Predict and Analyze Exchange Rate Fluctuations Accordingly Based on Quantile Regression Model and K-Nearest Neighbor

Dongyang Zhang¹,*, Xiaojie Wang¹, Liyuan Gao¹ and Yicheng Gong¹,²

¹Department of Mathematics and Statistics, Science College, Wuhan University of Science and Technology, Wuhan 430065, China
²Hubei Province Key Laboratory of Systems Science in Metallurgical Process, Wuhan University of Science and Technology, Wuhan 430065, China

*Corresponding author email: 1044600738@qq.com

Abstract. The exchange rate is one of the important indicators to measure macroeconomic fluctuations. The exchange rate of the US dollar against the RMB reflects the currency exchange rates of the world’s two largest economies. As a global emergency, the new coronavirus pneumonia (Corona Virus Disease 2019, COVID-19) has a significant impact on all aspects of society, so fluctuations in the exchange rate market cannot be ignored. This article uses the daily exchange rate of USD to RMB published on the official website of China Merchants Bank as the dependent variable representing the economic situation. The daily new confirmed cases of new coronavirus pneumonia confirmed by the National Health Commission, the number of new deaths each day, and the number of new cured cases each day are the influencing variables, and the Shanghai Interbank Offered Rate is used as the control variable. In order to explore the relationship between variables and predict and analyze exchange rate fluctuations accordingly, multiple linear regression and quantile regression are used for impact analysis. Comparing the prediction results of the multiple linear regression and the K nearest neighbor model, it is found that daily new confirmed cases, daily new cured cases, and market interest rates harm the exchange rate. At 0.5 quintile, the number of newly diagnosed cases per day increases by 1 unit, and the exchange rate of the U.S. dollar to the renminbi will correspondingly decrease by 0.00001%. At the 0.75 quantiles, new cured cases will change by one unit every day, and the exchange rate will change. The percentage is -0.00003%, the market interest rate changes by one unit, and the percentage of exchange rate changes are -0.037%, which is significant. And after comparison, it is found that the prediction effect obtained by the K nearest neighbor model is better than the multiple linear regression model. At the beginning of the epidemic, the exchange rate fluctuated by 1.78%. After the government took a series of prevention and control measures against major emergencies, the epidemic was effectively controlled, worries gradually eased, and the exchange rate has returned to a relatively stable state.

1. Introduction
The new coronavirus that broke out in December 2019 not only has an impact on China’s economy, but the level of concern among market participants about the spread of the epidemic is reflected in the fluctuations in foreign exchange prices. Ying Huizun [1] (2003) compared the number of overseas tourists received by Beijing during the SARS period and obtained the importance of predicting the accuracy of the time when the epidemic is under control for the rapid economic recovery. Zhang Wendou [2] (2012) evaluated the economic losses caused by the epidemic through the analysis of the past outbreaks of sudden infectious diseases. Liu Xueliang [3] (2020) analyzed the possible negative
impact of the existing new coronavirus on the national economic growth based on the impact of SARS on my country's macro economy. Li Xintong\(^4\) (2020) analyzed the situation of new coronavirus pneumonia at home and abroad and the adverse and beneficial effects of the epidemic on my country's economy. Alan Siu\(^5\) (2004) analyzed the spread of SARS in Hong Kong and discussed its impact on the economy.

Han Yujin\(^6\) (2020) took the data of Beijing and Shanghai as examples and analyzed the total retail sales of consumer goods, a macroeconomic indicator, to calculate the impact of the new coronavirus on this indicator. Liu Xueliang\(^3\) (2020) took SARS as an example, and analyzed the possible negative impact of the epidemic on my country's macro economy based on the impact of SARS on my country's macro economy in 2003. Ye Chenxi\(^7\) (2012) took SARS and Type A H1N1 as examples to estimate the economic losses caused by the epidemic to my country's tourism industry based on the trend line model equation. He Jinlei\(^8\) (2004) established a time series model for the 2003 SARS data and constructed a prediction formula. The study of economic prediction models for health emergencies can provide a reasonable and effective basis for my country's economic recovery and the government's macro-control of the market.

2. Data Pre-processing and Description

This paper selects China Merchants Bank's official website (http://fx.cmbchina.com/Hq/) from January 21, 2020 to August 25, 2020, as the explained variable. Newly, confirmed COVID-19 cases, new deaths, and new cured cases come from the national epidemic statistics released by the National Health Commission collected by the Kesci.com, (https://www.kesci.com/home/dataset/5e68da6ecdf64e002c97b3a8). Interest rate-related data is extracted from the Shanghai Interbank Offered Rate (http://www.shibor.org/shibor/web/DataService.jsp) announced by the China Foreign Exchange Trade System.

| Variable abbreviation | Variable definitions | Mean    | Standard deviation |
|-----------------------|----------------------|---------|--------------------|
| Exchange rate (UR)    | USD to RMB exchange rate | 7.034   | 0.063              |
|                       | Daily new confirmed cases | 393.200 | 1301.068          |
| Confirmed (Dnc)       | New deaths per day    | 15.790  | 35.759             |
|                       | Daily new cured cases | 372.600 | 720.860            |
| death (Ndp)           | Shanghai Interbank Offered Rate | 1.558   | 0.465              |
| cure (Ddc)            |                       |         |                    |
| interest rate (Shibor) |                      |         |                    |

It can be seen from Table 1 that the mean value of the nominal exchange rate of USD to RMB is 7.034 USD and the standard deviation is 0.063. The average daily number of newly confirmed cases was 393.200, and the standard deviation was 1301.068. The average number of new deaths per day was
15.790, and the standard deviation was 35.759. The average daily number of newly cured cases was 372.600, and the standard deviation was 720.860. The average daily interest rate is 1.558 cases and the standard deviation is 0.465.

With the rapid increase in the number of people infected with the new coronavirus, the exchange rate of the US dollar against the renminbi rose from a minimum of 6.903 on January 21, 2020, to a high of 7.026 on February 3, and the fluctuation range of the exchange rate reached 1.78 in two weeks. % (Figure 2), market participants’ worries about the virus are clearly reflected in the volatility of foreign exchange prices. Even if the virus is not a single factor in foreign exchange changes, it is also one of the main factors of this volatility.

3. Analysis of Exchange Rate Changes Under the Impact of the Epidemic

3.1. Multiple Linear Regression Analysis

In order to reduce the probability of multicollinearity, first of all, the nominal exchange rate (UR) of USD to RMB, the daily new confirmed cases (Dnc), the daily new death cases (Ndp), and the daily new cured Spearman correlation analysis was performed on five variables of the case (Ddc) and daily interest rate (Shibor), and the analysis results are shown in Figure 2.

![Figure 2. Spearman analysis chart of factors affecting the exchange rate.](image)

After Spearman correlation analysis, combined with the degree of interaction between the variables, finally three variables Dnc, Ddc and Shibor were selected from the four independent variables for multiple linear regression analysis, so the multiple linear regression model between the exchange rate and the influencing factors for:

$$\log(UR) = l_0 + l_1Dnc + l_2Ddc + l_3Shibor + \varepsilon$$  \hspace{1cm} (1)

Where is the regression coefficient and is the random error. The coefficients of each explanatory variable in the model are obtained by the OLS method, and the influence of Dnc, Ddc and Shibor on exchange rate can be further analyzed according to the regression results. The results are shown in Table 2.

| Explanatory variables | Coefficient estimate | 95%CI               |
|-----------------------|----------------------|---------------------|
| c                     | 1.968***             | [1.965 ~ 1.971]     |
| Dnc                   | -6.996e-07.          | [-7.735e-07 ~ 6.257e-07] |
| Ddc                   | -3.523e-06***        | [-4.867e-06 ~ -2.179e-06] |
| Shibor                | -1.040e-02***        | [-1.243e-02 ~ -8.362e-03] |
| R2                    | 0.389                |                     |
| Adjusted R2           | 0.3805               |                     |
| F-statistic           | 45.43***             |                     |
It can be seen from Table 2 that the significance level is $\alpha = 0.05$, and the statistics of the model F test, $P < 0.05$ indicates that the linear correlation between the label variable and the independent variable of the regression model as a whole is significant. In order to analyze the impact of explanatory variables on the exchange rate of USD to RMB at different levels in a more stable and comprehensive analysis. We constructed a multiple quantile regression model for further analysis.

### 3.2. Multiple Quantile Regression Analysis

This paper uses the "quantreg" package of R-4.0.0 to perform quantile regression, and performs regression model analysis on the 3 quantile points at 25%, 50%, and 75%. The quantile regression results are shown in Table 3 Show.

#### Table 3. Analysis of multiple quantile regression results.

| Explanatory Variables | $q=0.25$ | $q=0.5$ | $q=0.75$ |
|-----------------------|----------|----------|----------|
| Coefficient estimate (95% CI) | 7.163 [7.135 ~ 7.175] | 7.162 [7.144 ~ 7.185] | 7.143 [7.129 ~ 7.159] |
| $D_{nc}$ | 0.000 [-1e-05 ~ 0.000] | -1e-05 [-2e-05 ~ -1e-05] | -1e-05 [-2e-05 ~ -3e-05] |
| $D_{dc}$ | -2e-05 [-4e-05 ~ -1e-05] | -2e-05 [-2e-05 ~ -1e-05] | -3e-05 [-3e-05 ~ -2e-05] |
| Shibori | -0.101 [-0.107 ~ -0.085] | -0.080 [-0.094 ~ -0.077] | -0.037 [-0.051 ~ -0.025] |
| Coefficient estimate (95% CI) | 7.163 [7.135 ~ 7.175] | 7.162 [7.144 ~ 7.185] | 7.143 [7.129 ~ 7.159] |
| Coefficient estimate (95% CI) | 7.163 [7.135 ~ 7.175] | 7.162 [7.144 ~ 7.185] | 7.143 [7.129 ~ 7.159] |

When the quantile point is 0.5, the number of newly diagnosed cases per day increases by one unit, and the exchange rate of USD to RMB will correspondingly decrease by 0.00001%, with a significant coefficient. In addition, when the quantile point is higher, the number of newly cured cases changes by one unit every day, and the greater the percentage of exchange rate changes. When the quantile is 0.25, the number of newly cured cases changes by one unit every day, and the percentage of exchange rate changes is -0.00002%, with a significant coefficient. When the quantile is 0.75, the number of newly cured cases changes by one unit every day, and the percentage of exchange rate changes is -0.00003%, with a significant coefficient. The appreciation of the renminbi will weaken the international competitiveness of exported goods and increase domestic employment pressure. Taking effective measures to reduce the emergence of new cases and control the growth of cumulative confirmed cases is crucial to ensuring the stable development of my country’s economy.

The regression coefficients of market interest rates are significant at all three quantiles and all are negative. This result shows that market interest rates have a significant negative impact on exchange rates. And when the quantile points are 0.25 and 0.75, the market interest rate changes by one unit, and the exchange rate changes by -0.101% and -0.037% respectively. The higher the quantile point, the smaller the negative impact of market interest rates on the exchange rate.

### 4. Exchange Rate Prediction Model under the Impact of the Epidemic

#### 4.1. Forecast Based on A Linear Regression Model

The software environment for all prediction models in this paper, that is, the running operating system environment is mac OS; the processor is 1.4 GHz Intel Core i5, and the installed memory (RAM) is 8.00GB. Use Jupyter Notebook in Anaconda3 to run the algorithm program on a stand-alone basis.

Creating a linear regression model. Although there are tunable hyperparameters, they have limited significance for the data selected in this paper. Therefore, this paper does not perform hyperparameter tuning on the linear regression model. Since the prediction in this paper is a regression problem, $R^2 (r2\_score)$ is selected as the model performance index.

The formal expression of the function is shown in equation (2).

$$ UR = -5.141e - 06* D_{nc} + 1.041e - 05N_{dp} - 2.505e - 05D_{dc} - 7.304e - 02Shibor + 7.1588799254768825 $$  

From equation (2), we can get the linear relationship between UR and Dnc, Ndp, Ddc and Shibor. The prediction effect of the multiple linear regression model constructed according to formula (2)
corresponds to $R^2$ of 0.389, mean square error of 0.002, and root mean square error of 0.049. Among them, part of the forecast data is shown in Table 4.

Table 4. Comparison of partial prediction results of multiple linear regression model with real values.

| Actual value | Predictive value | Actual value | Predictive value |
|--------------|------------------|--------------|------------------|
| 6.90335      | 7.02038183       | ………………   | ………………       |
| 6.90275      | 7.04624006       | 6.92165      | 7.0006905        |
| 6.93475      | 6.9786158        | 6.9216       | 6.99863681       |
| 6.94465      | 7.00292054       | 6.9227       | 6.99278221       |
| 6.94465      | 7.00148047       | 6.9227       | 6.99996393       |
| 6.94465      | 6.97896158       | 6.9216       | 6.9982389        |
| ………………   | ………………       | 6.91455      | 6.99982389       |
| 6.94465      | 7.00138326       | 6.91225      | 6.98510228       |
| 6.94465      | 7.00138326       | 6.91225      | 6.98510228       |
| ………………   | ………………       | 6.91455      | 6.99982389       |

To make the prediction results clearer, this paper draws a histogram of the residuals of the prediction results of the multiple linear regression model constructed by equation (2), as shown in Figure 3 below.

![Figure 3. Histogram of the residuals of the multiple linear regression model.](image)

It can be seen from Figure 3 that the overall residual distribution is close to a normal distribution with a mean value of 0.

4.2. Prediction Based on the K-nearest Neighbor Model

Since the k-nearest neighbor model is based on the distance between samples, in order to prevent a certain independent variable from having an excessive influence on the distance calculation, it is first necessary to normalize the independent variable. The optimized hyper parameters are the number of neighbors (n_neighbors): 1, 3, and 7, and the distance norm (p): 1 and 2. At the same time, creating a grid search hyper parameter tuner. The hyper parameter tuning results are shown in Figure 4.

![Figure 4. K-nearest neighbor hyper parameter tuning diagram.](image)

In Figure 4, the abscissa represents the number of neighbors, and the ordinate represents the average accuracy of cross-validation. The blue line corresponds to a distance norm (p) of 2, and the orange line corresponds to a distance norm (p) of 3. According to Figure 5, it can be seen that in the interval [1,7],
the line with the distance norm (p) of 1 has been in a downward state, and the interval [1,3], it is in a slowly decreasing state, and 3 is the turning point. In the interval [3,7], it is in a state of rapid decline. In the interval [1,3], the line with the distance norm (p) of 2 is always rising, and 3 is the turning point. In the interval [3,7], it is in a rapid decline. Therefore, the optimal parameter combination is selected as the distance norm (p) is 1 and the number of neighbors (n_neighbors) is 1.

The prediction effect of the K-nearest neighbor model corresponds to $R^2$, the mean square error is 0, and the root mean-square error is 0. Among them, part of the forecast data is shown in Table 5 below.

| Actual value | Predictive value | Actual value | Predictive value |
|--------------|------------------|--------------|------------------|
| 6.90335      | 6.90335          | ..............| ..............    |
| 6.90275      | 6.90275          | 6.92165      | 6.92165          |
| 6.93475      | 6.93475          | 6.9216       | 6.9216           |
| 6.94465      | 6.94465          | 6.9227       | 6.9227           |
| 6.94465      | 6.94465          | 6.9227       | 6.9227           |
| 6.94465      | 6.94465          | 6.91455      | 6.91455          |
| ..............| ..............    | 6.91225      | 6.91225          |

5. Conclusion

In this paper, the exchange rate is used as a label variable representing economic conditions, and the relevant variables of the new crown pneumonia epidemic are used as influencing variables to predict changes in the exchange rate. Among the two selected prediction models, the K nearest neighbor algorithm has high prediction accuracy. In predicting the economic situation of future health emergencies, the K-nearest neighbor model will be more recommended.

Acknowledgments

This work was supported by two funds: 1. Hubei Province College Students’ Entrepreneurial Training Plan Program (201810488097) 2. Hubei Province Key Laboratory of Systems Science in Metallurgical Process (Y201906).

References

[1] Huizun Ying, Wei Luo. Analysis of the Impact of SARS on Economy[J].Journal of Mathematical Communications,2004(07):25.
[2] Wendou Zhang, Zhenghu Zu, Qing Xu. Research progress in the evaluation of economic losses in sudden infectious disease outbreaks[J]. Military Medicine, 2012(10): 797-800.
[3] Xueliang Liu, Xiaojing Zhang. Epidemic Impact and Economic Growth: An Empirical Analysis of SARS and the Potential Impact of New Coronary Pneumonia[J]. Industrial Economic Review, 2020(04): 05-25.
[4] Xintong Li, Yi Fang. Research on the impact of the new crown pneumonia epidemic on my country’s economy and countermeasures[J]. Journal of Changchun Finance College, 2020(04): 78-87.
[5] Alan Siu, Y.C. Richard Wong. Economic Impacts of SARS: The Case of Hong Kong.[J]. Asian Economic Reviews,2004(03):62.
[6] Yujin Han. How much impact will the new crown pneumonia epidemic have on the economy? Based on the perspective of the total retail sales of consumer goods[J]. Industrial Economic Review, 2020 (02): 13-23.
[7] Chenxi Ye, Shaoli Xu. Quantitative research on the impact of the epidemic on my country's tourism industry, taking SARS and H1N1 influenza as examples [J]. Journal of Leshan Normal University, 2012(03): 78-81.
[8] Jinlei He, Fuying Zhao, Xiaoyan Li, SARS Economic Forecast Model[J].