

# Analysis and Forecast of the Development of the Pet Industry in China and Globally Based on ARIMAX and Regression Models

Zhenjia Chang, Yuan Wei\*

College of Science, Liaoning Technical University, Fuxin, China

\*Corresponding author: 3146614525@qq.com

**Abstract.** This paper systematically analyses the development trends of the pet industry in China and globally based on the ARIMAX model and various regression models. The study first collects data on the number of pets in China and influencing factors such as GDP and population. It then uses the ARIMAX (p,d,q) model combined with the AIC criterion to determine parameters, constructs a multi-factor regression model to analyse key influencing factors, and predicts trends. For the global market, data from major regions were collected, and linear extrapolation and multiple linear regression models were used to analyse the mechanisms underlying key factors and predict demand scale. At the supply-demand level, linear and exponential regression models were employed to explore the interdependent relationships between domestic demand, export volumes, and the global market. Throughout the study, the cross-application of ARIMAX models and various regression models systematically revealed the key influencing factors and trends in this field, with model fitting results validating the reliability of the research conclusions.

**Keywords:** ARIMAX model; AIC criteria; multiple linear regression model; exponential regression model.

## 1. Introduction

The scale of China's pet industry continues to expand, while pet markets in major global economies also exhibit differentiated growth characteristics. However, existing research on the multi-factor driving mechanisms of pet industry development, cross-regional market linkage patterns, and quantitative analysis of future trends still requires further exploration, particularly in terms of systematic modelling that integrates time series and spatial dimensions [1].

This study takes the pet industry in China and globally as its research object, employing ARIMAX models and econometric methods such as multiple linear regression and exponential regression to construct a multi-dimensional analytical framework [2-3]. The research first uses the ARIMAX (p,d,q) model combined with the AIC criterion to optimise parameters, constructing a multi-factor regression model to analyse key influencing factors and predict trends [4]. Second, it employs linear extrapolation to predict baseline indicators and uses multiple linear regression models to analyse the driving factors behind regional differences. Finally, linear and exponential regression models are employed to explore the interplay mechanisms between domestic demand, export volume, and global market demand in China [5]. Through model cross-validation, the study systematically identifies the key influencing factors, regional differences, and future trends in the pet industry, providing data support for policy formulation and market strategy optimisation.

## 2. China's Pet Population Forecast

To accurately predict the development trends of China's pet industry, this paper collects data on the number of pet cats and dogs from 2009 to 2023, as well as data on influencing factors such as GDP, GDP per capita, population, urbanisation rate, and elderly population (aged 60 and above). Given that the industry's long-term expansion is closely related to the cumulative changes in the number of households, the population growth index can reflect the cumulative growth effect. Therefore, when

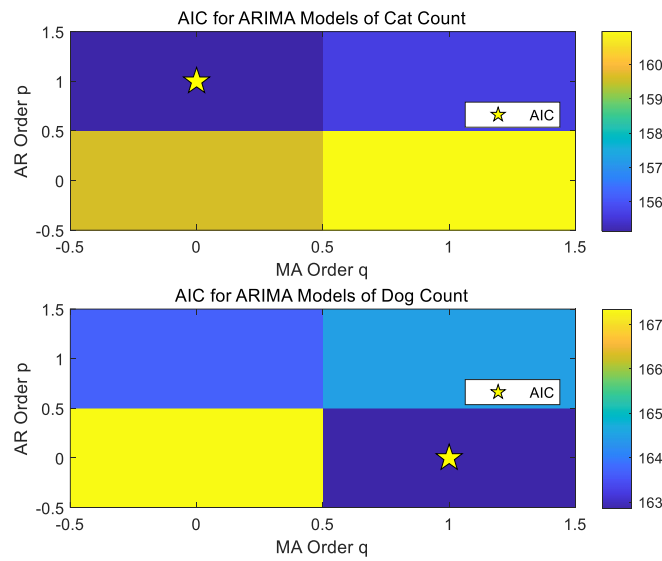
analysing long-term trends, it is recommended to use this index to effectively capture changes in the number of households and measure market demand.

## 2.1. Model Selection and Establishment

After preprocessing the data, this paper set up the ARIMAX (p,d,q) model for prediction. Firstly, due to the limited amount of data (15 data points), it is reasonable to limit the maximum order of AR and MA to 1 to avoid overfitting and calculation instability [6]. Then, for the selection of time parameters, AIC criterion is adopted:

$$AIC = -2\ln(\hat{L}) + 2k \quad (1)$$

Where  $\hat{L}$  is the maximum likelihood estimate of the model;  $k$  is the number of parameters of the model, that is, the number of influencing factors, including  $p$ ,  $q$  and constant terms. As shown in Fig. 1,  $AIC = 155.14$  when using cat data; When using dog data  $AIC = 162.86$ .



**Fig. 1** Calculation result of AIC

Available at this time  $p_{cat} = 0, q_{cat} = 1, p_{dog} = 1, q_{dog} = 0$ .

In addition, this study also conducted a correlation analysis of data on influencing factors such as GDP per capita, total population, urbanization rate of permanent resident population and elderly population (over 60 years old). The results showed that all four influencing factors had a correlation coefficient greater than 0.82 with the number of pet dogs and cats, indicating a strong correlation.

Therefore, this paper established a multi-factor regression model to explain the influence of external factors on the number of pet dogs and cats:

$$y_{Multi}(t) = \beta_0 + \beta_1 \cdot PGDP(t) + \beta_2 \cdot PGI(t) + \beta_3 \cdot Urban(t) + \beta_4 \cdot EPolu(t) + \varepsilon(t) \quad (2)$$

Where  $y_{Multi}(t)$  represents the number of pet dogs and cats in year  $t+2009$ ,  $PGDP(t)$  represents the per capita GDP in year  $t+2009$ ,  $PGI(t)$  represents the Population growth index in year  $t+2009$ ,  $Urban(t)$  represents the urbanization rate in year  $t+2009$ ,  $EPolu(t)$  represents the number of elderly population in year  $t+2009$ , and  $\varepsilon(t)$  is the random error.

Then the least square method is used to calculate the external variable coefficient  $\beta_k$ :

$$\beta_k = \frac{\sum_t (X_k(t) - \bar{X}_k)(y'(t) - \bar{y}')}{\sum_t (X_k(t) - \bar{X}_k)^2} \quad (3)$$

The comparative importance of coefficient standardization:

$$\beta_k^{norm} = \beta_k \cdot \frac{\sigma(X_k)}{\sigma(y)} \quad (4)$$

The data results indicate that for both cats and dogs, the order of their four factors is  $\beta_2 > \beta_1 > \beta_4 > \beta_3$ .

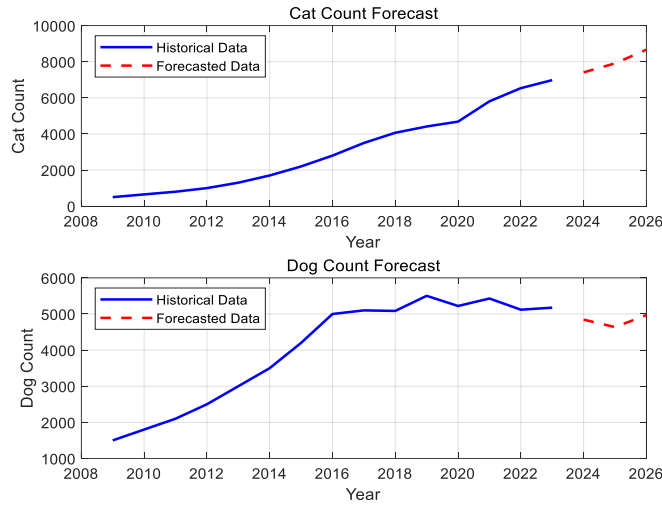
Therefore, the ARIMAX (p, d, q) model is established:

$$y(t) = c + \sum_{i=1}^p \phi_i y(t-i) + \sum_{j=1}^q \theta_j \varepsilon(t-j) + \sum_{k=1}^m \beta_k X_k(t) + \varepsilon(t) \quad (5)$$

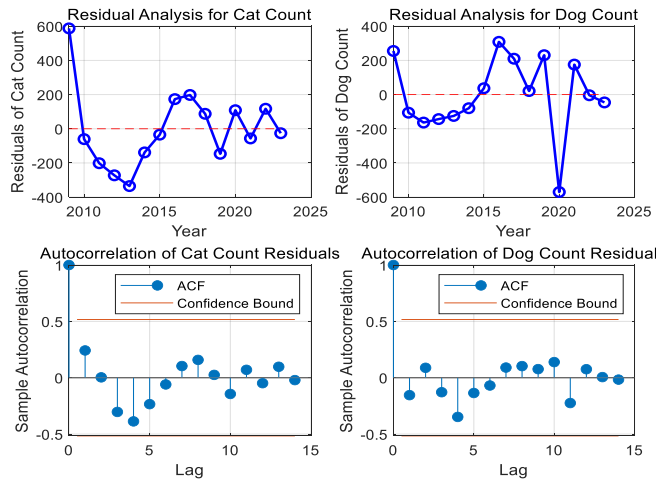
Where  $c$  is the constant,  $p$  is the autoregressive order,  $d$  is the difference order (determined to be order 1 by ADF test),  $q$  is the order of moving average.  $\sum_{i=1}^p \phi_i y(t-i)$  is the autoregressive part of the time series,  $\sum_{j=1}^q \theta_j \varepsilon(t-j)$  is the moving average part of the time series, and  $\sum_{k=1}^m \beta_k X_k(t)$  is the contribution of external variables (such as China's per capita GDP).

## 2.2. Prediction Results and Validation

This paper uses ARIMAX to forecast the number of pet dogs and cats from 2024 to 2026. The specific results are shown in Fig. 2.



**Fig. 2** Predicted values for cats and dogs



**Fig. 3** The residual map of cat and dog populations

From Fig. 3 above, it can be seen that the prediction results obtained in this paper are more appropriate. According to the search data, the number of cats and dogs in China accounts for about 82% of the total number of pets, so we can calculate the predicted value of the number of pets in China in the next three years: 2024 is 14945(10,000), 2025 is 15297(10,000), 2026 is 16628(10,000).

### 3. Global Pet Market Forecast for Major Countries

#### 3.1. Data Preprocessing

As for the changes of pet dog and cat data in the United States, France and Germany, this paper collected the data of their population and per capita GDP, hoping to analyze the influence mechanism behind the development of the pet industry through the data of these two dimensions:

Based on the above data, this paper calculated the growth rate of GDP:

$$GDP_{\text{Rate of increase}} = \frac{GDP_n - GDP_{n-1}}{GDP_{n-1}} \quad (6)$$

Where  $GDP_n$  and  $GDP_{n-1}$  represent the GDP of the last two years respectively.

According to the data this paper has collected,  $C_{catN} = 25\text{kg/piece}$ ,  $C_{dogN} = 50\text{kg/piece}$  and the annual food consumption of cats and dogs are related to the GDP growth rate, so the average annual food consumption of cats and dogs in the future is:

$$C_{cat} = C_{catN} \times (1 + k1 \times GDP_{\text{Rate of increase}}) \quad (7)$$

$$C_{dog} = C_{dogN} \times (1 + k2 \times GDP_{\text{Rate of increase}}) \quad (8)$$

Where  $k1$  is the adjustment coefficient of cat food consumption, the value is 0.1,  $k2$  is the adjustment coefficient of dog food consumption, and the value is 0.05.

Since this paper want to predict the number of pet cats and dogs in the future, and the number is related to the per capita GDP and population of different countries, this paper adopts the linear extrapolation method to predict the per capita GDP and population of different countries:

$$x_{i\text{year}} = x_{end} + (x_{end} - x_{start}) \times \frac{t_f - t_{end}}{t_{end} - t_{start}} \quad (9)$$

Where  $i$  is 2, the GDP of the country,  $i$  is 3, the population of the country,  $x_{end}$  is the  $x$  value of the last year,  $x_{start}$  is the  $x$  value of the starting year,  $t_f$  is the future year,  $t_{end}$  is the last year, and  $t_{start}$  is the starting year.

#### 3.2. The Establishment and Solution of the Model

##### 3.2.1. Prediction of future pet population

To predict the number of pets in the future, this paper used a multiple linear regression model with year, GDP, and population as independent variables, and the number of pet dogs and cats as dependent variables.

Based on the basic form of the multiple linear regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \quad (10)$$

Establish multiple linear regression equations for each country and pet dog and cat:

$$y_{p-c} = \beta_0 + \beta_1 \times \text{year} + \beta_2 GDP_{\text{country}} + \beta_3 Polu_{\text{country}} + \varepsilon \quad (11)$$

Where  $y_{p-c}$  represents the number of cats or dogs in different countries,  $\beta_0, \beta_1, \beta_2, \beta_3$  is the regression coefficient to be estimated,  $year$  represents the year,  $GDP_{country}$  represents the GDP of different countries,  $Polu_{country}$  represents the population of different countries, and  $\varepsilon$  represents the error term.

The regression parameters are estimated using the least squares method to minimize the residual sum of squares:

$$\min_{\beta_0, \beta_1, \beta_2, \beta_3} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i} - \beta_3 x_{3i}) \quad (12)$$

Using the above estimated parameters, the number of pets in the coming years is predicted:

$$\hat{y}_{year} = \hat{\beta}_0 + \hat{\beta}_1 x_{1,year} + \hat{\beta}_2 x_{2,year} + \hat{\beta}_3 x_{3,year} \quad (13)$$

Where  $x_{1,year}$  is the future year,  $x_{2,year}$  is the future GDP predicted using the extrapolation method, and  $x_{3,year}$  is the future GDP predicted using the extrapolation method.

Based on the prediction results, it can be seen that the size of the pet market shows a steady upward trend, but it is relatively declining in Germany. North America is the largest region in the pet industry market, accounting for 38% of the global market in 2023. Followed by Europe, the market share of about 30%. The Asia Pacific region ranks third with a market share of about 23%.

### 3.2.2. Forecast global pet food demand for the next three years

Based on the predicted pet population and adjusted annual pet food consumption, the future pet food demand is calculated:

$$D_{p-c} = y_{p-c} \times C_{p-c} \quad (14)$$

Where  $D_{p-c}$  represents a country's food requirements for cats or dogs,  $y_{p-c}$  represents the number of pets this paper forecast above, and  $C_{p-c}$  represents the adjusted annual pet food consumption.

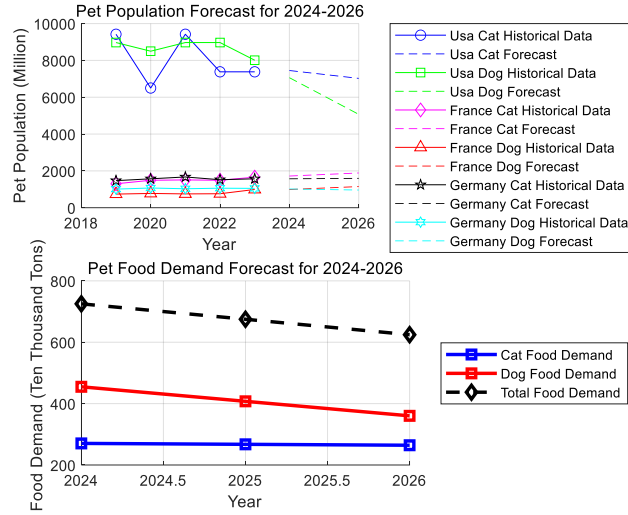
Therefore, the calculation formula of the total demand can be obtained as follows:

$$D_{total} = \sum_{country} (D_{cat-c} + D_{dog-c}) \quad (15)$$

The specific demand forecast results are shown in Table 1 and Fig. 4 below:

**Table 1.** Total demand forecast for pet food over the next three years

Year	Cat Food Demand (10,000 tons)	Dog Food Demand (10,000 tons)	Total Demand (10,000 tons)
2024	270.34	454.82	725.16
2025	267.36	407.49	674.86
2026	264.39	360.17	624.56



**Fig. 4** Predicted value

According to the population size and per capita GDP of the United States, France and German, their corresponding importance is calculated. By adding the corresponding degree to calculate the average, it can be calculated that the importance of population is 65.19%, and the importance of per capita GDP is 34.71%.

Therefore, this paper calculated the proportion of total global demand for pet food according to the proportion of population with greater importance:

$$\frac{\text{Population of the US, FRA and DE}}{\text{Global population}} = \frac{\text{total demand of US, FRA and DE}}{\text{total global demand}} \quad (16)$$

From this, the global demand for pet food can be calculated, as shown in Table 2.

**Table 2.** Global demand for pet food in the next three years

Year	2024	2025	2026
Total global cat and dog food demand (10,000 tons)	12066.78	12519.67	13779.63

According to the data this paper collected, 33 percent of animal owners worldwide own dogs and 23 percent own cats. The total global demand for pet food in the next three years is as follows: 2024 is 21547(10,000tons), 2025 is 22356(10,000tons), 2026 is 24605(10,000 tons).

#### 4. Supply and Demand Linkage Analysis: Two-way Influence between China and Global Markets

This paper modeled the number of pets in the next three years and the global demand for pet food in the next three years.

In this case, this paper model the relationship between the number of pets in China and the domestic demand in China as a linear function, using a linear regression model:

$$y = a \cdot x + b \quad (17)$$

Where  $y$  is China's domestic demand,  $x$  is the number of pets in China, and  $a, b$  are regression coefficients.

This paper solved for the optimal regression parameter by minimizing the sum of squares of error:

$$RSS = \sum_{i=1}^n (y_i - (ax_i + b))^2 \quad (18)$$

Taking the derivatives of  $a, b$  respectively and making the derivatives zero, the regression coefficients  $a=0.048218$  and  $b=641.0635$  are obtained. The estimate of the regression coefficients  $a$  and  $b$  is the optimal solution that minimizes the sum of squares of the errors.

This paper assumed the relationship between global demand and China's export volume as an exponential form, and then build an exponential regression mode:

$$y = a \cdot e^{b \cdot x} \tag{19}$$

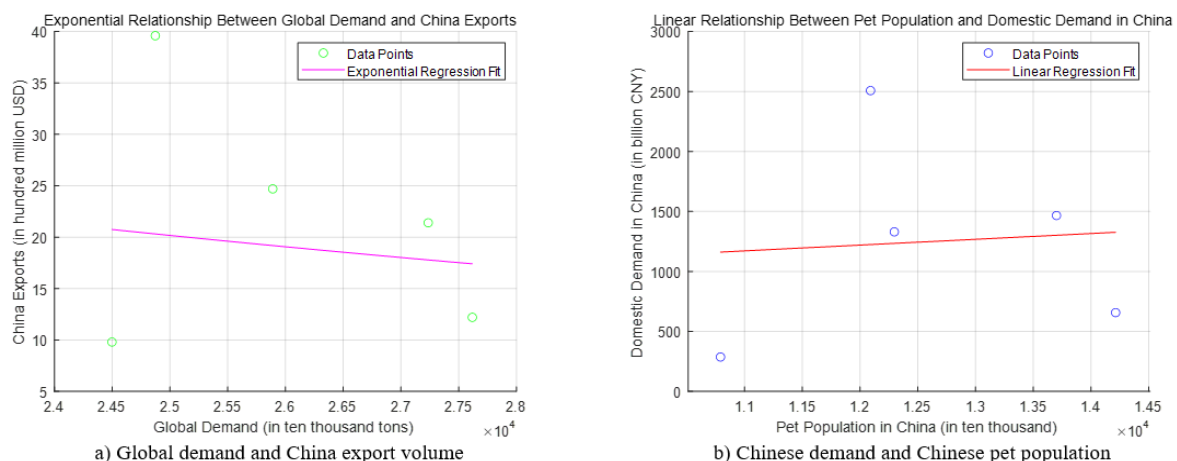
Where  $y$  is China's export volume,  $x$  is global demand, and  $b$  are regression coefficients. The regression coefficient of the exponential model  $a=82.3104$  and  $b=-0.000056255$  can be obtained by bringing in the data for solving.

**Table 3.** Forecast value of domestic and foreign pet industry in the next three years

Years	The Number of Pets in China (10,000)	The Global Demand for Pet Food (10,000 tons)	The Demand for Pet Food in China (10,000 tons)
2024	14945	21547	1361.7
2025	15297	22356	1378.6
2026	16628	24605	1442.8

According to the Table 3, it can be seen that from 2024 to 2026, the number of pets in China will show an increasing trend year by year, reflecting the continuous increase in the popularity of pet raising in China. The global demand for pet food is also increasing year by year, indicating that the global pet food market is constantly expanding. China's pet food demand is also rising year by year, reflecting the increasing number of pets in China and the upgrading of pet food consumption by pet owners and other factors, domestic demand for pet food is increasingly strong. However, the export volume of pet food in China has shown a year-on-year decline, which may be related to factors such as intensifying competition in the international market and expanding domestic market demand, making some enterprises shift their sales focus more to the domestic market. The total production of pet food in China is on the rise as a whole, indicating that the domestic pet food industry production scale is constantly expanding to meet the growth of domestic market demand and maintain a certain export scale.

As shown in Fig. 5, the fitting effect is better, so the reliability of the predicted data is higher.



**Fig. 5** The fitting results

To sum up, the number of pets in China, the domestic demand for pet food and the total production of pet food are growing, but the export volume of pet food has declined, while the global demand for pet food has continued to expand, reflecting that China's pet food industry has greater development potential in the domestic market, while facing certain competition and adjustment pressure in the international market.

## 5. Summary

This study systematically analysed the development trends and driving mechanisms of the pet industry in China and globally through the cross-application of ARIMAX models, multiple linear regression models, exponential regression models, and other models. The study first employs the ARIMAX (p,d,q) model combined with the AIC criterion to optimise parameters, constructing a multi-factor regression model to analyse key influencing factors and predict that the number of pets in China will continue to grow in the future. For the global market, linear extrapolation is used to predict baseline indicators, and a multiple linear regression model is employed to reveal the expansion trend and regional differences in global pet food demand. Finally, linear and exponential regression models are used to explore the interplay mechanisms between domestic demand, export volume, and global market demand in China. Through model cross-validation, the study systematically revealed the key influencing factors, regional differences, and future trends in the pet industry, providing data support for policy formulation and market strategy optimisation. Future research could further incorporate more regional data to deepen cross-scale driver analysis, providing more comprehensive theoretical support for the industry's globalisation strategy.

## References

- [1] Wang Bing. Analysis of the Current Status and Trends in China's Pet Market [J]. *Guizhou Animal Husbandry and Veterinary Medicine*, 2023, 47 (02): 6-8.
- [2] Wang Xiangqian, Wu Donglong. An Improved ARIMAX Method for Predicting Cargo Throughput: A Case Study of Tianjin Port [J]. *Operations Research and Management*, 2022, 31 (03): 138-144.
- [3] Xing, Yanchun, Gao, Tengfei. Application of the ARIMAX multivariate time series and BP neural network combination model in predicting the structure of resident consumption [J]. *Journal of Jilin Normal University (Natural Science Edition)*, 2021, 42 (03): 56-67.
- [4] Zhao Panpan, Zhang Mi, Jiao Libin. Prediction and Analysis of China's Tax Revenue Based on the ARIMAX Model [J]. *Journal of Higher Education Science*, 2024, 44 (05): 22-27.
- [5] Wang Min, Wu Yubo, Min Fan. A Multi-Classification Active Learning Algorithm Based on Multiple Clustering Algorithms and Multiple Linear Regression [J]. *Computer Applications*, 2020, 40 (12): 3437-3444.
- [6] Du Jie, Gao Shan, Jin Xinyue. Analysis of China's GDP Forecast Based on the ARIMAX Model [J]. *Journal of Fuyang Normal University (Natural Science Edition)*, 2020, 37 (01): 1-5.