share of the states is as follows, where *t* represents the predicted year

$$R_{i,float} = \frac{TPOPP_i \cdot E_{i,float,average}}{E_{i,t}}. \qquad (14)$$

(4) The target amount of renewable energy in the forecast year is shown in the following formula. The distribution of renewable energy is a manifestation of the promotion of renewable energy. This improves environmental indicators in the best energy profile.

$$RETCB_{i,t} = R_{i,total} \cdot TETCB_{i,t} \qquad (15)$$

2）The Establishment of Decomposition Model for the Planning Goal of Renewable Energy

We use the known data and forecast data to solve the above model and get the state renewable energy goals. We can see that the current state of clean energy profiles better in the future there will be greater energy allocation goals.

Table 10 Energy distribution targets in each state

Year	CA	AZ	NM	TX
2025	653365.5	149325.3	155681.6	343459
2050	686924.8	197577.7	182632.2	365891

As shown in the following figure, the goals of renewable energy in the states consist of fixed variable and floating variable. where the x-axis is a fixed variable, the bottom is the floating variable.

As shown in the following figure, the goals of renewable energy in the states consist of fixed variable and floating variable. where the x-axis is a fixed variable, the bottom is the floating variable.

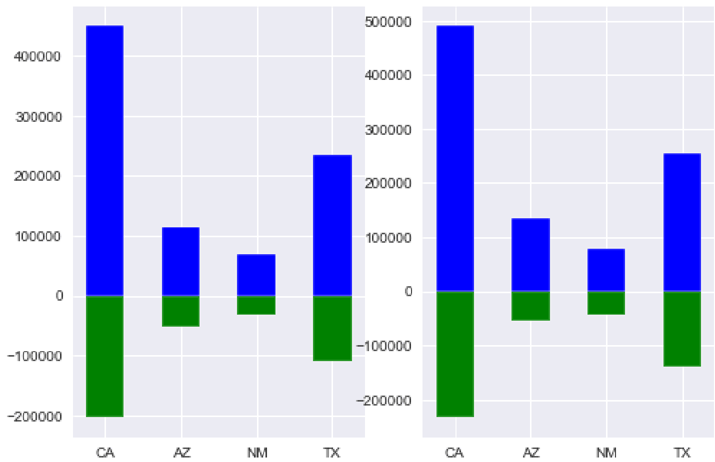

Figure 5 The picture variable and floating variable.

Allocate the energy goals that the states have already achieved to meet the 2025 and 2050 renewable energy goals. For sources of renewable energy, States can distribute them in a certain proportion to different industries which can product energy according to their respective development advantages. For the consumption of renewable energy, states can allocate them to different energy-consuming industries according to their own industry characteristics. The result of allocation according to the proportions of the various parts of 2009 is given below. The allocation of AZ, NM, and TX are in the appendix.

Table 11 Renewable Energy targets in various capacity industries in CA

Year	Wind Energy	Solar Energy	Geothermal Energy	Water Can	Biomass
2025	52269.24	26134.62	117605.79	248278.89	209076.96
2050	54953.984	27476.992	123646.464	261031.424	219815.936

Table 12 Renewable energy targets in various energy-consuming industries in CA

Year	Transport sector	Commercial sector	Industrial sector	Residential sector
2025	301332.1686	66773.9541	171965.7996	100422.2774
2050	316809.7178	70203.71456	180798.6074	105580.3418

5.2.2 Measures to achieve energy objectives

To achieve the aforementioned energy contract objectives, we propose the following initiatives.

1) Increase the production of various types of renewable energy. The states gradually increase the proportion of renewable energy in total energy and achieve diversified energy development, and according to local conditions, highlighting the advantages of development. For CA, the state should maintain a "best" energy profile and increase the proportion of renewable energy production in the total energy production of the whole society and improve social indicators. For AZ, the state should further increase renewable energy, to make up for the lack of renewable energy and the weakness of foreign power dependence. To maintain the traditional advantages of water to further develop solar energy and geothermal energy. For NM, renewable energy should be vigorously developed to reduce the CO_2 emissions caused by the depletion of non-renewable energy. Maintain the evolution of renewable energy profiles since 2000. On the basis of exerting its wind energy and bio-energy, changing the single mode of development. Development of solar and geothermal energy. TX Should improve the worst energy profile in 2009, reduce reliance on non-renewable resources, adjust the industrial structure and develop the tertiary industry accordingly. Based on the original biomass and nuclear energy, continue to play the advantages of wind energy.

2) Enhance cooperation between the states. To establish a state-help each other relationship. According to the energy profile of the evaluation results, establish mutual relationships between CA and TX, AZ and NM. If a state cannot complete a given renewable energy target, it can be shared with other states. If a state cannot complete its consumption target for renewable energy, it can work with other states to complete. Complement each other, mutual benefit.

3) Create a good external environment. Implement relevant encouragement policies and strengthen the investment and construction of infrastructures for the production of renewable energy. Incentive mechanisms for the use of renewable energy. Further expand the cooperation area of energy contract. Improve the technological level of renewable energy production and utilization. Reduce energy production costs and increase energy intensity. Developing recycling economy with resource saving. Do a good job of low carbon energy saving and environmental protection propaganda and education guide.

6 Model Testing and Sensitivity Analysis

6.1 Sensitivity Analysis of Analytic Hierarchy Process

When we use analytic hierarchy process to determine the weight of each indicator, the structure of paired comparison matrix is subjective. Because of that, we make sensitivity analysis of this, change the comparison matrix, and then calculate the score.

$$\begin{bmatrix} 1 & 1/2 & 1/3 & 1/2 \\ 2 & 1 & 2/3 & 1 \\ 3 & 3/2 & 1 & 3/2 \\ 2 & 1 & 2/3 & 1 \end{bmatrix}$$

Obtained, the maximum eigenvalue is 4, and the corresponding eigenvector is $u = [0.3922, 0.4714, 0.2098, 0.8742]$. Consistency index $CI = 0$, Consistency ratio $CR = 0 < 0.1$, so, we pass the consistency test. Thus, the weight vector is $\omega_j = [0.2014, 0.2420, 0.1079, 0.4487]$.

Table 13 The state's rating score

States	AZ	CA	NM	TX
Evaluation score	0.365904929	0.424528862	0.2014	0.128489593

Compare the results of the previous score, although the value of the evaluation scores are changed, the final evaluation score ranking order unchanged, and CA still has the best overview of clean and renewable energy.

6.2 Test of GBDT Prediction Model

We use the idea of cross-validation to test the correctness of the model, select the appropriate parameters. Here we use 5-fold cross-validation to test the robustness of the algorithm, which is a commonly used standard practice. The method randomizes the data into 5 parts, taking one of them each time and using the remaining 4 parts to predict the total energy consumption and the renewable energy consumption. Finally, the predicted results are obtained by means of averaging. Taking AZ's 2025 renewable energy consumption as an example, our cross-validation reduce our root mean square error from 5540.4 to 3224.7.

7 Strengths and Weaknesses

7.1 Strengths

（1）From the analysis of the data, the prediction, and then the proposal, we established the model before and after the association, layers of progressive

（2）GBDT Prediction Model has the idea of model merging, and adopts the method of Cross-validation, the model has good stability and robustness.

（3）Instead of being limited to the given data, we are looking for a lot of ancillary data to help us build our model.

7.2 Weaknesses

（1）The indicators we choose to build our models are based more on references, and there may be deep-seated indicators that do not uncover.

（2）Our model does not consider the import and export of energy impact, because the data only contains some of the indicators of import and export. We also can not get statistics from other ways.

（3）The amount of data rarely affects the accuracy of predictions.

8 Model Improvement

（1）We can help us improve the accuracy of our model predictions by getting statistics from 2009 to the present.

（2）We have found many dedicated energy forecasting models, such as the LEAP model and the message model. But because of the time factor and the data limit, we do not use these professional tools. We can consider using them in predictions or validations in the future.

9 Conclusion

By preprocessing and analyzing the data, we select the appropriate variables to reflect the change of energy profile. Then, we establish the evaluation model and CA is considered to have the “best” profile for use of cleaner, renewable energy. For the next problem, we establish the GBDT model to predict energy profiles for 2025 and 2050. Through the gradual decomposition of the goals of renewable energy, we have identified the energy goals that each state should achieve, and put forward three measures to help them reach this plan.

10 Reference

[1] https://www.eia.gov/state/seds/seds-technical-notes-complete.php?sid=US#Consumption
[2] https://baike.so.com/doc/7678009-7952104.html
[3] https://www.eia.gov/state/seds/data.php?incfile=/state/seds/sep_use/total/use_tot_CAcb.html&sid=CA
[4] Hua school, Wang Zhenglin. python machine learning [M] .Beijing: Publishing House of Electronics Industry, 2017.37-38.
[5] WANG Yanchen, DUAN Junsheng, WANG Yan. Improvement of Population Prediction and LOGISTIC Model [J]. Statistics and Decision, 2006, (22): 136-137.
[6] Wang Jin. Analysis of Primary Energy Production in China Based on ARMA Model [J]. Oriental Culture, 2011, (10): 256-256.
[7] Zhang Qianqian. Design of Prediction Model of Per Capita GDP of the United States Based on ARMA Algorithm [J]. Journal of Yunnan University for Nationalities (Natural Science Edition), 2016, (2): 190-194.
[8] Liu Zhen, Zhang Xiliang, Gao Hu, et al. An objective decomposition model based on the potential of renewable energy and energy consumption [J]. Renewable

Energy, 2011, (3): 78-83.
[9] LIU Zhen, ZHANG Xi-liang, ZHANG Da. Enlightenment of EU's Renewable Energy Target Decomposition on China's Provincial Planning [J]. China Mining, 2009, (9): 66-70

Appendix

Renewable Energy targets in various capacity industries ：

Table 14 AZ

year	Wind Energy	Solar Energy	Geothermal Energy	Water Can	Biomass
2025	433.04337	6824.16621	462.90843	91088.433	50770.602
2050	572.97533	9029.30089	612.49087	120522.397	67176.418

Table 15 NM

year	Wind Energy	Solar Energy	Geothermal Energy	Water Can	Biomass
2025	36686.664	681.32376	759.93804	6289.1424	41665.5684
2050	42309.792	785.75328	876.41712	7253.1072	48051.8352

Table 16 TX

year	Wind Energy	Solar Energy	Geothermal Energy	Water Can	Biomass
2025	188902.45	789.9557	1992.0622	10303.77	142878.944
2050	201240.05	841.5493	2122.1678	10976.73	152210.656

Renewable energy targets in various energy-consuming industries

Table 17 AZ

Year	Transport sector	Commercial sector	Industrial sector	Residential sector
2025	76812.93432	19710.9396	30357.83349	22443.59259
2050	101633.9689	26080.2564	40167.54641	29695.92831

Table 18 NM

Year	Transport sector	Commercial sector	Industrial sector	Residential sector
2025	63331.27488	16237.59088	57508.78304	18588.38304
2050	74294.77896	19048.53846	67464.33468	21806.28468

Table 19 TX

Year	Transport sector	Commercial sector	Industrial sector	Residential sector
2025	87959.8499	18615.4778	213562.8062	23355.212
2050	93704.6851	19831.2922	227511.0238	24880.588