Problem Chosen	2023HSB	Team Number
B	MCM/ICM Summary Sheet	MI00893

Color revolution All harm without any benefit

This paper focuses on how to assess the social risks of a country based on its national context, and then make judgements about future trends in society based on social risks and other factors, and make recommendations accordingly.

In response to question one, data from countries around the world were collected, taking into account five categories of indicators: economic, political, military, livelihood and cultural, with data collected on a total of 15 indicators. A total of 10 indicators were extracted from the raw data by **Data statute**, **Data enhancement** and correlation analysis to construct the social stability indicator system. The indicators are assigned weights using an **Optimized assignment model** and the causal relationships of the indicators are analyzed based on the weights.

In response to question two, a **Comprehensive evaluation model** was developed based on question one to assign rating values to countries around the world, **K-means clustering** was also performed to classify the social stability states into 3 categories. In terms of objective reality, the workflow for early warning of the risk of social imbalance is proposed, a system of early warning indicators for social stability is constructed, and **Early warning thresholds** for the corresponding indicators are set. On this basis, the **Support vector machine** technique is applied to build a social stability early warning model using MATLAB as a platform.

In response to question three, using Belarusian data for 1920-2020 as a sample, and data statutes, **Data enhancements**, data pre-processing to the original data, An **Example validation** of the social stability early warning model. According to the results, the risk of social imbalance in Belarus is medium alert.

In response to question four, data from Ukraine for 1920-2020 were used as a sample and the raw data were processed. Optimization of support vector machines is performed by selecting kernel functions, penalty coefficients and kernel function parameters and **Training the classifier**. According to the results, the risk of social imbalance in Ukraine is medium alert.

Based on the above analysis in this paper, the color revolution failed to change the status quo of economic poverty, political corruption and social unrest. We should stay away from "color revolutions" and break the attempts of the Western strategy of peaceful evolution, Forging ahead in an environment of mainstream ideological security.

Key words: Data statute, Data enhancement, Comprehensive evaluation model K-means clustering, Support vector machine

Contents

1 Introduction

1.1 Background

We are living a stable and happy life in China, but today the world still faces enormous challenges to stability: the outbreak of the new crown epidemic in 2020, the war between Russia and Ukraine, hegemony, etc. Since the end of the Cold War, there has been an increasing number of 'color revolutions' in which regime change has been achieved through non-violent means.According to statistics, over the past 30 years, so-called "non-violent revolutions" have overthrown more than 90% of all regimes that have fallen. Social stability[1] generally refers to the continuous functioning of the entire social system, including the economic, political and cultural systems of the state, in a coordinated and orderly manner and in dynamic balance.

1.2 Problem requirements

- Consider comprehensively the various areas that affect social stability, find the appropriate data and establish a system of indicators that affect social stability.
- This paper establishes an early warning model for social stability based on the dynamic equilibrium relationship between the indicators.
- Using Belarus as a sample, a reasonable assessment of its social stability is made through the Social Stability Early Warning Model.
- Using Ukraine as a sample, the Social Stability Early Warning Model identifies the main causes of regime change in the country
- To prevent color revolutions and maintain social stability, relevant recommendations are made.

2 Problem analysis

2.1 Analysis of Problem One

In this paper, to explore the construction of a social stability indicator system, suitable indicator data were sought in the economic, political, military, livelihood and cultural domains. The missing values are supplemented by data statute using cubic spline interpolation and non-linear fitting, five invalid indicators are removed, and the new dataset is tested. The weights of the indicators from the original and optimised datasets were calculated and visually represented. And causality analysis is carried out based on the weights.

2.2 Analysis of Problem Two

In this paper, to explore the construction of a social stability early warning model, countries around the world are scored using a combined TOPSIS and GRA evaluation model on the basis of problem one, followed by a K-means clustering algorithm to classify countries around the world into social stability types. A social stability early warning model is constructed based on the support vector institution, through the social risk early warning workflow, threshold setting of early warning indicators, data normalisation, selection of kernel functions and determination of classifier models.

2.3 Analysis of Problem Three

A Belarusian-based early warning model for social stability was applied to collect relevant data for data enhancement, data segmentation and normalisation. After comparison and testing, the optimal parameters of the training classifier are obtained. An early warning analysis of the risk of social imbalance in Belarus for 2011-2020 is also carried out. Finally, the reasons for Belarus' successful response to the color revolution are analysed.

2.4 Analysis of Problem Four

Based on the use of the Ukrainian social stability early warning model, relevant data were collected for data enhancement, data segmentation and normalisation. After comparison and testing, the optimal parameters of the training classifier are obtained. An early warning analysis of the risk of social imbalance in Ukraine for 2011-2020 is also carried out. Finally, the reasons for Ukraine's failure to cope with the color revolution are analysed.

2.5 Analysis of Problem Five

To prevent color revolutions and for people to live in peace and prosperity, relevant recommendations are made.

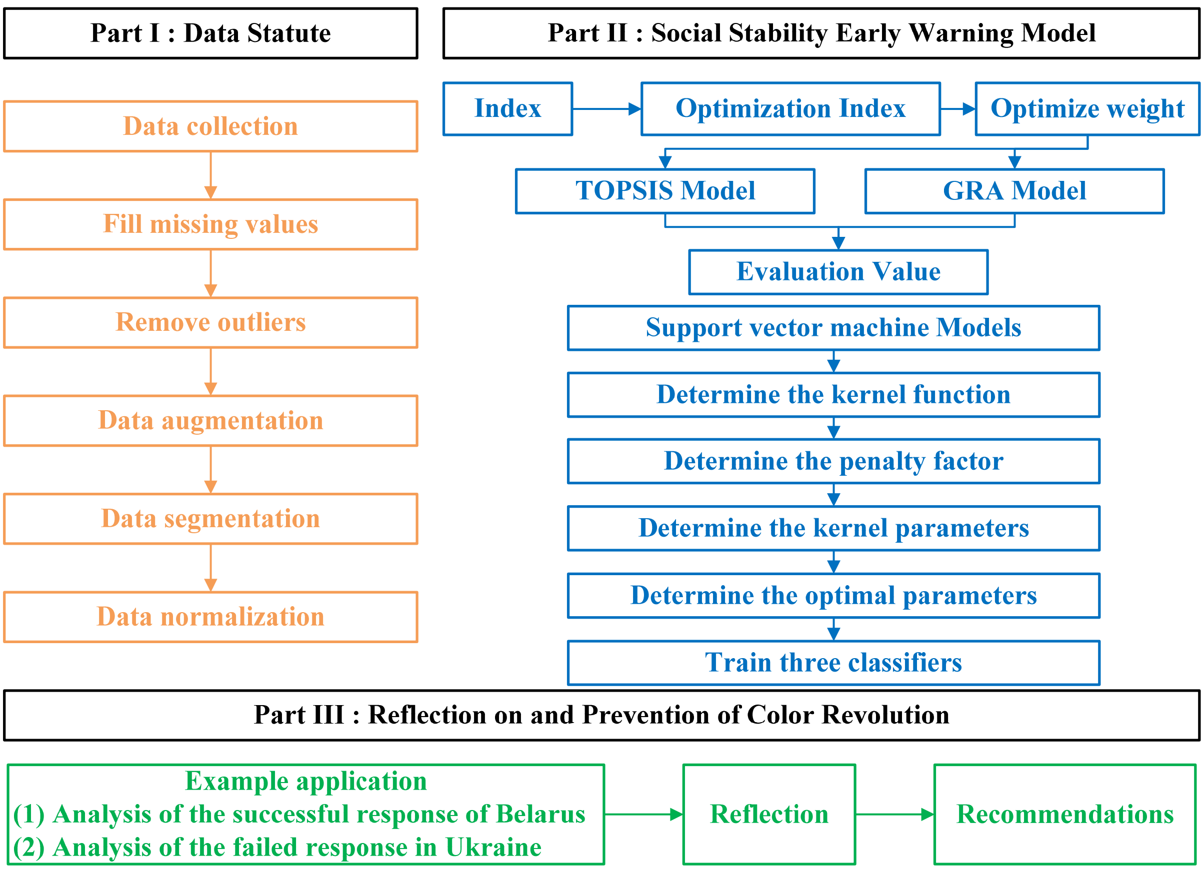


Figure 1 Diagram of the solution

3 Model assumptions

In the process of solving the problem, considering the actual situation and the need to simplify the calculation, the following assumptions are put forward:

(1) **The color revolution will be based on dynamic changes, abstracting representative features and patterns.** "Color revolutions" occur in a large number of countries with different circumstances, and the revolutionary situation varies from country to country, and the "color revolution" itself is a new means of peaceful evolution, which is constantly evolving and not a static method。

(2) **Consider the impact on social stability of the 5 domains factor only.** The color revolution affects all areas and aspects of society, but this paper is based on objective literature that downscales the areas of influence into 5 areas and validates them scientifically.

(3) **The data for the sample countries for 1920-2020 were collected without considering the effect of small probability events.** For example, geological disasters, traffic accidents, economic crises, systemic risks, production accidents, etc.

4 Symbol Description

Symbol	Description
s_i	Sample standard deviation
R	Correlation coefficient matrix
W^*	Portfolio weights
$\xi_i(k)$	Number of grey contacts
s_i^*	Distance of positive ideal solution
S_i	Relative proximity
$\mu_k^{(0)}$	Barycenter
M^+	Sample set
$f(x)$	Optimal classification function
ξ_i	Relaxation items
$Q(\alpha)$	Transformed optimization function
$K\left(x_i \cdot x\right)$	Inner product function
r_{ii}	Indicators
$K(x, y)$	Radial base core
T	Training set
C	Penalty Factor
γ	Kernel function parameters
$(\omega \cdot x)$	Regulating hyperplanes
$\operatorname{sgn}\left(w^* \cdot x + b^*\right)$	Optimal classification function

5 Building a system of social stability indicators

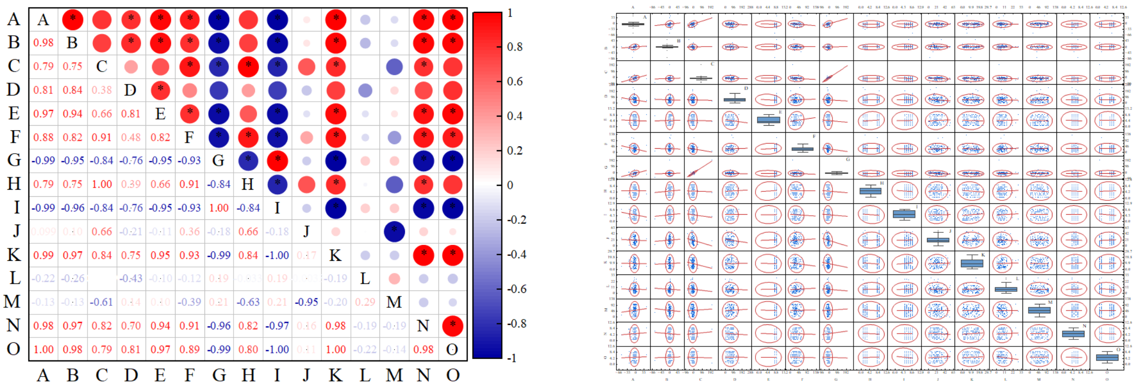
5.1 Data collection and pre-processing

The year 2020 was the year of the global outbreak of the new crown epidemic, which affected the economy, politics and culture of all countries to a certain extent and increased the instability of societies. Therefore, this paper searches and collates data on relevant indicators from the World Bank and World Trade Organization databases worldwide for that year.

Table 1 Social Stability Indicator System

Risk areas	Alarm indicators	Indicator Description
Economy	GDP growth rate Primary sector growth rate Commodity export rate	The economy plays a decisive role in the development of society as a whole. Examines the functioning of the economic base.
Politics	Central government debt Degree of political stability Total government expenditure	The government plays a leading role in the development of society and can guide economic, military and other developments.
Military	Military expenditure Armed Forces Terrorist attacks	The military is the final guarantee of social stability, against domestic reactionary forces and against foreign invaders.
Livelihood	Staff remuneration Gap between rich and poor Unemployment rate	People's livelihoods directly reflect the standard of living of the population and are the basis for social stability.
Culture	Literacy rate Cultural Market Cultural Resources	Culture is an important expression of a country's state of affairs and can reflect the smooth functioning of society.

The degree of political stability in the political sphere, armed forces and terrorist attacks in the military sphere. The qualitative analysis in this paper is based on a score of 1 to 9. The missing data from the quantitative analysis is filled by means of cubic spline interpolation [2] and non-linear fitting to obtain a complete dataset of relevant indicators for each country.



(a) Correlation graph (b) Matrix scatterplot

Figure 2 Raw data visualization processing

It can be seen from the figure that the degree of correlation between the indicators in figure (a) is relatively large, so that the data information of the indicators is not fully reflected (b) is relatively large, and each The linear relationship between the components is not obvious.

5.2 Data specification

Because there is a strong correlation between some indicators, it will cause information overlap and affect the objectivity of the calculation results. Therefore, the data set is further optimized through data reduction.

Step 1 Normalize the raw data. There are 15 index variables, and there are 266 evaluation objects in total. The value of the j index of the i evaluation object is a_{ij}. Convert each index value a_{ij} into a standardized index value $\tilde{a}_{ij}$, then

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j}, (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m), \tag{1}$$

In the formula: μ_j and s_j are the sample mean sample standard deviation of j indicator.

Step 2 Calculate the correlation coefficient matrix $\boldsymbol{R}$. $\boldsymbol{R} = \left(r_{ij}\right)_{m \times m}$, then.

$$r_{ij} = \frac{\sum_{k=1}^{n} \tilde{a}_{ki} \cdot \tilde{a}_{kj}}{n-1}, i, j = 1, 2, \cdots, m \tag{2}$$

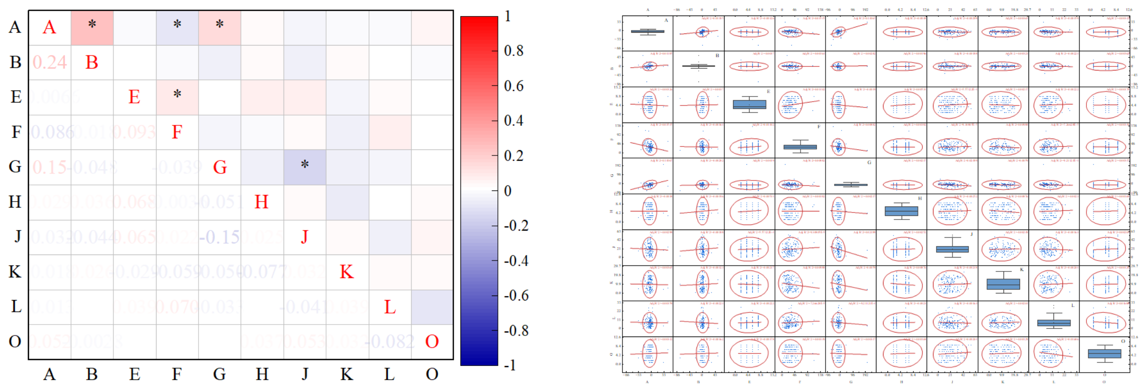
In the formula: $r_{ii} = 1$, $r_{ij} = r_{ji}$, r_{ij} is the correlation coefficient between the i indicator and the j indicator.

Step 3 Compute eigenvalues and contribution ratios. Calculate the correlation coefficient matrix $\boldsymbol{R}$ with an eigenvalue of $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$, and the corresponding eigenvector is $u_1, u_2, \cdots, u_m$, In the formula $u_j = \left[u_{1j}, u_{2j}, \cdots, u_{mj}\right]^T$.

Table 2 The eigenvalues of the matrix and their contribution

	1	2	3	4	5	6	7	⋯	15
Eigenvalue	2.13	1.69	1.64	1.24	1.14	1.06	0.97	⋯	0.084
Contribution rate	14.26	11.29	10.97	8.29	7.62	6.52	6.11	⋯	0.56

Through the attribute specification and numerical specification, it is displayed by the relevant information of the correlation coefficient diagram and the matrix scatter diagram.



(a) Correlation Heatmap (b) Matrix scatterplot

Figure 3 Dataset after data reduction

Two conclusions can be drawn from the visualization results in the figure:

- The correlation among the 10 indicators shown in Figure (a) is low, and the information reflected by each indicator has strong independence, so that the calculation results show objectivity.
- In figure (b), the boxplot in the oblique line indicates that the deviation between each component and its mathematical expectation is low, and the confidence ellipses in the upper right triangle and lower left triangle indicate that the linear relationship between the components is small. It is thus demonstrated that data reduction produces new datasets that are smaller and preserve data integrity.

5.3 Causality analysis

5.3.1 Build an Optimal Empowerment Model

This paper analyzes the objective weight, subjective weight, comprehensive weight and optimized weight from four aspects.

✧ **Objective weight**

The information weighting method is an objective weighting method. The objective weight of each index in the CRITIC method is calculated through the information contained in the index data.

Table 3 Algorithm steps of information empowerment method

Algorithm 2 CRITIC algorithm		
Step 1	Data standardization $X_{ij}=[f_j(i)-f_{\min}]/f_{\max}-f_{\min}$	
Step 2	Calculated linear correlation coefficient X_i and X_j	
Step 3	Calculate the amount of information contained in the index j	
Step 4	Calculate the normalized weight of each indicator $W_j=C_j/\sum_{k=1}^{n}C_k$	(3)

✧ **Subjective weight**

The ordinal relationship analysis method is a subjective weighting method improved on the basis of the analytic hierarchy process. While overcoming the shortcomings of the analytic hierarchy process.

Table 4 Algorithm steps of sequence relation analysis method

Algorithm 3 G1 algorithm		
Step 1	determine order relationship. Determine Relative Importance Ranking	
Step 2	Determine the relative importance between adjacent indicators.	
Step 3	Calculate Weight w_k , $w_n=(1+\sum_{k=2}^{n}\prod_{i=k}^{n}r_i)^{-1}, w_{k-1}=r_kw_k(k=1,\cdots,n)$	(4)
Step 4	Determining the weight set $w_k=(w_1,w_2,\cdots,w_n)^T$	(5)

✧ Integrated weight

The subjective weight and objective weight obtained by using the sequential relationship analysis method and the information weighting method are respectively taken as 50%, and the integration weight is obtained by adding them together.

✧ Game theory optimization weight

Game theory mainly studies the interaction between formulaic incentive structures. It is a mathematical theory that studies phenomena with the nature of struggle or competition. Based on this theory, the weight of indicators can be optimized.

Step 1 Build portfolio weights

The formula for calculating the combination weight W is constructed in the form of any linear combination.

$$W=\sum_{i=1}^{n} a_i W_i, \quad a_i>0 \tag{6}$$

$$W_i=\{w_{i1}, w_{i2}, \cdots, w_{im}\}^T \tag{7}$$

In the formula: W_i represents the weight set obtained by the i method.

Step 2 Optimizing Weights

According to the basic principles of game theory, the weights are optimized to minimize the deviation between the combined weights and the weights participating in the optimization.

$$\min\left\|\sum_{i=1}^{n}(a_i W_i - W_j)\right\|_2 \tag{8}$$

Step 3 Normalization

To ensure the consistency of the results, the weight coefficients are normalized.

$$a_i^* = a_i / \sum_{i=1}^{n} a_i, i=1,2,\cdots,n \tag{9}$$

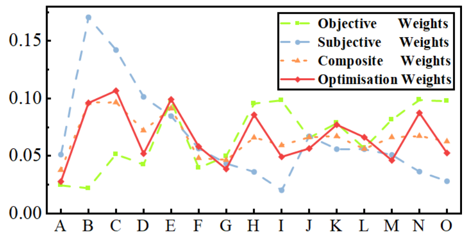
In the formula: a_i^* represents the normalization parameter corresponding to coefficient a_i.

Step 4 Determining Portfolio Weights

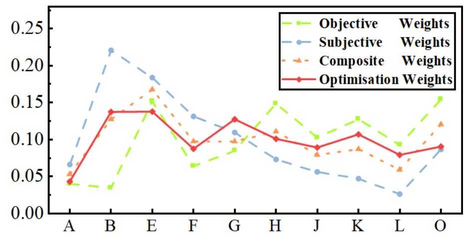
According to the calculation results, the optimized combination weight W^* is obtained.

$$W^* = \sum_{i=1}^{n} a_i^* W_i \tag{10}$$

Substituting the original data set and the optimized data set of the data specification into the optimization weighting model for calculation, the results are shown in the figure below.



(a) Original data set



(b) Optimize Dataset

Figure 4 The relative weight coefficient of the indicator

After the data set has been reduced, the volatility of the weight coefficient has decreased, and because the independence of each index is strong, the amount of information reflected by the weight has increased.

5.3.2 Weight-based causality analysis

The evaluation index system in this paper covers the fields of economy, politics, military, people's livelihood and culture, and optimizes it through correlation analysis. After consulting the reference [3], it is found that there is a certain causal relationship between these five factors.

Extract the optimization weights of the original data set and the optimized data set, and draw a double-layer pie chart for data comparison.

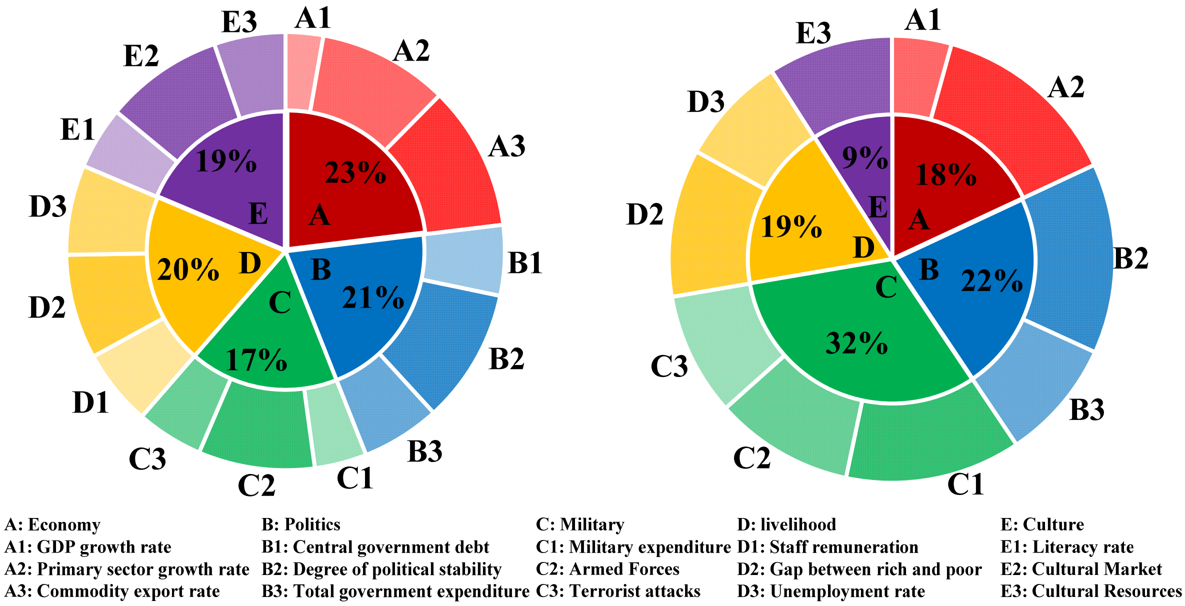


Figure 5 Optimize weight data comparison

- In the original data set, the weight values of each field are roughly the same, so it is impossible to make an accurate judgment. In the optimized data set, the weight values of various fields are different, so the optimized data set is analyzed.
- **Weight analysis of level II dimension indicators.** Among the five secondary indicators, military is the largest at 0.317, and cultural is the smallest at 0.091. It shows that among the above-mentioned five second-level dimension indicators, the military has the greatest impact on social stability, and the decline of the military level will also increase the possibility of causing social risks.
- **Level Ⅲ police situation index weight analysis.** Among the 10 police indicators, the primary industry growth rate index is the largest at 0.138, and the employee salary ratio index is the smallest at 0.043. Explain that the economy plays a decisive role in the whole process of social development.
- **Causality analysis.** The economy plays a decisive role in social stability, the military is a powerful guarantee for social stability, the government plays a leading role in social stability, and culture and people's livelihood are the superstructure on this basis.

6 Building an early warning model for social stability

6.1 Social Stability Scores of Countries Around the World

In the above, we collected the relevant indicator data of all countries in the country in 2020, and obtained the weights of each. In this question, we establish an optimal evaluation model based on the complementarity between the TOPSIS model and the GRA model. Its core formula is as follows:

✧ Calculate the **gray correlation coefficient**

$$\xi_i(k)=\frac{\min\limits_{1\le s\le m}\min\limits_{1\le t\le n}|x_0(t)-x_s(t)|+\rho\max\limits_{1\le s\le m}\max\limits_{1\le t\le n}|x_0(t)-x_s(t)|}{|x_0(k)-x_i(k)|+\rho\max\limits_{1\le s\le m}\max\limits_{1\le t\le n}|x_0(t)-x_s(t)|},i=1,2,\cdots,m \tag{11}$$

✧ Computes the **distances to positive** and **negative ideal solutions.**

$$s_i^*=\sqrt{\sum_{j=1}^{n}(c_{ij}-c_j^*)^2},i=1,2,\cdots,m \tag{12}$$

$$s_i^0=\sqrt{\sum_{j=1}^{n}(c_{ij}-c_j^0)^2},i=1,2,\cdots,m \tag{13}$$

✧ **Calculate the relative closeness** between the evaluation unit and the ideal solution

$$\begin{gathered}S_i^+=\alpha s_i^*+\beta s_i^0,S_i^-=\alpha s_i^*+\beta s_i^0\\ S_i=S_i^+/\left(S_i^++S_i^-\right)\end{gathered} \tag{14}$$

Substitute the indicator data of each country into the model to obtain the scores of each country and visualize them.

Figure 6 Social Stability Scores of Countries Around the World

As can be seen from the figure, according to the different shades of color, it can be roughly divided into 3 categories.

6.2 Classification of Social Stability Types Based on K-means Clustering

According to the social stability score obtained by the combined evaluation model, this paper collects the latitude and longitude of each country from the Google Earth landmark online website, and constructs a K-means clustering model for cluster analysis.

Step 1 Three centroids are randomly selected from 265 sample data as the initial cluster centers. The centroid is recorded as:

$$\mu_1^{(0)}, \mu_2^{(0)}, \cdots, \mu_k^{(0)} \tag{15}$$

Step 2 Define optimization goals.

$$J(c,\mu) = \min \sum_{i=1}^{M} \left\| x_i - \mu_{c_i} \right\|^2 \tag{16}$$

Step 3 Start the loop, calculate the distance from each sample point to the centroid, and assign the sample to which centroid is closer to whichever sample is closer, and get 3 clusters.

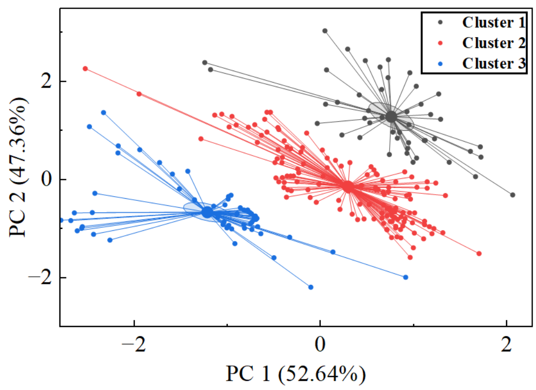
$$c_i^t < -\arg\min_k \left\| x_i - \mu_k^t \right\|^2 \tag{17}$$

Step 4 For each cluster, **calculate the average distance** of all sample points classified into the cluster as the new centroid.

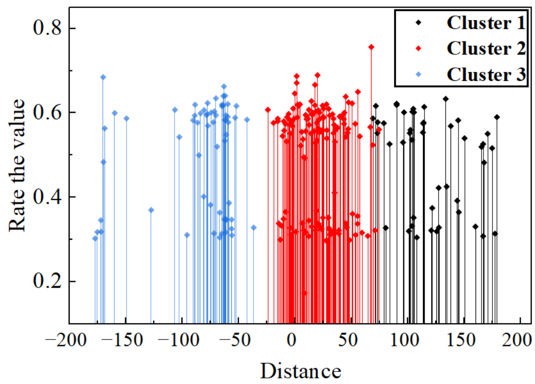
$$\mu_k^{(t+1)} < -\arg\min_u \sum_{i:c_i^t=k}^{b} \left\| x_i - \mu \right\|^2 \tag{18}$$

Step 5 Until J converges, that is, all clusters no longer change.

The running results of its K-means clustering model are shown in the figure below:



(a) Cluster diagram



(b) Group chart

Figure 7 Schematic diagram of K-means clustering

It can be seen from the figure that social stability can be roughly divided into three categories: stable, semi-stable, and turbulent, and all three categories have good data characteristics.

6.3 Early warning model of social stability based on SVM

6.3.1 The principle of support vector machine [4]

The support vector machine is a binary classification model, and its basic model is a linear classifier with the largest interval defined in the feature space. The learning strategy of SVM is to maximize the interval, which can be formalized as a problem of solving convex quadratic programming, which is also equivalent to the problem of minimizing the regularized hinge loss function. There are three cases of sample classification: linearly separable, linearly inseparable, and nonlinearly separable.

- **Linearly separable**

The theorem of support vector machine is: The necessary and sufficient condition for the training set T to be linearly separable is, The convex hull separation of the two types of sample sets M^+ and M^- of T; When the training set T is linearly separable, There is a unique gauge hyperplane $(\omega \cdot x)+b=0$, make

$$\begin{cases} \left(\omega \cdot a_i\right)+b \geq 1, & y_i=1 \\ \left(\omega \cdot a_i\right)+b \leq -1, & y_i=-1 \end{cases} \tag{19}$$

Let the classification training sample set be $\left(x_i, y_i\right), i=1, \cdots, n, x \in R^d, y \in \{+1,-1\}$.

When the equation satisfies the following two constraints, the classification surface is the optimal classification surface:

$$\begin{cases} \min \Phi(w)=\dfrac{1}{2}(w \cdot w) \\ y_i\left[\left(w \cdot x_i\right)+b\right]-1 \geq 0, i=1, \cdots, n \end{cases} \tag{20}$$

Take the partial derivatives of w and b separately and set them equal to 0, The original problem is transformed into a dual problem. At this time, the optimization function of the optimal classification is:

$$Q(\alpha)=\sum_{i=1}^{n} \alpha_i-\frac{1}{2} \sum_{i=1}^{n} \alpha_i \alpha_j y_i y_j\left(x_i \cdot x_j\right) \tag{21}$$

Correspondingly, the optimal classification function can be obtained as:

$$f(x)=\operatorname{sgn}\left(w^* \cdot x+b^*\right)=\operatorname{sgn}\left[\sum_{i=1}^{n} \alpha_i^* y_i\left(x_i \cdot x\right)+b^*\right] \tag{22}$$

- **Linear inseparability**

Adding a slack term $\xi_i \geq 0$ to condition $y_i(wx+b)-1 \geq 0$ becomes $y_i(wx+b)-1+\xi_i \geq 0$. Change the goal to find the minimum of $1/2\|w\|^2+C\left(\oplus_{i=1}^{n} \xi_i\right)$, which is to compromise the least misclassified samples and the largest classification interval, so as to obtain the generalized optimal classification surface, where C controls the degree of punishment for misclassified samples.

- **Nonlinearly separable**

A nonlinear transformation is performed on the original feature space, the nonlinear problem is transformed into a linear problem in a high-dimensional space, and then the optimal classification surface is determined in the high-dimensional space. The nonlinear transformation only needs to replace the inner product operation $\left(x_i \cdot x\right)$ in the optimized function expressions in formulas (11) and (12) and the optimal classification function expressions with the inner product function $K\left(x_i \cdot x\right)$, and the optimized function expression after transformation is :

$$f(x)=\operatorname{sgn}\left(w^* \cdot x+b^*\right)=\operatorname{sgn}\left[\sum_{i=1}^{n} \alpha_i^* y_i K\left(x_i \cdot x\right)+b^*\right] \tag{23}$$

The support vector machine uses the linear classification surface of the high-dimensional space to correspond to the nonlinear decision function of the original feature space, without increasing the computational complexity. Therefore, the obtained solution has a certain degree of scientificity and generalization sex.

6.3.2 Social Stability Early Warning Model

(1) Social Risk Alert Workflow

Determine the source of the risk. Determine which side of the field is responsible for the risk of social unrest. Social stability is multifaceted and interconnectedThis is a key aspect of social risk early warning. Risks can be classified as no risk, medium risk and serious risk, and correspondingly, social stability early warning can be divided into no, medium and serious alert levels. Comprehensive judgement and rational decision-making.

(2) Threshold setting for early warning indicators

The social stability indicator system constructed according to Question 1 contains indicators that correspond to the components of the input sample in the support vector machine, which is one dimension of the support vector machine. The imbalance warning thresholds of the social stability early warning element indicators are set using the following three methods:

- Based on internationally recognized safety limits settings
- Set according to relevant national conditions and policies.
- Developed in accordance with local conditions and data realities.

Early warning thresholds for the risk of social unrest are developed using countries around the world as examples, and the threshold bands are divided into no-alert, medium-alert and heavy-alert intervals.

Table 5 Warning thresholds for each alert level of the indicator

Index	Nature	No alarm threshold	Medium alarm threshold	Heavy alarm threshold
GDP growth rate	Forward	≥ 1.89	$[-10.98, 1.89)$	< -10.98
Primary sector growth rate	Forward	≥ 7.42	$[-5.55, 7.42)$	< -5.55
Degree of political stability	Forward	≥ 7.99	$[1.81, 7.99)$	< -1.81
government expenditure	Forward	≥ 48.69	$[14.34, 48.69)$	< -14.34
Military expenditure	Moderate	$[1.2, 7.2]$	$(-8.4, 1.2) \cup (7.2, 10.3)$	$\leq 8.4 V \geq 10.3$
Terrorist attacks	Moderate	$[6, 7]$	$(4, 6) \cup (7, 8)$	$\leq 4 V \geq 8$
Staff remuneration	Forward	≥ 33.45	$[8.41, 33.45)$	< 8.41
Gap between rich and poor	Negative	≤ 2.81	$(2.81, 16.93]$	> 16.93
Unemployment rate	Negative	≤ 2.71	$(2.71, 13.59]$	> 13.59
Cultural Resources	Forward	≥ 8.07	$[2.07, 8.07)$	< 2.07

In the modelling application, the sample data was divided into three parts for use: firstly, 20% of the data was used as example test data for the social stability early warning model, 70% of the data was used as training data for the social stability early warning model and 10% of the data was used as simulation test data for the social stability early warning model.

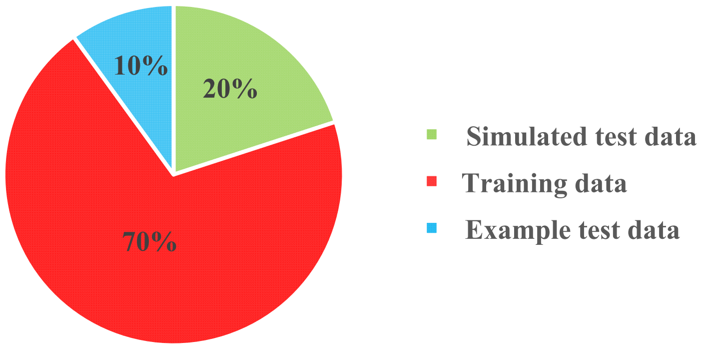


Figure 8 Sample Data Segmentation

(3) Data normalisation

The training set data and the test set data are normalised to eliminate the differences in the scale of the indicators, so that the role played by each indicator is relatively balanced. In the social stability early warning model indicator system, different normalisation methods are adopted due to the nature of the indicators of each element being positive, negative and moderate.

Positive indicators i.e. indicators with higher values are better indicators.

$$r_{ij} = \frac{a_{ij} - \min\left(a_{ij}\right)}{\max\left(a_{ij}\right) - \min\left(a_{ij}\right)} \tag{24}$$

Inverse indicators i.e. indicators with smaller values are better indicators.

$$r_{ij} = \frac{\max\left(a_{ij}\right) - a_{ij}}{\max\left(a_{ij}\right) - \min\left(a_{ij}\right)} \tag{25}$$

The moderate indicator is the indicator with the best value within the interval $\left[q_1^j, q_2^j\right]$.

$$r_{ij} = \begin{cases} 1 - \dfrac{\max\left(q_1^j - a_{ij}, a_{ij} - q_2^j\right)}{\max\left[q_1^j - \min\left(a_{ij}\right), \max\left(a_{ij}\right) - q_2^j\right]}, & a_{ij} \notin \left[q_1^i, q_2^j\right] \\ 1, & a_{ij} \in \left[q_1^i, q_2^j\right] \end{cases} \tag{26}$$

(4) Select the kernel function.

Commonly used kernel functions[5] are radial basis kernels, polynomial kernels, linear kernels, sigmoid kernels, etc. polynomial kernel, linear kernel, sigmoid kernel, etc. Among them, the feature space corresponding to the radial basis kernel function is infinite dimensional, and the finite samples in the feature space are definitely linearly divisible, so the general classification algorithm can be used. The radial basis kernel is used in the social stability early warning model after the comparison trial calculation, and its expression is:

$$K(x,y) = \exp\left(-\gamma \| x - y \|^2\right) \tag{27}$$

(5) Determining the classifier model

The data for each country in the world was brought into the vector machine and the results were obtained as shown in the figure below.

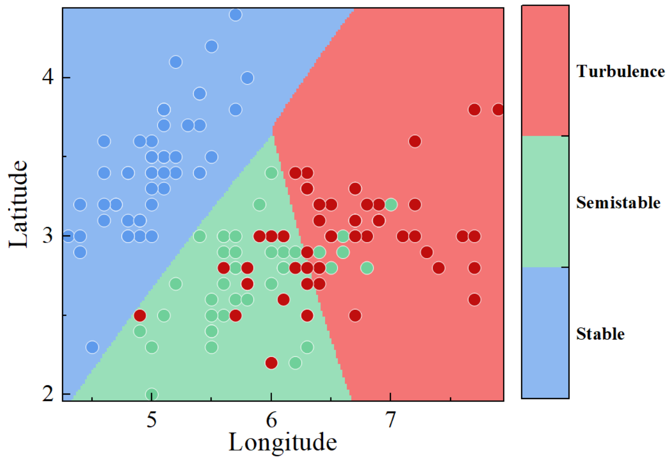


Figure 9 Support vector machine run results

After comparison and testing, the final four classifier models identified for the social stability early warning model were: no-alarm classifier, medium-alarm classifier and heavy-alarm classifier.

7 Early warning model of social stability based on Belarus

7.1 Data collection and preprocessing

Belarus has experienced three color revolutions since the beginning of the 21st century, but there has been no change of government, so this article takes Belarus as a sample. The sample data of the social stability early warning model consists of two parts. 1. Simulated sample data generated equidistantly within the indicator early warning threshold interval. It is specifically described that 15 simulation samples are generated for each warning threshold interval. Since the threshold interval of the moderate index is divided into 5, and the threshold interval of the positive index and the negative index is divided into 3, and the samples are generated in conjunction with the moderate index, a total of 75 standard simulation samples of 5×15 can be generated. 2. Correlation of 100 years in Belarus from 1920 to 2020. Subsequently, Perform data enhancement, data segmentation and normalisation.

2011-2020 is the example test data, 91 sets of Belarusian historical real data from 1920-2010 and 50 simulated samples randomly selected are combined as training data, and the remaining 25 simulated samples are used as simulated test data. The training set contains 141 sample points, The training set is denoted as $T = \{(x_1, y_1), \cdots, (x_{141}, y_{141})\} \in (X \times Y)^{141}$, Among them, $x_i \in X = R^{10}$ is composed of 10 indicators such as GDP growth rate and the gap between rich and poor, $y_i \in Y = \{1, -1\}, i = 1, \cdots, 141$. That is, the input includes 10 features, when the output is $y_i = 1$, it means there is risk, and when the output is $y_i = -1$, it means no risk.

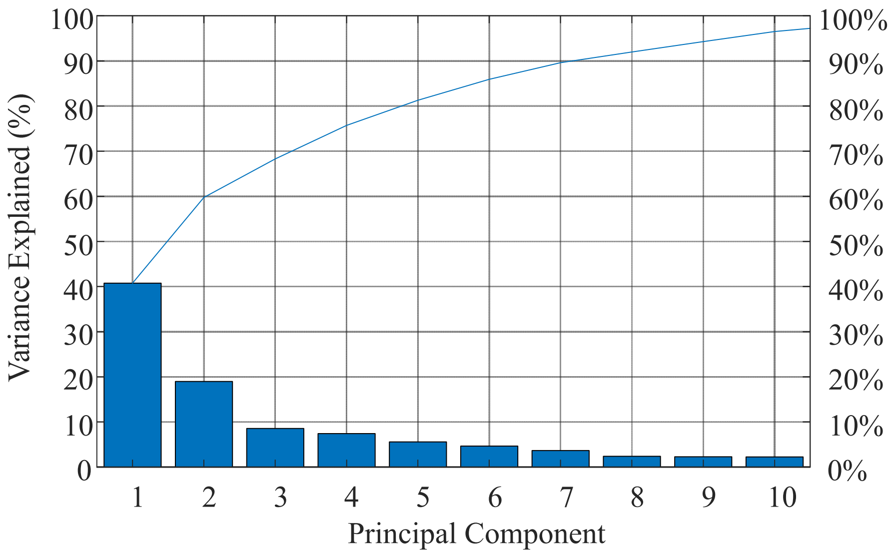


Figure 10 Pareto chart of raw data

The Pareto chart arranges the relative frequency values of the 10 indicators in a descending manner from left to right, so as to clearly show the data characteristics among the indicators.

7.2 Preferences

This paper uses MATLAB as the research platform, and calls LIBSVM -- A Library for Support Vector Machines library.

7.2.1 Choose kernel function

After comparison and calculation, the social stability early warning model adopts the radial basis kernel, and its expression is:

$$K(x,y)=\exp\left(-\gamma\| x-y\|^{2}\right) \tag{28}$$

7.2.2 Cross-validation selection penalty coefficient C kernel function parameter γ

In the radial basis kernel support vector machine, the selection of kernel function parameter γ value is directly related to the classification performance of the support vector machine: If γ is small, the misclassification rate of training samples is 0, but the generalization recognition ability is low; if γ is large, the correct classification rate of new samples is also very low. The same is true for the penalty coefficient C, if it is too large or too small, it will reduce the generalization recognition ability of the support vector machine.

7.2.3 Determining the optimal parameters

In the support vector machine, the cross-validation selection of parameters is carried out through the Meshgrid function, the upper limit of iterations is set to 10,000, and the search range of parameters is gradually narrowed, so that the accuracy of parameters is continuously improved.

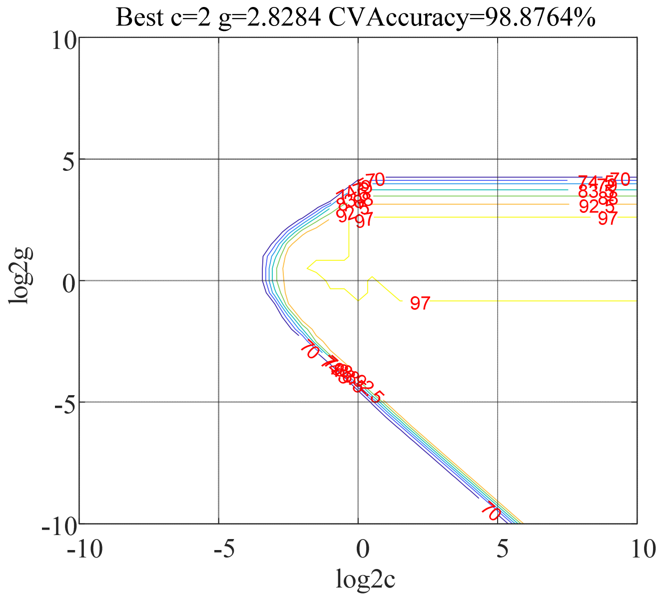


Figure 11 Optimal parameter selection

According to the results of parameter cross-validation selection, the classification accuracy rate of training samples and test samples output by different parameter models is compared, and the parameter with the highest sample classification accuracy rate is the optimal parameter.

7.2.4 Training a classifier

In this paper, the 1-against-rest multi-classification method is used to train the classification data separately to obtain the models of each classification. The 1-against-rest multi-classification method is to construct a two-class classifier with other classes for each class, Therefore, the trained Class i classifier can separate Class i samples. The social stability early warning system is divided into three early warning categories, and three classifiers need to be trained.

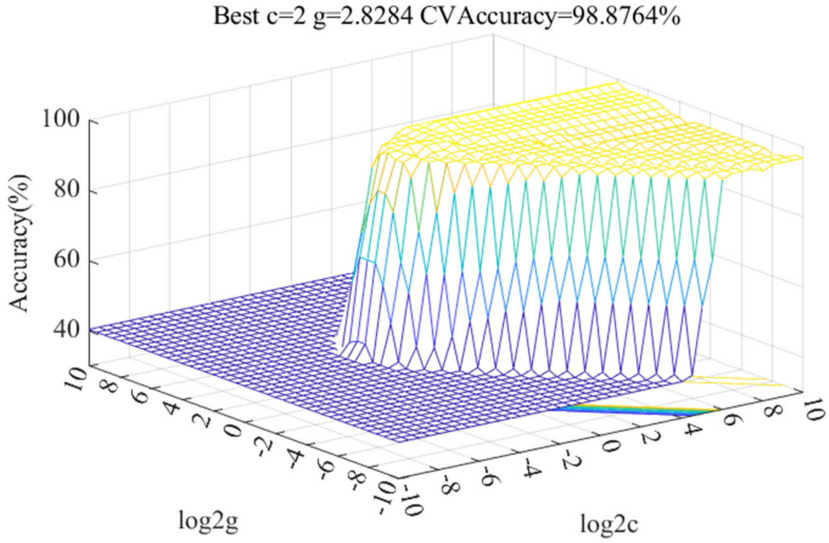


Figure 12 Classifier classification effect

After comparison and testing, the optimal parameters of the three classifier models finally determined by the social stability early warning model are as follows:

✧ **Non-alarm classifier:** C =2, γ =1, The classification accuracy rate of training samples is 98.88%, and the classification accuracy rate of 25 simulated samples is 93.71%.

- ✧ **Central alarm classifier:** C=94, γ=1, The classification accuracy rate of training samples is 92.41%, the classification accuracy rate of 25 simulated samples is 91.47%.
- ✧ **Heavy alarm classifier:** C=38, γ=0, The classification accuracy rate of training samples is 96.12%, the classification accuracy rate of 25 simulated samples is 95.83%.

7.3 Example application

Based on the established social stability early warning model, an early warning analysis of the social imbalance risk in Belarus from 2011 to 2020 is carried out.

Table 6 Belarus 2011-2020 early warning and comprehensive assessment comparison

Year	Alert level	Score value	Compared	Year	Alert level	Score value	Compared
2011	Central Police	0.52	Conform	2016	Central Police	0.57	Conform
2012	No police	0.36	Almost	2017	Central Police	0.26	Almost
2013	Central Police	0.26	Conform	2018	No police	0.21	Conform
2014	Central Police	0.67	Conform	2019	Central Police	0.37	Conform
2015	Heavy police	0.82	Conform	2020	Central Police	0.32	Conform

According to the early warning results of the model, the risk of social imbalance from 2011 to 2020 is roughly moderate, which is basically consistent with the comprehensive evaluation results of the combined evaluation model. According to the analysis of the data, a certain proportion of indicators have deviated from the normal value over the years, reaching the early warning threshold of each level. The early warning results show that Belarus has faced the risk of social imbalance at the middle police level in recent years, and its social risk should be improved.

7.4 Analysis of Belarus' successful response to the Color Revolutions

Each of the five major areas of Belarus is analysed quantitatively, with corresponding thresholds, and the reasons for its successful response to the color revolution are then analysed.

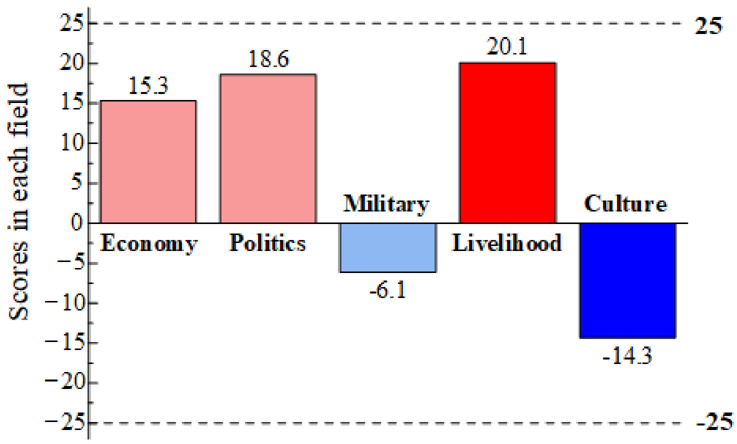


Figure 13 Corresponding ratings for each area in Belarus

As the graph shows, Belarus has positive economic, political and livelihood values, and although its military and culture are negative, they are not below the threshold.

- Economy: The main economy of Belarus is not state-controlled and there are no interest groups or oligarchies affecting the country. As a result, the Belarusian economy is generally well developed and the standard of living of the population is relatively high.
- Politics: Lukashenka has been re-elected as the first President of Belarus since 1 July 1994, during which time Belarus has maintained political stability and achieved sustained economic and social development.
- People's livelihood: people's living standards are relatively high, with the Gini coefficient ranging from 0.26 to 0.30 since 1995, and the gap between rich and poor in society at a "relatively average" and "relatively reasonable" level.

8 An early warning model of social stability based on Ukraine

8.1 Data collection and preprocessing

The 'Orange Revolution' in Ukraine refers to the series of protests and political events that took place across the country in 2004-2005, surrounding the 2004 presidential elections in Ukraine, which resulted in regime change due to serious corruption and electoral fraud The data was collected for 100 years from 1920 to 2020,Perform data enhancement, data segmentation and normalization.

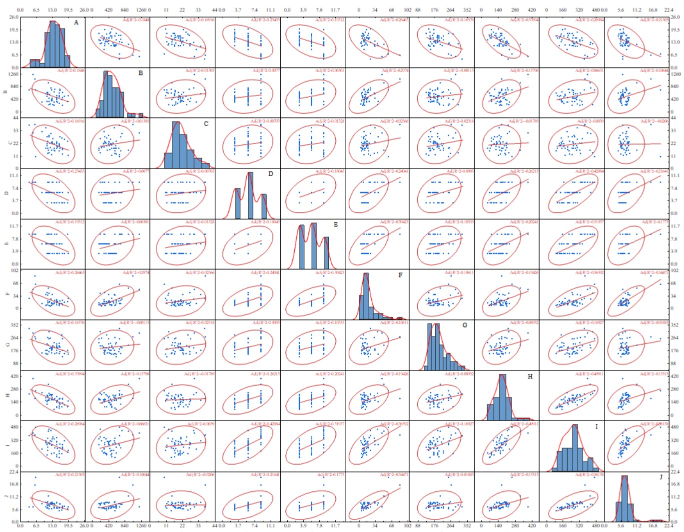


Figure 14 Matrix scatter plot of relevant data for Ukraine, 1920-2020

As can be seen from the graph, the average deviation between the components and their mathematical expectations is small and the linear relationship between the components is not significant. The period 2011-2020 for Ukraine will be the example test data.

8.2 Parameter selection

In this paper, MATLAB is used as the research platform and the LIBSVM - A Library for

Support Vector Machines library is called. This is followed by: 1. selecting the kernel function, 2. selecting the penalty coefficient C and kernel function parameter γ based on cross-validation, 3. determining the optimal parameters, and 4. training the classifier.

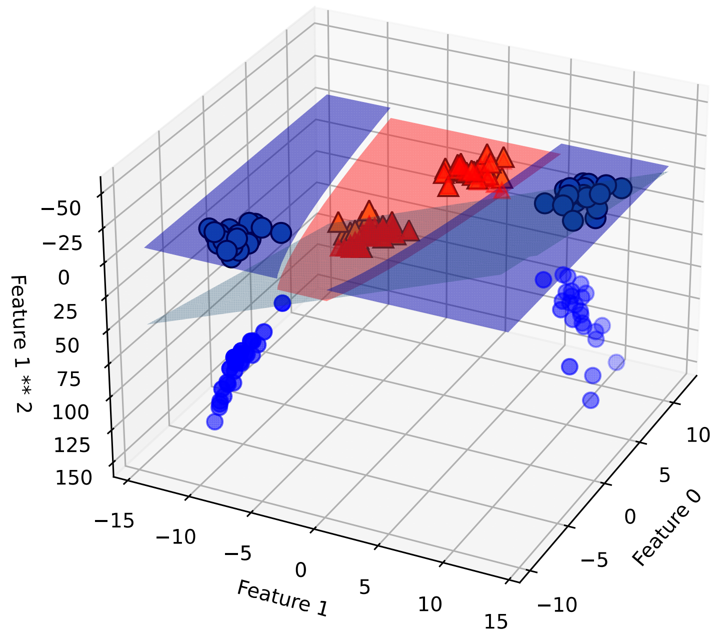


Figure 15 Classifier classification effect

After comparison and testing, the optimal parameters of the final three classifier models for the social stability early warning model are shown below:

- ✧ **Unpoliced Classifier:** C=4, γ=0, The correct classification rate was 92.48% for the training sample and 95.12% for the 25 simulated samples tested.
- ✧ **Medium Alarm Classifier**: C=86, γ=1, The correct classification rate was 94.27% for the training sample and 93.28% for the 25 simulated samples tested.
- ✧ **Heavy Alarm Classifier :** C=14, γ=0, The correct classification rate was 98.45% for the training sample and 92.41% for the 25 simulated samples tested.

8.3 Result analysis

Early warning analysis of the risk of social imbalance in Ukraine for the period 2011-2020 based on the established early warning model of social stability.

Table 7 Early warning and overview comparison for Ukraine 2011-2020

Year	Alert level	Score value	Compared	Year	Alert level	Score value	Compared
2011	Central Police	0.82	Almost	2016	Central Police	0.91	Almost
2012	Heavy police	0.76	Conform	2017	Heavy Police	0.73	Conform
2013	Heavy Police	0.64	Almost	2018	Heavy police	0.85	Conform
2014	Central Police	0.71	Almost	2019	Central Police	0.95	Almost
2015	Heavy police	0.85	Conform	2020	Heavy Police	0.79	Conform

According to the model's early warning results, the risk of social imbalance in 2011-2020 is roughly a heavy warning, which is largely in line with the results of the portfolio evaluation model overview. Analysis of the data shows that a certain percentage of indicators deviated from normal values in all years and reached the warning threshold for each level. The results of the early warning indicate that Ukraine is at risk of social imbalance of a heavy warning level in recent years and that the Russian-Ukrainian war should be ended as soon as possible to establish stability in the country.

8.4 Analysis of Ukraine's failed response to the color revolution

The areas of White Ukraine were analysed separately and quantitatively, with corresponding thresholds, and thus the reasons for their failure to respond to the color revolution.

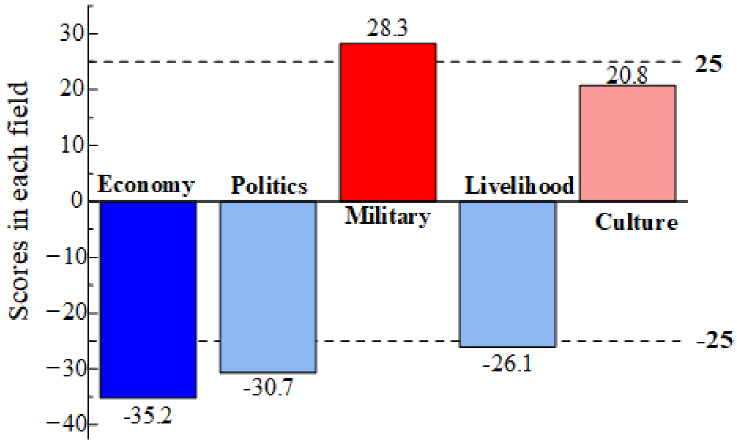


Figure 16 Corresponding scores for each field in Ukraine

As the chart shows, Ukraine is on the verge of economic, political, military and livelihood collapse.

- Economy: Since independence, along with the transition process, the Ukrainian economy has fallen into a serious crisis.
- Politics: Hopes are pinned on a change of government due to the incompetence of the government and the corruption of politicians.
- Military: US and European countries supply Ukraine with large amounts of weapons and war resources.
- People's livelihoods: people have not seen a significant improvement in their standard of living and are strongly dissatisfied with their current situation.

9 Preventing color revolutions and people living in peace

In an era of peaceful development, "color revolutions" are still emerging in waves and are prevalent in countries such as the Commonwealth of Independent States. The "color revolutions" advocate democracy and happiness for the people, the overthrow of authoritarian regimes and the establishment of a democratic system closer to Western civilisation, but in essence they are a peaceful evolution for the US and others.

Economy: To improve domestic economic construction, further liberate and develop the productive forces and raise the living standards of the people. It is only by doing so that the consensus of the people of a socialist country can be maximised. Economic development is the only way to provide more employment opportunities and to enable people to live and work in peace and happiness.

Politics: political party reform to consolidate the position of the ruling party. This requires that, on the one hand, as the ruling party, it must put the people at the centre. With the welfare of its people as its main objective, it must always be aware of the changes in all aspects of the region it leads, keep abreast of the times and promote social progress and the all-round development of people, thus gaining a broad mass base.

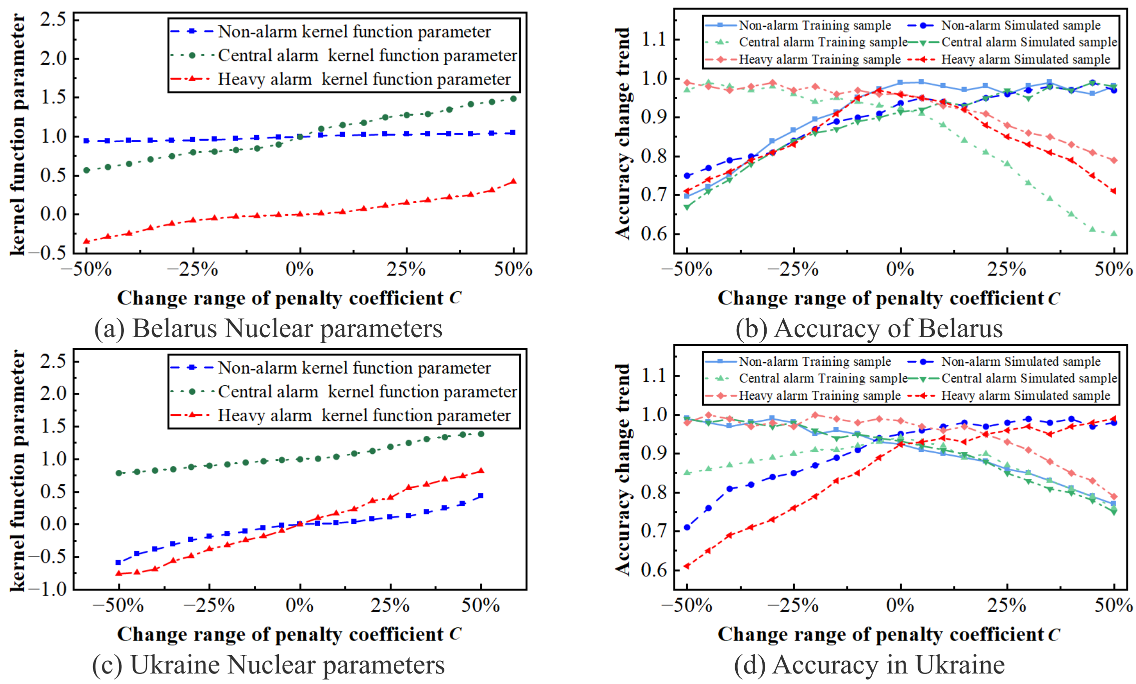
People's livelihood: strengthening democratic politics and expanding people's political participation. Only by putting the fundamental well-being of the people at the forefront of everything we do can we achieve the protection of their fundamental rights and promote social justice and equity.

On this basis, we will stay away from "color revolutions", break the Western strategy of peaceful evolution, take the initiative and move forward in an environment of mainstream ideological security, and move further and further along the road to socialism with Chinese characteristics and closer to the national rejuvenation of the Chinese nation.

10 Model Analysis and Testing

10.1 Sensitivity Analysis

In this paper, in our early warning model of social stability based on support vector machines, Both the penalty coefficient C and the kernel function parameter γ are the result of machine training and are somewhat unstable. Let the value of the penalty factor C vary from -50% to +50%, each time by 5%, in order to study the impact of fluctuations in this factor within a certain range on the accuracy of the social stability early warning model developed.



(a) Belarus Nuclear parameters (b) Accuracy of Belarus

(c) Ukraine Nuclear parameters (d) Accuracy in Ukraine

Figure 17 Diagram of sensitivity analysis

As can be seen from the graph, changes in the value of the penalty factor C do not have a significant effect on the kernel function parameter γ and the accuracy of the results. The magnitude of change does not fluctuate significantly, i.e. the social stability early warning model developed is not sensitive. In a sense, therefore, it suggests that the social stability early warning model we have developed has a degree of reliability and scientific validity.

11 Model Evaluation and Further Discussion

11.1 Analysis of Model Advantages

(1) The early warning model based on the support vector machine learning method is computationally simple, highly accurate and has a good generalisation capability: the prediction accuracy of the algorithm and its generalisation capability are greatly improved by finding a balance between error and penalty factors, giving the model good predictive performance.

(2) The TOPSIS algorithm avoids the subjectivity of the data, does not require an objective function, does not have to pass a test, and can well portray the comprehensive impact strength of multiple impact indicators.

(3) Data augmentation is a technique that artificially extends the training data set by allowing limited data to produce more equivalent data, thereby overcoming the shortcomings.

(4) The 1-against-rest multiclassification method chosen for this paper does not require retraining of all SVMs, but only retraining, which is relatively fast.

11.2 Analysis of Model Shortcomings

(1) During data pre-processing, there may be artificial and non-objective exclusion of valid data due to the incomplete nature of some of the data, which may cause the forecast to deviate from the actual data.

(2) The research in this paper has fixed a certain degree of "color revolutions", and to a certain extent abstracted representative features and patterns.

11.3 Model improvement

The support vector machine model developed in this paper can be further used to separate hot and cold data. By splitting the working set into two parts, with the hot data first and the cold data second, the SVM mostly iterates over a limited number of hot data after distinguishing between hot and cold, and occasionally, if not, it iterates over all of them once and then goes back to the hot and cold iterations, thus improving performance.

11.4 Model promotion

Early warning models based on the support vector machine learning method for non-linear problems, multi-feature problems and small samples. Support vector machines based on Gaussian radial basis kernel functions can map samples that cannot be classified in low-dimensional space to high-dimensional space for classification, which is a good solution to the problem of constructing high-dimensional models with a limited number of samples.

12 References

[1] Niu Gang. The expert seminar of the group standard "Social Stability Risk Assessment Specification for Major Projects" compiled by China Consulting Association was successfully held[J]. China Engineering Consulting,2022(12):33.

[2] WANG Weijun, JING Hao, GOU Na, YANG Jinhao. Research on feature set analysis method based on data statute of multiple choice problem[J]. Journal of the University of Electronic Science and Technology (Social Science Edition),2018,20(01):66-70.

[3] Li Shuzhuo, Wang Xiaoxuan. Risk management of imbalanced society under sustainable social development[J]. Studies on socialism with Chinese characteristics,2022(01):86-94.

[4] Dai S. L., Wang T. H., Zhou H. Y.. A fuzzy multicore support vector machine based on weighted martingale distance[J]. Computer Science,2022,49(S2):302-306.

[5] Chen R, Sun YF, Cheng DAGO, Guo Q, Chen YQ, Shi CQ, Sui YIC, Zhang YZ, Zhang YZ. Implementation and optimization of OpenCL kernel functions in TensorFlow [J]. Journal of Computer Science,2022,45(11):2456-2474.

13 Appendix

Appendix : MATLAB program for support vector machines

MATLAB the code SVM Model

```
addpath('LSSVMlabv1_8_R2009b_R2011a')
% Initializer
clear, clc
close all
X = xlsread('1');
Y = xlsread('2');
gam = 10;
sig2 = 0.5;
[alpha,b] = trainlssvm({X,Y,'c',gam,sig2,'RBF_kernel'});
Y_hat = simlssvm({X,Y,'c',gam,sig2,'RBF_kernel','preprocess'},{alpha,b},X);
disp(' Actual value  Predictive value ')
disp([Y, Y_hat])
% Drawing
figure('Units' , 'normalized' , ...
    'name' , 'Predicted Value Type vs. Actual Type' ,...
    'position' , [0.25 0.25 0.5 0.5] , ...
    'nextplot' , 'add');
plot(1: 30, Y, 'bo', 1:30, Y_hat, 'r*', 'linewidth', 1.0)
xlabel('Sample number'); ylabel('Type'); legend('Expected type' , 'Forecast type')
title({'LS-SVM Model Training Results' , ['Accuracy = ', num2str( sum(Y ==
Y_hat)/30 * 100 ), '%']})
```