Memo

[**1 Introduction** 1](#_Toc506249085)

[**1.1 Statement of the Problem** 1](#_Toc506249086)

[**1.2 Assumptions** 1](#_Toc506249087)

[**2 The Energy Profile of the Four States** 1](#_Toc506249088)

[**3 Selecting the Representative Variables** 4](#_Toc506249089)

[**4 Addressing the Provided Data** 5](#_Toc506249090)

[**5 CR Statistic Model Establishment** 6](#_Toc506249091)

[**5.1 CR Model Definition** 6](#_Toc506249092)

[**5.2 CR Statistic Application on Four States** 9](#_Toc506249093)

[**6 Future Data Forecasting with four models** 12](#_Toc506249094)

[**6.1 Using ARIMA model to predict prices** 12](#_Toc506249095)

[**6.2 Using Grey Theory to predict production** 14](#_Toc506249096)

[**6.3 Using Neural Network Algorithm to Predict Consumption** 15](#_Toc506249097)

[**6.4 Using Cobb–Douglas Production Function to Predict GDP** 16](#_Toc506249098)

[**7 The Common Future: Adjusting Prices to Make the Four States Cleaner and More Renewable** 17](#_Toc506249099)

[**7.1 Feasibility of government’s control of price changes** 17](#_Toc506249100)

[**7.2 The relationship between energy demand, energy output and energy prices** 18](#_Toc506249101)

[**7.3 Linear Programming Operation and Interpretation of Results** 20](#_Toc506249102)

[**7.4 Sensitivity Analysis** 20](#_Toc506249103)

[**Reference** 22](#_Toc506249104)

[**Appendices** 24](#_Toc506249105)

[**Appendix** **A CR statistic program** 24](#_Toc506249106)

[**Appendix** **B ARIMA model: Predict price** 25](#_Toc506249107)

[**Appendix** **C Grey Theory: Predict Production** 26](#_Toc506249108)

[**Appendix D Cobb–Douglas Production Function: Predict GDP** 27](#_Toc506249109)

[**Appendix** **E Linear Programming Operation** 28](#_Toc506249110)

**1 Introduction**

**1.1 Statement of the Problem**

Conventional energy sources such as oil and coal are depleting on the planet, and conventional energy reserves in the world can only last for half a century, for example, oil can only be mined for 41 years, natural gas can only be mined for 67 years. The largest amount of coal It can only be mined for 192 years. At the same time, the use of conventional fossil fuels has caused serious pollution to the planet. Therefore, comprehensive consideration from various aspects, the development of clean and renewable energy is imminent.

This article is about an evaluation and prediction about the energy profile in the four states in the U.S. We aim to design a measure system of the energy cleanliness level based on several statistic methods. To solve the problem, we will use the Time Series Analysis, Linear Programming, Grey Theory, Artificial Neural Network and some other methods to determine the optimal strategy in evaluating and predicting the energy consumption and cleanliness level.

**1.2 Assumptions**

A1. Policies within and between states will not change in short time

A2. The economies of all states will not fluctuate significantly in the future and will always maintain a stable operation so that the prices of various resources will not change significantly.

A3. Suppose the energy supply in these four states is largely satisfied by the energy output between the four states, that is, a small system is formed between the supply and demand of energy in the four states. Geographically speaking, the four states are adjacent, so this assumption is in fact quite justified.

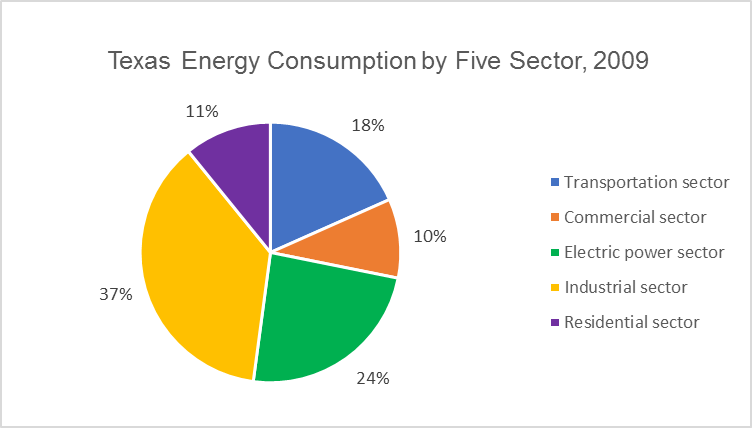
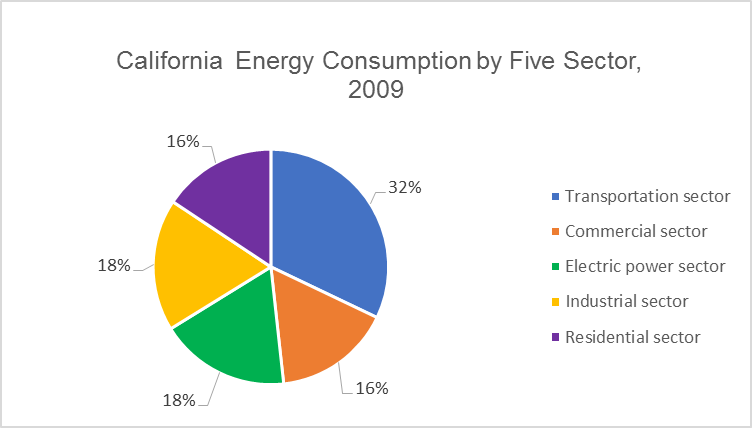
**2 The Energy Profile of the Four States**

To investigate what the energy profile of the four states, we decide to draw several photographs of the four states of various aspects.

First, energy consumption of the four states’ five sectors including transportation sector, commercial sector, electric power sector and industrial sector are drawn which are shown blow.

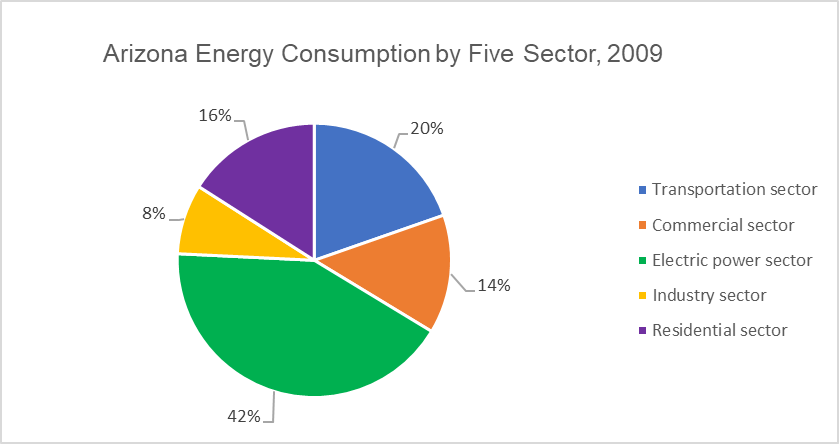
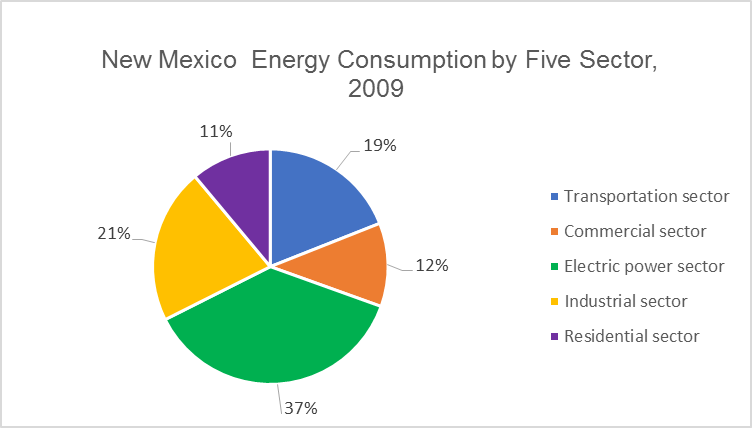
There are several characteristics which can be recognized from the four graphs. The transportation sector in the Arizona has the highest ratio among the four states. It is easily to be understood because there are four roads of Interstate Highway System, No.8, No.10, No.17, No.40, which go across Arizona. For Arizona, energy consumed by the transportation sector is also the second highest among the five sectors, accounting for as much as 20% of the energy consumed by the power sector. The Arizona economy relied on copper, cotton, cattle, citrus and climate for its early career, almost in the primary industry. After the Second World War, the manufacturing industry developed rapidly and has dominated the economy.

For California, the sector with the highest energy consumption is the transportation sector, with densely-populated highway networks and second-largest length in the country. While several other departments have a relatively average ratio of energy consumption. With its Mediterranean climate and favorable geographical conditions, California has the largest

Mediterranean climate and agricultural area in the world. California's agricultural output ranks the top in the nation for many years and its exports are far ahead. California is economically active, well-developed manufacturing sector, complete departments, a wide range of products, output and employment number ranks first among the country. Arizona has aerospace, electronics and other emerging sectors and oil refining, petrochemicals, arms, food processing, paper making, printing and so on. Mining oil and gas exploration accounts for 2/3 of the total output value.

*Figure 1: Texas energy consumption Figure 2: California energy consumption*

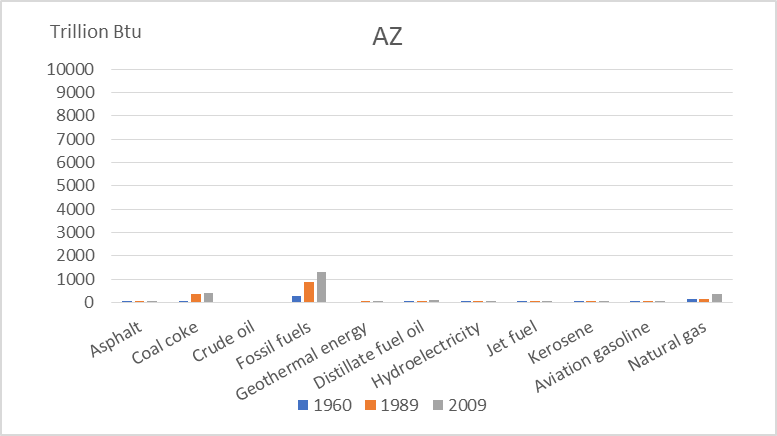
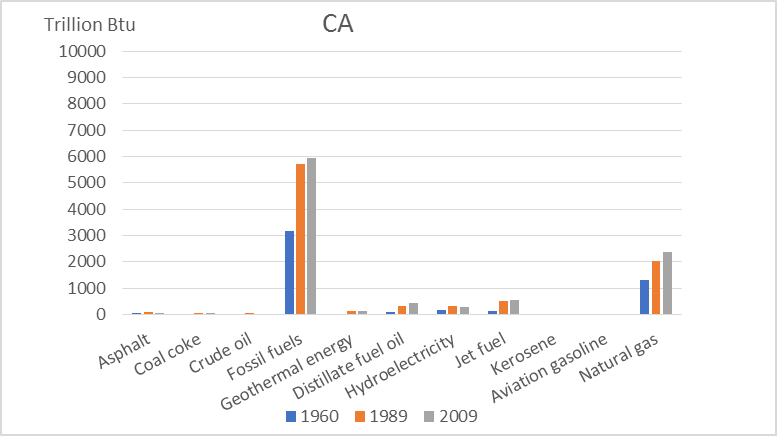
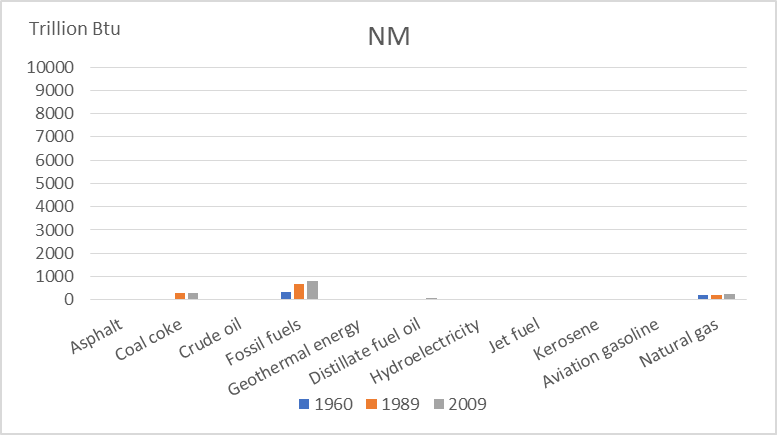
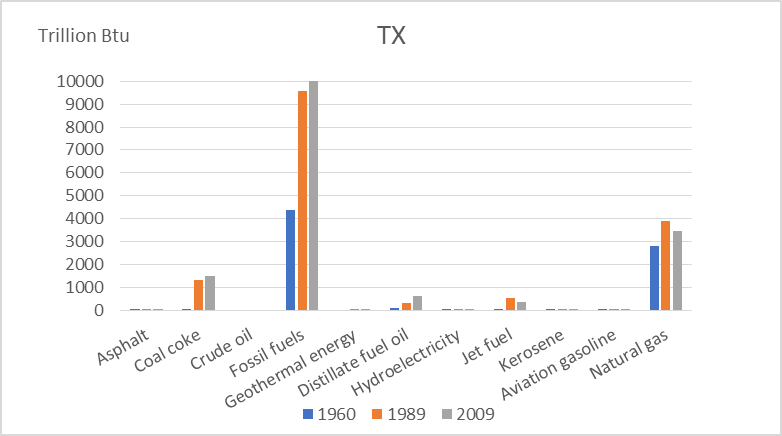
*by five sectors, 2009 by five sectors, 2009*

The main energy consumption in New Mexico came from the electrical sector, as the main industries in New Mexico are electronics and electrical equipment, food and related products, transportation vehicles, wood and wood products and machinery (excluding electronics).

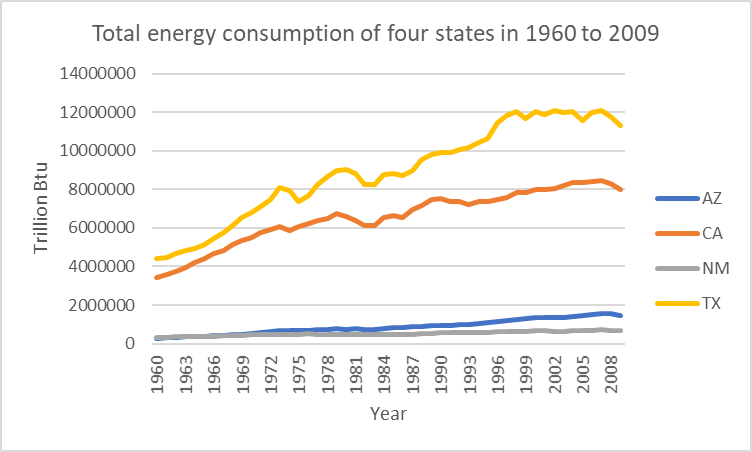
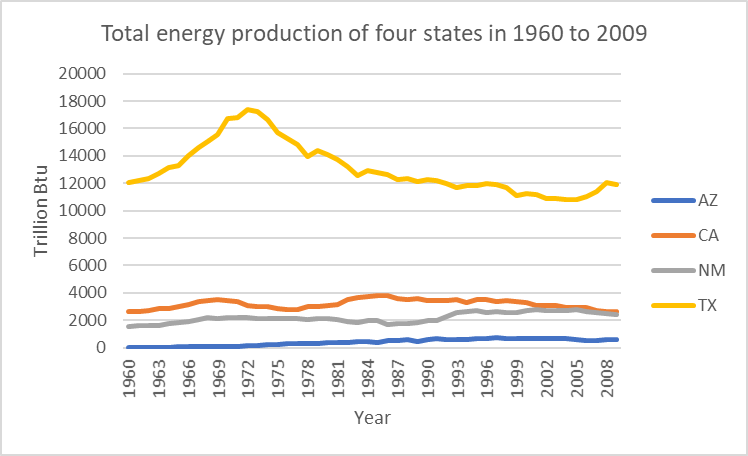
*Figure 3: Arizona energy consumption Figure 4: New Mexico energy consumption*

*by five sectors, 2009 by five sectors, 2009*

Texas's energy consumption is dominated by industry, due to its reputation for the energy and petrochemical industry and it has the largest energy and chemical state in the United States. Oil and gas production accounted for 1/3 and 1/4 of the national output, respectively, refining capacity accounted for 1/4 of the United States. Since the 1980s, Texas has pursued an economic diversification policy. Although oil and natural gas are still the major industries in Texas, their share has dropped, and the high-tech industries have been developing rapidly. Houston, San Antonio and the Dallas Triangle have become the second "Silicon Valley", Houston area has been developed into largest medical training and treatment center of the United States and even of the world. This explains why Texas's energy consumption comes mainly from the industrial sector.

Second, we selected three years that are representative: 1960, 1989 and 2009 to draw the energy composition of the four states’ consumption in these three years as follows:

*Figurer 5, 6, 7, 8: Four states’ energy consumption in 1960, 1989 and 2009*

For the first, we can figure out that Texas consumes more fossil fuel than the other three states among the four states, especially in 1989 and 2009. The reason is that the Texas tends to traditional industries and oil-rich industry. From the trend of each state, the growth rate of energy consumption increased more rapidly in the period from 1960 to 1989 than in the period from 1989 to 2009, which also reflected that 1960-1989 was a period of rapid industrial development in the United States. Thirdly, we analyzed the trend of the total cost of energy consumption and total energy consumption of the four states from 1959 to 2009 over the past 50 years and put it in a chart as follows:

*Figure 9: Total energy production of four Figure 10: Total energy consumption of four*

*states in 1960 to 2009 States in 1960 to 2009*

As we can see, energy consumption trends of all four states are gradually increasing in general, which shows that the development of the United States over the past 50 years determines the increase in energy consumption. Similarly, we can see that the energy consumption in Texas is always the largest of the four, which is reflected in the two previous figures, which proves the correctness of the conclusion again.

**3 Selecting the Representative Variables**

To evaluate whether a state's energy system is clean and renewable, we set up a model to rate the state's usage of the energy. For a comprehensive consideration of the price, cleanliness, and other characteristics of each energy source, a comprehensive set of scoring criteria needs to be established. First, we screen out the data to be representative of the energy involved in the discussion of the back issues, the first is the traditional energy as follows:

|  |  |
| --- | --- |
| NaMES | UNITS |
| ARTCB: Asphalt and road oil total consumption | Billion Btu |
| AVTCB: Aviation gasoline total consumption | Billion Btu |
| CLTCB: Coal total consumption | Billion Btu |
| DFTCB: Distillate fuel oil total consumption | Billion Btu |
| JFTCB: Jet fuel total consumption. | Billion Btu |
| KSTCB: Kerosene total consumption. | Billion Btu |
| LGTCB: LPG total consumption | Billion Btu |
| MGTCB: Motor gasoline total consumption | Billion Btu |
| NGTCB: Natural gas total consumption | Billion Btu |
| RTFCB: Residual fuel oil total consumption | Billion Btu |
| WWTCB: Wood and waste total consumption. | Billion Btu |

*Table 11: List of variables of traditional energy*

Then there is clean & renewable (CR) energy:

|  |  |
| --- | --- |
| NAMES | UNITS |
| GETCB: Geothermal energy total consumption. | Billion Btu |
| HYTCB: Hydroelectricity total production | Billion Btu |
| BMTCB: Biomass total consumption | Billion Btu |
| NUETB: Electricity produced from nuclear power. | Billion Btu |
| SOTCB: Photovoltaic and solar thermal energy total consumption | Billion Btu |
| WYTCB: Electricity produced from wind energy. | Billion Btu |

*Table 12: List of variables of clean & renewable energy*

Then the table below gives the price of each energy name and unit:

|  |  |
| --- | --- |
| NAMES | UNITS |
| ARTCD: Asphalt and road oil average price, all sectors | Dollars per million Btu |
| AVTCD: Aviation gasoline average price, all sectors | Dollars per million Btu |
| CLTCD: Coal average price, all sectors | Dollars per million Btu |
| DFTCD: Distillate fuel oil average price, all sectors | Dollars per million Btu |
| JFTCD: Jet fuel average price, all sectors | Dollars per million Btu |
| KSTCD: Kerosene average price, all sectors | Dollars per million Btu |
| LGTCD: LPG average price, all sectors | Dollars per million Btu |
| MGTCD: Motor gasoline average price, all sectors | Dollars per million Btu |
| NGTCD: Natural gas average price, all sectors (including supplemental gaseous fuels) | Dollars per million Btu |
| NUETD: Nuclear fuel average price, all sectors | Dollars per million Btu |
| RFTCD: Residual fuel oil average price, all sectors | Dollars per million Btu |
| WWTCD: Wood and waste average price, all sectors | Dollars per million Btu |

*Table 13: List of variables of energy price*

And we also select some useful data for follow-up studies:

|  |  |
| --- | --- |
| NAMES | UNITS |
| GDPRX: Real gross domestic product | Trillion chained (2005) dollars |
| TEGDS: Energy expenditures as share of current-dollar GDP | Percent |
| TETCB: Total energy consumption | Billion Btu |
| TPOPP: Resident population including Armed Forces | Thousand |

*Table 14: List of other useful variables*

**4 Addressing the Provided Data**

The provided datasets suffer from severe data repetitive as well as a lot of irrelevant data, which could undermine the reliability, complicate the problem or even interpretability of any results. We adopted several different methods to deal with the problem we faced.

1. There are two kinds of units used to describe the resumption of fuel, to be united in the way we are solving the problem, we choose “Billion Btu” as the unit we use after analyzing the units which is used more frequently.

2. We omit the variable that is all zero because the calculation can be simplified, and the results well be more accurate.

3. The level of consumption can be calculated according to the quantities that was consumed by different sections. What’s more, the prices level of the fuel is not stable. So, we decided to choose the quantity of fuel as the consumption level.

4. The data that used to describe the net export is omitted because the data about the import and export is given and we can calculate net export with it.

5. The data about ENTCP is preserved only when it contains the denaturant.

6. The data which is relevant to conversion rate is omitted.

7. We preserve the real GDP and omit the nominal GDP because the price level changes every year.

**5 CR Statistic Model Establishment**

**5.1 CR Model Definition**

After screening out the types of energy discussed later, we begin to create a scoring model. After considering and judging the advantages and disadvantages of various existing models and their accuracy, we decided to mimic the AIC**错误!未找到引用源。** Statistics to build an effective assessment system, say, CR (Cleaner and Renewable energy) statistics.

In the kth state of the year (k = 1, ..., 4), the consumption of all CR energy is xk1, ..., xkm, and the consumption of conventional energy is zk1, ..., zkn. The real GDP is Y and the energy expenditures as share of current-dollar GDP is , resulting in a CR energy score for each state that can be written as:

Here we let = 0.25.

Freedom=

Output is the sum of the regular energy and clean energy scores used by a state, where αkj and βki denote the scoring coefficients of conventional energy and clean energy. While xkj and zki indicate the amount of conventional and clean energy used in the area, respectively. There is a problem that the only factor in Input will have huge influence than the factors in the Output, resulting that the states with lower input will get higher . So, we add a correction index “” ( [0,1]) to verify the formula.

How to determine coefficients will be discussed later, we start with an explanation of the validity of this formula and how it serves to determine the degree of cleanliness of a state's energy first.

First, the amount of energy used in a state is constant, and the different energy mix in the state can lead to different scores for the state under the system, for example, when the state uses more clean energy and less for conventional energy, the Output value in this area will be smaller, and vice versa. When using the same amount of energy in the state, the higher the GDP generated in the state, the higher the energy GDP conversion rate ρ, and the smaller the value of Input in the state. And the reason why we add a logarithm before Output statistics is that for the same state, the increase of energy consumption per unit in the early stage of its economic development will bring greater profits. As the incremental energy per unit of energy increases in the late stages of development, the benefits will be diminished. Therefore, we take a logarithmic function to deal with the Output statistics.

The Freedom term shows that, in general, when a state uses multiple clean energy sources in combination, its CR statistic should be lower than using only one clean energy source. Similarly, when the type of conventional energy used in the area become less, its CR statistics becomes greater. Therefore, based on the above discussion, the smaller the CR statistic for a state, the clearer the energy use in the state.

The following discussion to determine the coefficient.

First determine the, to judge the degree of pollution of a conventional energy, we select the following aspects:

1. When burning this kind of energy, every 1 kg CO 2 emission increase, the score increases 1, every 1 kg CO 2 emission increase, the score increases 10**错误!未找到引用源。**[2][3][4]
2. The price of 1 million Btu of the energy is x dollar, then the coefficient increases x / 10
3. Evaluate the cost and difficulty of the infrastructure needed to build when use this kind of the energy source
4. Given score to the energy storage difficult
5. Evaluate the stability of access to the energy

The full mark of 3,4,5 is three, the score can be 0,1,2,3 points, then

=

Then we begin to decide the value of, we can also evaluate the traditional energy in several aspects.

1. Clean energy prices [6]

Take the median for each kind of energy cost x dollar when produced 1 MWh electricity then the coefficient will increase x/100

2. Evaluate the cost and difficulty of the infrastructure needed to build when use one kind of energy

3. Evaluate the difficulty of the storage of this kind of energy

4. Evaluate the stability of this kind of energy

The full mark of 3,4,5 is three, the score can be 0,1,2,3 points, then

Note: Compared with traditional energy sources, the first one is canceled here because no CO2, CO or other gases generated during the use of clean energy

=

Through the above criteria to determine the coefficient , we can see that for conventional energy, the larger its coefficient, the higher the degree of pollution of conventional energy. Similarly, for clean energy, the smaller the coefficient, the cleaner the clean energy is.

The table below shows the scores for the above items 3, 4, 5 for several conventional, and 2, 3, 4 for clean & renewable energy sources:



*Table 15: List of CR scores*

On the basis of the discussion above, we draw a picture to visually represent our CR model system:

Output

Input

Freedom

CR

-

log(

Conventional energy

consumption

CR energy

consumption

GDP

The energy expenditures as share

of current-dollar GDP

The number of Conventional

energy source

The number of CR

energy source

-

*Figure 16: The model of CR model system*

**5.2 CR Statistic Application on Four States**

Then we calculate the CR statistic of the four states from 1960-2009 and 50 years by R software based on the known data and scores. The results are as follows, and the CR.AZ, CR.CA, CR.NM, CR.TX mean the CR statistic of the four states respectively:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | CR.AZ | CR.CA | CR.NM | CR.TX |
| 1960 | 19.507 | 21.905 | 19.833 | 22.439 |
| 1961 | 19.627 | 22.011 | 19.871 | 22.511 |
| 1962 | 19.761 | 22.208 | 20.027 | 22.642 |
| 1963 | 19.930 | 22.370 | 20.105 | 22.740 |
| 1964 | 20.059 | 22.475 | 20.191 | 22.847 |
| 1965 | 20.342 | 22.631 | 20.249 | 22.961 |
| 1966 | 20.578 | 22.721 | 20.317 | 23.080 |
| 1967 | 20.676 | 22.870 | 20.396 | 23.177 |
| 1968 | 20.953 | 22.939 | 20.543 | 23.317 |
| 1969 | 21.168 | 23.098 | 20.627 | 23.463 |
| 1970 | 21.250 | 23.163 | 20.793 | 23.556 |
| 1971 | 21.383 | 23.290 | 20.891 | 23.690 |
| 1972 | 21.541 | 23.305 | 21.022 | 23.815 |
| 1973 | 21.839 | 23.528 | 21.197 | 24.036 |
| 1974 | 22.212 | 23.914 | 21.642 | 24.571 |
| 1975 | 22.373 | 24.082 | 21.803 | 24.830 |
| 1976 | 22.539 | 24.186 | 22.024 | 25.063 |
| 1977 | 22.871 | 24.363 | 22.312 | 25.262 |
| 1978 | 22.964 | 24.436 | 22.454 | 25.457 |
| 1979 | 23.367 | 24.764 | 22.775 | 25.789 |
| 1980 | 23.660 | 25.042 | 23.174 | 25.982 |
| 1981 | 23.835 | 25.197 | 23.559 | 26.251 |
| 1982 | 23.825 | 25.147 | 23.413 | 26.307 |
| 1983 | 23.801 | 25.105 | 23.404 | 26.290 |
| 1984 | 23.881 | 25.191 | 23.325 | 26.290 |
| 1985 | 22.955 | 25.161 | 23.382 | 26.237 |
| 1986 | 22.852 | 24.897 | 23.171 | 25.978 |
| 1987 | 22.938 | 24.992 | 23.290 | 26.014 |
| 1988 | 23.068 | 25.041 | 23.287 | 25.044 |
| 1989 | 23.025 | 25.162 | 23.425 | 25.153 |
| 1990 | 23.232 | 25.247 | 23.637 | 25.234 |
| 1991 | 23.212 | 25.202 | 23.532 | 25.207 |
| 1992 | 23.226 | 25.196 | 23.572 | 25.245 |
| 1993 | 23.331 | 25.186 | 23.625 | 25.313 |
| 1994 | 23.358 | 25.231 | 23.584 | 25.298 |
| 1995 | 23.357 | 25.238 | 23.403 | 25.298 |
| 1996 | 23.524 | 25.308 | 23.706 | 25.481 |
| 1997 | 23.556 | 25.369 | 23.837 | 25.504 |
| 1998 | 23.498 | 25.295 | 23.699 | 25.399 |
| 1999 | 23.614 | 25.401 | 23.756 | 25.439 |
| 2000 | 23.874 | 25.746 | 24.111 | 25.796 |
| 2001 | 23.895 | 25.816 | 24.162 | 25.812 |
| 2002 | 23.803 | 25.539 | 24.007 | 25.705 |
| 2003 | 24.013 | 25.846 | 24.233 | 25.910 |
| 2004 | 24.306 | 25.941 | 24.489 | 26.035 |
| 2005 | 24.634 | 26.123 | 24.798 | 26.244 |
| 2006 | 24.746 | 26.175 | 24.954 | 26.240 |
| 2007 | 24.819 | 26.226 | 25.033 | 26.322 |
| 2008 | 25.051 | 26.371 | 25.245 | 26.497 |
| 2009 | 24.539 | 26.053 | 24.705 | 26.155 |

*Table 16: Results of CR scores, four states, 1960-2009*

First analyze the changes in the energy profile of each state from 1960-2009 over these 50 years.

First of all, we make a longitudinal comparison. From the above calculation, we can see that the CR statistic of these four states gradually increased in the 49 years between 1960 and 2008. The previous modeling shows that the larger CR statistics is in a region, the less conspicuous its energy consumption structure is. Therefore, from big to small, the economic development of the states over the 49 years has led to an increase in the demand for conventional energy such as fossil fuels. Texas, known for its energy and petrochemical industries, is the nation's largest energy and chemical state. Crude oil and natural gas output accounted for 1/3 and 1/4 of the national output, refining oil accounted for 1/4 of the United States. Texas, the nation with the largest greenhouse gas emissions, is home to a larger number of thermal power plants, refineries and processing industries in the state, which explains why Texas's CR statistic is much higher the other three states. Because of its rich oil production and oil reserves, and its predominantly industrial-based industries, its demand for and consumption of conventional fuels, such as oil, is huge and its CR statistic is naturally the highest of the four states.

While the state of Texas is basically temperate and parts of the south are subtropical, so its climate is generally livable; while the state of New Mexico is a desert climate, the terrain of Arizona is very steep, and its geographical and climatic conditions are far worse than The other two states，which has played a role in restricting their industrial development and energy consumption. The climate in summer is warm and the winter is not very cold. Texas is located in the southern United States, its summer temperatures are basically above forty degrees Celsius, sooner or later the temperature difference is relatively large; in the spring and autumn, the temperature difference during one day can reach thirty degrees Celsius, the winter is also very cold. The climate in New Mexico is a typical desert climate, warm and dry. Thanks to the abundant sunshine and low humidity, the state's average annual precipitation is 380 millimeters, but rainfall in the mountains can reach 1,000 millimeters, while in the driest areas it is only 200-250 millimeters. Due to the high altitude in winter, most of the snowy mountains. The temperature difference between day and night is quite large. The annual average temperature of 12 ℃, the highest temperature of 43 ℃, the lowest -2 ℃. The climate in New Mexico is a typical desert climate, warm and dry. Thanks to the abundant sunshine and low humidity, the state's average annual precipitation is 380 millimeters, but rainfall in the mountains can reach 1,000 millimeters, while in the driest areas it is only 200-250 millimeters. Due to the high altitude in winter, most of the snowy mountains. The temperature difference between day and night is quite large. The annual average temperature is 12 ℃, the highest temperature is 43 ℃, the lowest temperature is -2 ℃. Therefore, in contrast, the climate of California is better than that of Texas, while the climate of Arizona and New Mexico tends to be desert arid climate, which is more unfavorable to the development of the city. The population distribution is climate-related. The population of California in 2009 was 36.96 million, that of Texas in 2009 was 24.78 million, that of New Mexico in 2009 was 2 million and that of Arizona in 2009 was 6.59 million. It can be seen that the climate environment determines the population distribution. Therefore, the same conclusion can be drawn from the perspective of climate and population. One more interesting feature is the reduction of CR statistics across all four states from 2008 to 2009 and from 1985 to 1986. Although the trend of these two-time periods is the same, the causes are different. The first was between 1985 and 1986 when a huge financial crisis broke out in the United States. It originated from a strong dollar policy that induced speculative funds to flow back to the United States, causing the real estate and stock markets in the United States to skyrocket and crash. This is a perfect fit for the results of CR statistics.

Since 2008, the clean energy sector in the United States has undergone rapid development. Sudden emergence of gas power generation, wind power and solar power generation doubled: the annual wind power generation capacity increased by 3 times, large-scale solar power generation capacity increased by 40 times. There is a big change in the energy structure of power generation, thereby reducing the intensity of carbon dioxide emissions in the power industry. With the improvement of energy efficiency, energy consumption in the United States gradually decouples from economic growth. This also explains why the CR indices of all four states started to decline between 2008 and 2009. From the above analysis, we can conclude that long-term CR statistics not only reflect the clean energy consumption in the region, but also reflect the economic development trend of the area as a fairly valid and meaningful statistic.

In order to compare the states with the cleanest energy used in all four states in 2009, we can directly compare CR statistics for the four states in 2009. 24.539 In Arizona, 26.503 in California, 24.705 in New Mexico, and 26.155 in Texas. The smaller the CR statistics, the more energy-intensive the state is. Thus, in 2009, the cleanest state of energy consumption was New Mexico.

**6 Future Data Forecasting with four models**

To predict the energy consumption of all states between 2025 and 2050 and their cleanliness level of energy use. We combine four different approaches: time series, gray models, artificial neural networks and Cobb–Douglas production function to predict prices, productions, consumption and GDP respectively.

**6.1 Using ARIMA model to predict prices**

First, since the price of the previous year will have an impact on the price of the following year for energy sources such as oil and coal, that is, the price of energy lags, the theory of goods and prices in macroeconomics also revealed this rule. And this feature exactly coincides with the ARIMA model in time series, which is mainly used to fit a stationary time series. In a stationary series, the statistical properties of the series will not change over time. In addition, for any lag order k, the sequence autocorrelation does not change. Non-stationary time series can be transformed into stationary sequences through transformation and difference. To ensure that the variance is constant, the logarithm of the array to take a logarithmic or open square transform. Then the data is differentially transformed to remove the linear trend. Finally, we verify the stationary hypothesis by Augmented Dickey-Fuller statistical test using the adf.test () function in the tseries package in the R software.

In a p-order autoregressive model, every value in a sequence can be represented by a linear combination of p values before it：

Among them, is any time series of observations，is the mean of time series, is the weight， Random disturbance. In a q-order moving average model, each value in the time series can be represented by a linear combination of the previous q residues, namely:

Among them is the residual of the forecast，is the weight.

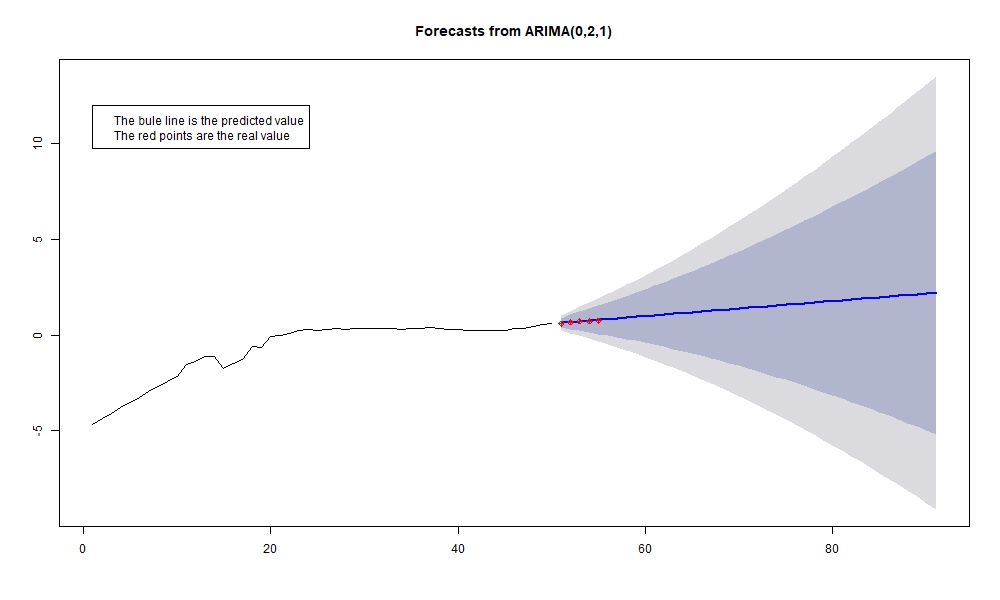
The mixture of these two methods is the ARMA (p, q) model, whose expression is:

At this point, each observation in the sequence is represented by a linear combination of p past observations and q residues.

The ARIMA (p, d, q) model implies that the time series is differentiated d times, and each observation in the series is represented by a linear combination of p past observations and q residues. The forecast is "error-free" or complete to make the final forecast.

The steps to create an ARIMA model include:

1. Make sure the time series is steady
2. Find a reasonable model (or several)
3. fitting model
4. Evaluate the model from statistical assumptions and forecast accuracy
5. Forecast

Next, let's take the coal price in Arizona as an example (Logarithmic transformations and second-order differences are performed in advance), to test the effectiveness of time series prediction:

*Figure 17: Results of coal price, Arizona*

Predict coal prices after the time series, the blue line is the forecast for the later prices, and the five red points are the known real prices [7]. The five points representing the actual prices fall near the forecast curve, proving that the forecasting curve is quite effective. We continue to take the time series ARIMA model to predict the latter part of the price of the remaining energy. However, we basically cannot pass the ADF test when forecasting the output and consumption, so we take other measures to make the forecast.

**6.2 Using Grey Theory to predict production**

To predict the production, the gray model is generally used [7].The following is a general introduction to the gray model:

In the formula, α, μ are the parameters to be identified, sequence is an accumulation of the original sequence which is:

,

Least-square principle is used to find the parameters to be identified ：

Among them,

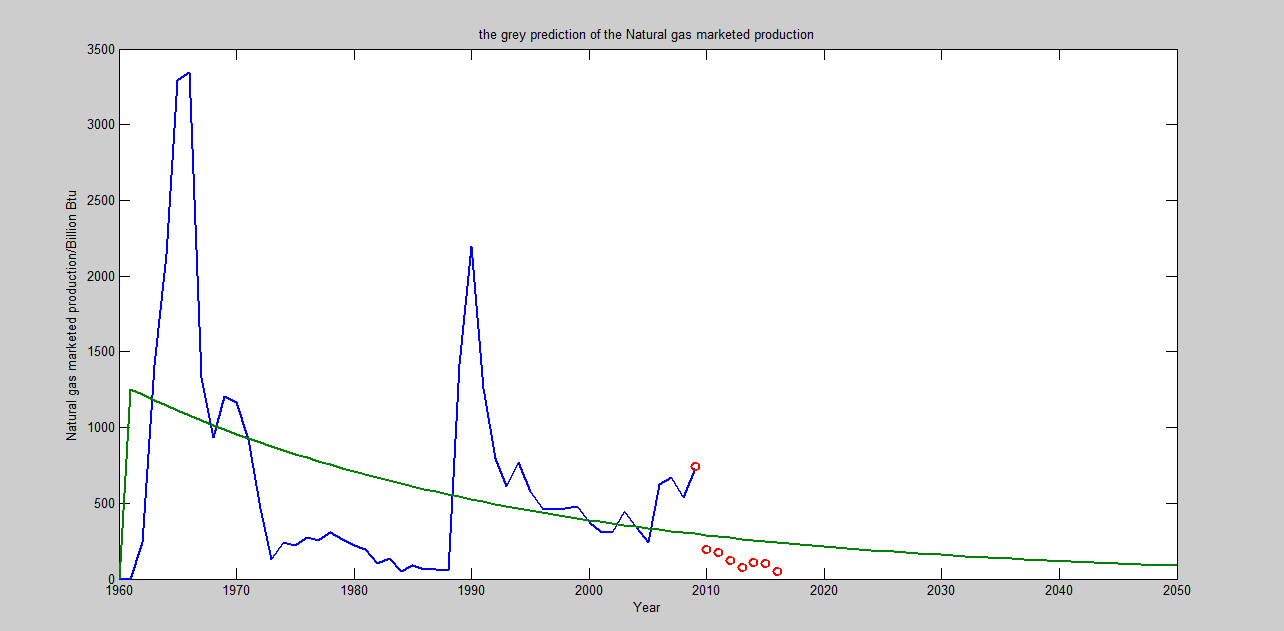
,

Differential equations resolution is：

The corresponding discrete equation is:

The method of posterior difference test model accuracy can be very good.

The following is an example of natural gas production, to illustrate the accuracy of gray theory prediction model in the production forecast.

Natural gas output predicted by the gray theory prediction model is as follows:

*Figure 18: The grey prediction of the Natural Gas Marketed Production*

The green curve represents the natural gas production predicted by the model, while the red point represents the actual natural gas production for the current year [9] it can be seen that these points are all distributed around the prediction curve, so the gray theory prediction model has a good prediction effect on energy yield forecasting.

**6.3 Using Neural Network Algorithm to Predict Consumption**

According to the study, the consumption of a certain energy can be predicted by its production, total energy consumption, the growth rate of GDP, and its price. So, we use the neural network algorithm to Predict Consumption.

The so-called neural network algorithm, namely artificial neural network is a kind of information processing system designed to imitate the structure and function of physiological neural network. Many artificial neurons are connected into a neural network with certain rules. The connections between neurons and the distribution of the weights of the connections are used to represent specific information. Neural network distributed storage of information, with high fault tolerance. Each neuron can calculate and process the received information independently and output the result. The network has the ability of parallel computing and real-time performance is very strong. Neurons are the basic computational units of a neural network, typically multiple inputs, one output nonlinear unit, and can have internal feedback and thresholds. Figure 1 is a complete neuron structure. Among them, X1，X2Xn is the input of neurons; V1, V2Vn is the strength of each connection with neurons, for example, weight,For the internal state of the feedback information; Wi is the threshold; f is the function of this neuron; Y is the output of neurons.

As shown in the figure is a three-layer artificial neural network model, we will be in this practical problem. . . Remember to enter, will. . . Recorded as output, the use of neural network algorithm for training. We find that the Artificial Neural Network can achieve good prediction results after continuous training.

Input layer Hidden Layer Output Layer

X1 Y1

X2

Y2

X3

*Figure 19: Artificial neural network schematic diagram*

**6.4 Using** **Cobb–Douglas Production Function to Predict GDP**

In economics and econometrics, the Cobb–Douglas production function is a particular functional form of the production function, widely used to represent the technological relationship between the amounts of two or more inputs (particularly physical capital and labor) and the amount of output that can be produced by those inputs. [11] Its form is：

where:

GDP = total production

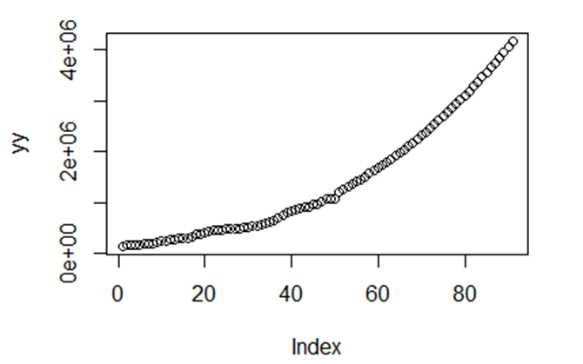
L = labor input (Here we use resident population instead)

K = capital input (Here we use real gross domestic product instead)

Take Texas's GDP forecast as an example. First, we obtain the previously estimated Texas population and total energy consumption by logarithm of the three:

R using the nls function，then the Rsquare is 0.9963. The result is below：

We can figure out that the fitted situation is perfect which explains the validity of the function.



*Figure 20: Texas's GDP forecast diagram using the**Cobb–Douglas Production Function*

**7 The Common Future: Adjusting Prices to Make the Four States Cleaner and More Renewable**

We hope that the adjustment will not only make each state cleaner, but also make them a closely united community and achieve the common-interest effect through cooperation measures. One solution is to minimize the sum of the CR stats for all four states. The simplest and most effective way to get the CR statistic is to transform all the traditional energy sources in each state into clean energy. The resulting CR statistic must be the smallest. However, such a program obviously does not meet the actual situation. First of all, the price of traditional energy sources such as coal is much lower than that of clean energy. Therefore, when considering options for clean energy and traditional energy, we should not blindly focus on clean energy. Second, in order to encourage businesses to use more clean energy, most governments use the same solution: Adopt the ladder price to the traditional energy. That is, when the demand and use of traditional energy sources is relatively small, the price of traditional energy sources is relatively low. However, as the use of traditional energy sources by companies exceeds a certain threshold, the government will raise the price of traditional energy sources to a level above that of clean energy sources and increase the usage of clean energy by enterprises. From here we can see that the price of energy will affect the amount of energy used. The price of energy and energy production are related.

Therefore, when we find the optimal solution that minimizes the CR statistic, we adopt the following idea: Discuss the feasibility of price adjustment by the government; then determine the relationship between energy demand and energy prices and determine the energy demand and the relationship between the supply of energy and the sum of the annual price best solution and the CR statistic. As an addition, we finally conduct a sensitivity analysis of the model by adjusting energy production.

**7.1 Feasibility of government’s control of price changes**

The above-mentioned price ladder, in all countries have specific examples, such as the United States, Japan, are implemented by connecting the ladder to control the power consumption of enterprises. Therefore, the price ladder can make the price fluctuations in a certain range adjustment. Similarly, the government can make the same kind of energy in different departments have different selling prices, such as the United States, China and some other countries, commercial and residential electricity prices are different. The government can use these measures to make the price fluctuate within a small range, but in general, in a market economy, the most important factor that determines the price is the size of the demand. Therefore, the government cannot adjust the price indefinitely and the scope of the adjustment is limited.

In summary, we set the annual price adjustment range, the kth state’s jth energy’s price in t year(t>2009). The forecast price is as described above , then the best price has an adjustable price range as

In other words, if a state is to increase the price of coal by 10% in the 10th year, it only needs to increase the price by 1% every year.

**7.2 The relationship between energy demand, energy output and energy prices**

The demand for energy decreases with the increase in price. Similarly, the output of energy in a region will determine the price of energy in the region. From the previous assumption, the energy supply in all four states is mainly from the total output of the four states, that is, the total demand EC of each energy source in the four states cannot exceed the total energy output EP of the four states：

(2)

The relationship between energy demand EC and energy price is , according to macroeconomic theory, we can easily have the following conclusions：

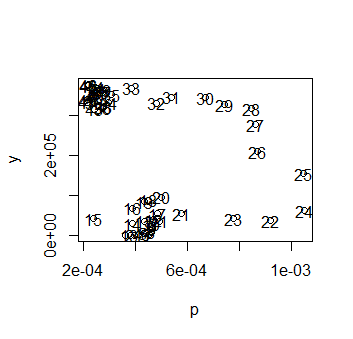
1. Price increases will make the demand decline, thus 。
2. When the price being very low, there is a great demand. At this moment, the price will increase rapidly, and the demand will decrease rapidly. At high prices, because other energy sources are substitutes for the energy source, the energy demand is very high Small, so at this time to further enhance the price, the demand will only reduce less. thus is a concave function.

Here we select the inverse proportion function as the fitting object, that is：

(3)

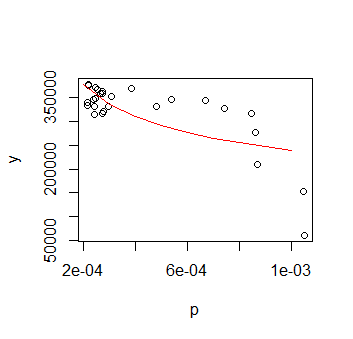
Among them, , Real energy price represents the true energy price of the year after eliminating the effects of inflation; Real dollar value represents of the U.S. dollar relative to the base year[12] ; and a, b are all parameters to be estimated. Since the price level is changing year by year, the value of one dollar fifty years ago is not the same as the value of one dollar today. Therefore, we must consider the impact of the inflation rate. Another reason for using this function is that the EC tends to be infinite when the real energy price is zero. In line with market rules.

In the actual fit, we find that there are two or more trend lines in many resources' relationship diagrams, such as the demand-price relationship of CA nuclear energy:



*Figure 21: Demand-Price Relationship Diagram for California Nuclear Energy*

The figures represent the corresponding years, such as 1 for 1960 and 50 for 2009. The reason for such a problem is likely to be that the state experienced a major increase in energy technology in a particular year, resulting in a significant reduction in costs, resulting in a rapid increase in demand, which is shown in the figure as a curve shift. Taken together, we think the earlier years are significantly different from the current ones, so considering the year similar to the current one will be more conducive to forecasting the future. Therefore, we discarded the first 23 years and then fitted it again. The results are as follows:



*Figure 22: Demand-Price Relationship Diagram for California Nuclear Energy,* *Fit renderings*

In the fitting process, we also encountered another problem: the previous year is in a position closer to the origin. This may be due to the underdeveloped energy technology of the state in the early days, so not only low prices and low yields but for the same reason as the previous one, we chose to discard earlier data for fitting.

In this way, we can represent the energy demand with its price and use its four-state total output as a constraint.

**7.3** **Linear Programming Operation and Interpretation of Results**

Calculate the results of the following linear programs in Lingo：

*St. ;*

*;*

*;*

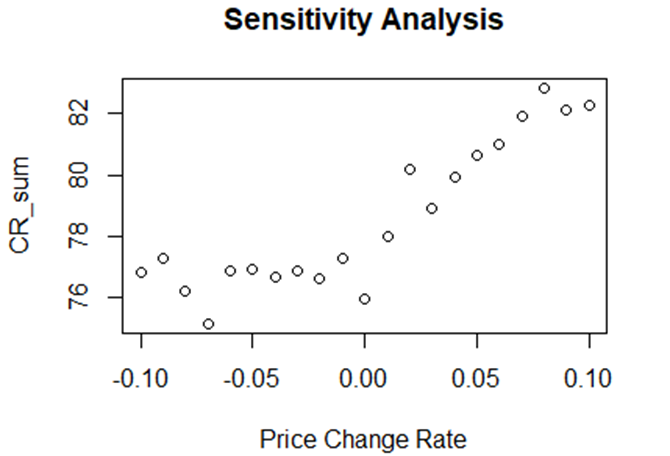
*;*

As can be seen from the results (attached to the appendix), when we adopt the global optimal solution given by LINGO software as the common use mode of four states ‘energy, the price of clean energy will not change or even decline, the price of traditional energy will rise, The sum of the CR statistic for each state decreases. Therefore, starting from the overall situation of the four states, as the economy continues to grow rapidly in subsequent years, its consumption of energy will surely continue to rise. The CR statistic will, theoretically, be year after year elevated. However, after adopting the solution we proposed, the CR statistic not only did not rise but declined, which shows that the rational allocation of energy can not only promote stable economic development, but also optimize the energy allocation structure and finally promote the economic and environmental The coordinated development. On the other hand, the change in the price of clean energy alone can make the price of clean energy unchanged or even decrease through the optimization of the energy structure. This shows that the actual price of clean energy has declined under the premise of inflation, which will attract more businesses and families to use clean energy. Likewise, under the optimal solution, the price of traditional energy will rise, which inhibits the use of traditional energy sources. So, to sum up, we can see that when we take the optimal solution given by us, the energy use in each state can still be cleaner under the premise of continuous economic development in each state.

**7.4 Sensitivity Analysis**

In the following, we determine the sensitivity of CR statistic to price changes by changing energy prices in comparison with the same period CR statistic.

Specifically, taking the 2025 data as an example, we can reduce the price of conventional energy by α% and increase the price of clean and renewable energy by α%. Then we bring it to the linear programming solution above to obtain the minimum CR, to see how the changes in market prices affect our CR values. The result is as follows:



*Figure 23: The price change rate*

It can be found from the figure that，in another word, clean and renewable energy As prices rise, the CR statistic becomes larger, indicating that on average, the four states become more polluting because the price increases for clean and renewable energy allow businesses and individuals to consume more conventional high-pollution energy sources. CR statistic is not only sensitive to the cleanliness of a state but also shows the degree of cleanliness under the cooperation of several states.

**Reference**

1. Akaike, H. (1974), "A new look at the statistical model identification", IEEE Transactions on Automatic Control, 19 (6): 716–723, doi:10.1109/TAC.1974.1100705, MR 0423716
2. Lingling Guo. Straw gasification energy saving and emission reduction safety technology [D]. Northeastern University,2011
3. 《Fossil fuel geology - oil, gas and coal》[J]. Marine oil and gas geology,1999(02):15
4. IPCC Fuel CO2, CO emission factor
5. International Association of Oil & Gas Producers <http://www.iogp.org/>
6. LAZARD Levelized Cost of Energy 2017 <https://www.lazard.com/perspective/levelized-cost-of-energy-2017/>
7. Wencai Wang,Tao Hou,Yidong Yang,Junfeng Yang. Coal Production Prediction and Comparative Analysis [J]. Modern mining,2012,27(03):5-6+9.
8. Electric Power Delivered Fuel Prices and Quality for Coal, Petroleum, Natural Gas Back to 1990, Arizona <https://www.eia.gov/state/search/#?1=101&3=21&a=true&2=184>
9. Arizona Natural Gas Production <https://www.eia.gov/dnav/ng/hist/n9050az2a.htm>
10. Jianliang Peng. Artificial Neural Network Method to Predict Coal Consumption [J]. Journal of China University of Mining and Technology,1997(03):65-69.
11. <https://en.wikipedia.org/wiki/Cobb%E2%80%93Douglas_production_function>
12. <https://zh.tradingeconomics.com/united-states/inflation-cpi>

Dear governor:

We are here to help you with evaluating the energy profile of your states and choosing proper contract for the four states. This letter will explain our CR evaluation system and our optimal advice to lowest the pollution for your states. Some colleges and universities recommended by our investment strategy will be pointed out at the end.

First of all, I will explain our CR(Cleaner and Renewable energy) evaluation system for you. The CR evaluation system is a kind of new system that we construct for the government to assess the pollution level of the four states. The CR statistic we constructed is the combination of the four-related value of the energy profile. The four statistics are the kind of energy, the price of the energy, the GDP of the states and the number of energy been used by the states. There is no need to write down the specific expressions of the CR statistics and the most important thing for the CR statistic is that the larger the value of CR statistic, the more polluted the sates is.

Second, I will show you the 2009 profile of the four states. It is obvious that Texas and California is the two which are more polluted than then other two states. Texas has the most production of shale gas among all the America which may explain the overuse of the traditional energy. What’s more, the California also consumes a large amount of fossil fuels that cause the CR value of it becomes relatively large. And the condition of New Mexico and Arizona is better than the former two states because the values of CR statistics are relatively smaller, which means the energy consumption of New Mexico and Arizona is cleaner.

Third, With the development of society, if there is no change in the energy structure, it is obvious that the pollution will be increasingly severe, so I will figure out a contract for your four states that can optimize your energy structure and minimize the emitted pollution of the four states. It is obvious that there exists a relationship between the energy consumption, energy price, and energy production, the government can change the consumption of each kind of energy by changing the price of it, so the target is to adjust the price properly to lowest the value of CR statistics of the four states by using the solution to linear programming problem. The result that can minimize the pollution of the is that set a threshold for the traditional energy. If the company uses more traditional energy than the threshold, the price of it will increase 7.33% and the price of the cleaner and renewable energy will decrease 2.82%.

By following the contract, the Arizona, New Mexico, California, Texas will all enjoy a bright future. I will explain it thoroughly to you. The CR statistic in 2050 of the four states is attached in the bottom of the letter when follow the contract. The results show that the value of CR statistic of the four states decrease in the following years which means the energy consumption has become less polluted even the energy consumption has become larger. So, the contact is perfect because it not only strikes a balance between the energy consumption and pollution but also optimate the energy structure.

So, we suggest sincerely that you follow the contact we put forward so that it can benefit a lot to the economy of your states and reduce pollution.

**Appendices**

**Appendix** **A CR statistic program**

Here is part of R simulation programs we used to build our CR model.

图片包含 屏幕截图

已生成极高可信度的说明

**Appendix** **B ARIMA model: Predict price**

Here is part of R simulation programs we used to predict price.

**图片包含 屏幕截图

已生成极高可信度的说明**

**Appendix** **C Grey Theory: Predict Production**

Here is part of MATLAB simulation programs we used to predict production.

1. %grey prediction
2. clc,clear;
3. syms a b;
4. c=[a b]';
5. production=[89677,99215,109655,120333,135823,159878,182321,209407,246619,300670];
6. B=cumsum(production);
7. n=length(production);
8. **for** i=1:(n-1)
9. C(i)=(B(i)+B(i+1))/2;
10. end
11. D= production;D(1)=[];
12. D=D';
13. E=[-C;ones(1,n-1)];
14. c=inv(E\*E')\*E\*D;
15. c=c';
16. a=c(1);b=c(2);
17. F=[];F(1)= production (1);
18. **for** i=2:(n+10)
19. F(i)=( production (1)-b/a)/exp(a\*(i-1))+b/a;
20. end
21. G=[];G(1)= production (1);
22. **for** i=2:(n+10)
23. G(i)=F(i)-F(i-1);
24. end
25. t1=1999:2008;
26. t2=1999:2018;
27. G
28. h=plot(t1,A,'o',t2,G,'-');
29. set(h,'LineWidth',1.5);

**Appendix D Cobb–Douglas Production Function: Predict GDP**

Here is part of R simulation programs we used to predict GDP.

图片包含 屏幕截图

已生成极高可信度的说明

**Appendix** **E Linear Programming Operation**

model:

sets:

qr/1..4/;

qt/1..17/;

se(qr,qt):ep,a,b,score,m,n,t,c,d;

!n represents the quantity demanded of the energy;

!m represents the actual price of energy;

!t represents a coefficient of the energy;

!c represents the emissions of the carbon dioxide;

!d represents the emissions of the carbon monoxide;

pr/1..4/:Y,rou;

endsets

data:

!Y represents the growth rate of GDP;

!rou represents Energy expenditures as share of current-dollar GDP;

Y=;

rou=;

a=;

b=;

score=;

ep=;

c=;

d=;

enddata

min=@log(@sum(se(i,j)|i#ge#1#and#i#le#1:(t(i,j)\*n(i,j))))/(0.75+0.25\*Y(1)\*rou(1)/100000000)+@log(@sum(se(i,j)|i#ge#2#and#i#le#2:(t(i,j)\*n(i,j))))/(0.75+0.25\*Y(2)\*rou(2)/100000000)+@log(@sum(se(i,j)|i#ge#3#and#i#le#3:(t(i,j)\*n(i,j))))/(0.75+0.25\*Y(3)\*rou(3)/100000000)+@log(@sum(se(i,j)|i#ge#4#and#i#le#4:(t(i,j)\*n(i,j))))/(0.75+0.25\*Y(4)\*rou(4)/100000000);

@sum(se(i,j):n(i,j))<15002328.4;

@for(se(i,j):ep(i,j)\*0.84<m(i,j));

@for(se(i,j):m(i,j)<1.16\*ep(i,j));

@for(se(i,j):n(i,j)=(a(i,j)/m(i,j)^b(i,j)));

@for(se(i,j):t(i,j)=c(i,j)\*10\*d(i,j)\*ep(i,j)/score(i,j));

end