Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2017, 4(4):157-162**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Atmospheric Ozone Prediction Based on ARIMA Model

Xinyu Xue*, Zhiyan Wang, Shuang Fang

Nanjing Normal University Taizhou College, NO.518, Dongfeng south Rd., Hailing District, TaiZhou, JiangSu, China

Abstract Ozone layer refers to the stratospheric stratosphere in the relatively high concentration of ozone, which plays an important role in absorbing ultraviolet light, heating the atmosphere and greenhouse gases. However, since the 1970s, the total amount of atmospheric ozone has been found to remain in a state of continuous decline. In this paper, we select the total amount of atmospheric ozone and ozone hole area as the index that we measure the level of atmospheric ozone indicators. Meanwhile, establish ARIMA time series model based on the continuity of the development of things to predict the next five years of ozone levels to predict ozone levels in the next fifty years. First, in order to eliminate the trend, seasonal and nonstationarity of the actual time series, we use Box-Jenkins method (difference method) to transform the sequence into a stationary sequence by the A-order difference operation. Secondly, based on the continuity of the development of things, we use the AIC and BIC criteria to establish the ARIMA model, and use Matlab software programming to solve the model parameters, calculate the original data 50 step prediction.

Keywords Ozone Total, Ozone Hole Area, ARIMA Sequence

Introduction

Ozone depletion has been observed from several phenomena since the late 1970s: a steady decline in the total amount of ozone in Earth's stratosphere, a much larger springtime decrease in stratospheric ozone — as known as ozone hole — around Earth's polar regions and springtime polar tropospheric ozone depletion events.

There are three main effects of atmospheric ozone layer. The first is the protective effect. The ozone layer can absorb the sun in the wavelength of 300μm below the UV, the protection of human beings and animals and plants on earth from short-wave ultraviolet damage. The second is the heating effect. Ozone absorbs ultraviolet light in the sun and converts it into heat to heat the atmosphere, which has an important effect on the circulation of the atmosphere. The third is the role of greenhouse gases. The high degree of ozone distribution and change is extremely important to surface temperatures.

It is believed that the primary cause of ozone depletion is the presence of chlorine-containing source gases, including CFCs and related halocarbons, nitrous oxide, which is often used as sprays, coolants and solvents. But it is difficult to break down, and it will remain in the environment for a long time. Furthermore, it is also proved that the solar activity, general atmospheric circulation and quasi-biennial oscillation also have an impact on the depletion of ozone layer [1].

In 1987, representatives from 43 nations signed the Montreal Protocol and agreed to freeze production of CFCs at 1986 levels and to reduce production by 50 percent by 1999. Since the adoption and strengthening of the Montreal Protocol has led to reductions in the emissions of CFCs, atmospheric concentrations of the most-significant compounds have been declining.

Based on the data of TOMS, OMI and FY-3A / B satellite satellites inversion of the total amount of satellite ozone and ozone hole area from 1979 to 2014, we further estimate and understand the total time variation of ozone characteristics from two perspectives. Establish ARIMA time series model to predict the next fifty years of ozone levels. First, in order to eliminate the trend, seasonal and nonstationarity of the actual time series, we use the difference methods to transform the sequence into a stationary sequence by the A-order difference operation. Secondly, based on the continuity of the development of things, we use the AIC and BIC criteria to establish the ARIMA model, and solve the model parameters, then calculate the original data 50 step prediction through Matlab.

Time series model

1. Box–Jenkins method

1.1 Stability identification

Based on the scatter plot of time series, autocorrelation function and partial autocorrelation function graph, the variance, trend and seasonal variation of ADF unit root were used to identify the station's stationarity.

1.2 Differential processing

The nonstationary sequence is stationary. If the data sequence is non-stationary and exists a certain growth or decline trend, we need to deal with the data differential, if the data is heterogeneous, the need for technical data processing, until the data after the autocorrelation function value and partial. The correlation function values are not significantly different from zero [2].

First-order difference：

$$\nabla X_t = X_t - X_{t-1} = (1-B)^2 X_t, \tag{2-1}$$

Second-order difference：

$$\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2} = (1-B)^2 X_t, \tag{2-2}$$

As usual, d-order difference

$$\nabla^d X_t = (1-B)^d X_t, \tag{2-3}$$

$$\nabla^d \equiv (1-B)^d = 1 - \binom{d}{1}B + \binom{d}{2}B^2 + \cdots + (-1)^{d-1}\binom{d}{d-1}B^{d-1} + (-1)^d B^d. \tag{2-4}$$

For nonstationary sequence $\{X_t, t = 0, \pm 1, \pm 2, \cdots\}$, If exists a positive integer d which makes:

$$\nabla^d X_t = W_t \tag{2-5}$$

Then call it can be stabilized by difference equation, and among them, $\{W_t, t = 0, \pm 1, \pm 2, \cdots\}$ is $ARIMA(p,d)$ Sequence, X_t is $ARIMA(p,d,q)$ Sequence.

If the observation samples of X_t are $X_1, X_2, \cdots, X_n$, after first order difference, the data reduce to $n-1$; after second order difference, the data reduce to $n-2$. As usual, after d -order difference, there are $n-d$ data. If use $\nabla^d X_t$ to Recover data, we need a given initial value such as $X_1, X_2, \cdots, X_d$.

If the initial value of $X_1, X_2, \cdots, X_d$ are known, by the function:

$$W_t = \nabla^d X_t, t = d+1, \cdots, n, \tag{2-6}$$

We can recover X_t. Then we will give the Recovery Formula when $d = 1, 2$ as following.

When $d = 1$:

$$X_t = X_1 + \sum_{j=1}^{t-1} W_{j+1} = X_k + \sum_{j=1}^{t-k} W_{j+k}, t > k \geq 1 \tag{2-7}$$

When $d = 2$:

$$X_t = X_2 + (t-2)(X_2 - X_1) + \sum_{j=1}^{t-2}(t-j-1)W_{j+2}$$
$$= X_k + (t-k)(X_k - X_{k-1}) + \sum_{j=1}^{t-k}(t-j-k+1)W_{j+k}, t > k \geq 2 \quad (2\text{-}8)$$

1.3 Model selection and Model establishment

According to the recognition rules of time series model, the corresponding model is established. If the partial correlation function of the stationary sequence is truncated and the autocorrelation function is smearing, it can be concluded that the sequence is suitable for the AR model. If the partial correlation function of the stationary sequence is smearing and the autocorrelation function is truncated, the deterministic sequence is suitable for the MA model. If the partial correlation function and the autocorrelation function of the stationary sequence are smearing, the sequence is suitable for the ARMA model [3].

Use the nonlinear least squares estimation genetic algorithm to fit the model fitting and parameter estimation by Matlab programming.

2. Model Test-- Diagnose whether the residual sequence is white

Denoted fitting residual error of the mode as $\widehat{\varepsilon}_t$, it is estimated value of ε_t .For example to the sequence $AR(p)$, let the estimation of unknown parameters as $\widehat{\varphi}_1, \widehat{\varphi}_2, \cdots, \widehat{\varphi}_p$,then there is equal to:

$$\widehat{\varepsilon}_1 = X_t - \widehat{\varphi}_1 X_{t-1} - \cdots - \widehat{\varphi}_p X_{t-p},\ t = 1, 2, \cdots, n \quad (\text{let } X_0 = X_{-1} = \cdots = X_{1-p} = 0).$$

Write for

$$n_k = \frac{\sum_{t=1}^{n-k} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t+k}}{\sum_{t=1} \widehat{\varepsilon}_t^2}, k = 1, 2, \cdots, L, \quad (2\text{-}9)$$

Among them, L is the autocorrelation function of the tail number of $\widehat{\varepsilon}_t$.

χ^2 on $Ljung - Box$ is

$$\chi^2 = n(n+2)\sum_{k=1}^{L} \frac{\eta_k^2}{n-k}. \quad (2\text{-}10)$$

Test hypothesis is $H_0 : \rho_k = 0$, when $k \leq L; H :$ to some $k \leq L, \rho_k \neq 0$ .When H_0 found, if the sample size n is sufficiently large, the distribution of χ^2 is similar to the distribution of $\chi^2(L-r)$, r is estimated number of model parameters.

χ^2 -method of calibration : Given significant level α quantile $\chi_\alpha^2(L-r)$,when $\chi^2 > \chi_\alpha^2(L)$,refuse H_0, considered ε_t as non white noise, model checking is not passed ; when $\chi^2 \leq \chi_\alpha^2(L-r)$, accept H_0, consider ε_t as white noise, model checking is passed.

Experiment and Result

Data from 1979 to 2014 were used to predict atmospheric ozone and ozone voids for 2015-2020. Due to TOMS technical reasons, the lack of 1995 measured data, so 1995 will also be regarded as the target.

When we get a first order difference on the data, we get a smooth sequence $\{x_n\}$ and $\{y_n\}$. The ARIMA (1,2,3) model and the ARIMA (1,3,0) model are given according to the order of the AIC and BIC criteria [4]. The specific data is shown in the following table [5].

Table 1: Order Data of the Ozone Total Model

p	*q*	*AIC*	*BIC*
0	1.0000	309.6938	312.8045
0	2.0000	306.4154	311.0814
0	3.0000	308.1632	314.3846
1.0000	0	304.9330	308.0437
1.0000	1.0000	306.9322	311.5982
1.0000	2.0000	307.8058	314.0272
1.0000	3.0000	309.7017	317.4784
2.0000	0	306.8427	311.5088
2.0000	1.0000	302.6782	308.8996
2.0000	2.0000	307.7101	315.4868
2.0000	3.0000	300.3861	309.7182
3.0000	0	306.0272	312.2486
3.0000	1.0000	307.5222	315.2989
3.0000	2.0000	304.5522	313.8843
3.0000	3.0000	305.9035	316.7909

Table 2: Order Data of the Ozone Hole Area Model

p	*q*	*AIC*	*BIC*
0	1.0000	192.0626	195.1733
0	2.0000	179.8020	184.4681
0	3.0000	179.5888	185.8102
1.0000	0	186.0399	189.1506
1.0000	1.0000	187.9943	192.6604
1.0000	2.0000	179.9197	186.1411
1.0000	3.0000	181.4263	189.2031
2.0000	0	187.9641	192.6301
2.0000	1.0000	183.4460	189.6674
2.0000	2.0000	181.4705	189.2473
2.0000	3.0000	182.8862	192.2183
3.0000	0	175.5137	181.7351
3.0000	1.0000	177.4926	185.2693
3.0000	2.0000	179.3614	188.6935
3.0000	3.0000	173.5177	184.4051

Establish models respectively as follows:

$$A(z)y(t) = C(z)e(t)$$

among them:

$$A(z) = 1 - 0.7674 z^{-1} - 0.2007 z^{-2}, C(z) = 1 - 0.4615 z^{-1} + 0.3133 z^{-2} - 0.09544 z^{-3};$$

$$A(z)y(t) = e(t)$$

among them:

$$A(z) = 1 - 0.4319 z^{-1} - 0.6101 z^{-2} + 0.02339 z^{-3}$$

The prediction results of both the total amount of atmospheric ozone and ozone hole area in the next ten years are shown in the following table.

Table 3: Global ozone total annual average predicted in 2015-2024

Year	**2015**	**2016**	**2017**	**2018**	**2019**
Value	124.8796	124.6624	123.0790	121.6844	120.5000
Year	**2020**	**2021**	**2022**	**2023**	**2024**
Value	119.4956	118.6440	117.9219	117.3095	116.7904

Table 4: Polar ozone hole area predicted in 2015-2024

Year	2015	2016	2017	2018	2019
Value	21.9150	25.1183	22.4763	23.8357	24.1094
Year	**2020**	**2021**	**2022**	**2023**	**2024**
Value	22.7589	24.5381	23.1591	23.7732	24.0413

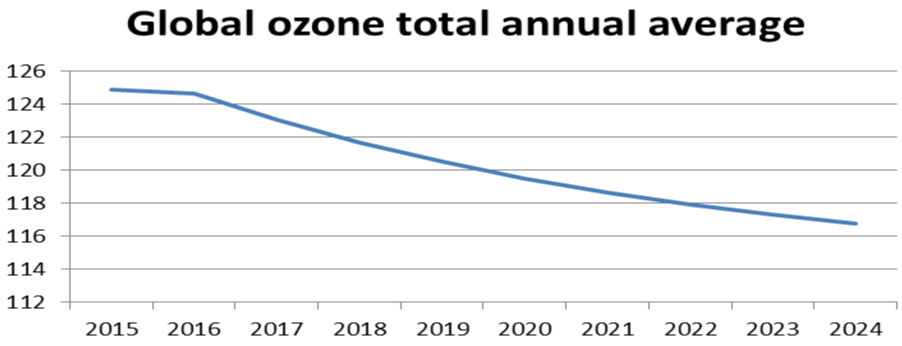

Figure1 Prediction of Global Ozone Total Ozone Annual Average

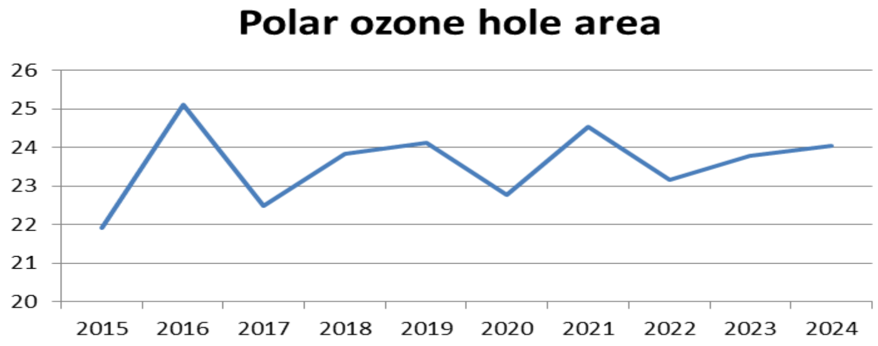

Figure 2: Prediction of Polar Ozone Hole Area

Do $Ljung-Box$ testing for the residual vector in the model. The result is $h=0$. The test is passed.

Conclusions

Observing the results we have obtained above, we can see that the total amount of ozone in the next 10 years still shows a slow decrease trend. The area of ozone is still fluctuating and fluctuating. Although ozone-depleting substances are under control, ozone recovery is still a long-term adjustment process. People still need to work together to minimize the use of chlorine-containing source gases. At the same time, through the test of the model, we found that the establishment of ARIMA model based on the data from 1979 to 2014 for the next 10 years of atmospheric ozone level short-term forecast is feasible.

In fact, there are many factors that affect the ozone content, including natural factors and human factors, making the relationship between ozone and weather more complex, such as lag, fluctuations, interaction. And the monitoring data accuracy, reliability, abnormal data and other issues, will affect the ozone content of the forecast analysis.

References

[1]. Cheng Guangxing., & Shao Jun. (1999). Chemical destruction of the ozone layer and its countermeasures. Chemistry Bulletin,9:45-47

[2]. Si Shoukui., & Sui Zhaoliang. (2016). Mathematical modeling algorithm and application. National Defense Industry Press

[3]. Zhao Xiaoxiao., & Zhu Ning., & Huang Liping.(2012).Modeling algorithm and empirical analysis based on the time series of the ARIMA model. Journal of Guilin University of Electronic Technology,32:410-415

[4]. Fan Yong., & Zhu Wenge., & Wang Liping.(2015). Predict mortality by age with trend ARIMA model. Insurance Studies,5:23-30

[5]. Yang Zhao., & Ye Zhonghui., & You Aiguo., & Guo Yirui., & Zhang Xiaoxiao., & Liang Shuying., & Xin Jin., & Wang Chongjian.(2013). Application of multiple seasonal ARIMA model in prediction of tuberculosis incidence, 29:469-472