Forecasting Particulate Matter in Seoul Through ARIMA Model

Time Series Data Analysis and Forecasting in Business

(Capstone Design)

Prof. Myung Suk Kim

20180825 Yujeong Shin

20190846 Seongwon Ryu

20190825 Euijin Shin

Contents

I. Problem Motivation

Clean air is a basic human right, as World Health Organization (WHO) stated[1]. On September 22nd, 2021, WHO has released revised version of Global Air Quality Guidelines (AQGs), which has first been updated in 16 years after they reported in 2005[2]. AQGs are the guidelines that gives recommendations on air quality guideline levels for six major air pollutants, particulate matter ($PM_{2.5}$ and PM_{10}), ozone (O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2) and carbon monoxide (CO)[3]. The revised AQGs has stricter limit values of the air pollutants than the old version as providing clear evidence of the damage air pollution inflicts on human health, at even lower concentrations than previously understood. WHO stated that AQGs have an ultimate purpose of improving public health by reducing air pollution since there are about 7 million premature deaths attributed to the air pollution and the air pollution gives a significant threat to the public health by causing various serious non-communicable diseases like heart attacks and stroke. We can be conscious of the seriousness and harmfulness of polluted air from how WHO reacts to and deals with the issue. As the world has been highly industrialized and the highly industrialized society has emitted a great amount of air pollutants from a variety of sources, the issue regarding air pollution has been on the rise in many places on the earth and many people in the world have been suffering from the impact of air pollution. Korea is also one of the heavily polluted nations. According to the report from Organization for Economic Cooperation and Development (OECD) that predicts the social cost in 2060 incurred by the air pollution [4], Korea would have the premature deaths of 1100 per 1 million people if it maintains the present condition, which is higher margin of increase than other OECD countries like Russia and Canada. As suggested above, the air pollution is one of the most significant social issue related to public health and all the people on the earth are likely to be influenced by the polluted air, thus, it is necessary to pay close attention to the issue.

---

[1] (2021. 9. 22), World Health Organization, What are the WHO Air quality guidelines? Improving health by reducing air pollution, retrieved from https://www.who.int/news-room/feature-stories/detail/what-are-the-who-air-quality-guidelines

[2] (2021. 9. 22), World Health Organization, 「WHO global air quality guideline」

[3] (2021. 9. 22), World Health Organization, Q&A for WHO Global Air Quality Guideline, retrieved from https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines

[4] (2016. 06. 09), OECD, 「The economic consequences of outdoor air pollution」

2005 V.S. 2021 WHO air quality guidelines (AQGs)

Preventable PM2.5 deaths avoided if new AQGs met globally: ~80% Source: WHO

Pollutant	Averaging Time	2005 AQGs	2021 AQGs
PM2.5 µg/m³	Annual	10	5
	24-hour	25	15
PM10 µg/m³	Annual	20	15
	24-hour	50	45
Ozone (O3) µg/m³	Peak Season*+	-	60
	8-hour**	100	100
Nitrogen dioxide (NO2) µg/m³	Annual	40	10
	24-hour*	-	25
Sulfur dioxide (SO2) µg/m³	24-hour	20	40
Carbon monoxide (CO) mg/m³	24-hour*	-	4

「2005 VS 2021 WHO air quality guidelines(AQGs)[5]」

II. Literature Review

WHO defined air pollution as the contamination of the indoor or outdoor environment by any chemical, physical, or biological agent that modifies the natural characteristics of the atmosphere[6]. In the report published by IQAir*, 2020 World Air Quality Report, it ranked the regions and the cities in the world only by the amount of fine particulate matter ($PM_{2.5}$)[7]. The reason why it used fine particulate matter as a criteria pollutant is because fine particulate matter is known as the most harmful pollutant to human health due to its prevalence and far-reaching health risks. It also added that the exposure to $PM_{2.5}$ can lead to negative health effects like cardiovascular disease, respiratory illness, and premature mortality since the microscopic size of $PM_{2.5}$ allows the particles to penetrate through the lungs and further be absorbed deep into the blood stream through inhalation.

Particulate matter, also called PM, is a mixture of solid and liquid particles in the air that are small enough not to settle out on to the Earth's surface under the influence of gravity, classified by aerodynamic diameter[8]. Particles with an aerodynamic diameter of equal or less than $2.5m^3$ called $PM_{2.5}$ or fine particulate matter and of equal or less than $10m^3$ called PM_{10} or particulate matter. $PM_{2.5}$ is

[5] (2021. 12. 02), New WHO air quality guidelines will save lives, IQAir, retrieved from https://www.iqair.com/blog/air-quality/2021a-WHO-air-quality-guidelines

[6] World Health Organization, retrieved from https://www.who.int/health-topics/air-pollution#tab=tab_1

[7] (2021. 08. 17), IQAir & Greenpeace, 「2020 World Air Quality Report」

[8] (2021. 9. 22), World Health Organization, 「WHO global air quality guideline」

usually emitted from the sources like fossil-fuel powered motor vehicles, power generation, industrial activity, agriculture, and biomass burning.

World Meteorological Organization (WMO) released 'The Air Quality and Climate Bulletin' for the International Day of Clean Air for Blue Skies on 7th of September 2021. According to the report, global ambient air pollution mortality has increased from 2.3 million in 1990 to 4.5 million in 2019 and among 4.5 million, 4.1 million of mortality was dominated by particulate matter, which accounts for 92% of total mortality[9]. As we can find out from the reports, particulate matter is deeply linked to human lives and gives much more remarkable impact on them compared to other air pollutants.

Korea is a country that is severely suffering from air pollution caused by PM. According to the research conducted by Greenpeace, Korea has the annual average $PM_{2.5}$ concentration of $24.8\mu g/m^3$, which is the highest among OECD countries[10]. Also, in the report 'How's Life?' released by OECD in 2021, measuring the well-being of OECD nations, the statistics showed that Korea has much higher level of share of population exposed to $PM_{2.5}$ above $10\mu g/m^3$ than the average level of OECD countries.

*IQAir: IQAir is a Swiss-based air quality technology company that seeks to empower individuals, organizations and communities to breathe cleaner air through information, collaboration and technology solutions. IQAir's AirVisual global air quality information platform aggregates, validates and calibrates air quality data from a wide variety of sources, including governments, private citizens and organizations. The platform supports the free integration of sensor data from a variety of low-cost sensor and monitoring devices.

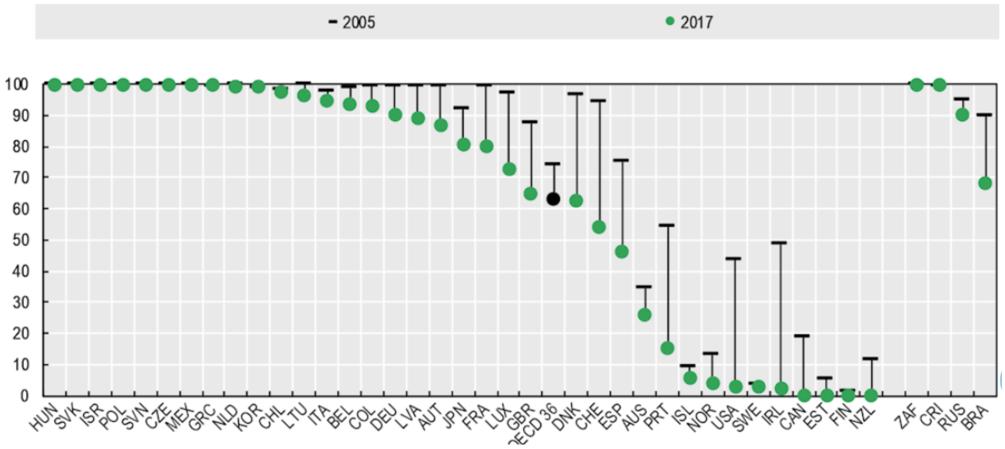
---

[9] (2021. 09. 07), World Meteorological Organization, 「WHO global air quality guideline」, retrieved from https://public.wmo.int/en/our-mandate/focus-areas/environment/air_quality

[10] (2020. 02. 25), Greenpeace, '한국, OECD 회원국 중 초미세먼지 최악의 국가', retrieved from https://www.greenpeace.org/korea/press/12092/korean-fine-dust-airvisual/

Figure 7.2. **Compared to 2005, fewer people are exposed to $PM_{2.5}$ above the WHO threshold level**

Share of the population exposed to $PM_{2.5}$ above 10 micrograms/m^3, percentage



「Share of the population exposed to $PM_{2.5}$ above 10$\mu m/m^3$, percentage[11]」

Korean government has implemented several policies for the purpose of improving the air quality by reducing PM level. There are two main policies exercised, 'Seasonal management system for PM' and 'Emergency Reducing Action for PM'. 'Seasonal management system for PM' is an intensive management measure to reduce the frequency and degree of high concentrations of PM by operating stronger preventive measures than usual from December to March, when high concentrations of PM are generated frequently[12]. 'Emergency Reducing Action for PM' is the policy implemented to reduce PM level in very short period by reducing the emissions of pollutants from cars or factories when the high concentration of PM within Seoul City lasts for certain period[13]. In addition, Korean government spent tremendous amount of its national budget in improving the polluted air. According to the report from Korean Ministry of Environment, 4,500 million dollars of the budget was assigned to the policy for enhancing the air quality called 'The Green New Deal', which accounts for over 30% of the total budget assigned for the Ministry of Environment as a whole[14]. The purpose of 'The Green New Deal' policy is

---

[11] OECD, How's Life? 2020: Measuring Well-being, retrieved from https://www.oecd-ilibrary.org/sites/9870c393-en/1/3/7/index.html?itemId=/content/publication/9870c393-en&_csp_=fab41822851fa020ad60bb57bb82180a&itemIGO=oecd&itemContentType=book

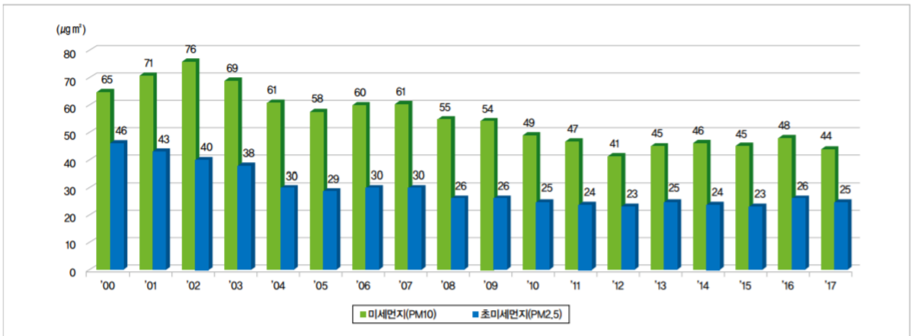
[12] Seoul Metropolitan Government, 미세먼지 계절관리제, retrieved from https://cleanair.seoul.go.kr/information/seasonal

[13] Seoul Metropolitan Government, 미세먼지 비상저감조치, retrieved from https://cleanair.seoul.go.kr/institution/prkplcContent

[14] (2020. 09. 01), Korean Ministry of Environment, '2021년 환경부 예산…그린뉴딜 선도하며 환경 안전망 강화', retrieved from

to reduce greenhouse gas emissions and PM level by encouraging usage of green mobility such as electric cars and hydrogen fueled cars. As mentioned above, Korean government keep trying to remove PM from air in many ways.

〈그림 1〉 연도별 미세먼지와 초미세먼지의 농도 변화(서울시 전체 평균, '00~'17)



출처: 서울특별시 2012–2017.

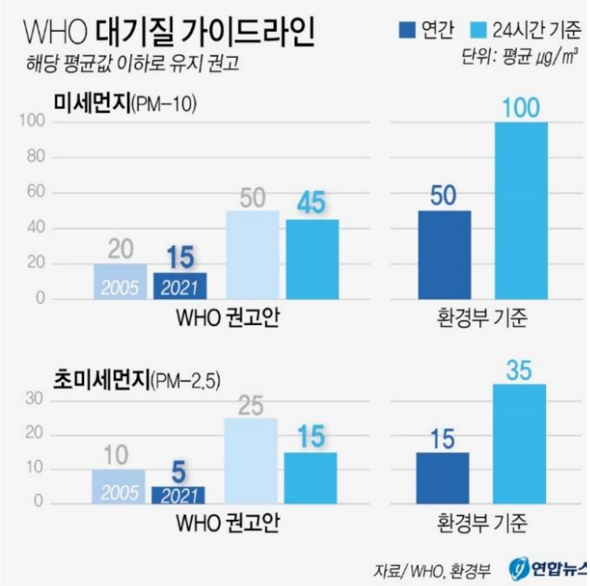
「Change in annul PM_{10}, $PM_{2.5}$ concentration in Seoul」

As a result of these efforts, the absolute level of PM has decreased in fact. The graph above showed Change in annual PM_{10}, $PM_{2.5}$ concentration in Seoul City. The graph indicates that the amount of PM_{10} and $PM_{2.5}$ definitely have a tendency of decrease in the long term, which is a positive signal this far. However, we cannot merely interpret the data as positive since there are two points that we should consider carefully. Firstly, even though the level of PM in Seoul City has been reduced, it is still much higher than the international standards suggested by WHO. Secondly, according to the research[15], the decrease trend of $PM_{2.5}$ is quite flat since 2017, which means that there are no noticeable improvements despite a lot of effort of the government. As the level of PM is deeply connected to people's health and have direct impact on them, we have to estimate the future amount of PM as precisely as possible, so that we can use the data to figure out the factors that interrupt the improvements of air quality and to effectively distribute the national budget to maximize the effect.

---

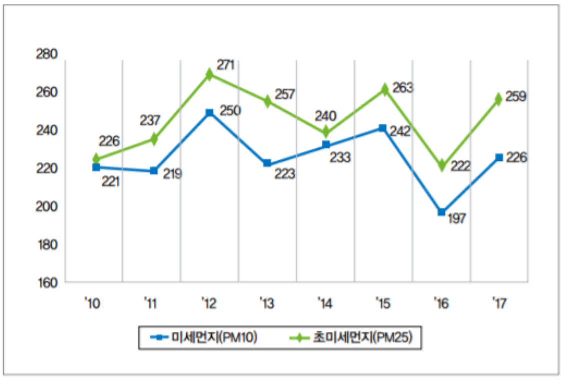
http://www.me.go.kr/home/web/board/read.do?boardMasterId=1&boardId=1393940&menuId=286

[15] (2019. 06), 이수지&김호, 「미세먼지가 건강에 미치는 영향」



「**Comparison between air quality guidelines of WHO and Korean Ministry of Environment**[16]」

〈그림 2〉 연도별 미세먼지 WHO 24시간 권고기준 이하 일수 추이



출처: 서울특별시보건환경연구원 2012-2019

「**Annual trend of the number of days below the WHO 24-hour recommendation for PM**[17]」

III. Statement of Research Objectives

The main goal of our study is to examine the trend in the concentration of PM in Seoul and forecast the future tendency by making a predictive model. From the literature review, we found out that PM in

[16] (2021.09. 23), retrieved from https://www.yna.co.kr/view/GYH20210923001200044

[17] (2019. 06), 이수지&김호,「미세먼지가 건강에 미치는 영향」

Seoul has tended to decrease in the long-term since 2002, but has been stagnant since 2017. Such particular matter is a factor greatly affects human's health, such as increasing cardiovascular diseases and respiratory-related diseases. So, we should continuously pay attention to and manage in the future. Accordingly, we intend to conduct research that could help identify and prepare for future trends in PM. Furthermore, we hope our model contribute to trace the factors that disturb the reduction of PM and efficiently utilize the national budget on environment sector.

IV. Description of Data and Applied Method

i. Description of Data

The monthly average amount of Particulate Matter(PM10) in Seoul

The data used in this paper is the monthly average level of PM10 in Seoul City. We managed the data collected for last 5 years and a half, from January 2015 to June 2021. The reason we decided to use the data of PM10 is that Seoul Metropolitan Government generally gives more attention to the level of PM10 when they measure and evaluate the quality of air and the citizens in Seoul are also more aware of the danger of PM10

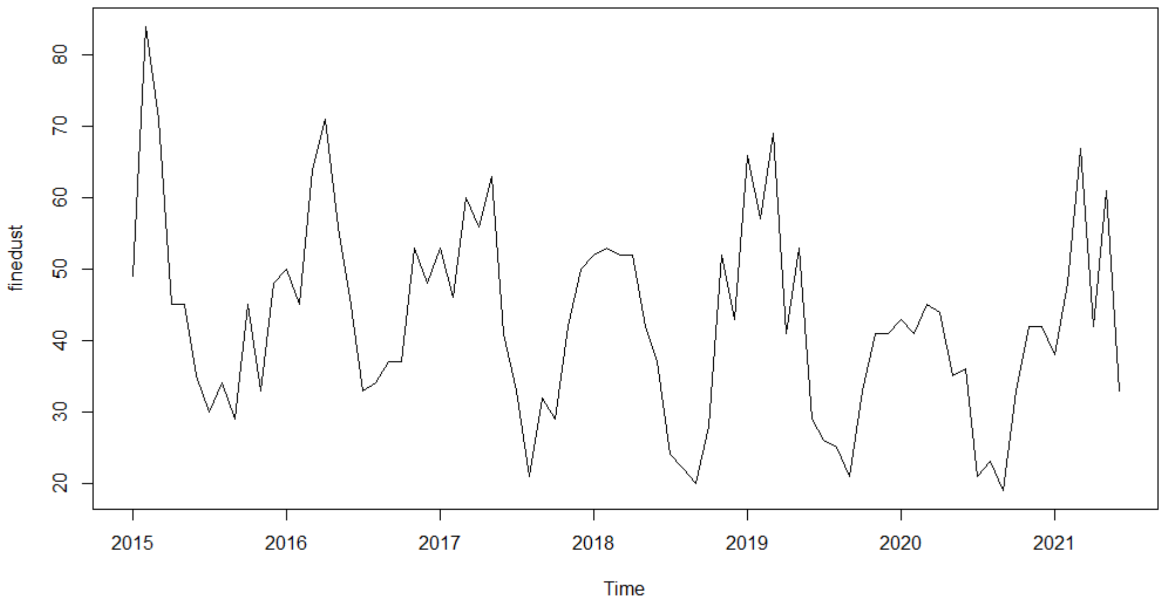


Figure 1 The Level of Air Pollution caused by Particulate Matter(㎍/m^3)

The graph above shows the data in time-series plot. The dependent variable is the level of air pollution caused by PM_{10} per month and the independent variable is the passage of time period between 2015 to 2021. From this graph, we could find that the level of air pollution caused by PM slightly decreases with the seasonal trend as time goes by.

```
> summary(finedust)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  19.00   33.00   42.00   42.82   52.00   84.00
```

Table 1 Summary Statistics

The summary of the data provides overall description such as mean value of 43, median value of 42 and also the minimum value of 19 and maximum value of 84.

ii. Additive Decomposition Model

Regression statistics								
Multiple R	0.328591121							
R^2	0.107972125							
Adj_R^2	0.097093736							
standard error of model	13.38774867							
n	84							
ANOVA								
	d.f	SS	MS	F-stat	Significance F			
Reg	1	1778.943586	1778.943586	9.925378429	0.002275662			
Residual	82	14697.00879	179.2318146					
Total	83	16475.95238						
	coefficient (OLS)	s. e. of coefficients	t-stat	p-value	lower 95%	upper 95%	lower 95%	upper 95%
Intercept	49.75674125	2.947726677	16.87969975	2.06962E-28	43.89277387	55.62070863	43.89277387	55.62070863
Time	-0.189794472	0.060243476	-3.15045686	0.002275662	-0.309637937	-0.069951007	-0.309637937	-0.069951007

Table 2 Regression Output

The result of the regression output using Excel is provided above. The adjusted R squared value was about 0.0971, which indicates that the explanatory power of the model is quite low. As the coefficient of x variable is negative, we could realize again that the influence of PM is decreasing as time goes by. Also, through the result of P-value for both F-test and T-test which are below 0.05, we noticed that estimation for both model and coefficient are statistically significant in level 5%.

➔ Air Pollution Caused by PM-10($\mu g/m^3$) = 49.75674.. - 0.19* Time

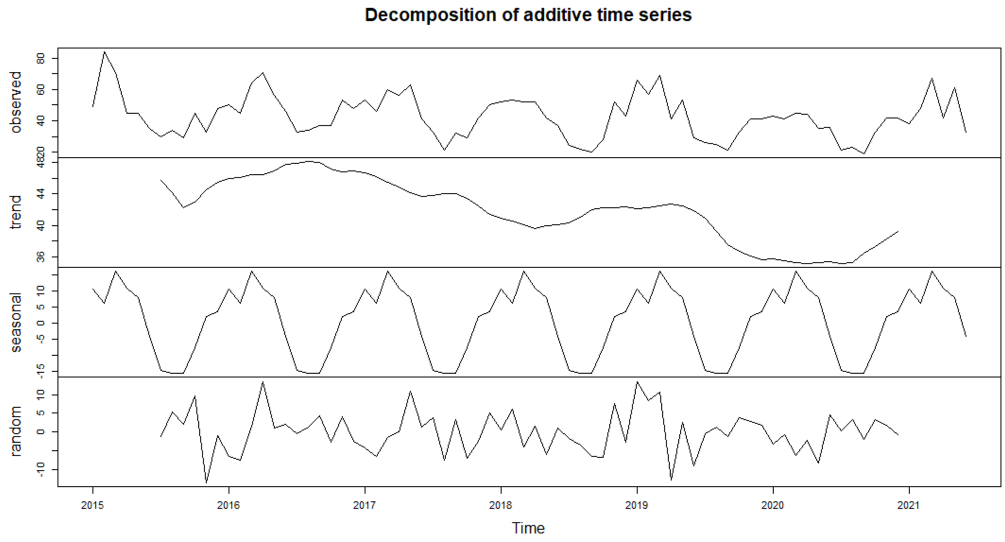


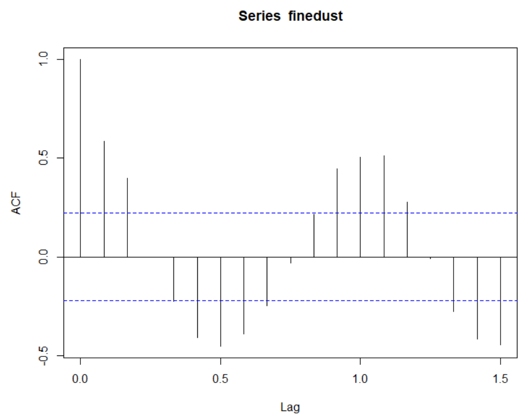
Figure 2 Summary of Additive Decomposition Model

The graph above is a decomposition of additive time series of the air pollution by PM_{10}. The first plot is the actual observed values. The second plot shows the trend of the data. We can observe that the level of air pollution is decreasing in macro scope. The third plot is the seasonal plot which shows clear seasonality. From the last one, we could see some stationarity. However, it was not certain enough to conclude that the model is stationary. So, we decided to check the stationarity of the model first. This led us to select seasonal ARIMA model.

In-Sample RMSE	Out-Sample RMSE
13.31762	11.99305

Table 3 Forecasting Result of Decomposition Model

iii. Multiplicative Seasonal ARIMA Model



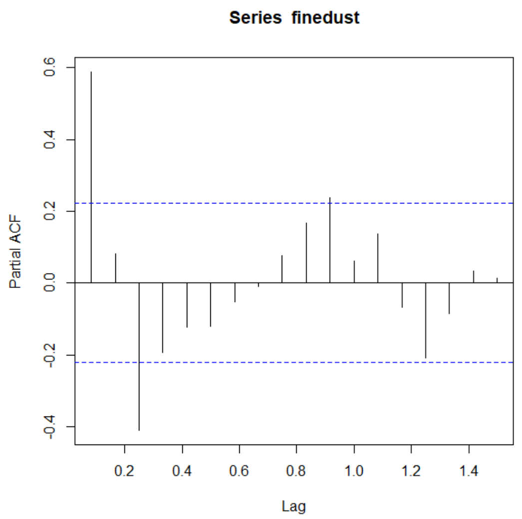


Figure 3 ACF & PACF Plots

Before constructing forecasting model, we performed autocovariance and partial autocovariance function test first. In ACF and PACF test results, we observed some values exceeded the threshold presented in blue line, which was not enough to judge the stationarity. So additional KPSS test was conducted to confirm the stationarity with numerical value.

```
> kpss.test(finedust)

        KPSS Test for Level Stationarity

data:  finedust
KPSS Level = 0.23763, Truncation lag parameter = 3, p-value = 0.1
```

According to the above result of KPSS test. Since the p-value is higher than 0.05, the data is finally proven to be stationary.

```
Series: finedust
ARIMA(0,0,0)(1,1,0)[12] with drift

Coefficients:
         sar1    drift
      -0.6024  -0.1507
s.e.   0.1119   0.0657

sigma^2 estimated as 92.8:  log likelihood=-244.85
AIC=495.69   AICc=496.08   BIC=502.26

Training set error measures:
                  ME     RMSE    MAE       MPE    MAPE      MASE      ACF1
Training set 0.25219 8.726168 6.3188 -1.677446 15.1757 0.7596371 0.1691545
```

With the data confirmed as stationary, we moved on to the next step of fitting seasonal ARIMA model. Using auto.arima function, we got (0,0,0)(1,1,0)[12] as a parameter. In addition, because ACF and PACF of the former model was not clear enough, we analyzed the residuals of the model to find out if the forecasting model is stationary.

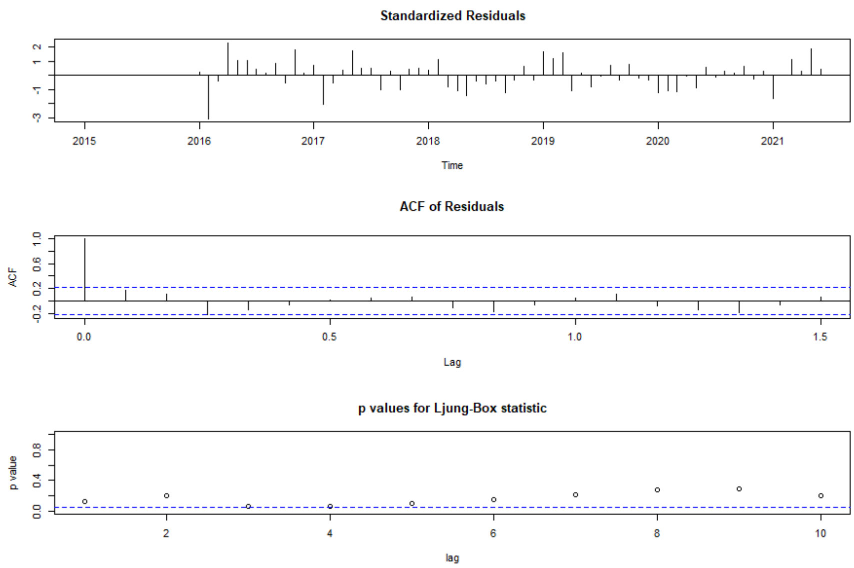


Figure 4 Diagnostic Plots for time-series Fits

The plot of standardized residuals shows the linearity and homogeneity of variance. Second plot shows that there is no autocorrelation between residuals. Also, according to Ljung-Box statistic, we can observe that all the p-values are over 0.05, which means there is no autocorrelation issue.

The forecasting result from July, 2021 to December, 2021 is presented below.

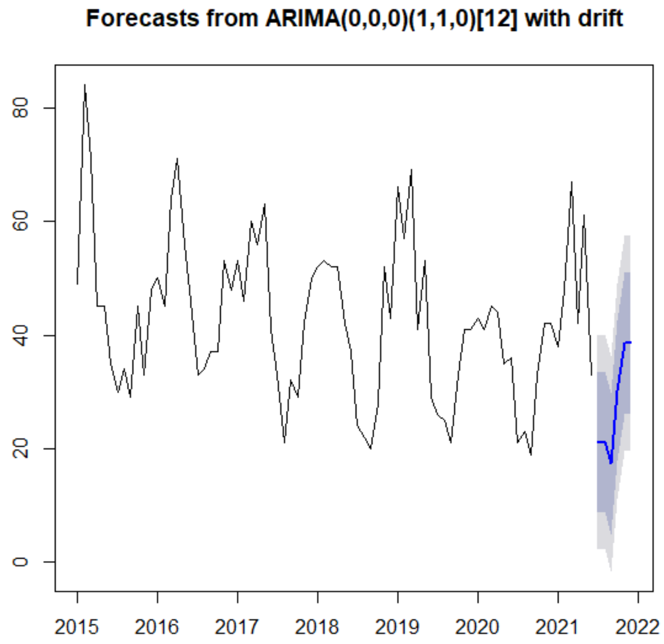


Figure 5 Forecasting Plot based on Multiplicative Seasonal ARIMA Model

In-Sample RMSE	Out-Sample RMSE
13.24035	9.706892

Table 4 Forecasting Result of Multiplicative Seasonal ARIMA Model

iv. Conclusion

This paper aimed to construct a forecasting model for air pollution caused by Particulate Matter with past data. Through examining several statistical methodologies and function as decomposition and seasonal ARIMA, this paper obtained the initial research objective by successfully predicting future tendency with the forecasting model. Whether the prediction model is stationary was verified several times at various stage, and as a result, the prediction appeared to be decreasing, reflecting the seasonal trends of past data.

V. Expected Original Contribution

According to the model built in this paper, the concentration of PM_{10} in the future will continue to decrease like the previous trend, but the decline trend is not steep. Also, it has a clear seasonality, as it is found that the concentration is particularly high in spring which is from December to March. It is expected that the model built in this paper can help the government effectively distribute their budgets related to fine dust reduction policies in the future and can prove the implementation of the seasonal management system should be maintained in the future.

Based on the trend we examined above, we conducted additional research on the cause of trend. First, the decrease in the overall PM level was expected to be derived from the reduction policy. On the other hand, the reason for the slow decline of PM is examined that China's contribution to particular matter didn’t decrease and the average wind speed in Seoul decreases so that the time PM staying in the air got longer. However, the causes we mention are only a part of it and there are various other reasons. As such, further researches and investigations would be required for this study since there are many factors affecting PM and cannot be concluded as one.

References

- (2021. 9. 22), World Health Organization, What are the WHO Air quality guidelines? Improving health by reducing air pollution, retrieved from https://www.who.int/news-room/feature-stories/detail/what-are-the-who-air-quality-guidelines
- (2021. 9. 22), World Health Organization, 「WHO global air quality guideline」
- (2021. 9. 22), World Health Organization, Q&A for WHO Global Air Quality Guideline, retrieved from https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines
- (2016. 06. 09), OECD, 「The economic consequences of outdoor air pollution」
- (2021. 12. 02), New WHO air quality guidelines will save lives, IQAir, retrieved from https://www.iqair.com/blog/air-quality/2021a-WHO-air-quality-guidelines
- World Health Organization, retrieved from https://www.who.int/health-topics/air-pollution#tab=tab_1
- (2021. 08. 17), IQAir & Greenpeace, 「2020 World Air Quality Report」

- (2021. 9. 22), World Health Organization, 「WHO global air quality guideline」
- (2021. 09. 07), World Meteorological Organization, 「WHO global air quality guideline」, retrieved from https://public.wmo.int/en/our-mandate/focus-areas/environment/air_quality
- (2020. 02. 25), Greenpeace, '한국, OECD 회원국 중 초미세먼지 최악의 국가', retrieved from https://www.greenpeace.org/korea/press/12092/korean-fine-dust-airvisual/
- OECD, How's Life? 2020: Measuring Well-being, retrieved from https://www.oecd-ilibrary.org/sites/9870c393-en/1/3/7/index.html?itemId=/content/publication/9870c393-en&_csp_=fab41822851fa020ad60bb57bb82180a&itemIGO=oecd&itemContentType=book
- Seoul Metropolitan Government, 미세먼지 계절관리제, retrieved from https://cleanair.seoul.go.kr/information/seasonal
- Seoul Metropolitan Government, 미세먼지 비상저감조치, retrieved from https://cleanair.seoul.go.kr/institution/prkplcContent
- (2020. 09. 01), Korean Ministry of Environment, '2021년 환경부 예산…그린뉴딜 선도하며 환경 안전망 강화', retrieved from http://www.me.go.kr/home/web/board/read.do?boardMasterId=1&boardId=1393940&menuId=286
- (2019. 06), 이수지&김호, 「미세먼지가 건강에 미치는 영향」
- (2021.09. 23), retrieved from https://www.yna.co.kr/view/GYH20210923001200044