TOPSIS and VAR-Based Predictive Modeling for the Pet Sector: Integrated Evaluation and Forecasting

Yibo Wang[1], Shuai Xia[1], #, Yu Zhao[2, *, #]

[1]International College, Chongqing University of Posts and Telecommunications, Chongqing, China,400065

[2]Communication and Information Engineering College, Chongqing University of Posts and Telecommunications, Changsha, China, 400065

*Corresponding author: zhaoyu@cqupt.edu.cn

#These authors contributed equally.

Abstract. Investigating the dynamic evolution and consumer-driven demands of the pet sector is pivotal for mapping industry trajectories, uncovering latent market opportunities, aligning with user expectations, refining product portfolios, designing adaptive marketing tactics, and accelerating sectoral innovation. This not only helps businesses and investors make scientific decisions but also fosters the sustained and sound development of the pet sector. Therefore, this paper proposes a comprehensive evaluation and forecast model for the trends in development and market demand of the pet sector based on the TOPSIS and VAR algorithms. Firstly, this paper selects five appropriate indicators by pet category and establishes a comprehensive evaluation model for the Chinese cat and dog market using TOPSIS and the entropy weight method. It then conducts k-means++ clustering based on the development scores of the cat and dog market in recent years. Secondly, using the Spearman correlation method, indicators with a correlation coefficient greater than 0.9 are identified as strongly correlated factors for the development of China's pet industry. Based on ARIMA, the values of these strongly correlated factors over the next three years are predicted, and the development situation of China's pet industry market in the coming three years is calculated, with the factor showing the largest increase reaching 35%. Then, the VAR model is used to apply positive shocks to tariff policy factors, monetary policy factors, and international trade factors, observing the quantitative changes in the impacted variable, which is China's pet food industry, to formulate feasible strategies to promote the growth of China's pet food sector. Finally, this paper discusses and analyzes the established model, comprehensively evaluating its advantages and disadvantages.

Keywords: TOPSIS, k-means++, ARIMA, VAR.

1. Introduction

Over the past decade, pets have gone from being cuddly companions to a booming business [1]. Pet owners' demands are becoming increasingly diverse. Competition within the industry is intensifying, and there is a noticeable trend towards branding and chaining. Concurrently, technological innovations are reshaping the industry landscape, introducing novel avenues for growth. Studying the development trends and market demand of the pet industry is becoming increasingly important.

Currently, many scholars both domestically and internationally have conducted research on the pet industry. ZHANG W et al. [2] deeply analyzed the macro environment of the pet market through the Political, Economic, Sociocultural, and Technological (PEST) analysis framework. They captured the pulse of the pet market by collecting and analyzing a large amount of market data, such as sales volumes of pet food, market demand for pet supplies, and the growth of pet medical services. Valdez J W et al. [3] systematically studied the past, present, and future development trends of the reptile trade using Google Trends, a powerful data tool. By inputting keywords related to reptiles, they analyzed the changes in search popularity of these keywords over different time periods. Xiao Y et al. [4] focused their research on the online pet food market. They designed detailed questionnaires to inquire about consumers' expectations and evaluations of pet food in terms of quality, price, flavor, packaging, etc. They also analyzed factors that influence consumers' actual purchasing behavior, such as brand awareness, promotional activities, user reviews, etc. Schleicher M et al. [5], on the other hand, conducted more focused research. Through an online survey, they deeply investigated the decisive factors for pet owners when purchasing pet food. The survey covered basic information about pet owners, the types and health status of pets, consumption habits of pet food, and more. By analyzing this data, they revealed the factors that pet owners are most concerned about when purchasing pet food, such as the health, nutrition, quality, freshness, and ingredients of pet food.

Drawing on the studies carried out by the previously mentioned scholars, it can be seen that most scholars studying the pet industry both domestically and internationally adopt qualitative rather than quantitative research methods. Building on the research of these scholars, this paper utilizes techniques such as k-means++ to classify the factors influencing the development of the pet industry into several indicators. Leveraging a hybrid analytical framework integrating TOPSIS prioritization, ARIMA forecasting, and VAR-based interdependency modeling, this research empirically evaluates growth trajectories and consumer-driven demands within the companion animal industry. The outcomes offer actionable insights for crafting adaptive sectoral policies across both domestic and global markets.

2. Description of applied methods

2.1. Entropy weight method for determining weights

The Entropy Weight Method (EWM) is a commonly used weighting method for measuring the degree of value dispersion in decision-making. The greater the degree of dispersion or differentiation, the more information that can be derived. Accordingly, higher weights should be assigned to indices with greater dispersion, and vice versa [6]. The Entropy Weight Method assigns metric-specific weights proportional to their statistical dispersion, wherein indicators demonstrating greater variability are allocated elevated significance in the evaluation framework.

First of all, this paper needs to carry out positive and normalized processing on the indicators of cat market and dog market to ensure the non-negative data

$$z_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

where Z_{ij} denotes the standardized metric derived by scaling the raw data between the observed extremes $x_{\min}$ and $x_{\max}$ for all criteria.

For each indicator *j* in year *i*, the probability p_{ij} is derived by normalizing its value relative to the sum of all indicators, forming the basis for entropy-based weight allocation

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{n} z_{ij}} \tag{2}$$

Calculate the information entropy e_j of the j-th cat and dog market index. Subsequently, the utility metric d_j is derived to counteract diminishing informational significance in pet market metrics. This adjustment stems from the inverse relationship between information entropy and data richness: higher entropy levels correspond to reduced discriminative capacity within indicators. By incorporating d_j, the framework ensures a constructive quantification of market dynamics.

$$\begin{gathered} e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln\left(p_{ij}\right) \\ d_j = 1 - e_j \end{gathered} \tag{3}$$

Finally, the entropy weight w_j of each index is obtained by normalization

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j} \tag{4}$$

2.2. TOPSIS Method for Scoring

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) employs a dual-objective optimization framework to rank constrained alternatives. This multi-attribute decision-making approach prioritizes options by concurrently reducing their proximity to the theoretical optimal solution and enhancing their deviation from the least desirable outcome [7]. In this example, it involves constructing a year where all indicators reach their optimal levels, and then measuring the proximity of actual years to this idealized year. The closer the actual year is to the ideal, the better the development of the cat and dog market in that year. Identify the maximum value for each column, which corresponds to each cat and dog market indicator, and denote it as $z_i^+ \quad (i = 1, 2, \cdots, m)$, forming a vector

$$Z^+ = \{z_1^+, z_2^+, \cdots, z_m^+\}$$

The vector represents the ideal development situation of the cat and dog market in a certain year. Similarly, identify the minimum value for each column, which corresponds to each cat and dog market indicator, and denote it as $z_i^- \quad (i = 1, 2, \cdots, m)$, forming a vector

$$Z^- = \{z_1^-, z_2^-, \cdots, z_m^-\}$$

This vector represents the least ideal development situation of the cat and dog market in a certain year, where each normalized indicator reaches its minimum value.

The distance D_i^+ between the i-th year and the ideal year is defined by the formula

$$D_i^+ = \sqrt{\sum_{j=1}^{m} \left(z_j^+ - z_{ij}\right)^2} \tag{5}$$

The distance between the i-th year and the unsatisfactory year is defined as D_i^-, and the calculation formula is

$$D_i^- = \sqrt{\sum_{j=1}^{m} \left(z_j^- - z_{ij}\right)^2} \tag{6}$$

Define the score of the first year as *S*, calculated by the formula:

$$S_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{7}$$

Obviously, *S* ranges between [0,1]. Proximity of *S* to unity reflects stronger alignment of the analyzed period with a theoretically optimal state, serving as a benchmark for idealized temporal conditions, and the cat and dog market performed better in that year. Conversely, when *S* is closer to 0, it indicates that the year is farther from the idealized year, and the cat and dog market performed worse in that year.

2.3. K-means++ Algorithm

The K-means++ algorithm is an extension of the k-means algorithm. Instead of randomly initializing the centers as in the traditional k-means algorithm, K-means++ selects initial cluster centers based on the principle of maximizing the distance between them[8]. In this modeling, we employ the K-means++ clustering algorithm to classify the past five years into three categories: rapid development, stable development, and slow development. We set the value of K to 3 for clustering. The steps of the K-means++ algorithm are as follows:

Step1) The primary cluster centroid c_i is initialized by uniformly sampling a single observation from the input dataset X;

Step2) Select the remaining clustering centers:

For each observation x_i in dataset X, compute the Euclidean metric $d(x_i, c)$ relative to the preassigned cluster centroid c, then record the minimum value as $d_{\min}$;

The assignment likelihood $P(x)$ for each observation to serve as the subsequent centroid is computed. The data instance exhibiting the highest likelihood value, derived from the probability distribution, is then designated as the new cluster representative

$$P(x)=\frac{d_i(x)^2}{\sum_{x\in X} d_i(x)^2} \tag{8}$$

Step3) Repeat the above process until all k clustering centers are identified.

Step4) Calculate the Euclidean distance between each sample and each cluster center:

$$D_{ij}=\sqrt{\sum_{k=1}^{m}(x_{ik}-K_{jk})^2} \quad j=1,2,\cdots,K; i=1,2,\cdots,n \tag{9}$$

there are m indicators and n samples. The samples with the minimum distance from each cluster center are compared and classified as the class of the cluster center.

Step5) After one generation selection, the cluster center is updated, and the new cluster center becomes

$$K_j^{'}=\frac{1}{n'}\sum_{x_i\in K_j} x_i \tag{10}$$

Step6) Repeat Step4 and Step5 until the clustering center does not change.

2.4. Application of Spearman correlation coefficient

Spearman's rank-order correlation serves as a non-parametric statistical tool to evaluate monotonic associations among ordinal or non-normally distributed datasets. Given that the predictors driving China's companion animal sector exhibit interval/ratio scaling without assuming normal distributions, this method is adopted to systematically assess interdependencies between market growth metrics and explanatory variables. Let X 、 Y be two sets of independently and identically distributed data, each with a sample size of N . X_i, Y_i represent the i -th value of the two sets of random variables, respectively, $i=1,2,......N$.

Firstly, sort the sets X 、 Y in either ascending or descending order simultaneously to obtain two ranked sets x , y , where the elements x_i , y_i represent the ranks of X_i, Y_i in their respective sets. Define the set d as the differences between the ranks of corresponding elements in the sets X and Y . The formula for calculating each element in the set d is as follows

$$d_i = x_i - y_i \tag{11}$$

Spearman's correlation coefficient r_s is calculated as follows:

$$r_z = 1-\frac{2\sum_{i=1}^{N} d_i^2}{N(N^2-1)} \tag{12}$$

2.5. ARIMA model

The ARIMA model combines autoregressive (AR), integrated (I), and moving average (MA) methods to handle time series data with trend or seasonal components. The ARIMA model has demonstrated excellent performance in predicting the next lag of time series with precision and accuracy[9]. The basic idea of the ARIMA model is to treat time series data as a combination influenced by factors such as random trends, seasonality, and cyclical fluctuations. By applying differencing, non-stationary time series data are transformed into stationary series, which are then used for forecasting through autoregressive and moving average models. The formula for the ARIMA model is as follows:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \text{蝌}_{t-1} + \theta_2 \;_{t-2} + \cdots + \theta_q ?_{t-q} + e_t \tag{13}$$

where Y_t represents the target variable at time t, c is the constant term, ϕ_t is the parameter for the autoregressive part, θ_t is the parameter for the moving average part, and $\grave{o}_i$ is the error term.

2.6. Vector Autoregression(VAR)

The Vector Autoregression (VAR) model is widely used in time series analysis and has been extensively studied in the literature[10]. As a multivariate extension of autoregressive (AR) frameworks, the Vector Autoregression (VAR) methodology is frequently applied to model interdependencies within temporal datasets and evaluate the ripple effects of stochastic shocks across interconnected variables. The model's architecture is governed by two critical hyperparameters: the dimensionality N and the temporal depth k. The general expression for a VAR model with N variables and k lags is:

$$Y_t = c + \prod_1 Y_{t-1} + \prod_2 Y_{t-2} + \cdots + \prod_k Y_{t-k} + u_t, \quad u_t \sim N(0, \Omega) \tag{14}$$

where, $Y_t = (y_{1t}, y_{2t}, \cdots y_{Nt})'$, $c = (c_1 \quad c_2 \quad \cdots c_N)'$

$$\prod_J = \begin{bmatrix} \pi_{11.\boldsymbol{j}} & \pi_{12.\boldsymbol{j}} & \cdots & \pi_{1N\boldsymbol{j}} \\ \pi_{21.\boldsymbol{j}} & \pi_{22.\boldsymbol{j}} & \cdots & \pi_{2N.\boldsymbol{j}} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{N1.\boldsymbol{j}} & \pi_{N2.\boldsymbol{j}} & \cdots & \pi_{NN.\boldsymbol{j}} \end{bmatrix}, \quad \boldsymbol{j} = 1, 2, \cdots, \boldsymbol{k} \tag{15}$$

$$u_t = (u_{1t} - u_{2t} - \cdots u_{Nt})',$$

Y_t is an $N \times 1$ vector of time series. c is an $N \times 1$ vector of constants. $\Pi_1, \cdots, \Pi_k$ are all $N \times N$ parameter matrices. $u_t \sim N(0, \Omega)$ is an $N \times 1$ vector of random error terms, where each element is non-autocorrelated. However, these elements, i.e., the random error terms corresponding to different equations, may be correlated with each other.

In Equation 1, the lag period is k, so it is referred to as the VAR(k) model. In practice, it is usually desirable for the lag period k to be sufficiently large to fully reflect the dynamic characteristics of the constructed model. Insufficient lag length induces pronounced residual autocorrelation, thereby compromising the consistency of coefficient estimations. Conversely, excessive lag orders escalate model complexity by expanding the parameter space while diminishing available degrees of freedom, ultimately undermining the statistical robustness of the inferred relationships. Therefore, a balance needs to be struck between the lag period and degrees of freedom. Generally, there are several methods for selecting the lag period k in a VAR model:

(1) LB statistic

LB stands for Likelihood Ratio, and its statistic is defined as:

$$LB = -2\left[\log L(k) - \log L(k+1)\right] \sim \chi(N^2) \tag{16}$$

If extending the temporal depth k in a Vector Autoregression framework fails to produce a statistically meaningful improvement in the log-likelihood value — specifically when the Ljung-Box (LB) test statistic falls below the predefined significance threshold — the augmented lags provide no explanatory power and should be excluded from the final specification.

(2) AIC criterion

$$AIC = \log\left[\frac{\sum_{k=1}^{T}\hat{u}_k^2}{T}\right] + \frac{2k}{T} \tag{17}$$

where $\hat{u}_k^2$ represents the sum of squared residuals, T represents the sample size, and k represents the maximum lag order. The principle for selecting the lag order is to choose the k that minimizes the value of AIC as k is increased.

(3) SC criterion

The formula for the SC criterion is:

$$SC = \log\left[\frac{\sum_{t=1}^{r}\hat{u}_t^2}{T}\right] + \frac{k\log T}{T} \tag{18}$$

The principle for selecting the lag order using the SC criterion is the same as that for the AIC criterion, which is to choose the k that minimizes the value of SC as k is increased.

(4) BIC criterion

$$BIC = \log\left[\frac{\sum_{t=1}^{T}\hat{u}_t^2}{T}\right] + \frac{k}{T}\ln T \tag{19}$$

The meanings of the parameters and the method for selecting the lag period are the same as those for the AIC criterion.

3. Results

3.1. The entropy weight method independently assigns weights to each indicator in the feline and canine markets

The weights of the five indicators for the cat market are obtained as Table.1:

Table.1. Weight of five indicators in the cat market

index	The number of pet cats in China	number of urban cat owners	retail sales of cat food	annual average consumption amount per pet cat	market size of cat food
weight	0.2412	0.1457	0.1714	0.1168	0.3249

The weights of the five indicators for the dog market are obtained as Table.2:

Table.2. Weight of five indicators in the dog market

index	Number of dog owners in the city	Retail sales of dog food	Dog food market size	Average annual consumption amount of a single pet dog	Number of pet dogs in China

weight	0.2974	0.1286	0.1017	0.1627	0.3096

3.2. The TOPSIS method is used to calculate the scores for the cat and dog markets over the past five years, respectively

The final evaluation scores for the cat and dog markets over the past five years are presented in Table.3:

Table.3. Evaluation scores of the cat and dog market in the past five years

Development of the cat market	Evaluation	Development of the dog market	Evaluation
2019	0.2716	2019	0.7588
2020	0.4249	2020	0.5553
2021	0.6482	2021	0.6753
2022	0.8482	2022	0.4562
2023	0.8071	2023	0.5545

3.3. The application of Spearman's rank correlation coefficient

This study identifies a framework of 26 metrics spanning five key dimensions to systematically assess the growth dynamics of China's companion animal sector. The five aspects are economy, society, business conditions, domestic pet ownership situation, and the pet industry situation. The 26 indicators include the growth rate of newly established enterprises, the growth rate of financing events, the number of pet cats in China (in ten thousand), the pet ownership rate in second-tier cities, per capita GDP, per capita disposable income, national GDP, the proportion of the population aged 65 and over, the crude marriage rate, the birth rate, etc.

Using MATLAB, the Spearman correlation coefficients between the 26 indicators across 5 aspects and the total marketing revenue of the pet market are calculated, as shown in Table.4. (Due to space limitations, this paper presents the correlation of the first five indicators with the development of China's pet sector):

Table.4. Spearman correlation coefficient

Spearman correlation coefficient					
index	GPT per capita	Proportion of popu- lation aged 65 and above	Number of pet dogs in China	Financing event growth rate	Growth rate of newly established enterprises
Correlation coefficient	0.9612	-0.8808	-0.7209	-0.4254	-0.174

And after removing the indicators with a correlation coefficient lower than 0.9 among the 26 indicators, we obtained the final 17 indicators for evaluating the China's development of pet sector. Due to space limitations, only the analysis of the first five indicators is presented in the main text:

I. Per capita GPT exhibits a robust positive association with the growth of China's pet sector, as reflected by a Pearson's correlation value of 0.9833. This article speculates that this is mainly because as per capita GPT increases, the standard of living improves, and pets become members of more and more households, driving the rapid development of the pet industry.

II. This paper speculates that this is mainly because an increase in per capita disposable income enhances residents' willingness to own pets and their consumption capacity, thereby promoting the rapid growth of the pet industry.

III. A strong positive association is observed between the elderly demographic share (aged 65+) and the expansion of China's companion animal sector, highlighting aging populations as a key driver of market growth. This paper speculates that this is mainly because the increase in the proportion of the elderly population increases the emotional demand for pets, driving the rapid development of the pet industry.

IV. A statistically significant positive relationship (r = 0.974) is observed between marriage rates and the growth trajectory of China's pet care market, suggesting marital trends may influence consumer behaviors in this sector. This article speculates that this is mainly because a decline in the marriage rate leads to an increase in the proportion of empty-nest elders and single youths, which in turn generates more demand for pet companionship and promotes the development of the pet industry.

V. A robust positive association (r = 0.9676) is observed between declining birth rates and the accelerated growth of China's companion animal sector. This trend may be attributed to shrinking household sizes, which elevate the societal value of pets as surrogate family members, thereby catalyzing market expansion through increased demand for emotional companionship services.

This paper utilizes MATLAB to conduct the Spearman coefficient test, with the results presented in Table.5 (due to space limitations, this paper shows the p-value test results for the first five indicators and their correlation):

Table.5. P-value test

P-value test					
index	GPT per capita	Proportion of popu- lation aged 65 and above	Number of pet dogs in China	Financing event growth rate	Growth rate of newly established enterprises
Correlation coefficient	0.000	-0.000	-0.0000	-0.0000	-0.000

As shown in the Table.5, the p-value for any indicator and the development of China's pet industry is 0, so we accept the null hypothesis and can consider that there is a significant correlation between any remaining indicator and China's pet sector development.

3.4. ARIMA forecasting of the development of China's pet industry over the next three years

The forecasting results are shown in Table.6 (due to space limitations, only the forecasting results for the first five indicators are presented in this paper).

Table.6. Forecast results for the next three years

	2024	2025	2026
GPT per capita	96940	104860	112780
Per capita disposable income	39000	37000	35000
Proportion of population aged 65 and above	0.1345	0.14426	0.15549
Crude marriage rate	0.159	0.162	0.165
Birth rate	0.0687	0.0763	0.0927

3.5. K-means++ clustering to evaluate the situation of the pet market development

The clustering results for the development situation of the cat and dog markets, obtained using MATLAB, are shown in Table.7:

Table.7. Cluster results of the development of the cat and dog market

Cluster center			
	Development situation of the cat market	Development situation of the dog market	category
Cluster center 1	0.8276	0.7588	Rapid development
Cluster center 2	0.5366	0.6153	Steady development
Cluster center 3	0.2716	0.5053	Slow development

So the development situation of the cat and dog markets over the past five years is presented in Table.8:

Table.8. Results of the development of the cat and dog market in the past five years

Cat market development	evaluation	Cat market development	evaluation
2019	Slow development	2019	Rapid development
2020	Steady development	2020	Slow development
2021	Steady development	2021	Steady development
2022	Rapid development	2022	Slow development
2023	Rapid development	2023	Slow development

The clustering results for the situation of China's pet sector, obtained using MATLAB, are shown in Table.9:

Table.9. cluster center

Cluster center		
	Development situation of China's pet industry	category
Cluster center 1	0.8300	Rapid development
Cluster center 2	0.6540	Steady development
Cluster center 3	0.4450	Slow development

The evaluation of the pet market development situation through k-means++ clustering is shown in Table.10:

Table.10. Evaluation of the Development of China's Pet Sector in Past Five Years and the Next Three Years

An evaluation of the development of China's pet sector over the past five years and the next two years			
2019	Slow development	2023	Rapid development
2020	Slow development	2024	Steady development
2021	Steady development	2025	Steady development
2022	Steady development	2026	Steady development

Based on the evaluation indicators of China's pet industry development, this article concludes that the reasons for the increasingly better development of China's pet industry since 2019 are as follows:

i. Economic growth and improved consumption capacity have enabled residents to have more disposable income for pet consumption. At the same time, demographic changes such as aging, decreasing birth rates, and increasing singlehood have also driven the growth in pet demand.

ii. Improvements in the policy environment have provided a more relaxed environment for the development of the pet industry, while strengthened regulation has promoted healthy, orderly, and high-quality development of the industry.

iii. In addition, intensified competition within the industry has prompted companies to pay more attention to brand building and product research and development. The satisfaction of market segmentation and personalized demands has also improved the overall service quality and product competitiveness of the industry.

3.6. The solution of the VAR model

This paper selects EU tariff rates, the US dollar exchange rate, and China's foreign exchange reserves to study the impact of new foreign economic policies of European and American countries on China's pet food industry.

i.suitability test

Before conducting factor analysis, an adaptability test for the factor analysis method is performed using the correlation coefficient matrix among indicators and the KMO test. Table.11 presents the results of the KMO and Bartlett's test of sphericity:

Table.11. KMO test and Bartlett's test

KMO Test and Bartlett's Test of Sphericity		
KMO value		0.603
Bartlett's Test of Sphericity	Approximate Chi-Square	0.000
	df	190
	P	0.000***
Note: ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively		

From Table.11, it can be seen that the observed value of the Bartlett's test of sphericity statistic is 190, and the corresponding probability P-value is close to 0. The significance level is set at $\alpha = 0.01$. Since the probability P-value is less than the significance level, the null hypothesis should be rejected, indicating that the correlation coefficient matrix is significantly different from the identity matrix. At the same time, the KMO value is 0.603. According to Kaiser's criteria for KMO measurement, the original variables are suitable for factor analysis.

ii.Extract factors

This paper adopts the maximum variance method to extract factors and selects three factors from them. The explanation of the total variance of the original variables by the factors is shown in Table.12:

Table.12. Explanation of Total Variance

Total Variance Explained						
component	The variance explanation ratio before rotation			The variance explanation ratio after rotation		
	eigenvalue	Variance explanation ratio (%)	Cumulative variance explanation rate (%)	eigenvalue	Variance explanation ratio (%)	Cumulative variance explanation rate (%)
1	16.596	82.979	82.979	1637.445	81.872	81.872
2	2.953	14.763	97.741	308.621	15.431	97.303
3	0.253	1.266	99.007	34.081	1.704	99.007

From Table.12, it can be seen that the three extracted factors encompass 99.007% of the information covered by the original indicators, with minimal information loss, indicating ideal factor extraction results. Figure 1 presents the scree plot, which shows that the eigenvalue of the first three factors is very high, contributing significantly to the explanation of the original variables, while the explanatory power of subsequent factors is relatively small.

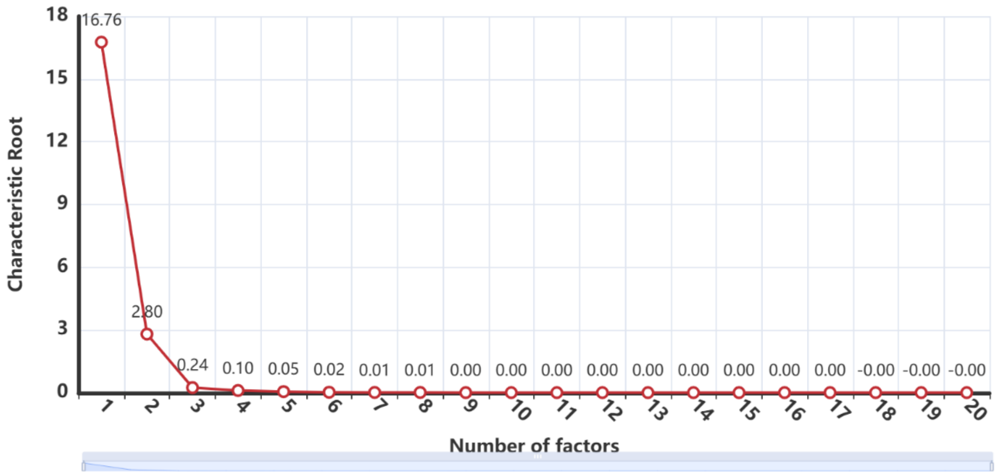
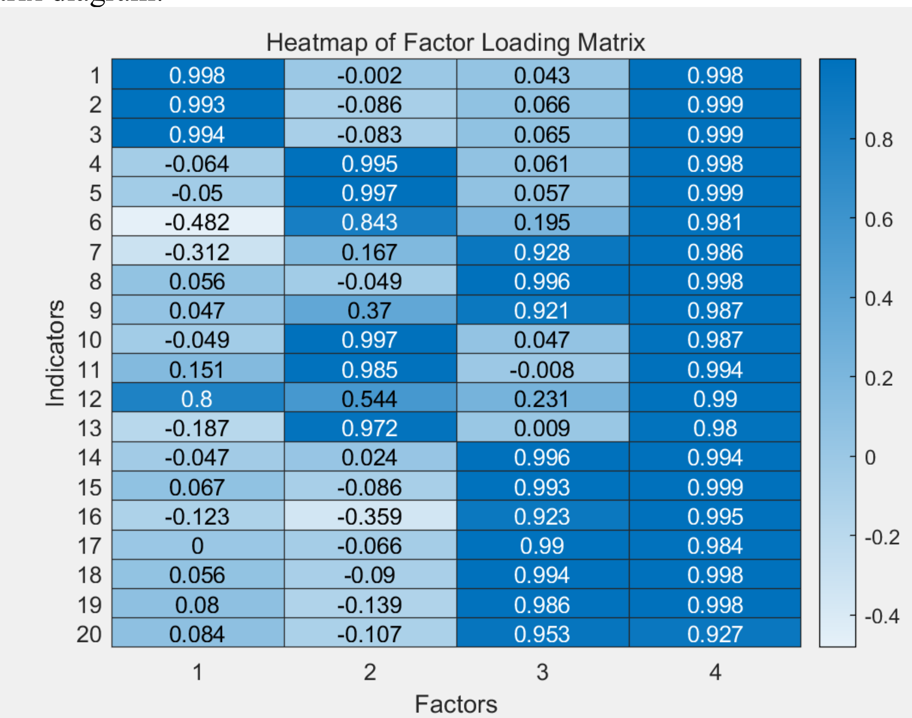


Figure 1. Scree plot

iii. Factor naming

Based on the maximum variance orthogonal rotation matrix, it can be seen that the main factor F1 is primarily composed of EU tariff rates, US tariff rates, and UK tariff rates. These three indicators have high factor loadings on F1, so F1 is named the "tariff policy factor". Similarly, F2 is named the "monetary policy factor", and F3 is named the "international trade factor". Figure 2 is the factor coefficient matrix diagram:



Indicators	1	2	3	4
1	0.998	-0.002	0.043	0.998
2	0.993	-0.086	0.066	0.999
3	0.994	-0.083	0.065	0.999
4	-0.064	0.995	0.061	0.998
5	-0.05	0.997	0.057	0.999
6	-0.482	0.843	0.195	0.981
7	-0.312	0.167	0.928	0.986
8	0.056	-0.049	0.996	0.998
9	0.047	0.37	0.921	0.987
10	-0.049	0.997	0.047	0.987
11	0.151	0.985	-0.008	0.994
12	0.8	0.544	0.231	0.99
13	-0.187	0.972	0.009	0.98
14	-0.047	0.024	0.996	0.994
15	0.067	-0.086	0.993	0.999
16	-0.123	-0.359	0.923	0.995
17	0	-0.066	0.99	0.984
18	0.056	-0.09	0.994	0.998
19	0.08	-0.139	0.986	0.998
20	0.084	-0.107	0.953	0.927

Figure 2. Coefficient matrix diagram of factors

The factor score coefficient matrix reflects the scoring of each main factor on every indicator.

(1) Stationarity Test

Using the ADF (Augmented Dickey-Fuller) unit root test method, the test results are shown in the Table.13:

Table.13. ADF unit root test results

variable	t	P	critical value 1%	5%	10%
F1	-6.189	0.000***	-3.964	-3.085	-2.682
F2	-4.321	0.000***	-4.012	-3.104	-2.691
F3	-5.753	0.000***	-3.964	-3.085	-2.682
The market size	-5.674	0.000***	-4.332	-3.233	-2.749

Note: ***, **, and * represent significance levels of 1%, 5%, and 10%

From Table.13, it can be seen that the absolute values of the ADF statistics for all variables are greater than the absolute values at the 5% critical level, meaning that the ADF statistic values fall within the rejection region, rejecting the null hypothesis. All three variables are stationary. Therefore, we can proceed to establish the vector autoregressive model.

(2) Solving the lag order for the VAR model

Table.14. Solving the Lag Order of Vector Autoregressive Model

lag order	logL	AIC	SC	HQ	FPE
0	-1959.889	219.694	219.89	219.714	2.5837696756525545e+95
1	-1809.479	217.333	218.299	217.383	2.656219932980063e+94
2	-1667.257	215.749*	217.449*	215.731*	1.0528965688167729e+94*

Based on the results of the four evaluation criteria: FPE, AIC, SC, and HQ, the lag order is ultimately selected as 2, meaning a VAR(2) model is established.

(3) Stability test of the VAR model

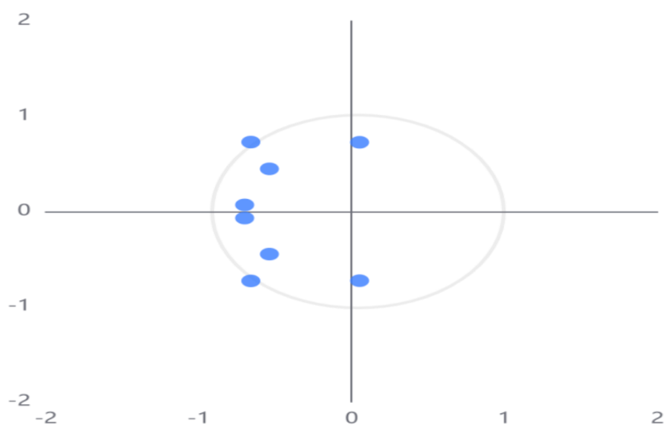
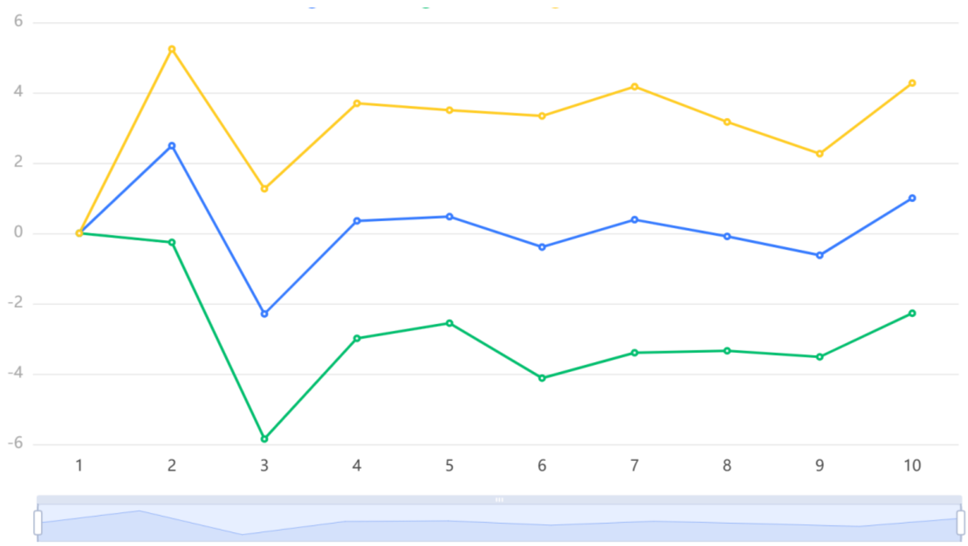


Figure 3. Stability testing of VAR model

From Figure 3, it can be seen that all the points lie within the unit circle, indicating that the VAR system is stable. Therefore, the model can proceed with impulse response analysis and variance decomposition.

(4) Establish the impulse response function



Explanation: Blue represents the response effect, green represents the 95% confidence interval (lower limit, LL), and yellow represents the 95% confidence interval (upper limit, UL)

Figure 4. Pulse response diagram from monetary policy

Figure 4 shows the response of China's pet food industry to a positive shock from monetary policy. It can be seen that after the shock, China's pet food industry rose by about 2.6 within 1-2 time units, then fell by about -2.3 within 2-3 time units, and subsequently rose slowly until it stabilized. The reasons for this are speculated as follows:

1. Ascending Phase

The depreciation of the RMB makes Chinese pet food more competitive in price on the international market, contributing to an increase in export volume. Customs data shows that during periods of RMB depreciation, both the export volume and export value of pet food have achieved significant growth.Domestic pet food enterprises, with most of their overseas orders settled in foreign currencies, benefit from exchange gains due to RMB depreciation, which in turn boosts their net profits.

2. Descending Phase

The depreciation of the Chinese yuan has led to an increase in the cost of importing raw materials and pet food, raising the production costs for enterprises.As import costs rise, some pet food brands that previously relied on imports may increase their prices or seek cheaper alternatives.

3. Stable Phase

As the government and the market gradually adjust and adapt, domestic pet food companies will find a new equilibrium point.As consumers' awareness of the quality and health of pet food continues to increase, they will pay more attention to the cost-effectiveness and brand influence of products.

The trends of the US dollar exchange rate and China's foreign exchange reserves are similar to the aforementioned results, so they will not be discussed individually in this paper.

4. Conclusions

This study develops an integrated analytical framework combining TOPSIS prioritization and VAR-driven dynamic analysis to evaluate growth patterns and consumer-driven requirements in the companion animal sector. This model combines the TOPSIS-entropy weight method and k-means++ clustering to qualitatively study the development of the pet market, enabling a good comparison between future development and past development. At the same time, the model adopts various mathematical modeling methods to accommodate different data types and trends, effectively addressing different issues with high flexibility. The VAR model has advantages such as comprehensiveness, systematicness, dynamic nature, predictability, flexibility, adaptability, and data-driven and empirical support in studying the impact of European and American foreign economic policies on the pet sector of development, helping to gain a deeper understanding of the internal laws and external influencing factors of industry development.

At the same time, this model also has some shortcomings. The model mainly focuses on tariff policies and changes in market demand, without considering other factors that may affect industry development, such as advancements in production technology and market competition. The VAR model in this paper can be improved to a SVAR model. By introducing contemporaneous relationships, the SVAR model reduces the correlation issues between error terms that may exist in the VAR model, thereby improving the accuracy of the model. The modeling approach can be extended to other industries that rely on exports and the international market (such as electronics and the automotive industry) to assess the comprehensive impact of policy changes and market demand on the industry.

References

[1] Salzman M. Pet trends: The state of the pet industry[J]. Vital speeches of the day, 2000, 67(5): 147.

[2] Zhang W, Cao H, Lin L. Analysis of the future development trend of the pet industry[C]//2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022). Atlantis Press, 2022: 1682-1689.

[3] Valdez J W. Using Google trends to determine current, past, and future trends in the reptile pet trade[J]. Animals, 2021, 11(3): 676.

[4] Xiao Y, Wang H H, Li J. A new market for pet food in China: Online consumer preferences and consumption[J]. The Chinese Economy, 2021, 54(6): 430-440.

[5] Schleicher M, Cash S B, Freeman L M. Determinants of pet food purchasing decisions[J]. The Canadian Veterinary Journal, 2019, 60(6): 644.

[6] Zhu Y, Tian D, Yan F. Effectiveness of entropy weight method in decision-making[J]. Mathematical problems in Engineering, 2020, 2020(1): 3564835.

[7] Olson D L. Comparison of weights in TOPSIS models[J]. Mathematical and computer modelling, 2004, 40(7-8): 721-727.

[8] Ren Q, Zhang D, Zhao X, et al. A novel hybrid method of lithology identification based on k-means++ algorithm and fuzzy decision tree[J]. Journal of Petroleum Science and Engineering, 2022, 208: 109681.

[9] Siami-Namini S, Tavakoli N, Namin A S. A comparison of ARIMA and LSTM in forecasting time series[C]//2018 17th IEEE international conference on machine learning and applications (ICMLA). Ieee, 2018: 1394-1401.

[10] Ding Y ,Zhu X ,Pan R , et al.Network Vector Autoregression with Time-Varying Nodal Influence[J].Computational E conomics,2025,(prepublish):1-27.