

# (Epidemiological Analysis of Varicella in Dalian from 2009 to 2019 and Application of Three Kinds of Model in Prediction Prevalence of Varicella) Study Design: Cross Sectional Study

**Tingting Cheng**

Dalian Medical University

**Yu Bai**

Dalian Center for Disease Control and Prevention

**Xianzhi Sun**

Dalian Medical University

**Yuchen Ji**

Dalian Medical University

**Fan Zhang**

Dalian Medical University

**Xiaofeng Li** (✉ [lxf\\_chen@dmu.edu.cn](mailto:lxf_chen@dmu.edu.cn))

Dalian Medical University

---

## Research Article

**Keywords:** varicella, epidemic characteristics, GM(1,1) model, Markov model GM(1,1)-Markov model

**Posted Date:** November 16th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1026752/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Objective** This study described the epidemic characteristics of varicella in Dalian from 2009 to 2019, explored the fitting effect of Grey model first-order one variable (GM(1,1)), Markov model, and GM(1,1)-Markov model on varicella data, and found the best fitting method for this type of data, to better predict the incidence trend.

**Methods:** In this study, the epidemiological characteristics of varicella from 2009 to 2019 were analyzed by epidemiological descriptive methods. Using the varicella incidence data from 2009 to 2018, predicted 2019 and compared with actual value. First made GM (1,1) prediction and Markov prediction. Then according to the relative error of the GM(1,1), made GM(1,1)-Markov prediction.

**Results:** This study collected 37223 cases from 2009 to 2019. The average annual incidence was 50.56/100000. Varicella occurred all year round, it had a bimodal distribution. The number of cases had two peaks from April to June and November to January of the following year. The ratio of males to females was 1.167:1. The 4 to 25 accounted for 60.36% of the total population. The age of varicella appeared to shift backward. Students, kindergarten children, scattered children accounted for about 64% of all cases. The GM(1,1) model prediction result of 2019 would be 53.6425, the relative error would be 14.42%, the Markov prediction result would be 56.2075, the relative error would be 10.33%, and the Gray(1,1)-Markov prediction result would be 59.508. The relative error would be 5.06%.

**Conclusions:** Varicella data had its unique development characteristics. The accuracy of GM (1,1) - Markov model is higher than GM(1.1) model and Markov model. The model can be used for prediction and decision guidance.

## 1. Background

Varicella is an acute infectious disease caused by the first infection with varicella-zoster virus (VZV). The incubation period is about 2 weeks, and some cases can reach 3 weeks. It is highly contagious from the day before the rash to full scab healing. The main transmission route is respiratory droplets or direct contact infection. Varicella patients are the only source of infection.

Varicella has a high infection rate in China. Varicella usually heals on its own. Individuals, infants, and adults with low immune function may be serious<sup>1</sup>. VZV is a teratogen and causes congenital varicella syndrome (CVS)<sup>2</sup> Moreover, VZV is a common cause of stroke<sup>3</sup>. Varicella has complicated complications and causing serious economic burden. Vaccines are the most economical method<sup>4</sup>.

The incidence of varicella in China ranks first among vaccine-preventable infectious diseases<sup>5</sup>. Studies found that children were vaccinated with varicella vaccine, the protective effect of the vaccine declined year by year, varicella would break out again<sup>6</sup>. Some severe breakthrough varicella can occur, but it does not seem to be common<sup>7</sup>. Research shows, the incidence of varicella has shown an overall upward trend<sup>8,9</sup>. Adults infected varicella easier and more dangerous. In many places, the age of varicella had

been found to shift backward<sup>10</sup>. Studies have found that adults in tropical and subtropical countries were more likely to suffer from varicella than in countries with mild climates<sup>11</sup>. Many factors like air pressure, temperature, humidity, rainfall, and other subtle factors can have a complex impact on the incidence of varicella<sup>12</sup>. There are differences in different regions, and the incidence of varicella is different. Find a model suitable for the varicella data in Dalian, explore the fitting effects of different models on the data, and provide a theoretical basis for varicella prediction and health decision-making.

The gray model was proposed by Professor Julong Deng in 1982<sup>13</sup>. It is a widely used data prediction model. It can suit various data types. It seeks valuable information by fully extracting and developing small samples and poor information data. And it generates predictions to demonstrate the direction of the system. Grey model first-order one variable (GM (1,1) ) is the core of the gray model and it was used widely. While in varicella prediction, the GM (1,1) was less used.

The Markov chain theory was proposed by a famous mathematics scientist A.A. Markov in 1906, The Markov model was maturely used in the economic field and played an important role in medical dynamics prediction<sup>14</sup>, Markov prediction means that in the process of data transfer, under the action of a certain factor, their state probability depends on the previous result, and the probability law of the Nth result depends only on the result of the (N-1)th experiment. And has nothing to do with the earlier results, the process is a random process<sup>15</sup>, Predict the state of the process at the next moment and the next few moments through the law of change between the states of the random process at different moments<sup>16</sup>.

The general experience of the Grey forecast dynamic model believes that its prediction accuracy for random, long-term, and volatile historical data is low, and it is mainly suitable for reflecting the overall development trend of the forecast<sup>17</sup>. While the Markov forecast is mainly based on the current state and the law of state transition predicts the possible state of the system in the future, generally through the transition probability matrix for prediction and decision-making. It can just make up for the limitation of the gray model. Therefore, try to use the Markov principle to fix the gray prediction value. Combine the advantages of two models to establish a Gray(1,1)-Markov prediction model<sup>18</sup>.

In this paper, two parts were included. First, an epidemiological description of varicella in Dalian from 2009 to 2019. Second, GM (1,1) model, Markov model, and GM(1,1)-Markov model were established to fit the incidence of varicella and to explore the fitting effects of different models on varicella data.

## 2. Materials And Methods

### 2.1. Materials

The data in this study was collected from Dalian Centers for Disease Control and Prevention (CDC). The varicella data from 2009 to 2019 includes the total infection number, age distribution, months distribution, and the occupation distribution. Among them, the occupation distribution in 2009 was

missed. The GM (1, 1) model and Markov model was built with data from 2009 to 2018. The GM (1, 1)-Markov model was built according to the result of the GM (1, 1) model.

## 2.2. Descriptive epidemiological analysis

The epidemiological characteristics of varicella in Dalian from 2009 to 2019 was described by time and population. Statistical charts were adopted to intuitively describe the epidemic characteristic.

## 2.3. GM (1,1)model

The data used from 2009 to 2018 as the original data, established a GM (1, 1) model and predicted the number of varicella in 2019, and analyzed the accuracy of the GM(1,1) model prediction results.

## 2.4. Markov model

The data used from 2009 to 2018 as the original data. K-mean cluster was adopted to divide groups. The state transition matrix was used to predict the 2019 group matrix and Markov model prediction value.

## 2.5. GM(1,1)-Markov model

The relative value between the predicted value of the GM(1,1) model from 2009 to 2018 and the true value was used as the original data. K-mean cluster was used to divide groups. The state transition probability matrix was calculated according to grouping. 2018, 2017, and 2016 were selected as the three most recent years from 2019. The three-step transition probability matrix was used to calculate the state interval of 2019. GM(1,1) model prediction value was combined to calculate GM(1,1)-Markov model prediction value.

## 2.5. Statistic software

The epidemiological analysis of varicella data was collected and analyzed by Excel 2010 software. R software 3.2.5 version was adopted to conducted GM(1,1) model prediction, the posterior difference ratio (C), and the probability of small error (P) were used to evaluate the effect of prediction of the GM(1,1) model. Markov model and GM(1,1)-Markov model was implemented with the help of MATLAB 2015a software. Relative error was selected as the index to evaluated the accuracy of the models.

## 3. Result

### 3.1. Epidemic characteristics

## 3.1.1. Time distribution

The total number of reported cases of varicella in Dalian from 2009 to 2019 was 37223. The incidence rate decreased and then increased on the whole. The lowest incidence rate was occurred in 2014(see Table1Figure1). Varicella occurred throughout the year. The number of cases was increasing from February to May and August to December. From April to June, and November to January of the following year, there were more cases(see Figure2). Varicella cases showed obvious seasonality.

**Table 1**

The cases and incidence rate of varicella in Dalian from 2009 to 2019

Year	Cases	Incidence rate(1/100000)
2009	3502	57.96
2010	3371	55.78
2011	3709	55.44
2012	2922	43.25
2013	2684	39.93
2014	2455	36.05
2015	3139	46.09
2016	2897	42.33
2017	3895	55.75
2018	4296	61.65
2019	4353	62.68

## 3.1.2. Population distribution

### 3.1.2.1. Age distribution

Varicella occurred any at age. The cases aged from 0 to 25 accounted for 92.46% of the total cases (see Table2). More cases aged from 5 to 10 (see Figure3).

**Table 2**

The varicella cases age distribution and proportion in Dalian from 2009 to 2019

Age group	Incidence number	Composition Ratio (%)	Age group	Incidence number	Composition Ratio (%)
0-	1144	3.07	10-	5953	15.99
1-	769	2.07	15-	5733	15.4
2-	639	1.72	20-	5144	13.82
3-	1092	2.93	25-	3165	8.5
4-	1601	4.3	30-	1788	4.8
5-	2264	6.08	35-	716	1.92
6-	1931	5.19	40-	165	0.44
7-	1679	4.51	45-	57	0.15
8-	1712	4.6	50-	22	0.06
9-	1596	4.29	55-	53	0.14

### 3.1.2.2. Gender distribution

The total cases of varicella in Dalian from 2009 to 2019 was 37,223, of which 20,043 cases were males (53.85%) and 17,180 cases were females (46.15%). The number of males was higher

than that of females, and the ratio of total male to female was 1.17:1. (Figure4)

## 3.2. GM(1,1) model

Using R software for GM(1,1) model prediction. The incidence for 2019 was 51.3/100000(Table3),the posterior difference ratio (C) was 0.987188,the probability of small error (P) was 0.3333.α=-0.0114 ,u=45.3757.Then the GM (1, 1) model was established as follows:

$$x^{(1)}(k + 1) = 4038.28e^{(0.52k)} - 3980.32$$

Table 3

GM(1,1) model prediction results with R software

## 3.3. Markov model

### 3.3.1. State division and state transition probability matrix::

Year	Actual value (1/100000)	R software prediction value(1/100000)
2009	57.96	57.96
2010	55.78	46.3
2011	55.44	46.83
2012	43.25	47.37
2013	39.93	47.91
2014	36.05	48.46
2015	46.09	49.01
2016	42.33	49.58
2017	55.75	50.15
2018	61.65	50.72
2019	62.68	51.3

Due to the amount of data was small, to ensure that each state had enough data, it was divided into 3 states, K-mean cluster was used to cluster data. The data distribution of each group could be determined. State division was: E1[36.05-41.13], E2[41.13-50.765], E3[50.765-61.65].

The state change from 2009 to 2018 is showed as E3-E3-E3-E2-E1-E1-E2-E2-E3-E3, state transition status and state transition probability matrix see table 4, Table 5.State transition matrix showed below:

$$P = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.33 & 0.33 & 0.33 \\ 0 & 0.25 & 0.75 \end{pmatrix}$$

**Table 4**

State transition from 2009 to 2018

State	E1	E2	E3
E1	1	1	0
E2	1	1	0
E3	0	1	3

**Table 5**

State transition matrix

State	E1	E2	E3
E1	0.5	0.5	0
E2	0.33	0.33	0.33
E3	0	0.25	0.75

### 3.3.2. Use the state transition probability matrix to predict:

With the help of MATLAB 2015a software, using the 2018 grouping matrix (0, 0, 1) and the state transition probability matrix to predict. The result of the 2019 grouping matrix was (0, 0.25, 0.75), the 2019 Markov model prediction was more likely in E3, so the value is: 56.2075

## 3.4. GM(1,1)-Markov model

### 3.4.1. State division and state transition probability matrix:

Relative value was used to divide the state, relative value=actual value/ R software prediction value. K-mean cluster was used to cluster data. The data distribution of each group could be determined. Divided all relative values into three states: underestimated, accurate, overestimated, namely E1[1.1,1.22], E2[0.82,1.1], E3[0.74,0.82].The status changed from 2009 to 2018 is E2-E1-E1-E2-E2- E3-E2-E2-E1-E1(Table6). State transition matrix see Table7

**Table 6**

State transition from 2009 to 2018

**Table 7**

State transition matrix

Year	Actual value	Predicted value	Relative value	State division
2009	57.96	57.96	1	E2
2010	55.78	46.3	1.2	E1
2011	55.44	46.83	1.18	E1
2012	43.25	47.37	0.91	E2
2013	39.93	47.91	0.83	E2
2014	36.05	48.46	0.74	E3
2015	46.09	49.01	0.94	E2
2016	42.33	49.58	0.85	E2
2017	55.75	50.15	1.11	E1
2018	61.65	50.72	1.22	E1
2019	62.68	51.3		

State	E1	E2	E3
E1	0.67	0.33	0
E2	0.4	0.4	0.2
E3	0	1	0

### 3.4.2. Three-step transition probability matrix to prediction varicella in 2019:

Select 2018, 2017, and 2016, the three most recent years from 2019. After one step (P1), two steps (P2), and three steps (P3), the state was transferred to 2019. Sum the column items, the maximum value was the state range, which the 2019 GM(1,1) model predicted value located in, see Table8.

$$P(1) = \begin{pmatrix} 0.67 & 0.33 & 0 \\ 0.4 & 0.4 & 0.2 \\ 0 & 1 & 0 \end{pmatrix} P(2) = P(1)^2 = \begin{pmatrix} 0.5809 & 0.3531 & 0.066 \\ 0.428 & 0.492 & 0.08 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$$

$$P(3) = P(1)^3 = \begin{pmatrix} 0.5304 & 0.3989 & 0.0706 \\ 0.4836 & 0.418 & 0.0984 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$$

**Table 8**

Status of the predicted values of the GM(1,1)model in 2019

Year	Initial state	transition steps	E1	E2	E3
2018	E1	1 (P1)	0.67	0.33	0
2017	E1	2 (P1)	0.5809	0.3531	0.066
2016	E2	3 (P3)	0.4836	0.418	0.0984
sum			1.7345	1.1011	0.1644

The maximum sum value was in state E1, That was the GM(1,1) model prediction value will be in the E1 state, the prediction value of 51.3 was been underestimated, and the GM(1,1)-Markov prediction value was  $51.3 \cdot (1.1 + 1.22) / 2 = 59.508$

### 3.5. Models results comparison

Three models results and comparison see Table9, Figure5 and Table10. GM(1,1)-Markov model fitted actual value better and had the lowest relative error of 2019.

**Table 9**

Models prediction results and actual value from 2009 to 2019

Year	Actual value	GM(1,1) value	Markov value	GM(1,1)-Markov value
2009	57.96	57.96	56.2075	55.6416
2010	55.78	46.3	56.2075	53.708
2011	55.44	46.83	56.2075	54.3228
2012	43.25	47.37	45.9475	45.4752
2013	39.93	47.91	38.59	45.9936
2014	36.05	48.46	38.59	38.0411
2015	46.09	49.01	45.9475	47.0496
2016	42.33	49.58	45.9475	47.5968
2017	55.75	50.15	56.2075	58.174
2018	61.65	50.72	56.2075	58.8352
2019	62.68	51.3	56.2075	59.508

**Table 10**

Models results and relative errors of 2019

Year	Actual value	GM(1,1)	Markov	GM(1,1)-Markov
2019	62.68	53.6425	56.2075	59.508
Relative error%		14.42	10.31	5.06

## 4. Discussion

There were 37223 varicella cases in Dalian from 2009 to 2019. The average annual incidence was 50.56/100000. The high incidence was corresponding with our country's overall situation<sup>5</sup>. Varicella cases showed an upward trend just like other places<sup>8,9</sup>. Varicella can occur all year round, mostly in spring and winter, seasonality is obvious. The two peaks of varicella cases were from April to June and November to January of the following year, showing a "bimodal distribution". This time the temperature change violently, people especially children who had low immunity prone to be infected. This time for children also studied at school. So, they can contact the outside world and other infected people easier.

In this paper, students, childcare, scattered children are the main disease groups. Most cases age from 5 to 10. The cases aged from 0 to 25 accounted for 92.46% of the total cases. High-age cases account for a large proportion. In many places, the age of varicella had shifted backward<sup>10,19,20</sup>. Adults infected

varicella became easy and common. Adult varicella is more harmful. Adults especially during pregnancy get more symptoms of varicella and more dangerous<sup>21,22,23</sup>. Many reasons like a vaccine, weather, climate, and others have a complication effect on the incidence of varicella<sup>11,12,6</sup>. We still need to find out the reason and solution to decrease the incidence.

The incidence ratio of varicella between men and women is 1.167:1, and the number of cases in men is higher than that in women. In many places, the varicella incidence rate, the male is higher than that of the female. And there is a statistical difference<sup>10,23,24</sup>. Student, childcare, scattered children accounted for 73.15% of the total of the cases. So varicella outbreaks were prone to occur in schools, kindergartens. These places require to Strengthen prevention and control.

The GM(1,1), Markov model, and Gray-Markov were used to predict the incidence of varicella in 2019. The predicted values were: 53.6425, 56.2075, and 59.508, and the relative errors of the true value in 2019 were: 14.42, 10.33, and 5.06. From the results, the GM(1,1) model predicted the varicella cases were not ideal, and the model fitted poorly. Consider two follow reasons: 1 The sample size was too small. 2 The sample size was too large. 3 The data was volatile. Concerning Tingting Zhang, Yanling Peng, and other documents<sup>25</sup>, they collected 8 years of data from 2011 to 2018 and conducted GM(1,1) model prediction. The prediction level is excellent, and the amount of data is small, which is not the main reason for the unsatisfactory prediction. Besides, the GM(1,1) model is famous for predicting "small samples". There are also articles showing that the GM(1,1) model can be developed with only 4 models<sup>26</sup>, and the sample size was small, which was not the main problem. Research showed that fewer and recent modeling data have a good forecast effect<sup>27</sup>. When the number of modeling data between 4 and 10. When the dimension was 4, the prediction accuracy was the highest. With the increase of dimension, the precision did not increase<sup>28</sup>. But in this study, the number of GM(1,1) modeling data should be the same as Markov modeling data. More data, especially much old data had a bad influence on GM(1,1) model. The accuracy of the Grey model was inevitably decreased. The GM(1,1) model had a good predictive effect for sequences with short sequences and an upward trend<sup>29</sup>. The prediction results of the GM(1,1) model were generally smooth curves, reflecting long-term growth trends<sup>30</sup>, Poor fit effect for fluctuating data. Therefore, considering the large fluctuation of the sample was another reason for the unsatisfactory GM(1,1) prediction of this experiment. Consider further processing to improve forecast accuracy with this kind of data type. In this study, we combined GM(1,1) model and the Markov model to improve accuracy.

Markov predicted value in 2019 was 56.2075, which was different from the actual value of 62.68. The status division had a greater impact. Theoretically speaking, the longer the historical data, the more state division, the higher the prediction accuracy<sup>31</sup>. Generally, when there was less historical data, the number of states should be less, so that each state had more data. When there were more data, more sample falls into each state. The more the state increased, the more accurate the obtained prediction interval would be. In this study, we had less data, less state, and the transition change situation could not be better reflected. The state interval span was large, and the data prediction could not be effective and accurate.

Besides, the state division based on the original data, the Markov model could not express the trend of continuous growth. Therefore, Markov prediction error would become larger and larger.

The Grey-Markov prediction result was better than the Grey model and Markov model prediction alone. The Gray model was widely used in predicting small sample data. It could describe the overall trend well, but it could not reflect data fluctuations well. The Markov model had a good predictive effect on fluctuating data. However, the data was less and the state was insufficient, which easily led to poor prediction results. And it could not describe monotonic change data well. Grey-Markov model forecasting can combine the advantages of the GM(1,1) model to predict small samples and the advantage of Markov's response to fluctuations, improving the prediction accuracy to a certain degree. GM(1,1)-Markov forecast achieved a relatively good forecast result with the small sample, volatility, and overall risen data. In further studies, GM(1,1)-Markov model can be used in similar medical data to predict disease development and formulate reasonable medical health measures. And we still need to continuously optimize the model and explore better ways to get better results.

## Conclusion

The incidence of varicella fluctuated greatly and showed an upward trend as a whole, and the age of onset shifted backward. GM(1,1)-Markov model can combine the advantages of small samples predicted by GM(1,1) model and the advantages of Markov model response fluctuation to improve the prediction accuracy. Varicella data was small and volatility, and good prediction results were obtained by using GM(1,1)-Markov model. GM(1,1)-Markov model can be used, improved and popularized in the prediction of similar types of data in the future.

## Abbreviations

P=the probability of small error, C= the posterior difference ratio CDC= Centers for Disease Control and Prevention, GM (1,1)= grey model first-order one variable, VZV= varicella-zoster virus

## Declarations

## Ethics approval and consent to participate

Not application

## Consent for publication

Not application

# Acknowledgements

Not application

# Availability of data and materials

The data that support the findings of this study are available from Dalian Centers for Disease Control and Prevention (CDC) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Dalian Centers for Disease Control and Prevention (CDC).

# Competing interests

The authors declare that they have no competing interests.

# Funding

Not application

# Authors' contributions

Author1.CT:Conducted the research, did data entry, processed and analysis the data and was a major contributor in writing the manuscript.

Author2.BY: Provided the original data.

All authors read and approved the final manescript.

# Acknowledgements

Not application

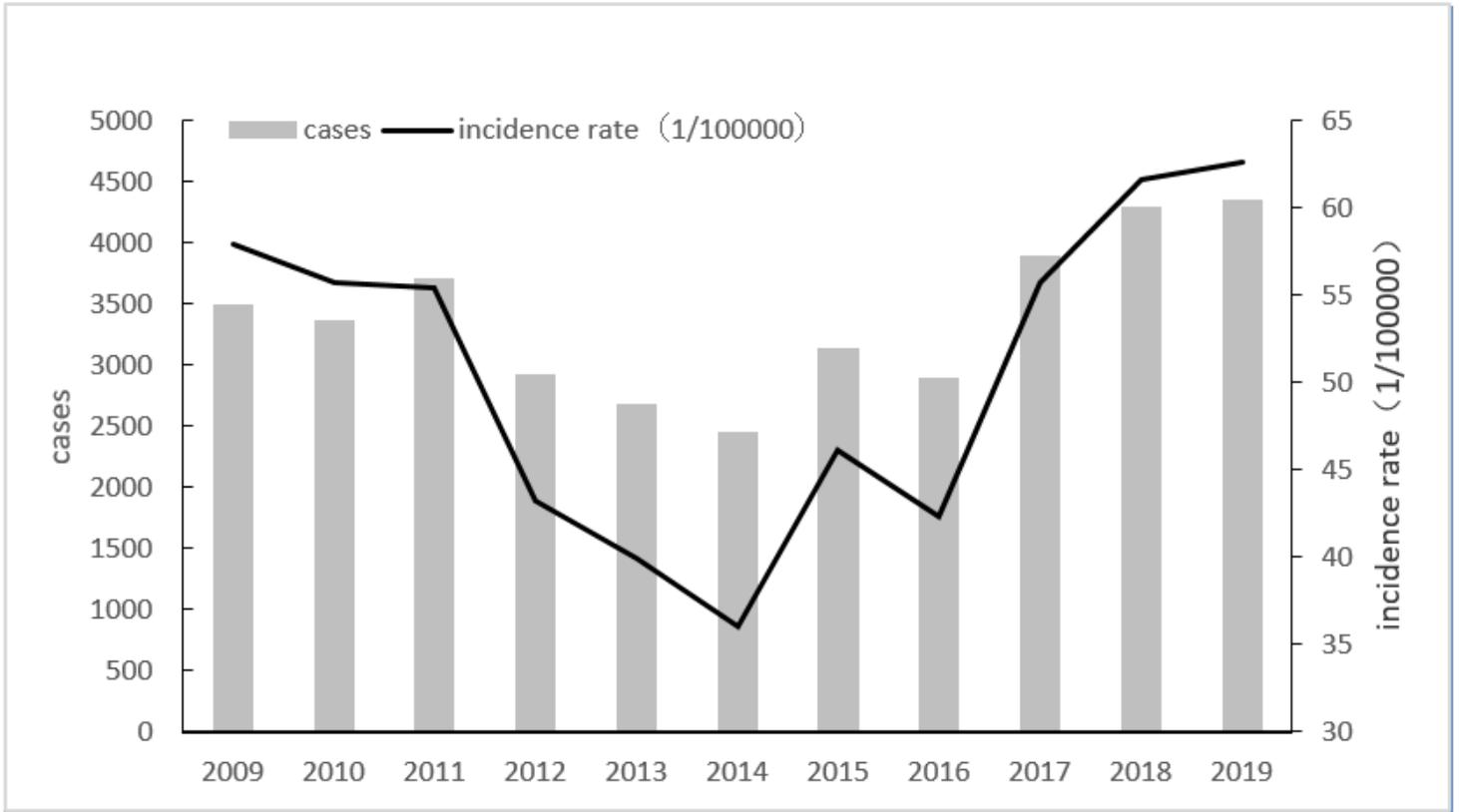
# References

1. Gershon, A. A. *et al.* Varicella zoster virus infection. *Nature reviews. Disease primers* 1, 15016, doi:10.1038/nrdp.2015.16 (2015).
2. Ahn, K. H. *et al.* Congenital varicella syndrome: A systematic review. *Journal of obstetrics and gynaecology: the journal of the Institute of Obstetrics and Gynaecology* 36, 563–566,

- doi:10.3109/01443615.2015.1127905 (2016).
3. Amlie-Lefond, C. & Gildea, D. Varicella Zoster Virus: A Common Cause of Stroke in Children and Adults. *Journal of stroke and cerebrovascular diseases: the official journal of National Stroke Association* 25, 1561–1569, doi:10.1016/j.jstrokecerebrovasdis.2016.03.052 (2016).
  4. Bernal, J. L., Hobbelen, P. & Amirthalingam, G. Burden of varicella complications in secondary care, England, 2004 to 2017. *Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 24, doi:10.2807/1560-7917.es.2019.24.42.1900233 (2019).
  5. Xu, Q. & Xu, A. Epidemiological analysis of varicella in China and advance in immunity prevention for varicella vaccine. *Chinese Journal of New Drugs* 21, 1093–1098 (2012).
  6. US children to get a second dose of chickenpox vaccine. *Chinese Journal of General Practitioners* 6, 603 (2007).
  7. Leung, J., Broder, K. R. & Marin, M. Severe varicella in persons vaccinated with varicella vaccine (breakthrough varicella): a systematic literature review. *Expert review of vaccines* 16, 391–400, doi:10.1080/14760584.2017.1294069 (2017).
  8. Jiang, C. Epidemiological characteristics of varicella patients in Xuzhou, 2010-2018. *Modern Preventive Medicine* 47, 2328-2331+2341 (2020).
  9. Pan, X., Ma, R., Fang, T., Xu, G. & Xu, H. Epidemiology of varicella in Ningbo city of Zhejiang province, 2012-2016. *Chinese Journal of Vaccines and Immunization* 23, 302–305 (2017).
  10. Zeng, J., Yu, X., Liu, J., Tan, A. & Xie, D. Analysis on epidemiological characteristics of varicella in Neijiang City of Sichuan Province from 2009 to 2018. *Journal of Medical Pest Control* 36, 400–403 (2020).
  11. Zhang, H. Diagnosis and treatment of chickenpox. *World Health Digest* 09, 229–230, doi:10.3969/j.issn.1672-5085.2012.18.229 (2012).
  12. Yang, Y., Geng, X., Liu, X., Wang, W. & Zhang, J. Association between the incidence of varicella and meteorological conditions in Jinan, Eastern China, 2012-2014. *BMC infectious diseases* 16, 179, doi:10.1186/s12879-016-1507-1 (2016).
  13. Deng, J. *Basic Methods of Grey System Chinese-English Second Edition*. (Huazhong Institute of Technology Press, 2005).
  14. Fu, B. & Wang, A. Application of Markov Chain in Market Forecast. *CO-Operative Economy & Science*, 56–57, doi:10.3969/j.issn.1672-190X.2017.01.027 (2017).
  15. Dang, S. L. Y. *Grey System Theory and Application Fifth Edition*. (Beijing: Science Press, 2010).
  16. Tang, T. & Li, L. Prediction of Container Throughput of Shanghai Port Based on Grey Markov Model. *Logistics Sci-tech* 43, 105–108,114 (2020).
  17. Du, B. *Research on Pilot Demand Quantity Prediction and Team Construction Based on Grey Markov Method*, DalianMaritimeUniversity, (2017).

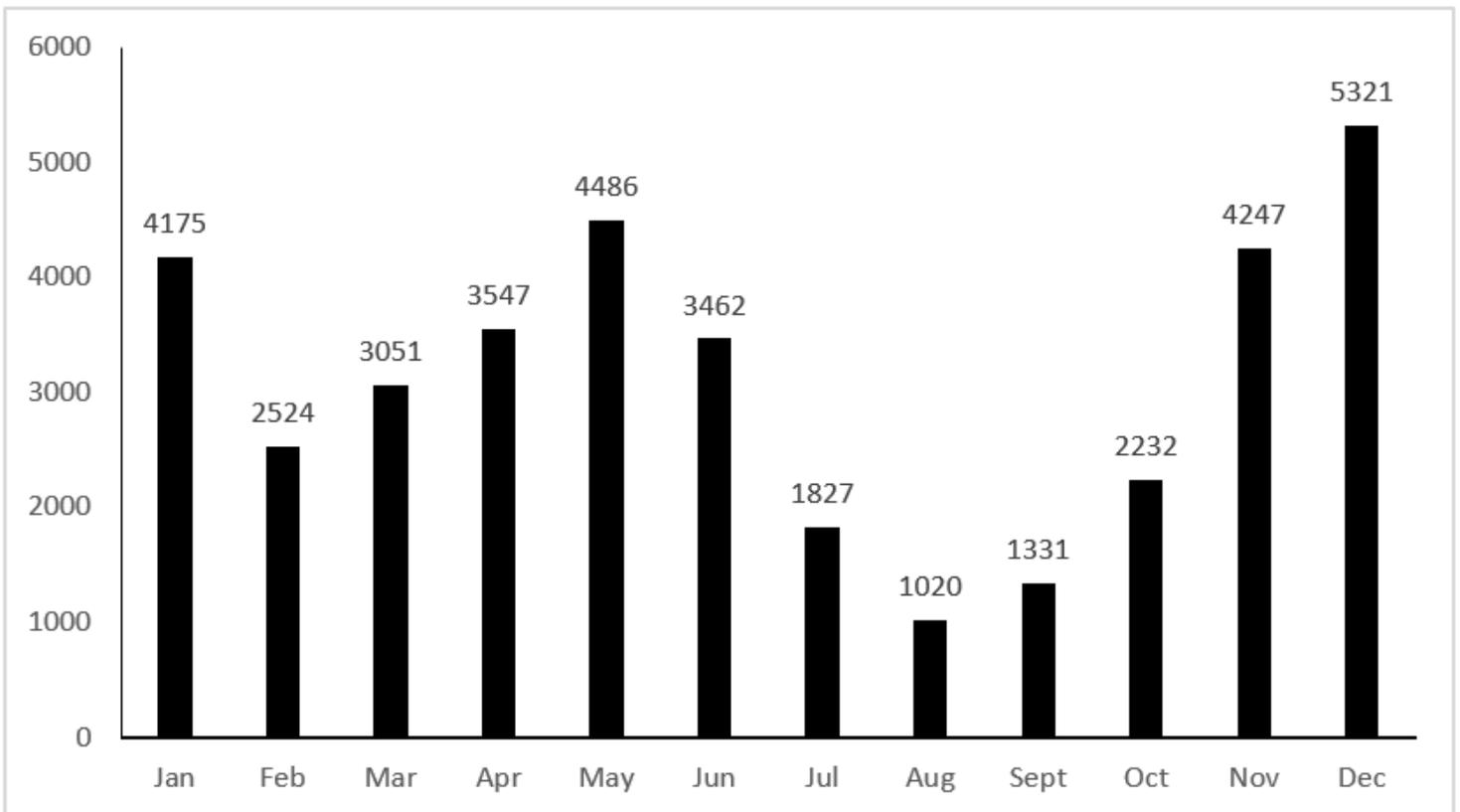
18. Song, X. Study on Coal Production Prediction Based on Gray Markov Model. *Mineral Engineering Research* 34 (2019).
19. Bai, S. Epidemiological characteristics and prediction of varicella in Shenyang from 2006 to 2018. *Journal of Tropical Medicine* 19, 1564–1566 (2019).
20. Li, X. *Analysis on the Epidemiologic Characteristics of varicella in Qingdao from 2007 to 2016*, Qingdao University, (2018).
21. LAFORET, E. G. & LYNCH, C. L. Multiple congenital defects following maternal varicella; report of a case. *The New England journal of medicine*, 534–537 (1947).
22. Song, Y. Consequences of varicella in pregnancy or perinatal period. *Chinese Journal of Experimental and Clinical Virology* 30 (2016).
23. Lv, T. & Yang, H. Varicella-zoster virus infection during pregnancy. *Chinese Journal of Medicine*, 18–21 (2008).
24. Peng, Y., Kong, D., Wang, P., Lu, S. & Yang, X. Epidemiological characteristics of chicken-pox in Wuhan city, 2009-2018. *Practical Preventive Medicine* 27, 330–332 (2020).
25. Zhang, T. Pre-hospital emergency treatment demand for cardiovascular and cerebrovascular disease in Urumqi based on grey model and time series model: Prediction and analysis. *Chinese Journal of Disaster Medicine* 7 (2019).
26. Wang, Y.-w., Shen, Z.-z. & Jiang, Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *Public Library of Science* 13 (2018).
27. Cheng, G. Impacts of Different Dimensions in Gray Model on Model Precision. *Standardization for Surveying and Mapping* 34 (2018).
28. Kai, Z. & Ya-dong, L. I. The mining subsidence based on the innovation GM(1,1)model. *Beijing Surveying and Mapping*, 31–33, doi:10.3969/j.issn.1007-3000.2013.02.008 (2013).
29. Yi, J. *et al.* Application of Three Kinds of Model in Forecasting Prevalence of Pulmonary Tuberculosis. *Chinese General Practice* 15, 1495–1497, doi:10.3969/j.issn.1007-9572.2012.13.022 (2012).
30. Shi, D. Study on prediction of occupational diseases based on grey GM(1,1)-Markov model. *Journal of Safety Science and Technology* 13 (2017).
31. Pei, Y. *Predictive analysis of bridge condition by improved Grey Markov model*, Hebei engineering university (2018).

## Figures



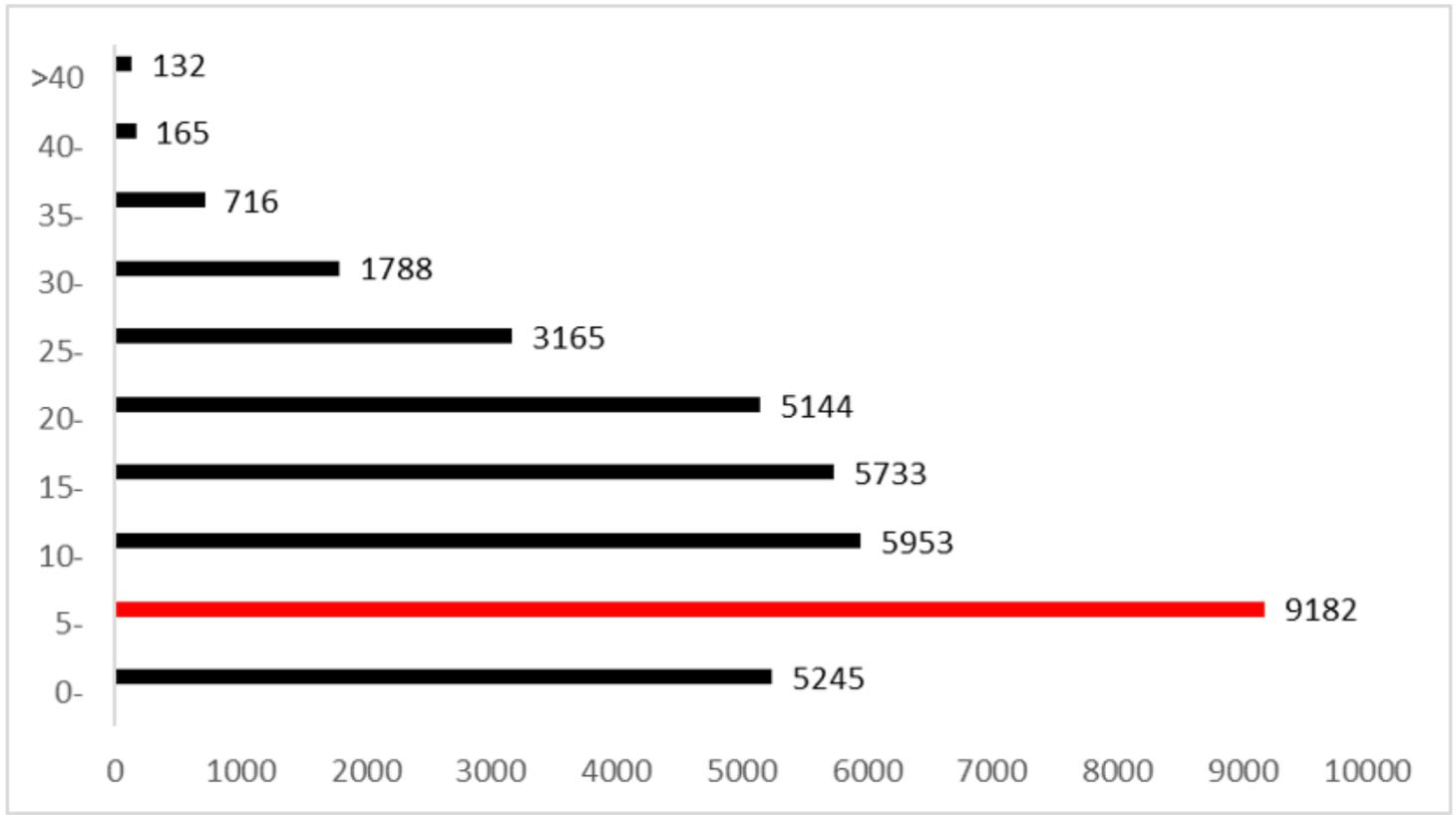
**Figure 1**

The cases and incidence rate of varicella in Dalian from 2009 to 2019



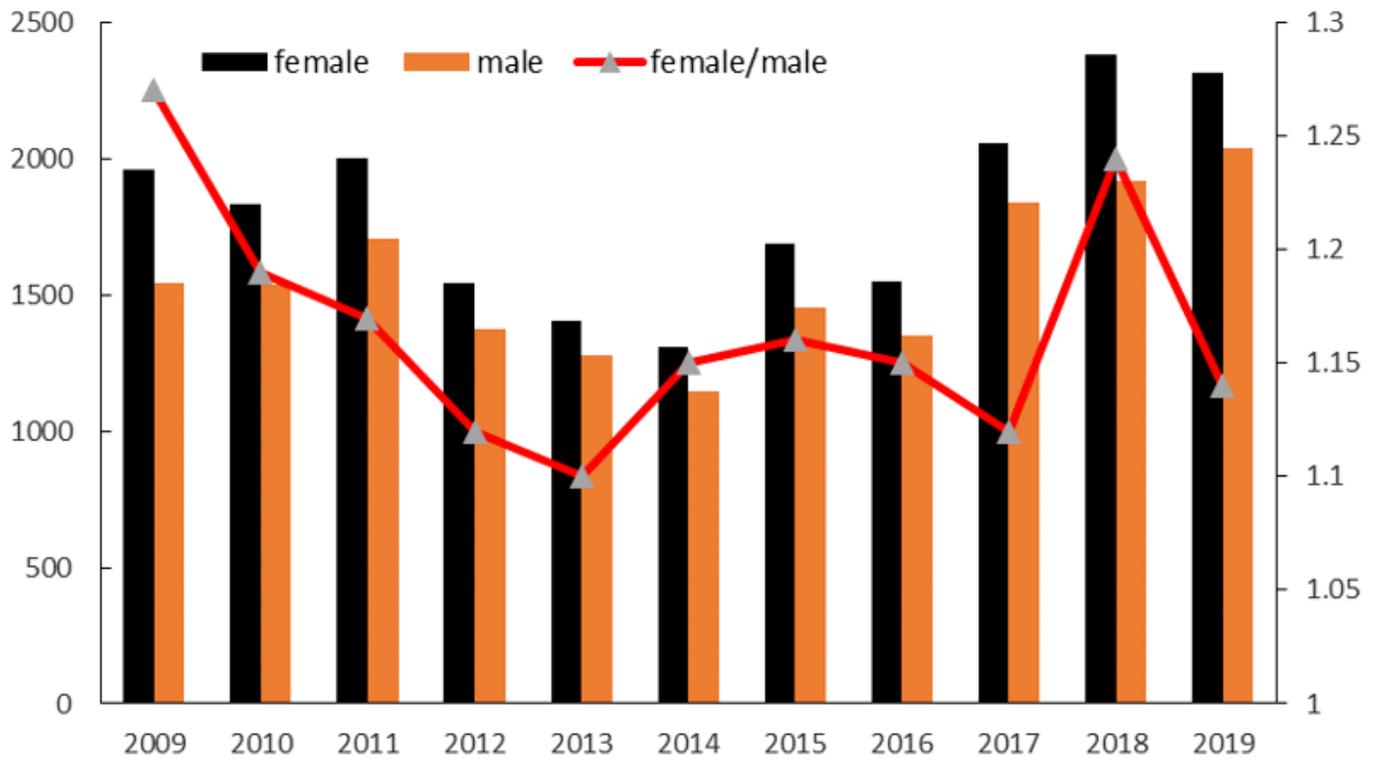
**Figure 2**

Month distribution of varicella cases in Dalian from 2009 to 2019



**Figure 3**

The varicella cases age distribution in Dalian from 2009 to 2019



**Figure 4**

Total cases and gender distribution of varicella in Dalian from 2009 to 2019

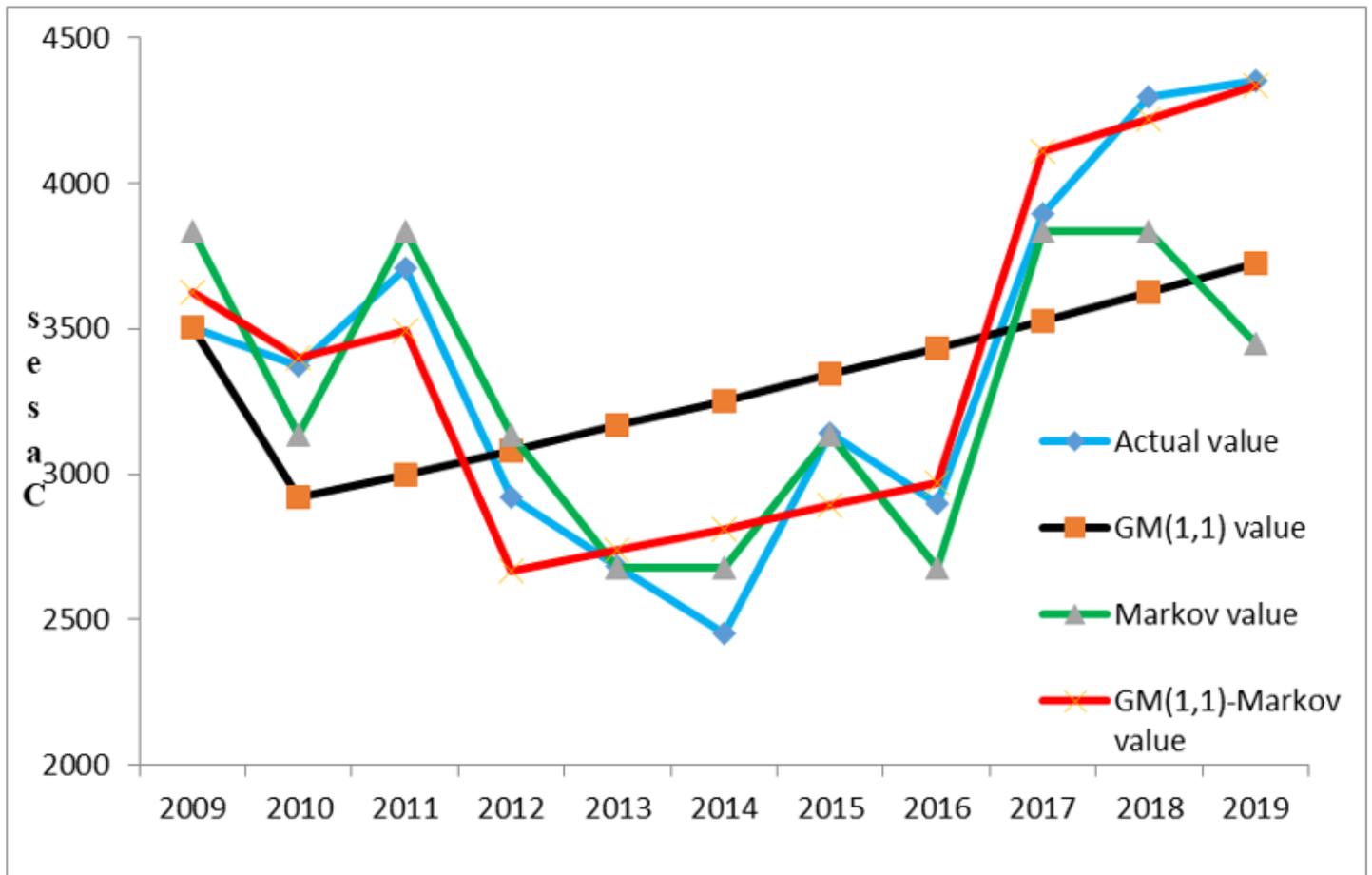


Figure 5

Comparison between prediction value and actual value

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)