Team Number :	apmcm2208959
Problem Chosen :	E

How Many Nuclear Bombs can Destroy the Earth?

The explosion of the first atomic bomb in the United States marked the beginning of the nuclear age in 1945. Such explosions and radiation can severely damage the ecological environment. Therefore, how to correctly understand and reasonably predict the development trend of nuclear weapons and discuss related issues is particularly important. In this paper, Analytic hierarchy process and Cluster analysis are used, time series prediction model and Neural network model are constructed obtain the quantitative values of future nuclear weapon possessors, possession and deterrence by Python and Excel.

For question 1, data processing results shows that 10 countries have had nuclear weapons, namely China, France, India, Israel, North Korea, Pakistan, Russia, United Kingdom, South Africa, United States. The stockpile of nuclear weapons in 2002 and 2022 is subtracted and sorted to find that the United States has the largest stockpile. Moreover, the number of nuclear weapon tests are traversed in five years, from 1962 to 1966 ,the number of 422.0 times was the highest. Based on the Analytic hierarchy process, the evaluation matrix is set with the increase and decrease of nuclear weapons and the number of test launches, and North Korea is regarded as the most active country. Finally, the time difference of each country was calculated and ranked, and the United States was the fastest country to change from "no weapons" to "possessing nuclear weapons".

For question 2,the time series prediction mathematical model, neural network training and Bayesian network are adopted, it is inferred that Iran will possess nuclear weapons in the next 100 years. The total of nuclear weapons are 5,979 in 2123.

For question 3, in question a, from the two aspects of destroying the earth surface and disintegrating the earth. Reasonable explosion radius is obtained by Simulated test explosion, and it is calculated that 1.51 million Czar bombs could destroy the earth surface. In question b, physical formula is used to deduce that the number of nuclear bombs that will collapse the Earth is about 1,075 trillion Czar bombs. The maximum destructive power nuclear bombs currently possess may be 50Mt, which is not enough to destroy the Earth. Finally, cities in the world are divided into 5 categories by Cluster analysis, and the number of nuclear bombs could destroy each type of city was obtained. Finally, it is concluded that the total number of nuclear bombs in the world may be limited to 2,948, and each country could have a maximum of 328 nuclear weapons.

For question 4, findings in this paper are described to the United Nations, and advice on testing, use and management of nuclear weapons is offered to all countries.

Key words: Analytic hierarchy process; Time series model; Neural network model; Cluster analysis; Bayesian Networks.

Contents

1 Introduction

1.1 Background

Nuclear weapons are very powerful, they can release a huge amount of energy in an instant, its maximum power is 100 million tons. The Czar Bomb is by far the most powerful nuclear bomb in the world. It is also the most powerful nuclear weapon ever built. In today's world, peaceful development is the theme of The Times, so how to correctly understand and reasonably predict the development trend of nuclear weapons and discuss its related issues has become the focus of attention.

1.2 Problem Requirements

Nuclear weapons have been a focus of international relations since the end of the Second World War. Over the past few years, various countries have conducted a large number of nuclear weapon tests in different fields. this paper mainly solves the following problems:

Question 1: a) Analyze which countries in a given list of 10 have ever possessed nuclear weapons. b) Which country's nuclear weapons stockpile changed the most between 2002 and 2022. c) The five years with the most nuclear weapons tests. d) Analyze which country has been the most active in nuclear weapons research in the last 10 years, based on the available data. e) From which country, out of a given list of 10, made the fastest transition from "not considering nuclear weapons" to "possessing nuclear weapons".

Question 2: a) Build mathematical models to predict the number of nuclear weapons and predict which countries will possess nuclear weapons in the next 100 years. b) Forecast the number of nuclear weapons over the next 100 years, the total number of nuclear weapons in 2123, and the number of nuclear weapons in each country.

Question 3: a) Collect relevant data, select appropriate indicators, build mathematical models, analyze and calculate at least how many nuclear bombs would be needed to destroy the Earth. b) Analyze the maximum destructive power of today's nuclear bombs and determine whether that power could destroy the Earth. c) Discuss what limits should be placed on the global nuclear Arsenal and determine the safe number of nuclear weapons for each nuclear-armed state.

Question 4: Prepare a non-technical article explaining the team's findings and making several recommendations for all countries.

2 Assumptions

(1) It is assumed that the predicted number of nuclear weapons of each country is only considered under objective circumstances and force majeure factors are ignored.

(2) Assuming that the explosion radius of the atomic bomb is determined, the explosion of the atomic bomb to the ground is approximately a circle, and the explosion range of the atomic bomb is the same in all countries.

(3) Consider only the effects of nuclear bombs on the Earth, ignoring the interference of other systems.

(4) Assuming that the data provided by relevant data websites are real and reliable, this study is based on real data.

3 Symbol Description

Symbols	Definitions
λ	Eigenvalue of maximum
CI	Coincident indicator
RI	Random consistency index
CR	Verification coefficient
E	Work
G	Gravitational constant
m	Earth mass
r	Earth radius
x_{ij}	The value of the ith row, j column matrix
y_{ij}	The value of the matrix in row i and column j after normalization

4 Problem Analysis

Before the problem analysis, we conducted data preprocessing and divided the official data into four Excel tables, i.e. proliferation.xlsx, position.xlsx, Phys. xlsx, and tests.xlsx.

4.1 Basic Data Analysis

a) In this problem, it is necessary to combine existing data, namely "possession", and use Python and Excel for discrimination.

b）Introduce a change index as the difference between the number of nuclear weapons in stockpile in 2022 and the number in stockpile in 2002. The difference is the change in each country's nuclear stockpile over the past 20 years. Find, in descending order, the countries that have reduced or increased their nuclear stockpiles the most over the past 20 years and chart them.

c）Use the Python dictionary to match the year with the number of experiments. The total value of nuclear weapons is calculated by traversing the years in five years. The iterative algorithm determines the size successively and assigns the maximum value to the larger value of each comparison, so as to obtain the five years with the largest number of experiments.

d）In this question, the method of hierarchical analysis is adopted. Firstly, the input discriminant matrix is constructed according to practical experience. The eigenvector and maximum eigenvalue need to be calculated, and the input discriminant matrix constructed is analyzed by using python's numpy library to return the maximum eigenvalue and its corresponding eigenvector. Find the maximum eigenvalue and carry out consistency test. After the input data, the data is normalized, the eigenvector of the discriminant matrix is used to calculate the weight of each index, and finally the result is obtained by matrix multiplication.

e）Use Python to count the year when 0 and 3 first appeared in each country, put them into Excel, and then use Excel to calculate the year difference. The year difference is sorted in ascending order, the first place is the country, and then visualization is performed.

4.2 Predict the Number of Nuclear Weapons

a) In this problem, the time series prediction method is applied to the model predicting the number of nuclear weapons. The model of predicting the number of nuclear-weapon states is fitted with Bayesian Networks to get the number of nuclear-weapon states in the next 100 years, and then the specific countries are deduced according to relevant literature.

b) Predict the trend of the number of nuclear weapons in the next 100 years and the total number of nuclear weapons in 2123 according to the model of question a). **Bayesian neural network** was used to fit the model predicting the number of nuclear weapons of each country. The known data, the forecast data obtained in question II (a), the data of a country's nuclear weapons in the past and the trend of the country's nuclear weapons in the global nuclear weapons were included into the input parameters, and the neural network training was carried out to obtain the relevant model. After the training is completed, the number of nuclear weapons of each country is predicted, and the visual results are output at last.

4.3 Protect Our Earth

a) Because of the study of how many nukes at least are needed, we assume that this nuke is the most powerful Czar bomb available. And there are two dimensions to this problem. One is the minimum number of nuclear bombs needed to destroy the surface of the earth, and the other is the minimum number of nuclear bombs needed to completely disintegrate the earth. For the first level, we use analysis software to test the explosion, select three representative sites for simulation, get a reasonable explosion radius, and calculate the minimum number of nuclear bombs needed to destroy the surface of the earth according to the explosion area of the nuclear bomb is greater than or equal to the area of the earth. On the second level, we use physical **formulas** and collected data to determine at least how many nuclear bombs it would take to completely disintegrate the planet.

b) Based on the data from the relevant models and simulations, we can calculate the maximum destructive power of the nuclear bomb at present, and then combine the current data to determine whether it can destroy the Earth.

c) In this problem, firstly, we find relevant population information and conduct Cluster analysis to classify five types of cities. Then through the test explosion software analysis of the destruction of various types of cities need how many nuclear weapons, calculate the total number of nuclear weapons, and finally summed up, equal to each nuclear state, is the minimum number of nuclear weapons that each country should maintain.

5 The Solution to Problem Number One

5.1 a）

Possession index is used to measure whether a country has ever possessed nuclear weapons. With the help of python, the "possession" index of each country for each year is iterated and tested. If the value is greater than 1, it indicates that a country has ever possessed nuclear weapons.

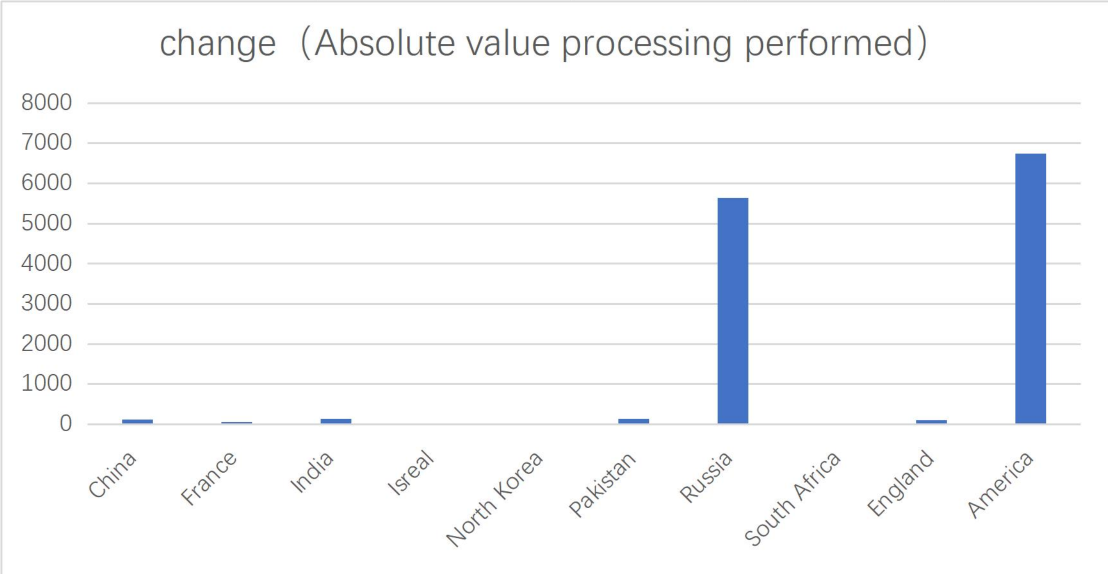
By crunching the numbers, we find out which countries have had nuclear bombs: China, France, India, Israel, North Korea, Pakistan, Russia, United Kingdom, South Africa, United States.

5.2 b）

Python was used to make a statistical analysis of the changes in the nuclear weapons reserves of each country in 2002 and 2022, and Excel was used to calculate the difference between the nuclear weapons reserves in 2022 and 2002, and finally a descending order was made, as shown in Table 1.

Table 1 :Changes in nuclear weapons stockpile

Country	Abbreviation	Year	Stockpile	Change
Pakistan	PAK	2022	165	139
India	IND	2022	160	137
China	CHN	2022	350	115
North Korea	PRK	2022	20	20
Israel	ISR	2022	90	14
South Africa	ZAF	2022	0	0
France	FRA	2022	290	-60
United Kingdom	GBR	2022	180	-100
Russia	RUS	2022	4477	-5637
United States	USA	2022	3708	-6749



Through data screening, calculation and sorting, it is found that the United States has the largest reduction in nuclear weapons reserves, and visual data display is carried out.

5.3 c）

In this case, you just need to use Python to go through five years, starting in 1945, and count the total number of nuclear weapons in each five years. After that, the five years with the largest number of experiments were obtained through sequential comparison, and visualization processing was carried out, as shown in Figure 1.

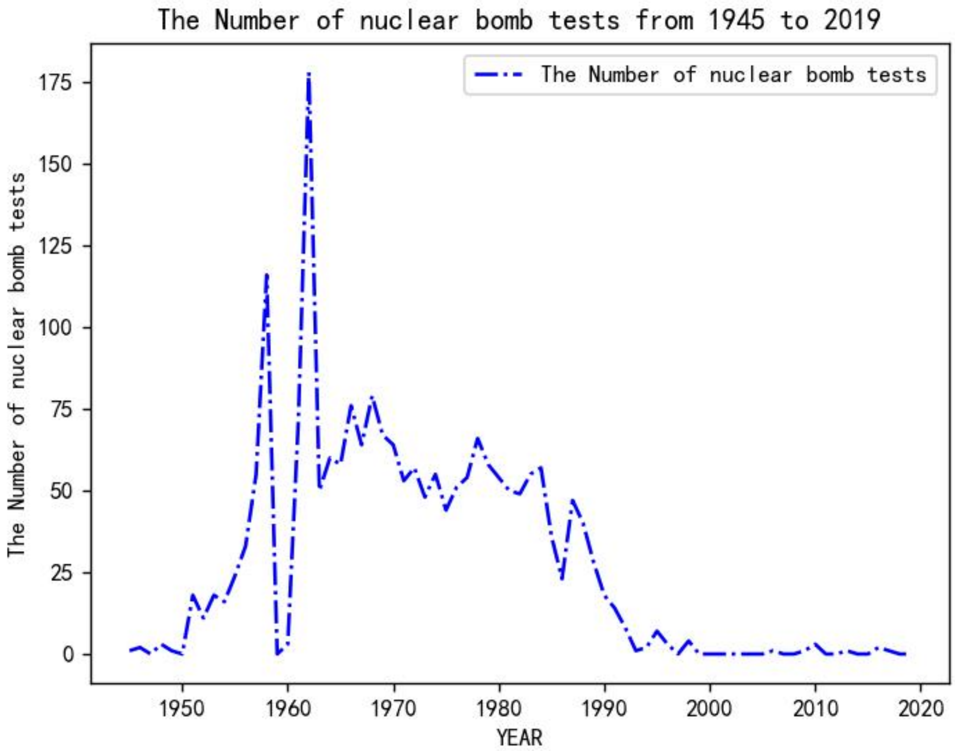


Figure 1: Number of nuclear tests

By using the iterative method, the maximum number of nuclear bomb tests in the five years from 1962 to 1966 was 422.0.

5.4 d）

In theory, the positive variables that can account for a country's nuclear tests are the number of nuclear weapons grown and the number of nuclear weapons tested. The number of nuclear weapons tests is an important indicator for a country to conduct nuclear weapons research, while the increase of nuclear weapons is an additional explanatory condition. Therefore, we choose the number of nuclear tests as an important factor in the country's nuclear weapons research activity and take the increase number of nuclear weapons into consideration. Finally, the degree of nuclear weapons research activity is scored through the Analytic hierarchy process.

This problem is solved by Analytic hierarchy process. According to the nature of the problem and the overall goal to be achieved, the Analytic hierarchy process (AHP) divides the problem into different component factors, and combines the factors at different levels according to the interrelated influence and subordination of the factors, forming a multi-level analysis structure model.

Firstly, data preprocessing was carried out, and the number of nuclear weapons of each country in 2012 and 2022 was obtained through Excel screening and sorting, and the number of experiments was counted to obtain relevant data.

Then an evaluation matrix is set, assuming that the evaluation matrix is

[[1,1/5],[5,1]], which indicates that the number of test launches of nuclear weapons is significantly more important than the number of nuclear weapons added. Then, the np.linalg.eig function in Python is used to calculate the eigenvalues and eigenvectors, and the row and column where the maximum eigenvalues are located and their corresponding eigenvectors are found, and the eigenvectors are treated as weight vectors. Finally, consistency test was carried out, and the consistency index was calculated by CI. The smaller the CI, the greater the consistency. The consistency index was defined as:

$$\mathrm{CI}=\frac{\lambda-\mathrm{n}}{n-1}$$

CI=0, there is complete consistency; CI is close to 0 and has satisfactory consistency; The larger the CI, the more serious the inconsistency.

To measure the size of CI, the random consistency index RI is introduced:

$$\mathrm{RI}=\frac{CI_1+CI_2+\ldots+CI_n}{n}$$

Considering that the deviation of consistency may be caused by random reasons, when verifying whether the judgment matrix has satisfactory consistency, CI and random consistency index RI need to be compared to obtain the test coefficient CR, the formula is as follows:

$$\mathrm{CR}=\frac{CI}{RI}$$

Generally, if CR<0.1, it is considered that the judgment matrix passes the consistency test; otherwise, it does not have satisfactory consistency. The test shows that the evaluation matrix is available.

Read data for processing, the first data normalization, so as to reduce the tempering.

$$y_{ij}=\frac{x_{ij}-\min\limits_{1\le i\le m,1\le j\le n} x_{ij}}{\max\limits_{1\le i\le m,1\le j\le n} x_{ij}-\min\limits_{1\le i\le m,1\le j\le n} x_{ij}}$$

Then the weight vector is used to calculate the evaluation between the two. The final results are shown in Figure 2.

```
0.0
consistency test passed
The scores of number1=0.8333333333333334
The scores of number2=0.7553910106521174
The scores of number3=0.8008573655494934
The scores of number4=0.7683813977656534
The scores of number5=0.9415432579890881
The scores of number6=0.7976097687711093
The scores of number7=0.5845674201091193
The scores of number8=0.7326578332034295
The scores of number9=0.0
```

Figure 2 ： Score of nuclear weapon development initiative of nuclear weapon states

Based on object scores, we know that North Korea, China and India are the top three active nuclear weapons developers. At the same time. At the same time, the relevant data are summarized to form Table 2. During the decade, only North Korea conducted nuclear tests, which further verified our results.

Table 2: Growth and decline of nuclear States and number of test launches

Country	Abbreviation	Number of nuclear bombs in possession in 2012	The number of nuclear bombs in possession by 2022	Object number	Nuke increment	The number of experiments in a decade
China	CHN	240	350	1	110	0
France	FRA	300	290	2	-10	0
India	IND	100	160	3	60	0
Israel	ISR	80	90	4	10	0
North Korea	PRK	0	20	5	20	4
Pakistan	PAK	110	165	6	55	0
Russia	RUS	4750	4477	7	-273	0
United Kingdom	GBR	225	180	8	-45	0
United States	USA	4881	3708	9	-1173	0

According to Table 2 and hierarchical analysis, North Korea is the most active in nuclear weapons research.

5.5 e）

Using Python, the status in position.xlsx was traversed to find out the time period in which 0 and 3 appeared for the first time in each country, and then subtracted each other. After that, the graph was derived and visualized.

Table 3 :Time spent in nuclear countries

country	consuming time	Year of acquisition
United States	7	1945
Russia	11	1949
United Kingdom	14	1952
France	22	1960
China	26	1964
Israel	29	1967
South Africa	41	1979

India	49	1987
Pakistan	49	1987
North Korea	68	2006

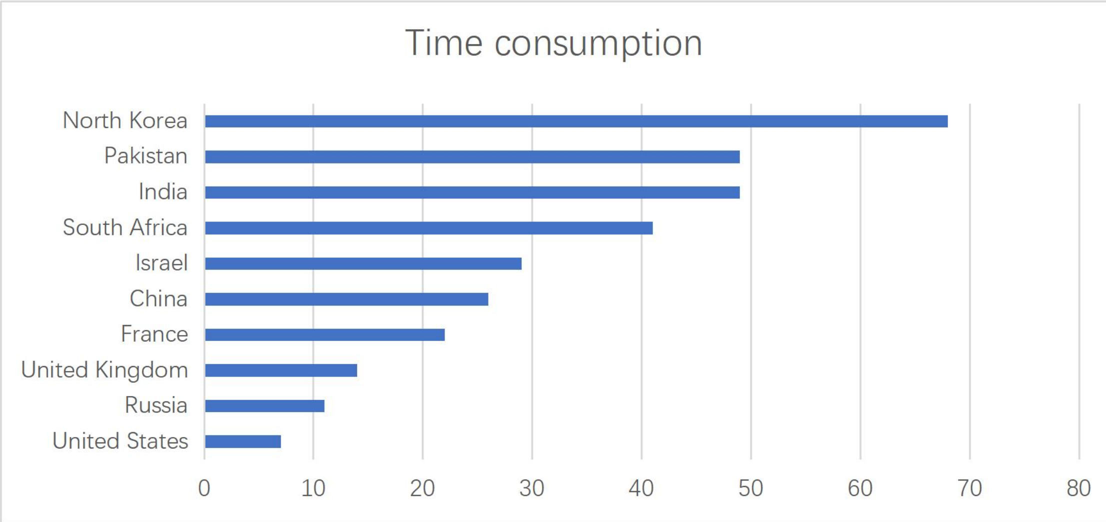


Figure 3 :Shows the time spent by nuclear countries in possessing nuclear weapons

It can be seen from Table 3 that the United States has the fastest change from "no weapons considered" to "nuclear testing, possessing nuclear weapons".

6 The Solution to Problem Number Two

6.1 Building Models to Predict the Number of Nuclear Weapons

The time series analysis model is established according to the following ideas. First, ADF test was conducted to analyze the p value, and to analyze whether it could significantly reject the hypothesis of sequence instability ($P<0.05$). If it is significant ($P<0.05$), it means that the null hypothesis is rejected and the series is a stationary time series; otherwise, it means that the series is an unstable time series. ADF test results are shown in Table 4.

Table 4:ADF test list

ADF test list							
Variable	difference order	t	P	AIC	critical value		
					1%	5%	10%
Stockpile	0	-2.164	0.220	1102.297	-3.519	-2.9	-2.587
	1	-2.139	0.229	1088.27	-3.519	-2.9	-2.587

2	-6.751	0.000***	1068.378	-3.522	-2.901	-2.588

Note: ***, ** and * represent the significance level of 1%, 5% and 10% respectively

The result of the sequence test shows that, based on the variable Stockpile, when the difference is second order, the significance P value is 0.000***, showing significance at the level, rejecting the null hypothesis, which indicates that the sequence is a stationary time series. The optimal difference sequence diagram is shown in Figure 4.

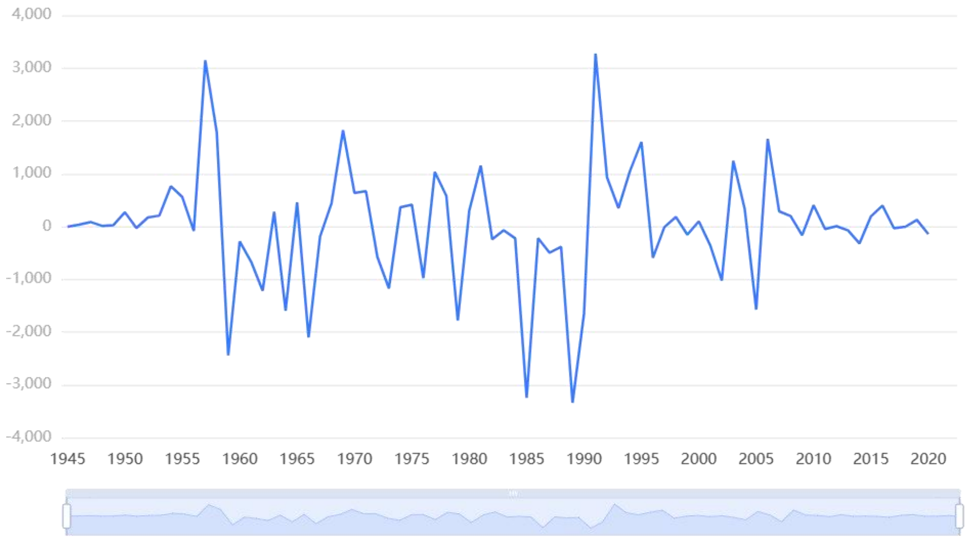


Figure 4: Optimal difference sequence diagram

Check the data comparison graph before and after the difference to determine whether it is smooth. At the same time, conduct partial (self) correlation analysis on the time series, and estimate its p and q values according to the truncated situation. The autocorrelation diagram of final difference data is shown in Figure 5, and the partial autocorrelation diagram of final difference data is shown in Figure 6.

Figure 5:Autocorrelation diagram of final difference data

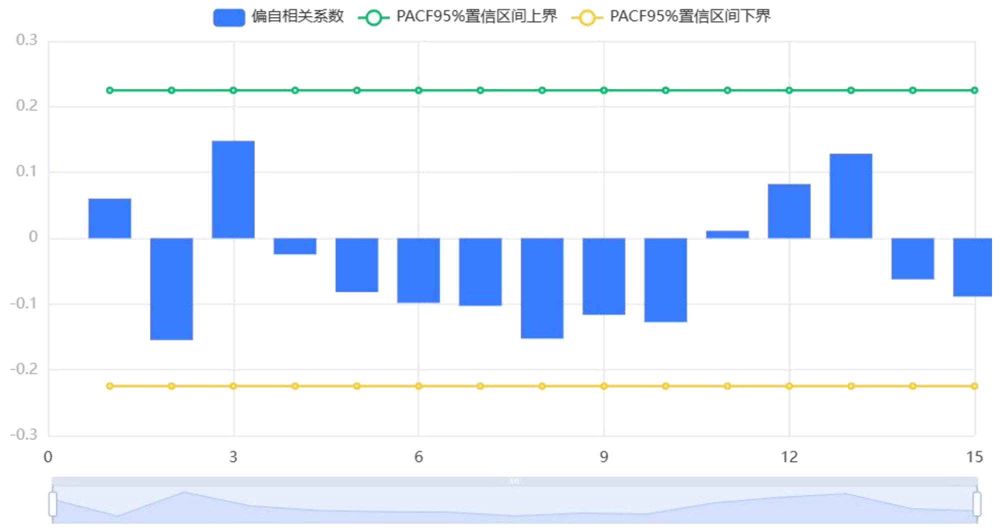


Figure 6: Partial autocorrelation diagram of final difference data

Figure 5 shows the autocorrelation graph (ACF), including the coefficients, upper and lower confidence limits. The horizontal axis represents the number of delays and the vertical axis represents the autocorrelation coefficient. Figure 6 shows the partial autocorrelation graph (PACF), including the coefficients, upper and lower confidence limits. The horizontal axis represents the number of delays and the vertical axis represents the partial autocorrelation coefficient.

It can be seen from FIG. 5 and FIG. 6 that both are trailing, and the original data is a stable time series after difference, so it is basically suitable for ARIMA.

The ARIMA model requires the model to have pure randomness, that is, the model residual is white noise. Check the model test table and analyze the model residual ACF/PACF diagram. According to the model parameter table, the model formula is obtained and the time series analysis diagram is comprehensively analyzed to obtain the order result of backward prediction.

Table 5: Test table of ARIMA model (0,2,0)

ARIMA model (0,2,0) test list		
Item	sign	value
he number of samples is	Df Residuals	75
	N	78
Q statisti	Q6(P value)	0.287(0.592)
	Q12(P value)	5.344(0.501)
	Q18(P value)	11.251(0.508)
	Q24(P value)	15.963(0.595)
	Q30(P value)	25.005(0.405)
Information Standard	AIC	1284.436
	BIC	1289.097
Information Standard	R^2	0.979

Note: ***, ** and * represent the significance level of 1%, 5% and 10% respectively

The system automatically finds optimal parameters based on AIC information criteria, and the model result is the test list of ARIMA model (0,2,0). Based on variable: Stockpile, the result analysis of Q statistic can be obtained: Q6 does not show significance at the level, and the hypothesis that the residual error of the model is white noise sequence cannot be rejected. Meanwhile, the goodness of fit R^2 of the model is 0.979, which indicates that the model performs well and basically meets the requirements.

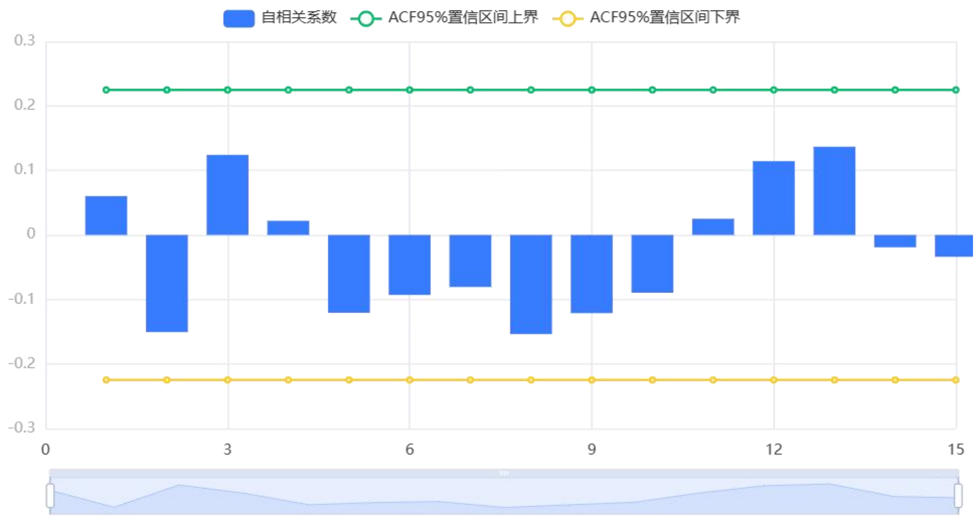


Figure 7: Residual autocorrelation diagram

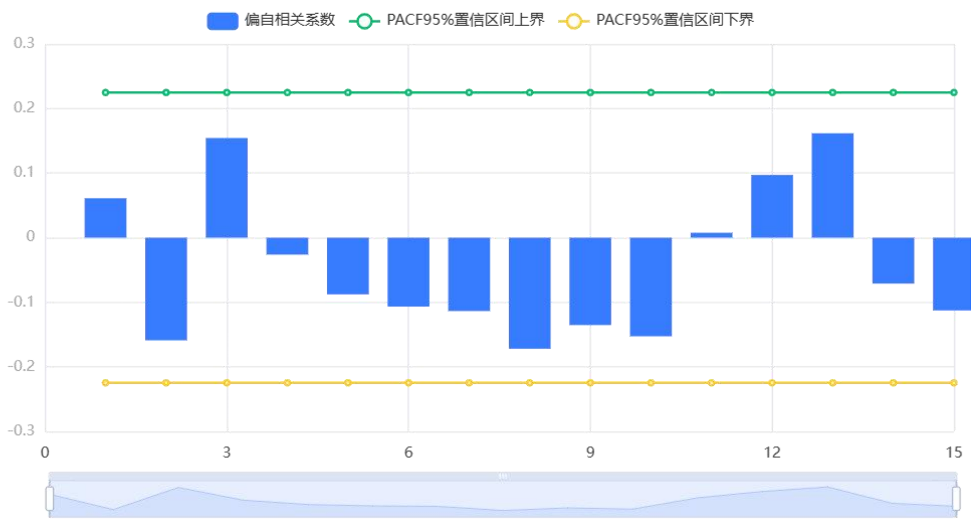


Figure 8 Residual partial autocorrelation

Figure 7 shows the residual autocorrelation (ACF) of the model, including coefficients, upper and lower confidence limits. The horizontal axis represents the number of delays and the vertical axis represents the autocorrelation coefficient. Figure 8 shows the residual partial autocorrelation (PACF) of the model, including coefficients, upper and lower confidence limits. The horizontal axis represents the number of delays and the vertical axis represents the partial autocorrelation coefficient.

It can be seen from Figure 7 and Figure 8 that the correlation coefficients are both within the upper and lower limits of the confidence region and are white noise sequences, which can be directly used for time series analysis.

Table 6 :Model parameters table

Model parameter table						
	The coefficient of	standard deviation	t	P>|t|	0.025	0.975
The constant is	-0.329	126.421	-0.003	0.998	-248.109	247.451

Note: ***, ** and * represent the significance level of 1%, 5% and 10% respectively

Table 6 shows the parameter results of this model, including the coefficient, standard deviation and T-test results of the model, which are used to analyze the model formula.

Based on the variable Stockpile, the system automatically finds optimal parameters based on AIC information criteria. The model result is the ARIMA model (0,2,0) test table based on 2 difference data. The model formula is as follows：

$$y(t)=-0.329$$

Through the above analysis, the fitting value and predicted value of the model are visually displayed, as shown in Figure 9.

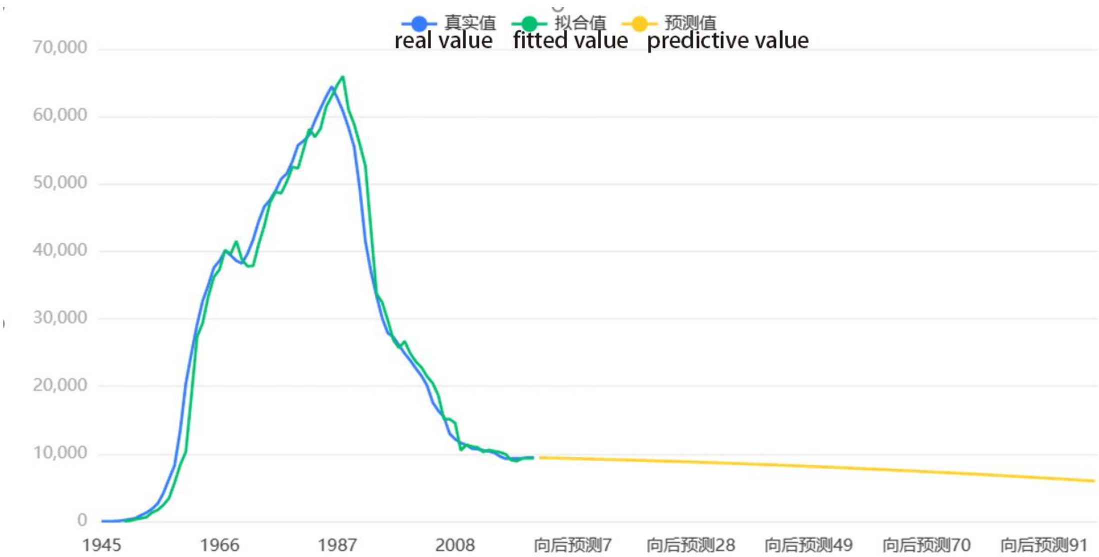


Figure 9 :Fitting and prediction of the Time series model

According to the prediction model, we can know the specific data. Due to the length of the data, the total number of nuclear weapons in 2123 will reach 5978.815789473637, which is about 5979.

6.2 Predict the Countries with Nuclear Weapons in the Next 100 Years

6.2.1 The Development of Models to Predict the Number of States with Nuclear Weapons

The specific algorithm steps are as follows:

1. Data should be processed first. Converts the given table to an array of cells and writes all the contents of Year into Possession to the new list.

2. Secondly, conduct model training. The default lmh function has a poor effect on training time series. Bayes regularization algorithm is adopted, which takes more time but has higher accuracy. A nonlinear self-combining temporal neural network was established, and the preparents function was used to initialize the data, and the data transformed by preparents was trained. 70% of the input vector and target vector were used for training, 15% for preventing overfitting, and 15% for testing, to test the error between the simulated results and the actual results. Then save the trained model.

3. Then make predictions. Read into the model, predict the past values for fitting, try the fitting effect. Then put the known data into the model in turn, iterate to predict the follow-up, and repeat the previous steps.

4. Finally, the forecast result is displayed. Visualize the resulting data, as shown in Figure 10。

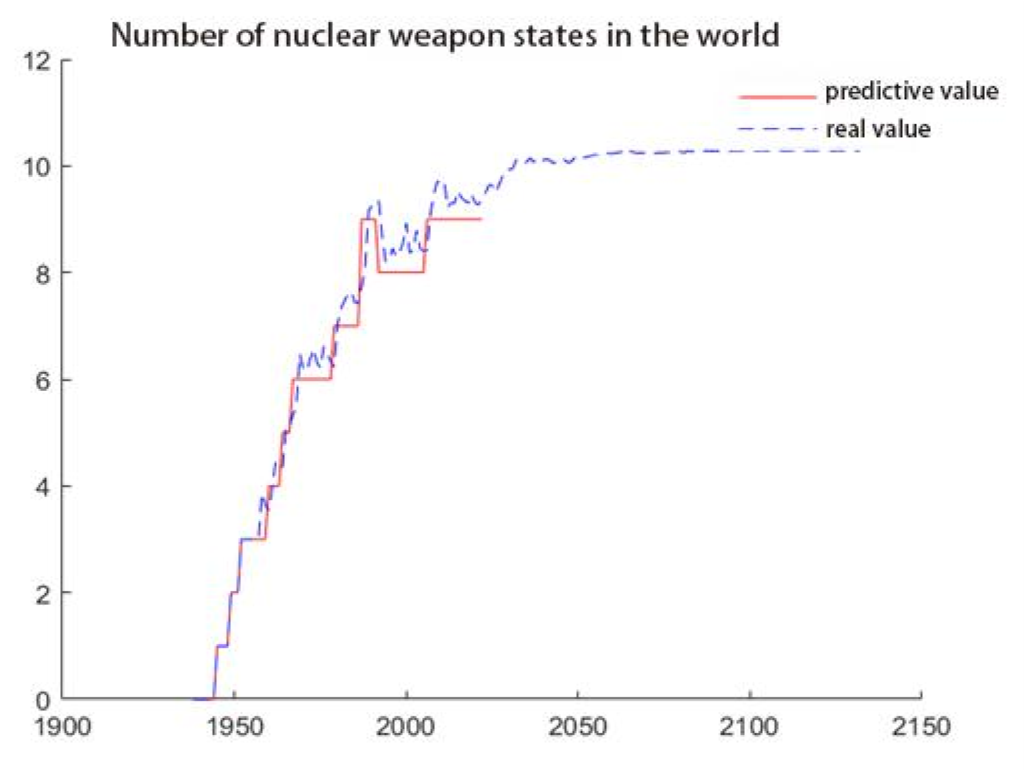


Figure 10: The number of countries possessing nuclear weapons worldwide

The model predicts that 10 countries will have nuclear weapons in the next 100 years. In the data given, 9 countries are known to possess nuclear weapons, so a new country needs to be analyzed in combination with relevant literature。

6.2.2 Infer Countries from the Literature

Through the literature, we believe that the next 100 years of nuclear weapons is likely to be Iran. According to the documentation, Iran's "Green Salt Project", high explosives testing and "Meteor program" since 2006 May involve the development of nuclear weapons. As of May 2015, a total of 19,138 centrifuges were installed at Iran's two fuel enrichment plants at Natanz and Fordow, of which 9,156 IR-1 centrifuges were in operation and the rest had been installed and were awaiting operation. In January 2021, Iranian government spokesman Ali Rabeei announced that Iran had begun to enrich uranium to 20%. Although Iran has a long way to go in terms of technology, there is a good chance that it will become a nuclear power within the next 100 years.

Therefore, based on the above predictions and known data, we believe that the ten countries with nuclear weapons in the next 100 years include Iran, the United States, Russia, the United Kingdom, China, France, India, North Korea, Pakistan and Israel.

6.3 Predict the Number of Nuclear Weapons in Each Country

Given that the number of future nuclear weapons changes in each country cannot simply be considered to be the same as the total number of nuclear weapons in the world (but the trend is roughly the same). In addition, due to the international situation and world background (the trend of decreasing nuclear weapons year by year, the mainstream status of world peace and development, and the constraints of the UN Charter and other factors), we believe that the data of a country's past possession of nuclear weapons, the overall number of nuclear weapons in the world and the trend of the proportion of nuclear weapons in the country should be included into the input parameters of the neural network. To make country-specific forecasts.

The specific algorithm steps are as follows:

1. Initialize data first. Convert the country label to a numeric number and turn it into a categorical array of type, that is, a categorical array. Then it counts the number of countries, and counts the number of nuclear weapons in each country and the total number of nuclear weapons in the world.

2. Conduct model training. The country's previous nuclear weapons possession, the global total number of nuclear weapons and the proportion of nuclear weapons in the country are taken into account. Given the input parameters, the above results of predicting the global number of nuclear weapons are retained. The Bayesian regularization algorithm is adopted to train each country, and the corresponding training model is retained.

3. Predict the number of nuclear weapons in each country. Read the training

model, then store the neural network for each country, read the total number of nuclear weapons for each country, and use the neural network to predict the number of nuclear weapons for each country in the future.

4. Visual presentation. Matlab is used to visualize the number of nuclear possession predicted by each country. See Figure 11 below.

The Blue one is predictive value, the Red one is real value。

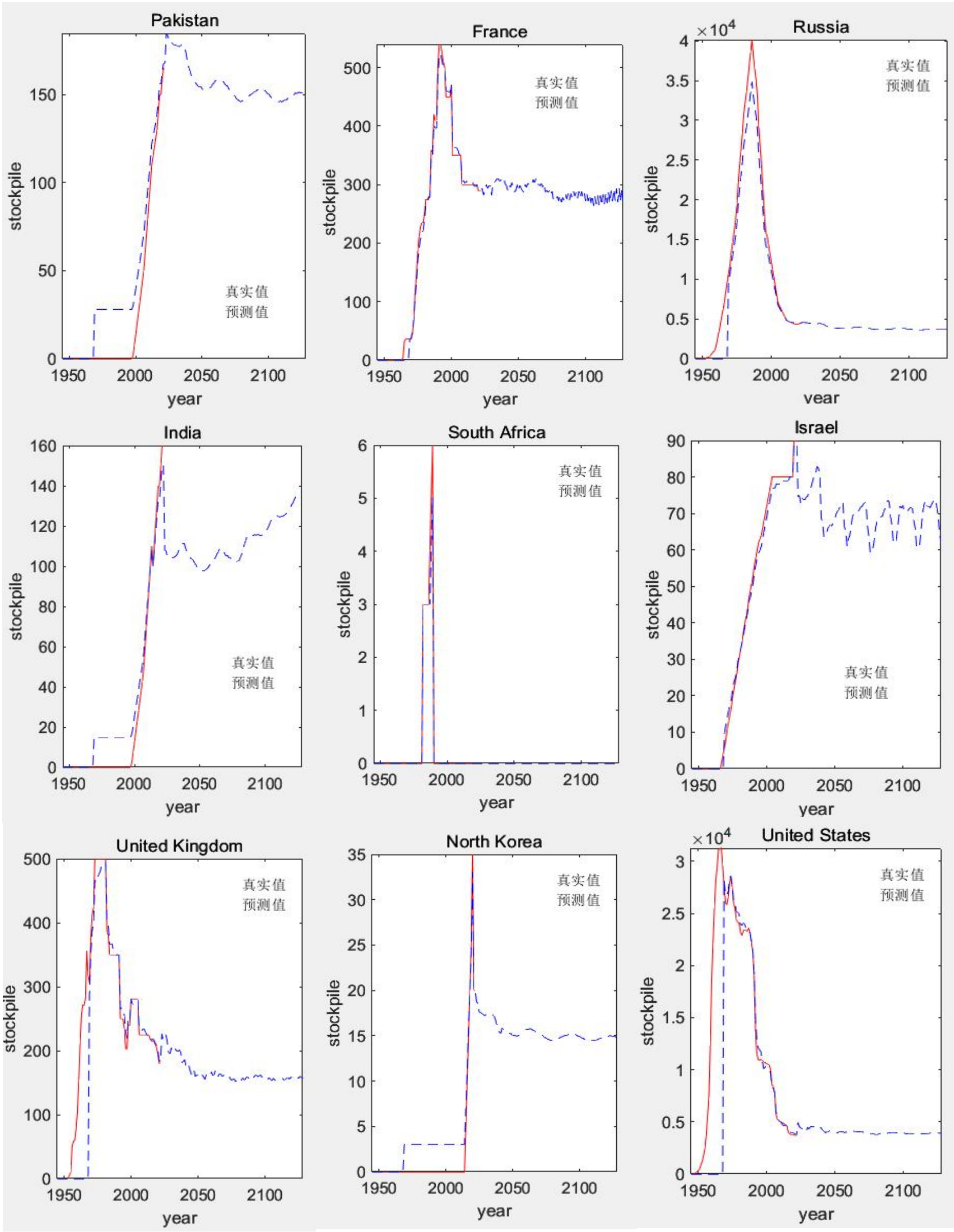


Figure 11 :Forecast of the number of nuclear weapons possessed by countries

7 The Solution to Problem Number Three

7.1 Study the Number of Nuclear Bombs Desroying the Earth

7.1.1 Aspect 1

In this question, because we are studying how many nuclear bombs are required to destroy the earth, we assume that the nuclear bomb used is the most powerful nuclear bomb in the world, namely The Tsar Bomba .According to the above assumptions, the explosion radius of the nuclear bomb is determined.Besides, the explosion of the nuclear bomb on the ground is approximately a circle, and the explosion range of nuclear bombs in all countries is the same. Therefore, we believe that the nuclear bomb explosion is not affected by external forces such as terrain, so the explosion point of the nuclear bomb and its surrounding areas are plain. It also follows the core idea of Aspect 1: if the explosion area of a nuclear bomb is greater than or equal to the earth's area, the earth's surface can be destroyed.Therefore, the test explosion point we take each time is the maximum killing radius, and any two explosion ranges do not overlap each other.

The plain cities (Tokyo, Moscow, São Paulo) are selected for simulation test blasting, and the average results are as follows:

（1） Fireball radius：4.62 km（67.1 km^2）。The correlation between the maximum size of nuclear fireball and ground damage depends on the explosion height.If it touches the ground, the amount of radioactive fallout will increase significantly.Everything in the fireball will evaporate effectively. Minimum explosive height of negligible sediment：4.16km。

（2） Radiant radius（500 rem）：5.05 km（80.2 km^2）。500 rem ionizing radiation dose possibly fatal; 15% of survivors will eventually die of cancer due to exposure.

（3） Heavy explosion damage radius（20 psi）：10.4 km（339 km^2）。Under the overpressure of 20 pounds per square inch, the severely constructed concrete buildings were severely damaged or demolished, and the death toll was nearly 100%.It is usually used as a benchmark for serious urban damage. The optimum blasting height to maximize this effect is 6.7km.

（4） Thermal radiation radius (third degree burn)：60.1 km（11350 km^2）.Third degree burns cover all layers of the skin and are usually painless because they destroy painful nerves.They can cause severe scars or disabilities and may require amputation. In this case, the 100% probability of third degree burn is 13.6cal/cm2.

Here, because we are talking about destroying the earth, we calculate the damage radius of heavy explosion as the benchmark, and then obtain the earth surface area of 510072000 km^2 by collecting data ， It can be seen from the simple division calculation that if the nuclear bomb explosion is seamlessly connected, 510072000/339=1504637, that is, about 1.51 million Tsar missiles can destroy the earth's surface.

7.1.2 Aspect 2

Matter is bound together by gravity and can only be separated by doing work. The energy that different parts of the same object combine together is called binding energy.The number of nuclear bombs needed to destroy the earth can be obtained by dividing the energy generated by the explosion of The Tsar Bomba from the combined energy. Derived from simple physical formula：

$$E=\frac{3}{5}\cdot G\cdot\frac{m^2}{r}$$

By bringing in the gravitational constant G, the earth mass m, and the earth radius r for calculation, it can be concluded that the work required to make the earth's matter move to an infinite distance $E=2.25\cdot 10^{32}$（J）。

At the same time, each kilogram of trinitrotoluene can produce 4184000 joules of energy. One ton of trinitrotoluene is equal to 4184000000 joules. The energy produced by The Tsar Bomba is 50 million tons of trinitrotoluene. It can be obtained through multiplication calculation that the energy produced by a Tsar Bomba is 2.092e+17 joules. After division calculation, it can be seen that 1075525812619502 Tsar Bomba are needed to completely destroy the earth.

7.2 Can Nuclear Bombs Destroy the Earth at Present

Taking the most powerful Tsar missile as an example, according to the above simulation test results, the damage radius of heavy explosion is 10.4 km.Under the over pressure of 20 pounds per square inch, the severely constructed concrete buildings were severely damaged or demolished, and the death toll was nearly 100%.The optimal blasting height to maximize this effect is 6.7 km, which requires millions of pieces to destroy the earth's surface. With the current scientific and technological level of mankind and the earth's resources, it is completely impossible to produce so many missiles, so it is not enough to destroy the earth.

7.3 The Limits of the Number of Nuclear Bombs to Be Studied

7.3.1 Cluster analysis

The population information is obtained by searching relevant data and the urban population is divided by Cluster analysis method. In order to facilitate the research, we draw lessons from the city size classification standard and divide it into five categories, namely small city, medium city, big city, supercities and megacities. SPSS was used for Cluster analysis, as shown in Figure 12.

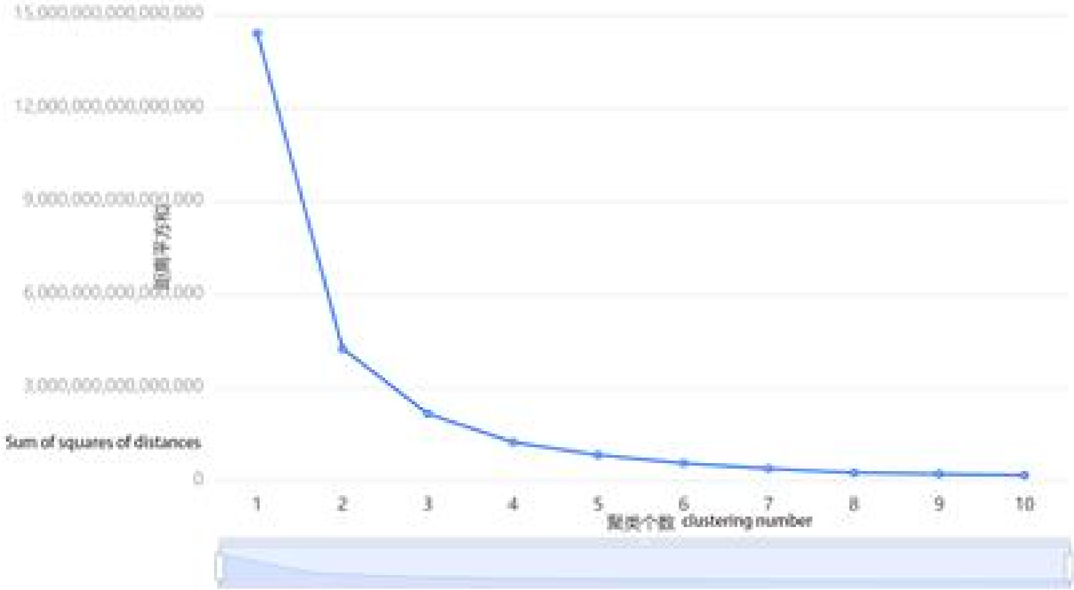


Figure 12: Comparison of cluster number

In this figure, the abscissa is the number of clustering, the ordinate is the loss function of K-mean clustering, and is the sum of squared distances of all samples to the category center, which is also the sum of squared errors (the smaller the value, the better the clustering effect). It can be seen from the figure that the effect is better after the number of clustering is 4.

Field variance analysis is then performed, as shown in Table 7.

Table 7: Field difference analysis

	Cluster class (mean ± standard deviation)					F	P
	Category 3(n=758)	Category1(n=143)	Category 5(n=43)	Category 2(n=19)	Category 4(n=9)		
Population Estimate	1049296.834±491403.487	3882370.629±1049850.611	8981976.744±1974328.608	17235000.0±2221086.421	27858222.222±5358298.722	3994.543	0.000***
Square Kilometers	305.439±335.098	946.315±1017.119	1975.814±1889.251	3208.105±2670.117	3588.222±2087.7	165.194	0.000***

Note: ***, ** and * represent the significance level of 1%, 5% and 10% respectively

For the variable Population Estimate, the significance P value is 0.000***, showing significance at the level. The null hypothesis is rejected, indicating that the variable Population Estimate has significant differences among the categories divided by Cluster analysis. For the variable Square Kilometers, the significance P value is 0.000***, showing significance at the horizontal level. The null hypothesis is rejected, indicating that the variable Square Kilometers has significant differences among the categories divided by Cluster analysis.

The frequency of each cluster category is analyzed according to clustering summary, as shown in Table 8.

Table 8 :Clustering summary

Clustering category	frequency	Percentage（%）
Clustering Category_1	143	14.712
Clustering Category_2	19	1.955
Clustering Category_3	758	77.984
Clustering Category_4	9	0.926
Clustering Category_5	43	4.424
total	972	100.0

The frequency of cluster category _1 was 143, accounting for 14.712%. The frequency of cluster category _2 was 19, accounting for 1.955%. The frequency of cluster type _3 was 758, accounting for 77.984%. The frequency of cluster category _4 is 9, accounting for 0.926%. The frequency of cluster category _5 was 43, accounting for 4.424%.

According to the data clustering category annotation, we can know which category each sample data falls into. Secondly, we can analyze the distance between each sample and the central point according to the clustering center coordinates. Finally, we can judge the sample category by sorting the central value, as shown in Table 9.

Table 9 :Coordinates of clustering center points

Cluster type	center value_Population Estimate	center value_Square Kilometers
1	3882370.629370628	946.3146853146851
2	17234999.999999996	3208.105263157894
3	1049296.8337730826	305.43931398416896
4	27858222.222222216	3588.2222222222217
5	8981976.744186044	1975.813953488372

According to the ranking of central values in the figure, we can judge that 4 is megacities, 2 is supercity, 5 is big city, 3 is small cities, and 1 is medium-sized cities.

Finally, the relevant indicators are evaluated. As shown in Table 10

Table 10: Evaluation indicators

The silhouette coefficient	DBI	CH
0.726	0.525	3994.54

For a sample set, its silhouette coefficient is the average of the silhouette coefficients of all samples. The value range of the silhouette coefficient is [-1,1]. In this case, the silhouette coefficient is 0.726, which means that the score is high and the clustering effect is good. The DBI metric is used to measure the ratio of the intra-cluster distance to the inter-cluster distance after any two clusters. In this problem, the index is small and the clustering effect is better. The CH index is obtained by the ratio of separation and tightness; in this case, the CH is large, the DBI is moderate, and the clustering effect is good.

7.3.2 Limiting Results

NUKEMAP by Alex Wellerstein simulates and calculates how many nuclear weapons are needed to destroy a city of each type, and finally sums up. However, our explosion radius all uses the heavy explosion damage radius, that is, the casualties generated by the heavy explosion damage radius. The nuclear weapons that destroyed each class of cities are shown in Table 11, and the number of each class of cities and their corresponding number of nuclear weapons are shown in Table 12.

Table 11: Summary of nuclear weapons destroyed in various types of cities

megacity	Czar Missiles（50Mt，249km^2，14）
supercity	Czar Missiles（50Mt，249 km^2，13）
Big City	Czar Missiles（50Mt，249 km^2，8）
Medium city	castle bravo（15MT，90.5 km^2，2）Czar Missiles（50Mt，249 km^2，3）
Small city	castle bravo（15MT，90.5 km^2，1）Czar Missiles（50Mt，249 km^2，1）

Table 12: The number of different cities and their total nuclear weapons

City category	Average population size	Mean urban area (cubic km)	quantity	total quantity
Small city	1049296.83	305.44	758	1516
Medium city	3882370.63	946.31	143	715
Big City	8981976.74	1975.81	43	344
supercity	17235000.00	3208.11	19	247
megacity	27858222.22	3588.22	9	126

Finally, it is calculated that in the case of severe explosion damage, only 2,948 missiles are needed to destroy the basic major cities of the world. Considering the enhancement of weapon killing capability brought by the progress of human technology, and the theme of peaceful development in the current era, it is better to have fewer restrictions on nuclear weapons in the future. Each of these countries should have a maximum of 328 nuclear weapons.

8 Study Findings and Recommendations

To understand the development of nuclear weapons over the past few decades and over the next one hundred years,we have analyzed the number of nuclear countries, the reserve of nuclear weapons and the number of nuclear weapon tests.We used the time series analysis model to forecast the data.By comparison, the number of nuclear bombs is gradually decreasing and tends to be stable, and the total number of nuclear weapons will reach 5979 in 2123.

Then we use Bayesian regularization algorithm to establish a nonlinear self combining time series network. It is found that there will be 10 nuclear weapon states in the next 100 years, of which Iran is the new nuclear weapon state.Finally, we used The Tsar Bomba to simulate bombing on the plain, obtained its maximum damage radius and concluded that 1.51 million tsarist bombs were needed to destroy the earth's surface, and deduced that 1075 trillion tsarist bombs were needed to completely disintegrate the earth according to the physical formula.Therefore, it is concluded that human beings cannot destroy the earth through nuclear bombs at present. Finally, using Cluster analysis and corresponding calculation, it is concluded that each country with nuclear weapons should have at most 328 nuclear weapons.

In this regard, we put forward the following suggestions:

(1) Effectively use and manage weapons systems without harming human interests and international security;

(2) Disarmament in accordance with the development of the international community and universal requirements;

(3) Promote test technology, accelerate nuclear testing for peaceful purposes, and promote the modernization of nuclear technology;

(4) Stop the proliferation of nuclear weapons through diplomatic, political, economic and even military means;

(5) Limiting the proliferation of nuclear weapons.

9 Strength and Weakness

9.1 Strengths

1.Analyze the Time series model used in the world nuclear weapons quantity change table, apply the domestic advanced modeling tool SPSSPRO cloud software, through unit root test and partial autocorrelation analysis, automatically optimize the parameters, and finally the goodness of fit R reaches 0.979, which is relatively stable.

2.In the process of predicting nuclear weapons in various countries, Bayesian regularization algorithm is used to train time series, establish neural network.Adopting this high-precision algorithm leads to better results.

3.In the third question, the definition of destroying the earth was considered comprehensively, and two kinds of results were analyzed. The destructive force of nuclear bombs was predicted through the U.S. simulated nuclear bomb website, with high accuracy.

4.In the 3C question, the Cluster analysis is more accurate, the profile coefficient and DBI are acceptable, and the classification is more obvious. At the same time, the model is relatively new and innovative.

5.The overall idea of the three questions algorithm is relatively simple. Some questions use Excel and other software to easily get results, and some questions use Python, MATLAB and other comprehensive methods to process data. The model and idea design are simple, practical and efficient.

9.2 Weaknesses

1.Due to the existence of the "black box" problem, the neural network is difficult to correct errors in the learning process, which requires more time. Although the algorithm has high accuracy, it requires a lot of time for training. Using MATLAB to run the algorithm takes more time, which may take about 1 minute.

2.The AHP method is adopted to judge the most active countries in nuclear weapons research. The judgment matrix of this algorithm is constructed subjectively by individuals, with a certain degree of subjectivity.

3.In the calculation of the third question of nuclear weapons, the factors considered in this model are relatively ideal, which reduces the universality and generalization ability of this model.

10 References

[1] Scientific Platform Serving for Statistics Professional 2021.SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com.

[2] Yu Ningli, Yi Dongyun, Tu Xianqin. Analysis of Autocorrelation and Partial correlation Functions in time Series [J]. Mathematical Theory and Applications,2007(01):54-57. (in Chinese)

[3] WANG Yan. Applied Time Series Analysis [m]. Beijing: China Renmin University Press, 2005.

[4] A scientific platform for the statistical profession 2021. SPSSPRO. (Version 1.0.11)[Web application]. Retrieved from https://www.spsspro.com.

[5] Saroj, The Story. Research on simple K-means and improved K-means Clustering [J]. Journal of Computer Engineering, 2016,6(7):279-281.

[6] Shi Jianbin, Zhu Jianyu. Analysis on the influence of the Joint Comprehensive Plan of Action on Iran's nuclear capability. International Forum,2016(3):1-6.

[7] Yang Chen, Chen Zhanjun. Iran has been able to build a nuclear bomb US Troop surge in the Gulf US Sanctions against Iran Unpopular, China Maintains trade ties with Iran despite pressure [J]. China,2019,(12):40-43.