Inference for Functional Data

by
MUNG Tsz
(19230680)

A thesis submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science (Honours)
in Mathematics and Statistics

at

Hong Kong Baptist University

Date: 29th, December 2022

ACKNOWLEDGEMENT

The methodology described in Section 1 and 3 were done jointly with Jin-Ting Zhang, Ming-Yen Cheng, Hau-Tieng Wu, Bu Zhou, Xue-hua Liang respectively. Part of the work presented in this thesis was done by Christian Acal1. Ana M. Aguilera, Annalina Sarra, Adelia Evangelista, Tonio Di Battista, Sergio Palermi. All other experiments described in this thesis were my own original work and were carried out by myself under the supervision of Professor M.Y. Cheng.

Signature of Student

_____Mung Tsz_____
Student Name

Department of Mathematics
Hong Kong Baptist University

Date: _29/12/2022_____

Functional one-way ANOVA approach for detecting changes in air pollution in Hong Kong during COVID-19 Pandemic

MUNG Tsz
(19230680)

Department of Mathematics

ABSTRACT

FANOVA has shown to be useful for tracking the changes in air quality throughout the course of both time periods. In view of this, we apply a test for the one-way ANOVA problem in functional data analysis by using the air pollutants data during COVID-19. The maximum point-wise F-test statistic during the time span in which functional responses are recorded is used to calculate the test statistic, which is called F_{max} . Nonparametric bootstrap is utilized, which is more widely applicable and easier to apply than parametric bootstrap, to approximate the critical value and the null distribution. The findings of this study indicate that the PM2.5 and PM10 concentration levels altered as Hong Kong implemented rigorous containment measures in accordance with COVID-19.

1 Introduction

As data collection technology advances, functional data analysis (FDA) is getting more sophisticated and popular. Functional data may provide more specific information about the underlying system than traditional data. By observing the random variable made over a continuous time, space, interval or a substantial discretization of it, the results can be functions, curves and surfaces which are used to represent functional data. The papers by Ferraty (2011), Ramsay and Silverman (2002, 2005), Zhang (2013), offer a wide range of viewpoints of functional data analysis and significantly advance FDA. The issues with functional data have frequently been researched, such as classification (Alonoso and Casado 2012), principal components analysis (Berrendero et al., 2011) and regression analysis (Cuevas et al., 2002). The application of FDA methods has been widely used in many fields, including biomechanics, psychology, medicine etc. One of those fields is air quality, which is the focus of this paper. Many studies in recent years have used FDA technique to analyze air quality. For example, Martı́nez Torres et al. (2020) suggested three approaches to test how well they work in locating pollution NO2 events and outliers. It was discovered that the functional data analysis method's continuous feature enhances the ability to spot outliers, making FDA more successful for assessing air pollution control techniques.

From 2020 to 2022, the shock of the coronavirus affects the economy, society and environment of many countries. Many articles have studied the impact of human activities on the changes in air quality during Covid-19. For instance, the analysis by Bao and Zhang (2020) revealed that travel restrictions implemented in 44 northern Chinese cities significantly reduced air pollutant emissions during the lockdown caused by Covid-19, especially NO2 and PM10. Also, they found a significant discovery in the robust correlation between sharp drops in human mobility and air pollution reduction. Sharma et al. (2020) found that PM2.5 had the greatest reduction overall among all pollutants in most areas of India. On the other hand, an interesting finding was the increase in O3 in most regions during Covid-19. A study (Wang et al., 2020) in China showed a significant result that the decline in anthropogenic emissions, particularly from the transportation and industrial sectors, was a factor in the decline in PM2.5 concentrations. However, there is insufficient evidence to show that air pollution is severe in some places. For example, some decrease in PM2.5 is caused by unfavourable meteorological conditions. Therefore, there are many factors affecting the air quality but not only human activities.

Globally government responses to COVID-19 include limiting restaurant hours, encouraging working from home, requiring online learning, forcing the closure of public services, and implementing lockdowns. Hale et al. (2021) gave an overview of the relationship between Covid-19 containment response policies and air pollutants changes in 1851 cities. They observed a greater reduction in NO2 at the onset of the COVID-19 pandemic due to different lockdown measures. They also found that the increased NO2 was associated with mild containment response policies. Although strict containment response policies reduce emissions from some major sources, such as the transport and industrial sectors, its impact on other sectors such as the residential, electricity and agricultural sectors is limited. The conclusion is that the response Covid-19 policies may help improve or worsen air quality.

Otmani et al. (2020) showed the significant decrease in PM10, SO2 and NO2 in Morocco due to the lockdown. Similar results caused by lockdown are also found in Brazil (Dantas et al., 2020) and Spain (Mesas-Carrascosa et al., 2020). Unlike the previous result, Acal et al. (2022) point out that NO2 decreased but PM2.5 and PM10 increased during the lockdown in Abruzzo (Italy). Some regions do not implement lockdowns, but most industrial and commercial activities have been forced to stop due to strict containment policies. Except for New York (Zangari et al., 2020), other countries or regions including the U.S. (Berman and Ebisu 2020), Malaysia (Abdullah et al., 2020) and California (Narger and Murphy 2020) have found that the air pollutants such as PM2.5 and NO2 significantly decreased.

In this study, we would like to investigate whether the concentration of air pollutants differs before and after implementing the strict containment response Covid-19 policies in Hong Kong, which is considered as a city with high population density and severe air pollution. The data was provided by Environmental Protection Interactive Centre. In the past three years, The Hong Kong government implemented the strictest containment policies when the fifth wave of the Covid-19 epidemic occurred, so we chose the period from October 9, 2021, to January 6, 2022, before the policies were released, and January 7, 2022, to April 6, 2022, during the implementation of the policies. Nitrogen dioxide (NO2), fine particulate matter (PM2.5), ozone (O3) and respirable suspended

particulates (PM10) are the target air pollutants studied in this paper. In this paper, we focus on the method for testing the functional one-way ANOVA problem (Cuevas et al., 2004; Zhang and Liang 2013; Faraway 1997), which is a classical problem for inference. A one-way functional ANOVA problem is often tackled by first identifying the differences between the mean group functions.

Let $x_{i1}(t), x_{i2}(t), \dots, x_{in_i}(t), i = 1,2, \dots, k$ represent k distinct sets of random function groups that are defined throughout the bounded and closed interval $\mathcal{T} = [a, b]$. Let $SP(\mu, \gamma)$ stands for a stochastic process with mean function $\mu(t)$, $t \in \mathcal{T}$, and covariance functions $\gamma(s,t), s,t \in \mathcal{T}$. Suppose $x_{i1}(t), x_{i2}(t), \dots, x_{in_i}(t), i = 1, 2, \dots, k$ are i.i.d. samples taken from the stochastic process $SP(\mu, \gamma)$. The purpose of this problem, which is often referred to as the one-way ANOVA problem for functional data, is to test the null hypothesis

$$H_0 : \mu_1(t) \equiv \mu_2(t) \equiv \cdots \equiv \mu_k(t), t \in \mathcal{T}, \tag{1}$$

versus the alternative that at least two of the mean functions are not equal. This problem is also known as k-sample testing problem for independent functional data.

For functional data, there are multiple current tests for the one-way ANOVA problem (1). It is possible to use the $\mathcal{L}^2$-norm-based test (Zhang and Chen 2007; Faraway 1997) for (1), which based on the the pointwise between-group variations. It is also possible to use the F-type test (Zhang 2011, Fan and Lin 1998) for (1), which is based on the pointwise within-group and between-group variations. Additionally, *globalization of the pointwise F-test* is also known as the GPF test (Zhang and Liang 2013), which is produced by globalizing the pointwise F-test (Ramsay and Silverman 2005; Zhang 2013). Due to nonparametric bootstrap, the Globalized Pointwise F-test (GPF test) requires less computation than the previously described two tests (Zhang and Liang 2013). The pointwise F-test (Ramsay and Silverman 2005) uses the test statistic

$$F_n(t) = \frac{SSR_n(t)/(k-1)}{SSE_n(t)/(n-k)}, \tag{2}$$

where $n = \sum_{i=1}^{k} n_i$ denotes the total sample size,

$$SSE_n(t) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}[\bar{x}_{ij}(t) - \bar{x}_i(t)]^2 \text{ and } SSR_n(t) = \sum_{i=1}^{k} n_i[\bar{x}_i(t) - \bar{x}(t)]^2 \tag{3}$$

are the pointwise within-group and between-group variations respectively, where

$$\bar{x}_i(t) = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}(t) \text{ and } \bar{x}(t) = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}(t), i = 1,2, \dots, k, \tag{4}$$

are the sample group mean functions and sample grand mean functions respectively.

The null distribution of pointwise F-test $F_n(t)$ uses the same critical value $F_{k-1,n-k}(\alpha)$ from the F-distribution for any prescribed level, as a result, we can test (1) at all points in the interval $\mathcal{T}$. Even when the pointwise F-test is significant for every t in $\mathcal{T}$ at the same significance level, the pointwise F-test make it time-consuming to test all $t \in \mathcal{T}$, and it does not ensure that (1) is significant overall. Zhang and Liang (2013) then introduced the GPF test employing the integral of the pointwise F-test over interval $\mathcal{T}$

$$T_n = \int_{\mathcal{T}} F_n(t)\, dt. \tag{5}$$

to address this problem. For these existing techniques in use, there are some thorough simulation studies. The GPF test used for (1), according to Zhang and Liang (2013), is comparable to the $\mathcal{L}^2$-norm-based test and F-type test in terms of power and size controlling. Additionally, Górecki and Smaga (2015) came to the conclusion that GPF is one of the tests that function effectively when solving one-way ANOVA problems after thoroughly stimulating and comparing the available methods.

Zhang and Liang (2013) globalize the pointwise F-test using its maximum value over $\mathcal{T}$ via bootstrapping

$$F_{max} = \sup_{t \in \mathcal{T}} F_n(t). \tag{6}$$

By establishing a threshold for the relevant statistics, the F_{max} approach can determine the areas of rejection. Ramsay and Silverman (2005) (p. 234) use the square root of $F_n(t)$ as test statistic and a permutation-based critical value, which share similar concept in the F_{max} test. Via intensive simulation studies by Zhang et al. (2019), the

F_{max} test is typically preferable to the GPF test in terms of power and size control. As a result, this paper mainly adopts F_{max} test to study the air quality dataset. Based on the p-values from F_{max} test, the result of reconstructed PM2.5 and PM10 are significant. F_{max} test is a relatively new technique for solving one-way ANOVA problem. Zhang and Liang (2013) briefly mentioned it in their paper but did not explore it in depth until Zhang et al. (2019) performed a simulation study in 2014 to contrast the power of F_{max} test with the other three tests. Few comparative research on F_{max} test and *Globalized Pointwise F-test* (GPF) for the functional one-way ANOVA problem were conducted after that. In order to assess the effectiveness of the GPF and F_{max} tests, this paper also applies the GPF test to the data on air quality. We found that they share the same conclusion from the result, that is both tests for PM2.5 and PM10 are significant while NO2 and O3 are not.

To derive the random expression of the F_{max} test statistic, Zhang et al. (2019) used parametric bootstrap (PB) method to approximate the null distribution. Despite the reliance of this method, the method is computationally inefficient. Therefore, Zhang et al. (2019) proposed a more efficient nonparametric bootstrap (NPB) method to approximate the finite sample null distribution. Also, they demonstrated that the asymptotic level of the F_{max} test derived from NPB is accurate. On the other hand, through extensive simulation studies of Smaga (2020), Box-type approximation and parametric bootstrap are not suggested if the sample size is small. He proposed that nonparametric bootstrap (NPB) can work correctly. The result from Zhang and Liang (2013) suggested that the F_{max} may have higher power than the GPF test. Thus, we can apply our dataset on this test by using NPB method.

The structure of the essay is as follows. The studied area, the details of studied period, the data, and the preliminary analysis of the air pollutants in this study are all described and shown in Section 2. The methodology about functional data reconstruction and F_{max} test is introduced in Section 3. By applying the F_{max} test, the result of air pollutants concentration described in Section 2 is given in Section 4. The paper is concluded in Section 5.

2 Air quality data and studied period

This section gives the background of the studied area and the reason of choosing the studied period. In order to make the test perform better, this section also introduces how to prepare reconstructed data.

2.1 Description of studied region

We have chosen three general stations: Tuen Mun, Tung Chung, Kwun Tong and two roadside station stations: Mong Kong and Causeway Bay, which are all shown in Fig. 1. Air pollution in Hong Kong is mainly from vehicle exhaust, ships and power plants. Hong Kong is currently facing two types of air pollution problems - roadside air pollution and regional smog problems. Roadside air pollution is mainly caused by the exhaust of diesel vehicles, especially trucks, buses and minibuses; while the regional smog problem is caused by pollutants emitted by vehicles, factories and power plants in Hong Kong and the Pearl River Delta region. Tuen Mun and Tung Chung, located in the Northwest District of Hong Kong, mostly face Shenzhen and the Pearl River Estuary. When northerly winds blow, these areas bear the brunt of pollutants from onshore industrial activities, which are also expected to be linked to local and cross-boundary vehicle emissions. Among them, respirable suspended particulates (PM10), fine suspended particles (PM2.5), and NO2 are the main pollutants. These geographic locations are more susceptible to regional air pollution. Mong Kok is located in the west of Kowloon, Hong Kong. Its commercial, retail and other economic activities are very frequent, so its population density is extremely high. The traffic in the area is developed, with minibuses, buses and MTR (East Rail Line, Tsuen Wan Line and Kwun Tong Line) going directly to various areas. Causeway Bay is located in the west of the north shore of the center of Hong Kong Island. It is the main commercial and entertainment venue in Hong Kong. Therefore, the urban traffic in the business district of Causeway Bay is very busy and surrounded by many high-rise buildings. Finally, Kwun Tong is the largest industrial area in East Kowloon, and there are a large number of residential buildings to the north of Kwun Tong Road to provide housing for people. In terms of road design, the roundabout on Hoi Yuen Road connects the entire industrial area with Kwun Tong Road, but the road is very narrow. The roundabout is a

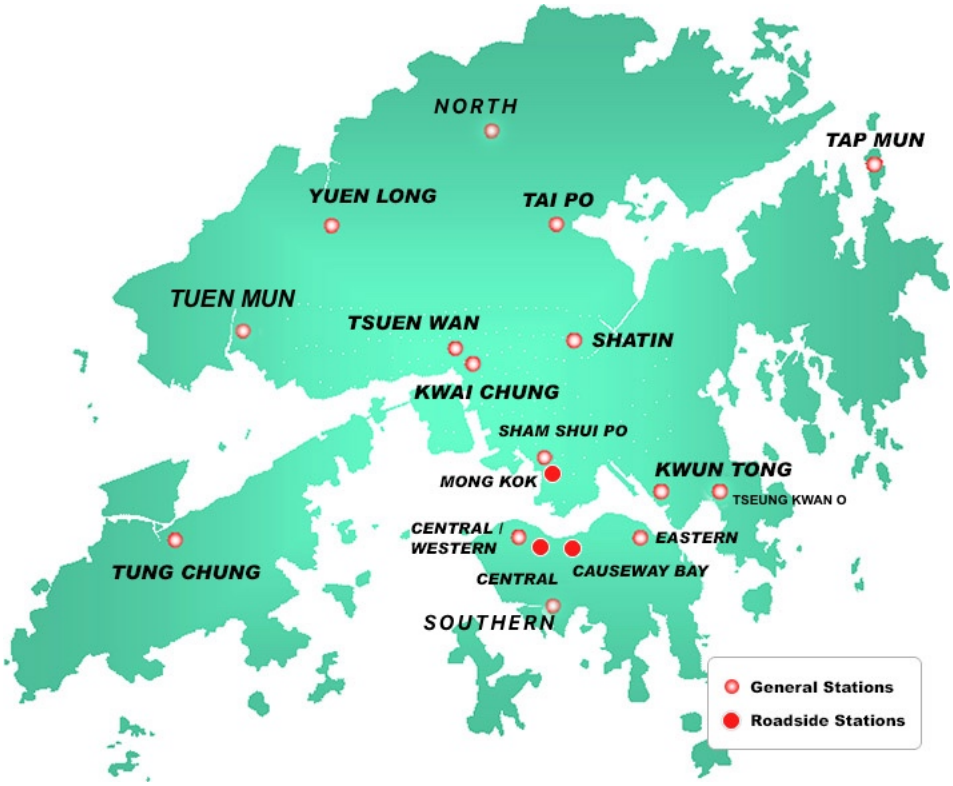

Fig. 1 locations of monitoring stations, graph from: Environmental Protection Interactive Centre.

traditional two-ring design. The inner lanes cannot be tangent. Many vehicles use the outer lanes, resulting in uneven utilization of the inner and outer circles. For convenience, some vehicles enter and exit the roundabout in a straight line, making the traffic extremely difficult confusion. During peak hours, Kwun Tong Road often encounters heavy traffic jams and continuously emits air pollutants such as nitrogen dioxide.

2.2 Data Description

According to the 2020 Hong Kong Air Pollutant Emission Inventory Report of the Environmental Protection Department (EPD), the six major pollutants are sulfur dioxide (SO2), nitrogen oxides (NOx), respirable suspended particulates (RSP or PM10) , fine suspended particles (FSP or PM2.5), volatile organic compounds (VOC) and carbon monoxide (CO). In this paper, we mainly focus on the air pollutants NO2, PM2.5, O3 and PM10. The concentration unit of each pollutants is ug/m3.

Until September 2022, Hong Kong has experienced five outbreaks of the epidemic, and the most serious one is the fifth wave of the epidemic, which is still ongoing from late December 2021 to late 2022. According to the Hong Kong Chief Executive's speech to the public on February 28, 2022 on stabilizing the fifth wave of the epidemic [7], from February 20 to 27, Hong Kong recorded 117,033 confirmed cases of the novel coronavirus, making it since the outbreak of the current wave of the epidemic on December 30 last year, the cumulative number of cases has reached 158,683, which is 12 times the total number of confirmed cases in Hong Kong in the past two years. Therefore, the Hong Kong government began to adopt more stringent social distancing measures than before to deal with the fifth wave of the epidemic on January 7, 2022. For example, listed premises regulated by Chapter 599F of the Hong Kong Laws must be closed (such as bars and pubs, and game machine centers , bathrooms, fitness centers, playgrounds, party rooms, etc.), cancelling a number of large-scale events, prohibiting dine-in after 6 pm in restaurants, etc. It was not until April 21 that Hong Kong's social distancing measures were gradually relaxed. So we take January 7th as the midpoint. In order to ensure that we have the same time interval, we chose 90 days before the tightening measures and 90 days after the tightening measures as the research time period, that is, from October 9, 2021 to January 6, 2022, and January 7, 2022 to April 6, 2022. Daily mean measurements of pollutants have been collected from October 2021 to April 2022, there are few missing values that are replaced by its mean value for simplicity.

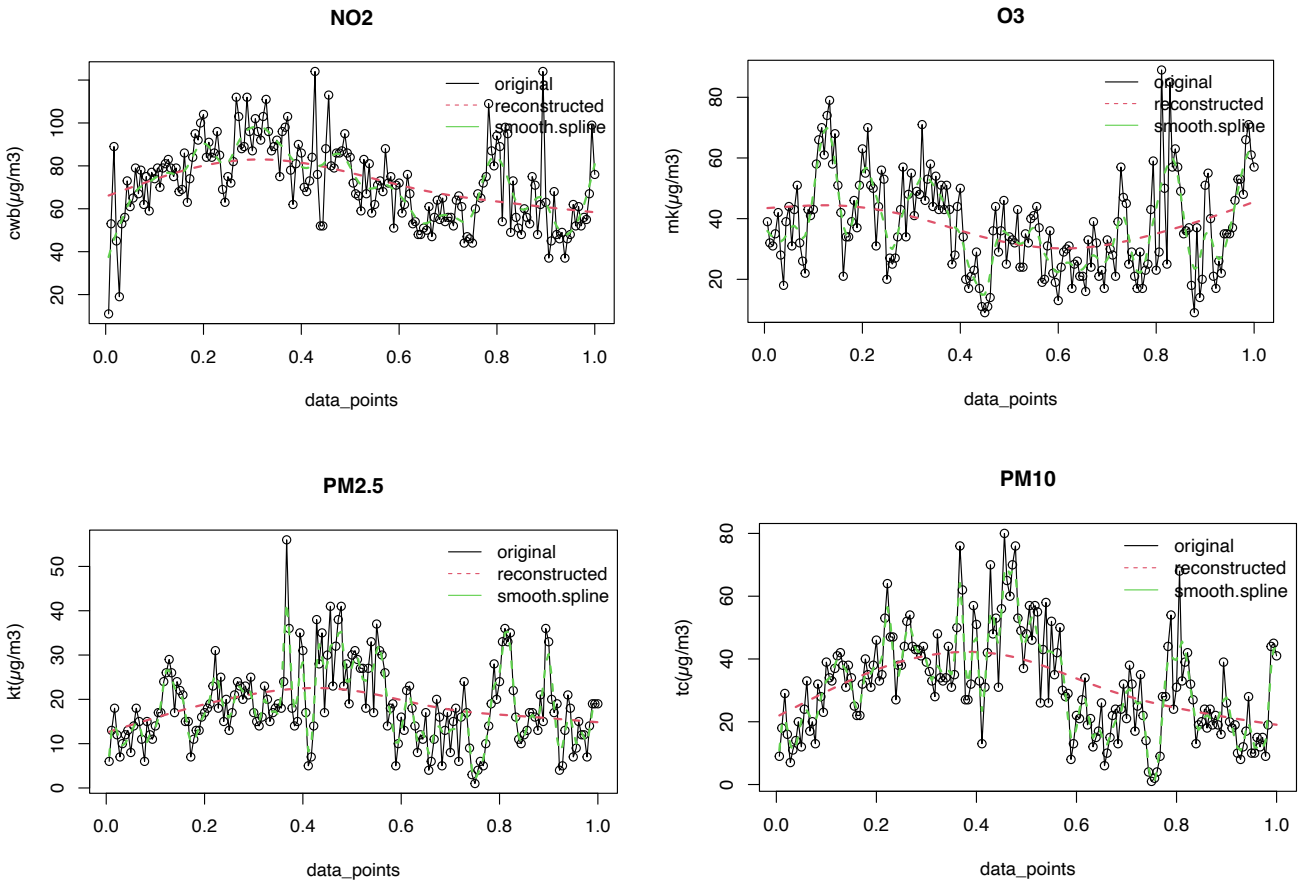

Fig. 2. Daily variation of each pollutants for stations before and during implementing stringent policies. Original raw data is shown in normal, black dash while the data after smoothing and reconstructed are shown in dotted, green dash and red dash respectively.

2.2 Preparation of the data

According to the study of Zhang and Liang (2013), the pointwise F-test and the F_{max} test may be affected by , which could, in certain cases, lead to misleading results. Figure 2 show the difference between the plot of raw data and data after smoothing. Raw data is shown in normal, black dash while the data after smoothing is shown in dotted, green dash. There is 180 data points, the first 90 represents the date from October 9, 2021 to January 6, 2022, before releasing stringent policies, and the next 90 represents the date from January 7, 2022 to April 6, after releasing stringent policies. We scale 180 days to the interval [0, 1]. The region abbreviations are cwb (Causeway Bay) , kt (Kwun Tong), mk (Mong Kok), tc(Tung Chung), tm (Tuen Mun), respectively. It is not difficult to notice that the curve of data after smoothing is less noisy than original curve. The overall concentrations of pollutants PM2.5 and PM10 from figure 3 are obviously decreased. As for the concentration of O3 and NO2, there are some fluctuations and further tests are needed, which will be explained in Section 4.

To better perform the F_{max} test, we first apply smoothing techniques to the functional data suggested by Zhang (2013), such as regression splines, smoothing splines, P-splines, or local polynomial smoothing (see Zhang 2013) (p. 25 -73). For simplicity, we use regression splines to reconstruct the air quality dataset from the observed discrete data. In this case, we use the smooth.spline() function in the stats package (R Core Team 2017) to achieve it. We choose λ as our smoothing parameter, which will be explained more later and in section 3. The optimal λ can be obtained by gcv() function in R (R Core Team et al., 2022). In regression spline smoothing (Zhang 2013), the local neighborhoods are $\tau_0, \tau_1, \dots, \tau_K, \tau_{K+1}$ and $\tau_r, r = 1, 2, \dots, K$ are interior knots, which divide the interval [0, 1] into K subintervals: $[\tau_r, \tau_{r+1}), r = 1, 2, \dots, K$. Zhang (2013) proposed that the regression spline is constructed by kth degree truncated power basis with $\tau_r, r = 1, 2, \dots, K$:

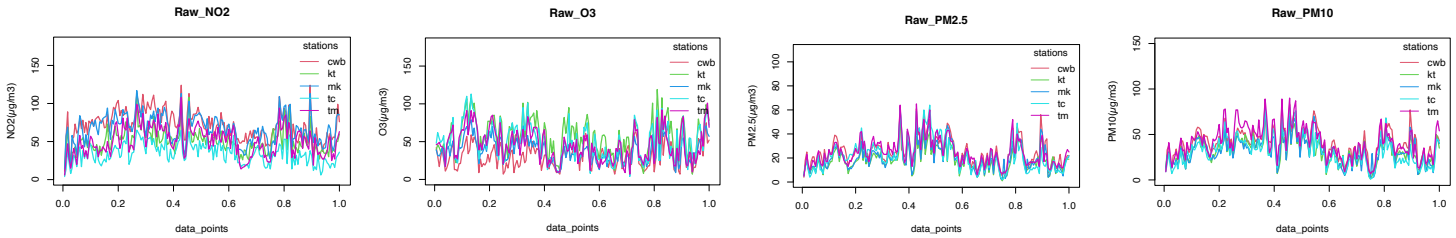

Fig. 3. Noisy raw air pollutants curves of the period of before and during implementing stringent policies. The colours of each solid line are red (Causeway Bay), green (Kwun Tong), blue (Mong Kok), turquoise (Tung Chung), magenta (Tuen Mun), respectively.

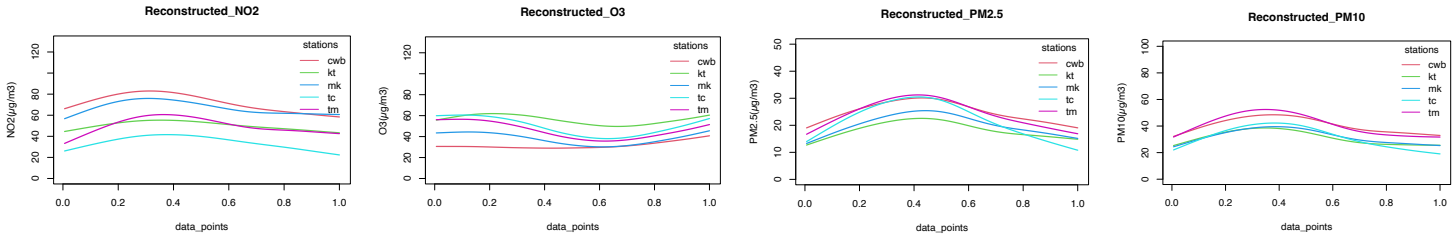

Fig. 4. Reconstructed air pollutants curves of the period of before and during implementing stringent policies. The colours of each solid line are red (Causeway Bay), green (Kwun Tong), blue (Mong Kok), turquoise (Tung Chung), magenta (Tuen Mun), respectively.

$$1, t, \dots, t^k, (t-\tau_1)_+^k\,, \tag{7}$$

where $w_+^k = [w_+]^k$ expresses the power k of the positive part of w. The truncated power basis of degree k we used in this smoothing technique is 3. Figure 2 shows three curves which are constructed by raw data, data after smoothing by regression splines technique, and reconstructed data of regression splines respectively. It is obvious that the red dash line is much smoother than the other two curve for each pollutants. The curve in figure 2 do not show all the stations. We observed that some smoothing splines are smooth but some are not when the data are only smoothing by regression splines technique. Therefore, we decided to choose a smoothing parameter λ to ensure every curve is smooth enough. In order to show their differences more clearly, the figure 3 and 4 display the curves of raw and reconstructed air pollutants concentration for each stations respectively. We noticed that the curves of raw data are nosier that may increase the measurement errors on the F_{max} test (Zhang and Liang 2013). As a result, we use the curve form by reconstructed data for further implementation, which has better performance (Zhang and Liang 2013). Afterwards, we discretize each individual function of the reconstructed functional data so that the implementation of the test can be applied to the discretized data.

3 Methodology and theoretical results

The functional ANOVA approaches for functional data can be tested in univariate and multivariate cases. In this paper, we only propose univariate samples, i.e. the level of each air pollutants in two different periods. This problem is also mentioned above, called one-way functional ANOVA problem. This section introduces the details of reconstruction technique and the derive of the test statistic, which is contributed by the methodology of Zhang (2012) and Zhang et al. (2019), respectively.

3.1. Reconstruct functional data

The functional data set of an underlying stochastic process is defined by Zhang (2013) as

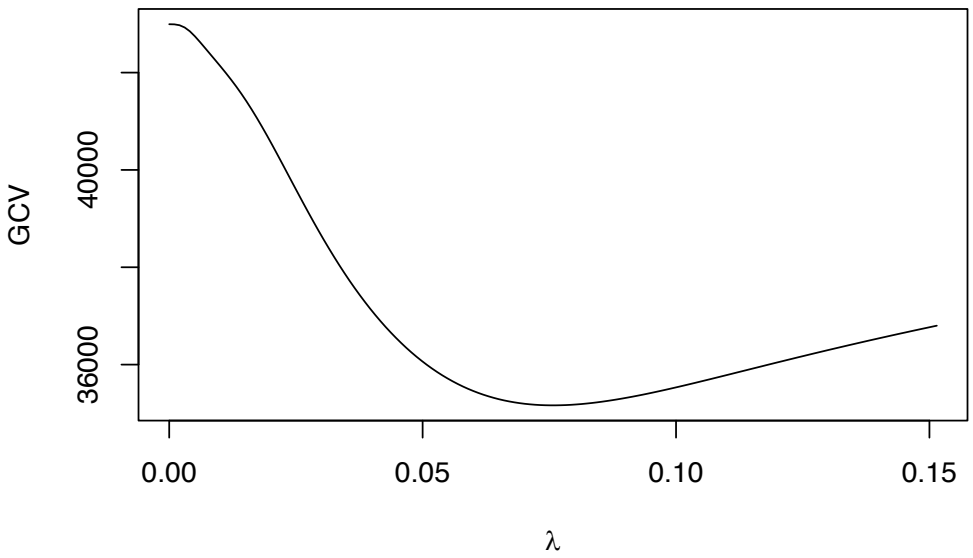

Fig. 5. Plot of GCV score versus smoothing parameter λ for regression smoothing spline approach fitting the curve of NO2 concentration in Causeway Bay.

$$x_i(t) = \mu(t) + \nu_i(t) + \epsilon_i(t), i = 1,2, \dots, k, \tag{8}$$

where $\mu(t)$ has denoted in introduction; $\epsilon_i(t)$ is the *i*th measurement error process; $\nu_i(t)$ denotes the *i*th individual subject-effect from $\mu(t)$. We assume $\nu_i(t)$ and $\epsilon_i(t)$ are independent copies of $\nu(t) \sim SP(0, \gamma)$and $\epsilon(t) \sim SP(0, \gamma_\epsilon)$ respectively, where $SP(\mu, \gamma)$ has defined in introduction as well. By Zhang (2013), we have $f(t) \underset{i.i.d.}{\sim} SP(\mu, \gamma)$ and the individual functions $f_i(t)$ is defined as

$$f_i(t) = \mu(t) + \nu_i(t),), \ i = 1, 2, \dots, k. \tag{9}$$

By Zhang (2013), the model can be rewrited as

$$x_{ij} = f_i(t_{ij}) + \epsilon_i(t_{ij}), \ i = 1, 2, \dots, k; \ j = 1, 2, \dots, n_i. \tag{10}$$

Based on the above model (10), the estimators of $f_i(t)$ denote as

$$\hat{f}_i(t) = a_i(t)^T x_i, i = 1, 2, \dots, k, \tag{11}$$

where $x_i = [x_{i1}(t), x_{i2}(t), \dots, x_{in_i}]^T$. The smoother matrix A_i can be defined as follows

$$A_i = [a_i(t_{i1}), a_i(t_{i2}), \dots, a_i(t_{in_i})]^T, \ i = 1, 2, \dots, k \tag{12}$$

to estimate each individual function. So that

$$\hat{x}_i = [A_i x_i,], i = 1, 2, \dots, k, \tag{13}$$

where $\hat{y}_i$ is the fitted response vector of $f_i(t)$ at the design time points t_{i1}, t_{i2}, t_{in_i}. Given in Zhang (2013), let δ express the smoothing parameter used in the individual function reconstructions $\hat{f}_i(t), i = 1,2, \dots, k$.The smoothing parameter δ can be bandwidth h, number of knots K and λ. It can be selected by the *Generalized cross-validation* rule (GCV) (Zhang 2013; Ruppert et al., 2003; Hastie and Tibshirani 1990) which is a popular method for choosing the smoothing parameter (Craven and Wahba 1979) define as

$$GCV_i(\delta) = \frac{\|y_i - \hat{y}_i\|^2}{(1 - n_i^{-1} \, tr(A_i)^2}, \tag{14}$$

where tr (**A**) represents the trace of the matrix **A**. We can choose an appropriate smoothing parameter for each of the individual functions $f_i(t)$ using the GCV rule mentioned above (14). By minimizing GCV(δ) (14) over a range of potential smoothing parameter candidates δ, the ideal smoothing parameter can be found (Zhang 2013;

Green and Silverman 1994). To be more specific, we choose λ proposed by Zhang (2013) and Ruppert et al. (2003) as the smoothing parameter to reconstruct the data, which can be expressed as

$$GCV(\lambda) = \frac{RSS(\lambda)}{(1-n^{-1}\, tr(A_\lambda)^2}, \tag{15}$$

where $RSS(\lambda) = \sum_{i=1}^{n}\{y_i - \hat{f}(x_i;\lambda)\}^2 = \|y_i - \hat{y}_i\|^2$. Therefore, the optimal λ will be obtain by minimizing the GCV rule (15). Figure 5 has shown a plot of how to obtain optimal smoothing parameter λ by GCV rule. The example is fitting the curve of NO2 concentration in Causeway Bay, it is discovered that λ = 0.0757 is the ideal smoothing parameter that minimizes the GCV rule. We than apply smooth.spline() in software R (R Core Team 2017) by using smoothing parameter λ = 0.0757 and finally get the reconstructed data. Other reconstructed air pollutants in each monitoring station can also deduce the rest from this.

3.2. Statistical model and deriving procedure

Our model and deriving procedure follow the method from Zhang et al. (2019) suggested for determining if two mean functions are equivalent. The deriving process and theorical properties comprise of the specific steps listed below.

STEP 1. According to Zhang et al. (2019), we first denote $GP_k(\mu, \Gamma)$ as a k-dimensional Gaussian process with mean functions vector $\mu(t), t \in \mathcal{T}$ and covariance functions matrix $\Gamma(\mathrm{s},\mathrm{t}) \in \mathcal{T}$. Let $\Gamma(\mathrm{s},\mathrm{t}) = (s,t)\, \boldsymbol{I_k}\ while\ \Gamma(\mathrm{s},\mathrm{t}) = diag[\gamma(s,t),\ \gamma(s,t),\ \ldots, \gamma(s,t)]$. As usual, $GP(\mu,\gamma)$ denotes a Gaussian process with mean function $\mu(t)$ and covariance function $\gamma(s,t)$.

STEP 2. To derive an NPB method to approximate the null distribution of the F_{max} test statistic, we consider the proposition 1 defined by Zhang et al. (2019). By Proposition 1, $\omega_i(t), i = 1, 2, \ldots, k-1 \underset{i.i.d.}{\sim} GP(0, \gamma_\omega)$ which are known except for $\gamma_\omega(s,t)$. The within-function correlation function $\gamma_\omega(s,t)$ can be estimated as

$$\hat{\gamma}_\omega(s,t) = \frac{\hat{\gamma}(s,t)}{\sqrt{\hat{\gamma}(s,s)\hat{\gamma}(t,t)}}, \tag{16}$$

where $\hat{\gamma}(s,t)$ is the pooled sample covariance function $\hat{\gamma}(s,t)$. Zhang and Liang (2013), who found that the F_{max} test might have more power than the GPF test, $\mathcal{L}^2$-norm-based test and F -type test when the within-function correlation $\gamma_\omega(s,t)$ in (16) is large or moderate. It is because the latter three tests tend to average down the information from the functional data. We can estimate the critical value of the F_{max} test statistic by the upper 100-percentile of random variable R_0 for any given significance level, and adopt the following NPB approach for both large and finite sample sizes.

STEP 3. Given in Zhang et al. (2019), first define the estimated subject-effect functions as $\hat{v}_{ij}(t)$ specified in condition A2 given by Zhang et al. (2019). Let

$$v_{ij}^*(t), j = 1,2, \ldots, n_i; i = 1,2, \ldots, k, \tag{17}$$

be k bootstrap samples randomly drawn from the estimated subject-effect functions. Then compute the $SSR_n^*(t)$ and $SSE_n^*(t)$ based on bootstrapped k samples as

$$SSR_n^*(t) = \sum_{i=1}^{k} n_i[\bar{v}_i^*(t) - \bar{v}^*(t)]^2,\ \ SSE_n^*(t) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}[v_{ij}^*(t) - \bar{v}_i^*(t)]^2\,. \tag{18}$$

After that, the F_{max} test statistic can be obtained as

$$F_{max}^* = \sup_{t\,\in\,\mathcal{T}} F_n^*(t),\ where\ F_n^*(t) = \frac{SSR_n^*(t)/(k-1)}{SSE_n^*(t)/(n-k)}, \tag{19}$$

Re-sampling the above bootstrapping process a large number of times. Then we may compute the bootstrap sample's upper 100α-percentile on F_{max}^*. According to Proposition 2, the NPB technique will produce roughly equivalent critical values for the F_{max} test statistic for large samples.

	cwb	kt	mk	tc	tm	Average mean	sd
Change in Mean (sd)							
NO2	-17.23	-7.28	-9.76	-7.24	-8.22	-9.95	15.10
O3	4.22	-6.87	-6.07	-11.62	-10.17	-6.10	9.64
PM2.5	-4.16	-2.86	-2.89	-8.97	-5.80	-4.93	2.80
PM10	-8.23	-9.23	-7.27	-12.48	-13.73	-10.19	5.07
% Change in Mean							
NO2	-21.39%	-13.52%	-13.56%	-19.18%	-15.14%	-16.56%	
O3	14.36%	-11.54%	-14.88%	-21.08%	-20.07%	-10.64%	
PM2.5	-15.32%	-14.37%	-13.27%	-33.67%	-21.12%	-19.55%	
PM10	-18.38%	-25.65%	-20.23%	-32.81%	-28.45%	-25.10%	

Table 1 Change in mean and % change in mean of pollutant concentration levels in each stations of Hong Kong. The standard deviation (sd) of each mean of each stations is shown in the right hand side. Acronyms of monitoring stations: cwb (Causeway Bay), kt (Kwun Tong), mk (Mong Kok), tc (Tung Chung), tm (Tuen Mun).

STEP 4. The relationship given by Zhang et al. (2019) between F_{max} test statistic described in (6) and the GPF test statistic T_n defined in (5) are related is as follows:

$$T_n = \int_{\mathcal{T}} F_n(t)dt \leq (b-a)\, F_{max}\,, \tag{20}$$

where we use the fact that $\mathcal{T}$ = [a, b]. It then follows that

$$P(F_{max} \geq C_\alpha^*) \geq P(T_n \geq (b-a)C_\alpha^*), \tag{21}$$

where C_α^* represents NPB critical value. It is possible for $(b-a)C_\alpha^*$ to be smaller than or equal to the upper 100α-percentile of the GPF test statistic T_n , but this is not guaranteed. Thus, the power of F_{max} test over the GPF test is not assured by (21). Via extensive simulation studies, Zhang et al. (2019) show that the NPB method approximates the F_{max} $test$ statistic reasonably well based on 5000 or 10000 bootstrap result. Also, the F_{max} works well with the number of design time points M = 80 - 150 and the number of bootstrap N = 5000 or 10000. Based on these result, we then apply it for the air pollutants dataset.

4 Results and Discussion

We now provide an example of how to apply the testing techniques that were previously described to determine whether the concentration of each pollutant has changed during the stringent policies implementation. We first hypothesize that there is no change on the mean concentration level of each pollutant before and after releasing the stringent policies. The hypothesis is shown as

$$\begin{cases} H_0: \mu_1(t) = \mu_2(t) \forall t \in [a,b] \\ H_1: \mu_1(t) \neq \mu_2(t) for some t \end{cases}, \tag{22}$$

where $\mu_1(t)$ and $\mu_2(\mathrm{t})$ denote the mean curves of regions of each pollutant under the period of before and after releasing the stringent measures. Since the 90 design time points are common for all the pollutant curves, so we use [1, 90] as our interval. The acquired environmental datasets were divided into two time periods of equal duration (90 days): i) before releasing stringent policies (February 1, 2020 – March 10, 2020) and ii) after releasing

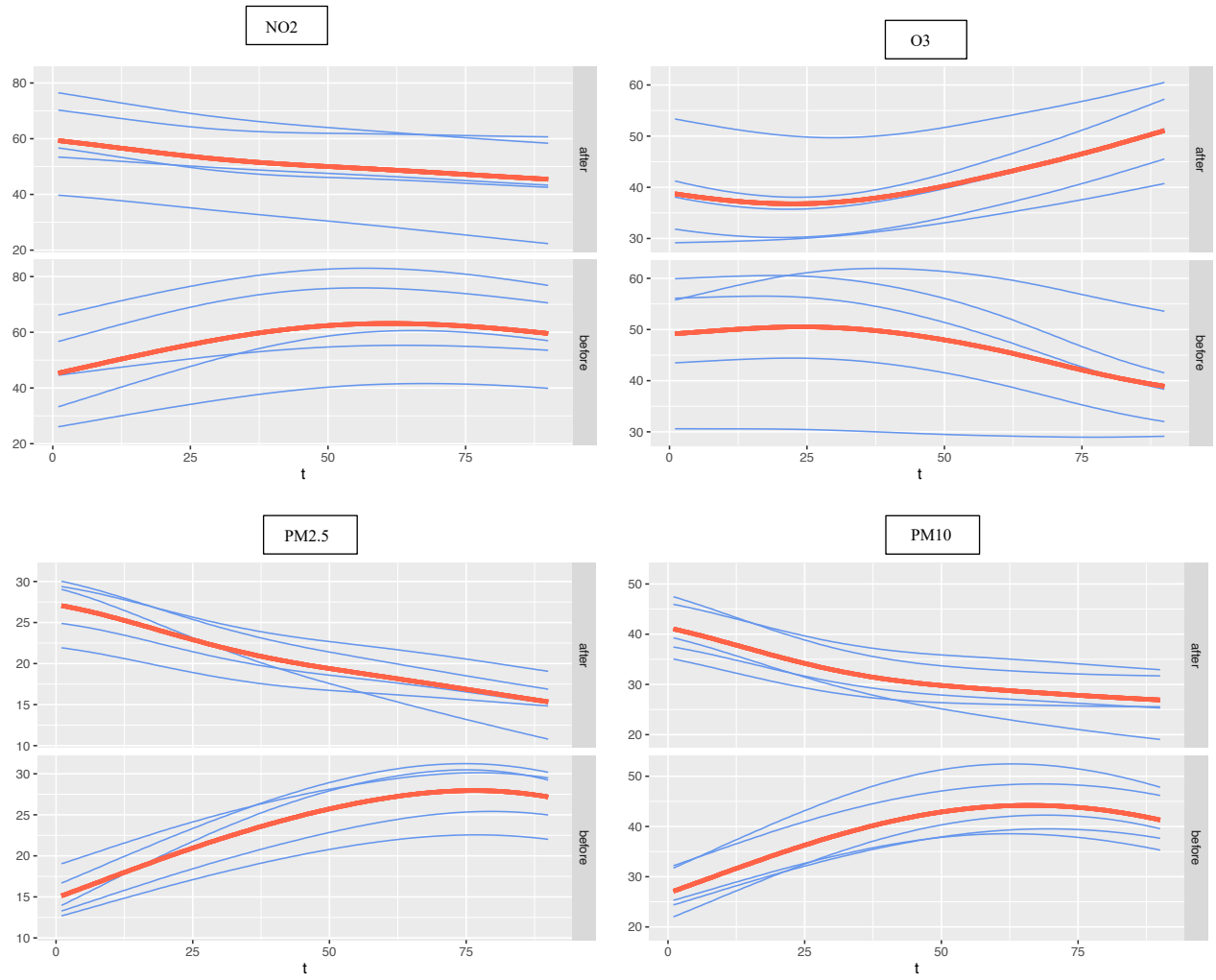

Fig.6. the concentrations level before (lower panel) and after (upper panel) starting to tighten Covid-19 prevention measures. The Blue curves represent the concentration level of pollutants of five selected regions, and the red curve stand for the mean concentration level of the five selected stations. The unit of y axis is μg/m3.

stringent policies in order to highlight the influence of restriction measures resulting from the COVID-19 on the air quality (January 7, 2022, to April 6, 2022).

Górecki and Smaga (2019) have proposed a R software package for analyzing one-way functional ANOVA problem. Based on the fdANOVA package, we have conducted a F_{max} test to test the equality of mean functions of individual groups. Based on the result in Section 3, we chose the design time points M = 90 and bootstrap number N = 10000 when running the code. Figure 6 shows the concentration levels before and after starting tighten Covid-19 prevention measures. The Blue curves represent the concentration level of pollutants of five selected regions, and the red curves stand for the mean concentration level of the five selected stations. We use number of experiment days instead of using date in the x-axis. The unit of y-axis is the concentration level μg/m3. In order to see the difference between the mean before and after them better, picture 6 gives a clearer comparison.

Before looking at the final result of F_{max} test, we first illustrate the mean differences of raw air pollutants NO2, PM2.5, PM10 and O3 concentrations before and during the implementation of Covid-19 policies in table 1. During the policy implementation period, the average change in percentage of the mean of PM10 among these five stations decreased the most which is 25.10%. A statistically significant declines of PM2.5, NO2 and O3 are also observed.

We now apply the F_{max} test to the reconstructed air pollutants data and use 8e-04 to denote 0.0008 for simplicity. The p-values and critical values of F_{max} test were determined by using 10000 bootstrap replicates. The

NO2
O3

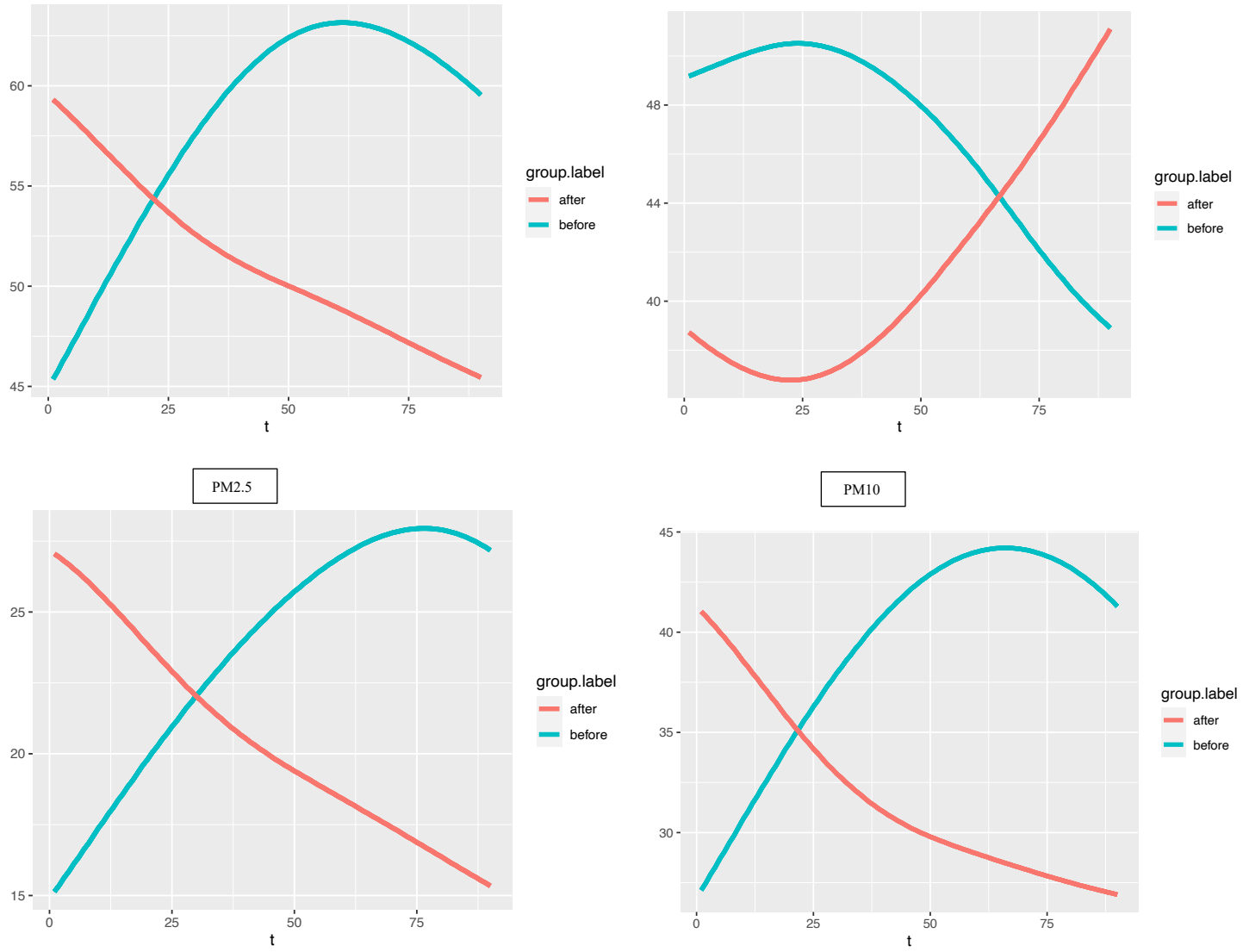

Fig.7. the mean concentrations level of 5 selected stations before and after starting to tighten Covid-19 prevention measures. The Blue curve represents the concentration levels before releasing stringent policies (February 1, 2020 – March 10, 2020, and the red curve stands for concentration levels after releasing stringent policies (January 7, 2022, to April 6, 2022)). The unit of y-axis is μg/m3.

resulting F_{max} critical values and p-values are displayed in Table 2. The main conclusion is that the mean concentration levels of PM2.5 and PM10 before releasing stringent policies are differ to those after releasing stringent measures. The results of reconstructed data may surprise you as some do not consistent with the results in table 1 and the results of raw data, which will be further discuss. For suspended particles PM2.5 and PM10, the decrease trends are obvious and shown in figure 6 and 7. The p-values of F_{max} test are also smaller than significance level $\alpha = 0.05$. We have enough evidence to reject the null hypothesis for PM2.5 and PM10, respectively. We state that there are significant differences in the mean curves of these two pollutant in before and during policy implementation periods. A possible explanation is that the reduced emissions from diesel vehicles and power plants in Hong Kong. For O3, we noticed that only the mean function of O3 had a tendency to decline and then rise during the execution of the Covid-19 regulation. We also observed that whether it is the raw O3 data in figure 3 or the reconstructed O3 data in figure 4, the downward trend of O3 is not very obvious, but only maintains an intermediate level. Therefore, it is not very surprising that the p-value of O3 is greater than 0.05, which means that we do not have enough evidence to reject the null hypothesis. This outcome, however, is opposite with the mean change in table 1 and the test results for the raw data in table 2. O3 is a complex air

	GPF test		F_{max} test	
[a,b] = [1, 90]	T_n	p-value	F_{max}	p-value
Raw data (NO2)	5.280204	0.009889093	0.005	0.005
Raw data (PM2.5)	39.74985	0	632.2577	3e-04
Raw data (O3)	9.793031	1.198868e-06	45.84086	0.0139
Raw data (PM10)	35.52352	0	8e-04	8e-04
Reconstructed data (NO2)	1.393391	0.3465451	2.468758	0.1971
Reconstructed data (PM2.5)	15.82685	2.458926e-06	36.55802	0.0044
Reconstructed data (O3)	2.257669	0.1852082	4.721588	0.1055
Reconstructed data (PM10)	12.11276	0.0001280236	20.62332	0.0073

Table 2 Critical values and p-values of the GPF and F_{max} tests for raw and reconstructed concentrations of each air pollutant.

pollution problem, as well as a regional one. It is not directly discharged from the pollution source, but is formed by a series of photochemical reactions of the precursor nitrogen oxides (NOx) and volatile organic compounds (VOCs) in the sun, of which nitrogen oxides and volatile organic compounds mainly come from factories and automobiles exhaust gas emissions. Thus, the possible reason for the change in O3 concentration is that when people go out less during the Covid-19 policy implementation period, exhaust emissions are also reduced, and so as O3 is reduced, which can be also supported by the mean O3 concentration in table 1. At the same time, local vehicles reduce NOx emissions, resulting in less NO reacting with O3, thus reducing O3 depletion and allowing more O3 to exist in the air, which explains the rising trend of the O3 mean function. The F_{max} test result of last air pollutant NO2 may bring surprise as its p-value is bigger than 0.05, which means we cannot state that there are significant differences in the mean curves of this pollutant in before and during policy implementation periods. We noticed that each reconstructed data curve in picture 4 shows a downward trend. At the same time, the mean concentration of NO2 in each region shown in table 1 also has a significant decrease, and the average mean concentration of NO2 in each region has a 16.56% decrease. However, these results are inconsistent with the final F_{max} test result, so we cannot conclude that the concentration of NO2 has decreased after the regulation execution. Therefore, we apply F_{max} test again by choosing only two stations with little difference between the two curves, the p-values obtained by F_{max} test are much smaller 0.05 as expected. The result is same when we choose another two stations with little difference to do the test. The possible reason for the difference of results is that the difference between each curve of each region is relatively large, that is, the standard deviation is relatively large, and the data in table 1 proves this. Table 1 has shown the standard deviation of NO2 mean concentration in each region. The results show that the standard deviation of NO2 data is relatively large as a whole, which means that there is a large difference between most values and their average value. Therefore, the mean function of NO2 may not be accurate enough to reflect the overall NO2 concentration level. It may explain why we get this unexpected result.

In addition, we have applied the F_{max} test to the raw data to find the difference. We noted that all p-values from raw air pollutants data are smaller than 0.05, which is not inconsistent with those from reconstructed data. In general, their p-values are smaller than the reconstructed data. It cannot show an improvement of the power of F_{max} test, which is suggested by Zhang ang Liang (2013)'s application on orthosis data.

Zhang et. al (2019) demonstrated that, in contrast to the GPF test, the F_{max} test performed on the ischemic cardiac ECG dataset is significant, which further verified that F_{max} test has higher power than GPF test. We have also applied GPF test to the reconstructed air pollutants data, see the results shown in table 2. In contrast to result of Zhang et al. (2019), both GPF test and F_{max} test revealed the significant results in PM2.5 and PM10. In general case, the p-values of GPF test are similar to or even smaller than F_{max} test (see Table 2), especially the p-values of

PM2.5 and PM10. Therefore, we cannot conclude which is preferred or has higher power in this case, maybe further applications on the other aspects are need.

5 Conclusions

Regarding the outbreak of covid-19 from 2020 to 2022, many articles have studied its impact on society, economy and environment. It was also mentioned at the beginning of this article that air quality has been used as the research and test object of many articles. The air quality of some countries has not improved but worsened, and the air quality of some countries has changed greatly. In this paper, the Fmax test for the one-way ANOVA problem is proposed as a new approach to more efficiently understand the effects of stringent policies on four key air pollutants measured at five monitoring stations. We found that PM2.5 and PM10 levels were significantly reduced during the implementation of the strict containment policy. These results are consistent with those of other published studies on this topic, such as Berman and Ebisu (2020), Abdullah et al. (2020), Naeger and Murphy (2020), and Otmani et al. (2020). Determining whether these differences are statistically significant is also crucial. In this regard, it has been demonstrated that functional analysis of variance is useful for tracking the development of air quality before and during the lockdown period as well as for evaluating the equality of the mean function across groups, individually according to the location of the measurement site. However, this experiment has limitations. We ignore other factors that contribute to the decline in pollutants, such as weather and government environmental policies. Also, the sample size of stations may be a factor misleading the final result, such as the result of NO2. The NO2 concentrations in some stations are higher due to regional problem so that large standard deviation of each stations are large. Therefore, selecting appropriate sample size may be needed. It is also recommended to further investigate multivariate analysis of variance (MANOVA) by evaluating multiple dependent variables simultaneously. Overall air quality has improved as Hong Kong's COVID-19 restrictions have reduced the number of people going out, limited factory operations and reduced vehicle use. But as measures are eased, pollutants may return to previous levels. The purpose of this article is not to suggest that the government tighten its measures, but to provide the government with insights. It is hoped that the findings of this study will be of interest to environmental agencies and governments involved in formulating policies to achieve air quality improvements and establish mechanisms to reduce pollution emissions.

Reference

[1] S. Abdullah, A. A.,Mansor, N. N. L. M. Napi, W. N. W. Mansor, A. N. Ahmed, M. Ismail and Z. T. A. Ramly, Air quality status during 2020 Malaysia Movement Control Order (MCO) due to 2019 novel coronavirus (2019-nCoV) pandemic, *Sci Total Environ,* 2020.

[2] C. Acal, A.M. Aguilera, A. Sarra, A. Evangelista, T. Di Battista and S. Palermi, Functional ANOVA approaches for detecting changes in air pollution during the COVID-19 pandemic, *Stoch Environ Res Risk Assess*, 2022, 36(4), 1083-1101.

[3] A.M. Alonoso and D. J. R. Casado, Supervised classification for functional data: A weighted distance approach. *Comput Stat Data Anal*, **56** (2012), pp. 2334–2346.

[4] R. Bao and A. Zhang, Does lockdown reduce air pollution? Evidence from 44 cities in northern China, *Sci Total Environ*, **731** (2020), 139052

[5] J. D. Berman and K. Ebisu, Changes in U.S. air pollution during the COVID-19 pandemic, *Sci. Total Environ.* **739** (2020), 139864

[6] J.R. Berrendero, A. Justel and M. Svarc, Principal components for multivariate functional data, *Comput Stat Data Anal,* **55**(2011), pp. 2619–2634

[7] *CE addresses public on stabilisation of fifth wave of epidemic early (with photo/video)*, 2022, https://www.info.gov.hk/gia/general/202202/28/P2022022800803.htm?fontSize=1

[8] P. Craven and G. Wahba, Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerische Mathematik* , **31** (1979), pp. 377–403.

[9] A. Cuevas, M. Febrero and R. Fraiman, Linear functional regression: the case of fixed design and functional response, *Can J Stat,* **30** (2002), pp. 285–300.

[10] G. Dantas, B. Siciliano, B. B. França, C. M. da Silva and G. Arbilla, The impact of COVID-19 partial lockdown on the air quality of the city of Rio de Janeiro, Brazil, *Sci Total Environ,* **729** (2020), 139085.

[11] J. Fan and S. -K. Lin, Test of significance when data are curves, *Amer Statist Assoc*, **93** (1998), pp. 1007–1021.

[12] J. Faraway, Regression Analysis for a Functional Response, *Technometrics*, **39** (1997), pp. 254–261.

[13] F. Ferraty, *Recent Advances in Functional Data Analysis and Related Topics*. Springer-Verlag, Berlin Heidelberg, 2011.

[14] T. Górecki and Ł. Smaga, A comparison of tests for the one-way ANOVA problem for functional data, *Comput Stat*, **30** (2015), pp. 987–1010.

[15] T. Górecki and Ł. Smaga, fdANOVA: an R software package for analysis of variance for univariate and multivariate functional data, *Comput Stat*, **34**(2019), pp. 571-597.

[16] P. J. Green and B. W. Silverman, *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall, 1994.

[17] T. Hale, N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar and H. Tatlow, A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker), *Nat Hum Behav,* **5** (2021), pp. 529– 538.

[18] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*, Chapman and Hall, 1990.

[19] J. Martı´nez Torres, J. Pastor Pere´z, J. Sancho Val, A. McNabola, M. Martı´nez Comesan˜a and J. Gallagher, A functional dataanalysis approach for the detection of air pollution episodes andoutliers: a case study in Dublin, *Ireland.Mathematics*, **8** (2020), 225

[20] F.-J. Mesas-Carrascosa, F. Pérez Porras, P. Triviño-Tarradas, A. García-Ferrer and J. E. Meroño-Larriva, Effect of lockdown measures on atmospheric nitrogen dioxide during SARS-CoV-2 in Spain, *Remote Sens,* **12** (2020), 2210

[21] A. R. Naeger and K. Murphy, Impact of COVID-19 Containment Measures on Air Pollution in California, *Aerosol Air Qual. Res,* **20** (2020), pp. 2025– 2034.

[22] A. Otmani, A. Benchrif, M. Tahri, M. Bounakhla, M. El Bouch and M. H. Krombi, Impact of COVID-19 lockdown on PM10, SO2, and NO2, concentrations in Sale City (Morocco), *Sci Total Environ*, **735** (2020), 139541.

[23] R Core team, M. Ribatet and R. Singleton, SpatialExtremes: Modelling Spatial Extremes, URL http://spatialextremes.r-forge.r-project.org/index.php, 2022.

[24] J. O. Ramsay, B. W. Silverman, *Applied functional data analysis: methods and case studies*, Springer-Verlag, New York, 2002.

[25] J.O. Ramsay, B.W. Silverman, *Functional Data Analysis, seconded*, Springer Series in Statistics, Springer, New York, 2005.

[26] R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/, 2017.

[27] D. Ruppert, M. P. Wand and R. J. Carrol, *Semiparametric Regression Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, 2003.

[28] S. Sharma, M. Zhang, A. J. Gao, H. Zhang and S. H. Kota, Effect of restricted emissions during COVID-19 on air quality in India, *Sci Total Environ*, **728** (2020), 138878.

[29] P. Wang, K. Chen, S. Zhu, P. Wang and H. Zhang, Severe air pollution events not avoided by reduced anthropogenic activities during COVID-19 outbreak, *Resour Conserv Recy,* **158** (2002), 104814.

[30] S. Zangari, D. T. Hill, A. T. Charette and J. E. Mirowsky, Air quality changes in New York City during the COVID-19 pandemic, *Sci. Total Environ,* **742** (2020), 140496.

[31] J.-T. Zhang and J. W. Chen, Statistical inferences for functional data, *Ann Statist*, **35** (2007), pp. 1052–1079.

[32] J.-T. Zhang, X. Liang, 2013, One-way ANOVA for functional data via globalizing the pointwise F-test, *Scand J Stat*, 41(2013), pp. 51–71.

[33] J.-T. Zhang, *Analysis of Variances for Functional Data*, Chapman and Hall, London, 2013.

[34] J.T. Zhang, Statistical inferences for linear models with functional responses, Statist Sinica, **21** (2011), pp. 1431–1451.

[35] J.T. Zhang, M.Y. Cheng, H.T. Wu and B. Zhou, A new test for functional one-way ANOVA with applications to ischemic heart screening. *Comput Stat Data Anal,* **132** (2019), pp. 3–17.