



# Functional ANOVA approaches for detecting changes in air pollution during the COVID-19 pandemic

Christian Acal<sup>1</sup> · Ana M. Aguilera<sup>1</sup> · Annalina Sarra<sup>2</sup> · Adelia Evangelista<sup>2</sup> · Tonio Di Battista<sup>2</sup> · Sergio Palermi<sup>3</sup>

Accepted: 24 July 2021  
© The Author(s) 2021

## Abstract

Faced with novel coronavirus outbreak, the most hard-hit countries adopted a lockdown strategy to contrast the spread of virus. Many studies have already documented that the COVID-19 control actions have resulted in improved air quality locally and around the world. Following these lines of research, we focus on air quality changes in the urban territory of Chieti-Pescara (Central Italy), identified as an area of criticality in terms of air pollution. Concentrations of NO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> and benzene are used to evaluate air pollution changes in this Region. Data were measured by several monitoring stations over two specific periods: from 1st February to 10th March 2020 (before lockdown period) and from 11st March 2020 to 18th April 2020 (during lockdown period). The impact of lockdown on air quality is assessed through functional data analysis. Our work makes an important contribution to the analysis of variance for functional data (FANOVA). Specifically, a novel approach based on multivariate functional principal component analysis is introduced to tackle the multivariate FANOVA problem for independent measures, which is reduced to test multivariate homogeneity on the vectors of the most explicative principal components scores. Results of the present study suggest that the level of each pollutant changed during the confinement. Additionally, the differences in the mean functions of all pollutants according to the location and type of monitoring stations (background vs traffic), are ascribable to the PM<sub>10</sub> and benzene concentrations for pre-lockdown and during-lockdown tenure, respectively. FANOVA has proven to be beneficial to monitoring the evolution of air quality in both periods of time. This can help environmental protection agencies in drawing a more holistic picture of air quality status in the area of interest.

## Abbreviations

ANOVA Analysis of variance  
AQI Air quality index

CO Carbon monoxide  
GDP Gross domestic product  
FANOVA Functional analysis of variance

---

✉ Annalina Sarra  
asarra@unich.it

Christian Acal  
chracal@ugr.es

Ana M. Aguilera  
aaguiler@ugr.es

Adelia Evangelista  
adelia.evangelista@unich.it

Tonio Di Battista  
tonio.dibattista@unich.it

Sergio Palermi  
s.palermi@artaabruzzo.it

<sup>2</sup> Department of Philosophical, Pedagogical and Economic-Quantitative Sciences, University G. d'Annunzio, V.le Pindaro, 42, 65127 Pescara, Italy

<sup>3</sup> Agency of Environmental Protection of Abruzzo (ARTA), V.le G. Marconi, 51, 65127 Pescara, Italy

<sup>1</sup> Department of Statistics and O.R. and IMAG, University of Granada, Granada, Spain

FDA	Functional data analysis
FPCA	Functional principal component analysis
MANOVA	Multivariate analysis of variance
NMHC	Non methane hydrocarbons
NH <sub>3</sub>	Ammonia
NO <sub>2</sub>	Nitrogen dioxide
NO <sub>x</sub>	Nitrogen oxides
O <sub>3</sub>	Ground level ozone
PM <sub>2.5</sub>	Particulate matter including only particles less than 2.5 µm in diameter
PM <sub>10</sub>	Particulate matter including only particles less than 10 µm in diameter
PC	Principal component
SO <sub>2</sub>	Sulphur dioxide
UB	Urban background
UT	Urban traffic type
VOC <sub>s</sub>	Volatile organic compounds
WHO	World Health Organization

## 1 Introduction

After the discovery of the first case in Wuhan (China) in December 2019, the current outbreak of COVID-19, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has dramatically affected all the countries (Wang et al. 2020). On January 30 2020, the World Health Organization (WHO) declared worldwide public health emergency and in March 11 2020, due to widespread global infection, the WHO authorities categorised the new Coronavirus as pandemic (WHO 2020). To contain the virus and save lives, governments around the world have been taking a range of actions and measures, such as social and travel restrictions. More specifically, coronavirus pandemic has forced nations under partial or complete lockdowns, resulting in prohibition of unnecessary commercial activities in people's daily lives; prohibition of any types of gathering by residents; restrictions on private (vehicle) and public transportation. Different studies have already documented the effects of COVID-19 lockdown measures on many aspects of human activities, such as transportation (Mogaji 2020), renewable and sustainable energy (Hosseini 2020), health risk assessment (Gautam and Trivedi 2020; Ambade et al. 2021; Zoran et al. 2020; Gupta et al. 2021), tourism (Sigala 2020), commodity markets (Rajput et al. 2021). Certainly, COVID-19 has severe negative impact on the world activities as well as on local economy. In the major economies across the globe, lockdown will directly affect the Gross Domestic Product

(GDP) of each country. On the contrary, efforts to restrict transmission of the SARS-CoV-2 have had outstanding effects on the ecosystems which are being greatly recovered. In many cities where lockdown measures have been implemented, the decline in economic activities, the non-functioning of industries, the drop in road transport, have contributed to mitigate air pollution. Bherwani et al. (2020) monetarily quantified the overall benefit due to pollution reduction over the potential economic loss sustained by local governments, as an outcome of lockdown.

Several researchers around the world reported that there is a considerable reduction of air pollution level across geographies. For instance, the findings of many studies showed a substantial enhancement in the air quality in the lockdown period globally (Lal et al. 2020; Dutheil et al. 2020; Gope et al. 2021; Venter et al. 2020; Gautam 2020b).

A visible improvement in air quality parameters was recorded in different areas of China and India (see, among others, Wang et al. 2020; Bao and Zhang 2020; Li et al. 2020; Gautam and Trivedi 2020; Mahato et al. 2020; Sharma et al. 2020; Agarwal et al. 2020; Gautam et al. 2021). The study of Kanniah et al. (2020) highlighted that large emission reduction in transportation and anthropogenic activities, resulted in a significant reduction in air pollution levels in the urban regions of Malaysia as well as in a sizeable reduction aerosol optical thickness concentration during March-April 2020 in comparison with the same period in 2019 and 2018.

Kerimray et al. (2020) analyzed the effect of the lockdown from March 19 to April 14, 2020, on the concentrations of air pollutants in Almaty, Kazakhstan, finding out a substantial difference between concentrations of air pollutant recorded before and during lockdown.

Otmani et al. (2020) gave evidence that the government decisions in response to COVID-19 had an important impact on the air pollution in Salé city (Morocco). Findings from that study showed that the difference between the concentrations recorded before and during the lockdown period were 75, 49 and 96% for PM<sub>10</sub>, SO<sub>2</sub> and NO<sub>2</sub>, respectively.

Berman and Ebisu (2020) investigated the changes in levels of air pollutants across USA during COVID-19 pandemic. The authors reported a significant reduction on NO<sub>2</sub> (up to -25.5%) and an overall decline in PM<sub>2.5</sub>, compared with pre-lockdown phase. Zabran-Monserrate and Ruano (2020) studied the effects of quarantine policies on air pollutants concentrations in Quito, Ecuador. Using parametric methods, they detected a significant reduction of NO<sub>2</sub> and PM<sub>2.5</sub> since the introduction of lockdown measures. However, there was a noticeable growth in ozone levels. There were reports that documented a decline in NO<sub>2</sub>, NO<sub>x</sub> and an increase in surface O<sub>3</sub> at São Paulo in

Brazil during the quarantine period (see, for instance, Nakada and Urban 2020; Dantas et al. 2020). The similar trends of reducing air pollution and increasing air quality due to introduction of lockdown were observed in large parts of Europe, such as France, Germany, Spain, and Italy (Sicard et al. 2020; Collivignarelli et al. 2020; Zambrano-Monserrate et al. 2020; Tobías et al. 2020). An overview of selected studies on air pollution and COVID-19 is provided in the Appendix. In this study, we focus on investigating the possible effects of the lockdown due to the COVID-19 pandemic on air quality in the Pescara-Chieti urban area, Abruzzo (Italy), identified as an area of criticality in terms of air pollution. Data of monitoring stations of the regional air quality network managed by the Regional Agency for the Environmental Protection (ARTA) of Abruzzo have been collected and examined. We compared data from 1st February to 10th March 2020, before the beginning of the main limitations on personal mobility, with data from 11th of March to 18th of April, during the adoption of lockdown restrictions. Measured concentrations of  $\text{NO}_2$ ,  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$  and benzene were used to evaluate air pollution changes. Commonly, strategies used in monitoring air quality refer to descriptive statistics, box-plots, autocorrelation analysis and spatio-temporal models. Unfortunately, in the monitoring of environmental pollutants, the temporal observations of the different pollutants for the different stations have not always been referred to the same instants of time. As a result, the implementation of classical statistical procedures might be problematic. Besides, for interpretative purposes, it is convenient to rely on statistical methods able to capture the speed and acceleration of pollutants variation over time. For these reasons, in our research, to overcome the weakness of classical statistical procedures and to effectively detect to what extent extreme changes in human behaviour after the quarantine policies adopted by the Italian Government have affected air quality, we followed an approach based on functional data analysis (FDA). During the two last decades, it has emerged an important literature in this methodological framework. A comprehensive introduction to the foundations and applications of FDA can be found in Ramsay and Silverman (2002, 2005), whereas nonparametric functional methods are summarized in a monograph by Ferraty and Vieu (2006). It is well known that FDA extends the classical multivariate techniques to data whose observations are functions, usually curves, with the advantage of reducing a large number of discrete observations highly correlated for each curve to a functional form that conserves all relevant information.

Recently, the use of FDA methods for environmental data has received attention. For instance, Escabias et al. (2005), proposed a functional principal component logistic regression to estimate the risk of drought in terms of time evolution in

temperatures in Canada. Gao and Niemeier (2008) use functional methods to model the dynamics of diurnal ozone and nitrogen oxides cycles, their interrelationships, and the multilevel spatio-temporal variations of ozone and nitrogen oxides measurements from Southern California. A functional model for forecasting the time evolution of a binary response from discrete time observations of a continuous time series, is introduced by Aguilera et al. (2008) to predict the risk of drought in a future period of time from monthly observations of *El Niño* phenomenon. Martínez et al. (2014) and Martínez Torres et al. (2020) implement a model based on functional analysis to detect outliers in air quality samples, with the overall aim to achieve a better solution for the air quality control. Likewise, Sancho et al. (2014) expand the concept of functional outliers to the set of control-chart techniques. Additional applications that prove the advantages of FDA in environmental research can be found in Ocaña-Peinado et al. (2008); Valderrama et al. (2010); Park et al. (2013); Shaadan et al. (2013); Escabias et al. (2013); Hörmann et al. (2015); Aguilera-Morillo et al. (2017); Gautam and Trivedi (2020).

Following these research streams, also in this study, rather than simply considering the data as vectors to apply classical multivariate analysis methods, which may lead to a loss of useful information, we explicitly exploit the functional form of environmental data. It is worth noting that from a physical point of view the processes involved in the production of pollutants, such as  $\text{NO}_2$  and particulate matter ( $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$ ) are of continuous time type and thus turn out to be appropriate for the analysis of functional data. As a result, the use of models that account for the continuity of the whole trajectories along time seem to be more natural. The FDA paradigm makes it possible to work with the entire time spectrum of pollutants time series. In doing so, the FDA approach might bring additional information to be recovered from the data than in the vectorial approach, by looking at the smoothness of underlying functions and its derivatives. Furthermore, in contrast to most other methods commonly used to model trends in time series data, the functional techniques make non-parametric assumptions and there is no concern about correlation due to repeated measures. Our goal in this paper is to ascertain whether the level of each pollutant has changed during the lockdown period. In other terms, we want to test the equality of mean functions related to each pollutant in two different periods of time: before and during lockdown days. The theoretical framework involves the use of FDA tools for repeated measures, and in particular, the analysis of variance. In the literature there are not many works related to this matter for the field of FDA. In this work, the statistics proposed by Martínez-Cambor and Corral (2011) and Smaga (2020) to test the equality of two mean functions are extended by assuming a basis

expansion of the sample curves. On the other hand, in order to check the differences between the temporal evolution of all pollutants in terms of the location of measuring stations, a novel approach for multivariate FANOVA for independent measures is introduced. This is based on multivariate functional principal components analysis of the sample curves of all pollutants and the problem is reduced to testing multivariate homogeneity or MANOVA (gaussian data) on the vectors of the most explicative principal components scores. This new approach solves the problems of high dimension and multicollinearity that affect the basis expansion approach developed in Gorecki and Smaga (2017) when the number of basis coefficients for each functional variable is large. The relevance of this research work lies in its contribution to extending the parametric and nonparametric functional homogeneity testing procedures proposed by Aguilera et al. (2021) for univariate functional data to the multivariate case.

In addition to this methodological development, this study provides empirical suggestions for local environmental authorities. The comparison of air pollutants allows to study whether the level of each pollutant decreased during the quarantine days. The suitable analysis of a different behaviour of pollutants at measuring sites provides ground to decide on whether to keep recording measurement at different locations or redesign the local surveillance network, with the overall aim to reduce air pollution.

The paper is organized as follows. Section 2 is dedicated to illustrate the studied area, the monitoring stations where air quality data have been collected and some explorative analysis of pollutants used for this research. Section 3 introduces the theoretical framework. In Section 4, our method is applied for checking differences between air quality data collected at the monitoring stations placed in the urban area of Chieti-Pescara. Finally, Section 5 concludes the paper.

## 2 Air quality data and studied period

This section gives details about the studied area with respect to its geographic characteristics and the monitoring stations used to collect air quality data. In addition, the dataset employed to derive insights into research problem has been analyzed with descriptive statistics and visualization tools. For high quality graphics, we used the “*openair*” R package (Carslaw and Ropkins 2012).

### 2.1 Description of studied region

In this work, we closely examine the air quality of the metropolitan area of Chieti-Pescara, situated in the Abruzzo region, along the Adriatic coast of central Italy. The Chieti-Pescara metropolitan area (Fig. 1) is a territory,

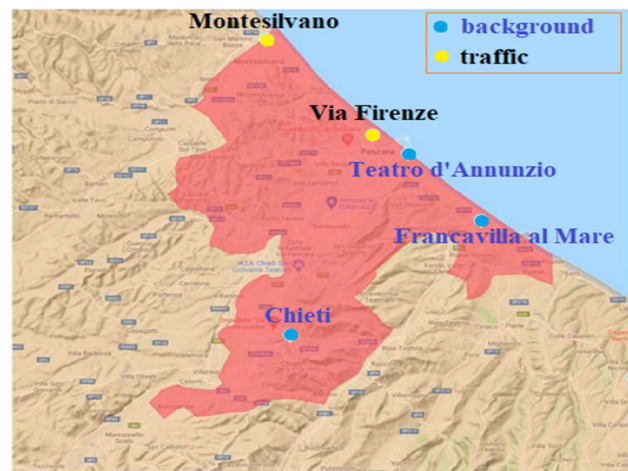


Fig. 1 Abruzzo and Chieti-Pescara metropolitan area

identified according to a functional criterion, formed by six municipalities, Pescara, Montesilvano, Chieti, Francavilla al Mare, San Giovanni Teatino, Spoltore, covering a total area of 159.33 km<sup>2</sup>, and accounting for around 281,101 inhabitants at 31/12/2019.

The configuration of Chieti-Pescara urban area is limited by the sea, in the North-East, and by hilly reliefs in the South-West. The central city is formed by the two provincial capitals: Chieti, not in a central position for the municipalities of the province, and the city of Pescara, which are extremely close to each other (approximately 12 km). Pescara city, located on the centre of a metropolitan area (on the coast), is the administrative and commercial heart of Abruzzo and in a few decades, it has become the most populated city of the region, with 120,000 inhabitants. It developed on a flat territory, with a surface of 33.62 km<sup>2</sup>, whose urban area develops around the terminal stretch of the homonymous river and a restricted coastal area.

The Chieti-Pescara conurbation is characterized by a system of infrastructures, which is one of the strongholds of the Abruzzo: significant and industrial sites are located around this pole. However, the progressive growth of the industrial activity, the increased road travel, the urban expansion, make the metropolitan area the locus of growing environmental concerns, for the rising levels of energy and resource consumption, greenhouse gas emissions and air quality pollution. For these reasons, the conurbation of Chieti-Pescara has been identified as an area deserving mitigation measures to reduce air pollution by the “Plan for the Protection of Air Quality”, drafted by the Abruzzo Region, in accordance with current Italian legislation (Legislative Decree 155/2010). Also, in the Plan, there is the inventory of the main sources of polluting emissions (updated to 2012), which in this area largely sees the contribution of non-industrial combustion plants (mainly domestic heating plants) as regards particulate emissions



(78.5% of the total for PM<sub>10</sub>, 88.8% for PM<sub>2.5</sub> and 88.4% for benzene), while for nitrogen oxides road traffic prevails (49.7%), with industrial combustion plants in second place with 23.6%.

## 2.2 Data

This study tracks four pollutants over two specific periods: from 1st February to 10 th March 2020 (before lockdown period) and from 11 st March 2020 to 18 th April 2020 (strict lockdown period). The analyses include measures of NO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> and benzene obtained from the automatic reporting platform, run by Regional Agency for the Environmental Protection of Abruzzo Region (ARTA). These variables are measured in micrograms per cubic meter ( $\mu\text{g}/\text{m}^3$ ) and information are obtained from five monitoring sites. The spatial location of all five monitoring stations is shown in Fig. 1. The air monitoring stations of Pescara (Via Firenze) and Montesilvano are designed as *Urban Traffic* (UT) and are located roadside, where the pollution level is most influenced by traffic emissions from neighboring roads with medium-high traffic intensity; conversely, air quality data collected from the monitoring stations of Pescara (Teatro d'Annunzio), Chieti and Francavilla are deemed *Urban Background* (UB), located where the pollution level is not influenced mostly by emissions from specific sources and are representative of the population average exposure. Hourly measurements of pollutants have been collected from February to April 2020.

## 2.3 Descriptive statistics and graphical analysis

After the implementation of strict lockdown measures starting from 11 st March 2020, air pollution of the urban area of Chieti-Pescara has witnessed a substantial improvement. Table 1 highlights the net and percentage variations of each pollutant, in each monitoring site, before and during the lockdown. Early evidence of significant reduction of NO<sub>2</sub> (air pollutant mainly ascribable to fossil fuel combustion) as a consequence of lockdown, both in Asia and Europe, was showed by the data collected by NASA and ESA satellites (Gautam 2020a). In our studied area it can be also noticed that NO<sub>2</sub> has shown the most significant declining trend. In particular, the concentrations of this pollutant were approximately 50% lower compared to the previous average, that of pre-lockdown period. On the other hand, we recorded an increase of PM<sub>10</sub> and PM<sub>2.5</sub> concentrations during the lockdown weeks, whereas benzene levels dropped in the traffic measuring stations and increased in the background monitoring sites. A trend analysis of 24-h daily average data for the four pollutants was also considered for the above stated periods in all

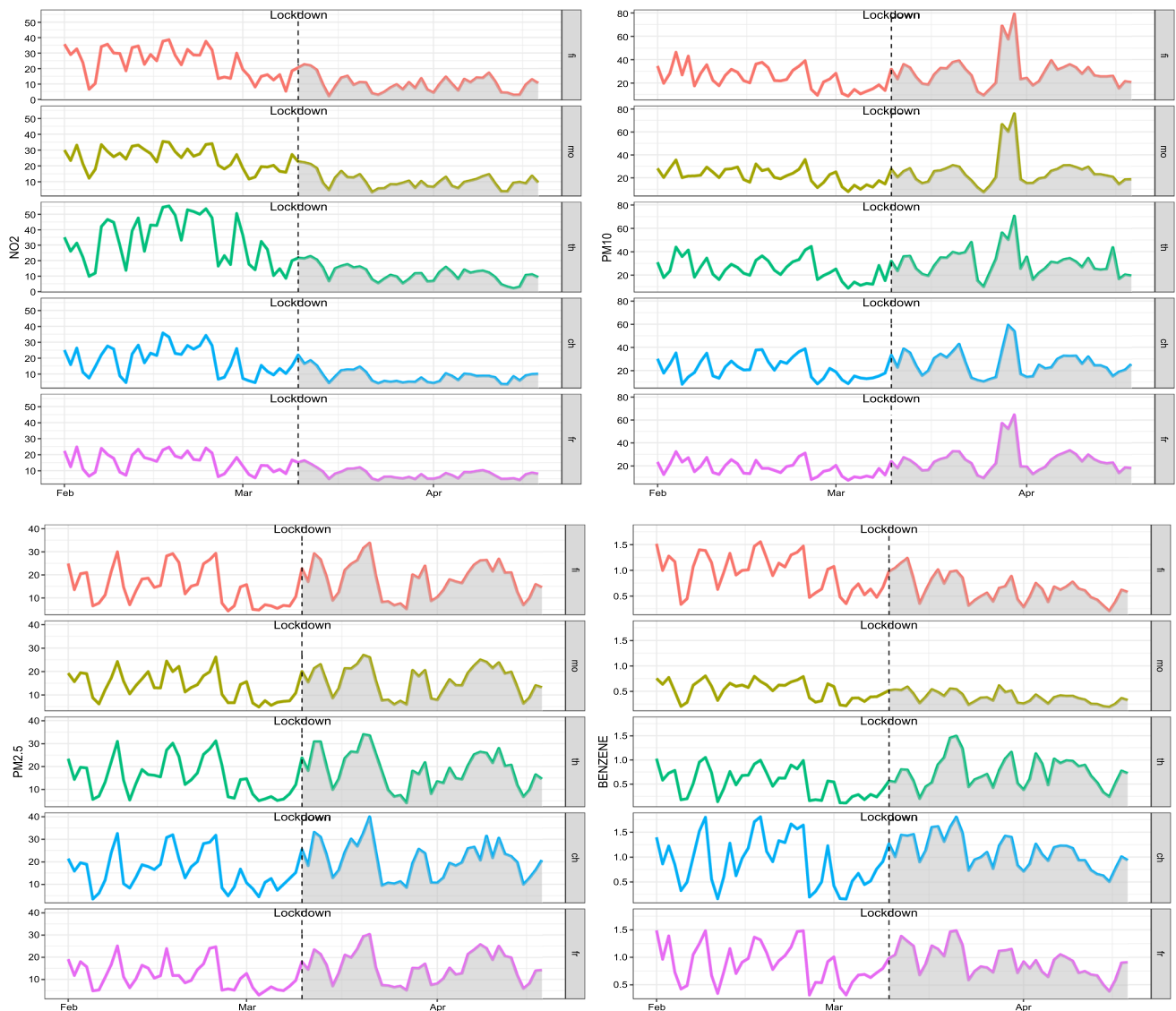
**Table 1** Net and % variation of pollutants concentration levels in the urban area of Chieti-Pescara

	UT		UB		
	fi	mo	th	ch	fr
<i>Net variation</i>					
NO <sub>2</sub>	-13.9	-14.7	-21.2	-10.3	-7.6
PM <sub>10</sub>	5.1	3.7	5.7	4.3	7.3
PM <sub>2.5</sub>	2.9	2.2	3.1	4.4	4.1
Benzene	-0.31	-0.15	0.22	0.18	0.04
<i>% variation</i>					
NO <sub>2</sub>	-57.9	-58.7	-65.2	-54.8	-49.1
PM <sub>10</sub>	20.5	16.8	22.0	19.4	40.8
PM <sub>2.5</sub>	19.0	15.6	19.8	26.7	34.4
Benzene	-32.57	-27.56	40.06	19.63	4.27

Acronyms of monitoring stations: *fi* Via Firenze, *mo* Montesilvano, *th* Teatro d'Annunzio, *ch* Chieti, *fr* Francavilla al Mare

monitoring stations to better understand the impact in the levels of pollutants accumulation amid the lockdown period.

Figure 2 allows to capture the changes in concentrations of four pollutants for the pre-lockdown and during-lockdown period. The reduction of NO<sub>2</sub> during the lockdown is clearly visible and marked in all monitoring sites and is due to the collapse of vehicular flows, even if differences in magnitude exist depending on the stations. As observed by several researches in other areas worldwide (Wang et al. 2020; Donzelli et al. 2020), the particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>) seems to be rather independent from the measures adopted during the COVID-19 nation-wide lockdown: our data show that background and traffic stations undergo an increase, whereas benzene levels dropped in the UT stations and increased in UB ones. This apparently strange behaviour of particulate matter can be explained taking into account the peculiarity of this pollutant. Airborne particulate matter can be regarded as a complex mixture of solid particles and liquid droplets generated by a wide variety of natural or anthropogenic sources, with different size and chemical composition. The fine fraction (PM<sub>2.5</sub>) arises mainly from combustion (primary PM<sub>2.5</sub>) or gas-to-particle conversion processes (secondary PM<sub>2.5</sub>), while the coarse fraction (PM<sub>10</sub>) arises mainly from traffic-linked mechanical processes (road dust resuspension, brake and tyre wear emissions) (Grigoratos and Martini 2014) or comes from natural sources (marine aerosol, wind-blown soil, pollens). This implies that the monitoring sites might be under the effect of multiple emission sources, not linked to urban traffic. In this regard, we must note that in the lockdown period the average temperature in the area under study was only slightly higher than in the pre-



**Fig. 2** Daily variation of pollutants for stations before and during lockdown occurred on 10th March

lockdown period ( $11.1\text{ }^{\circ}\text{C}$  vs  $10.8\text{ }^{\circ}\text{C}$ , measured in Pescara at ARTA weather station), despite the fact that lockdown occurred in the first part of spring, while the previous period occurred in February and the first decade of March. This fact, together with the prolonged stay of the people at home, could have raised the emissions related to domestic heating. Besides, a pertinent amount of  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  variability has a meteorological origin. As known (Querol et al. 2004; Galindo et al. 2011; Wang et al. 2020), meteorological conditions play key roles in physico-chemical processes governing formation and transport of airborne pollutants and their variability may have nullified the reductions caused by the drop in emissions related to road traffic. Considering, for example, the concentrations of  $\text{PM}_{2.5}$  in the “Teatro d’Annunzio” station, we note that the average value in the month of February and the first decade

of March 2019 was  $23.6\text{ }(\mu\text{g}/\text{m}^3)$ , while in the same period of 2020 (the “pre lockdown period” of our analysis) the average was  $15.3\text{ }(\mu\text{g}/\text{m}^3)$ . This is just the effect of inter-annual variability, due to the different meteorological scenarios. As an extreme example of variability induced by weather conditions, we must mention a notable long-distance transport event of natural dust from the Central Asia (Aral Sea) occurred at the end of March (29–31 March), which produced a sudden increase of the  $\text{PM}_{10}$  level (clearly visible in the graphs of Fig. 2). Concerning the benzene concentrations, as stated earlier, it appears that this pollutant exhibits very different behaviours in background stations compared to traffic ones. In the latter, there is a decline, albeit contained, due to the reduction of vehicular flows during the lockdown, while in the former there is stability or even slight increases, probably due to domestic

heating systems, particularly those fuelled by wood or pellets. To get a comprehensive understanding of how lockdown policies have affected air pollution, we also look at the weekly concentrations of each pollutant at background and traffic stations before and during the restriction periods. From Figs. 3 and 4, it is evident that on Sunday, the traffic during the lockdown phase is virtually zero, therefore the concentrations of  $\text{NO}_2$ , pollutant specifically linked to vehicle emissions, reduce more than in the other days of the week. It is worth noting that “Teatro d’Annunzio” measuring station seems to be affected by traffic emissions in an anomalous way (despite being a background site). Also the inspection of weekly concentrations reveals that the impact of restrictions measures on  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  is the most complex of the four pollutants studied: we are not able to detect consistent patterns with vehicular flows even if the small sample size, only five weeks and half, could affect the empirical findings. More in detail, among the background monitoring stations, “Teatro d’Annunzio” results more subject to the natural component of  $\text{PM}_{10}$ , probably due to marine aerosol: this station is

about 200 m from the sea, with no buildings in the way. Regarding the weekly benzene concentrations, the comparison between the two traffic monitoring sites indicates that “Via Firenze” is probably subject to the emissions arising from combustion (mainly domestic heating) compared to those due to traffic road. The results of FANOVA (see subsection 4.3) seem to point out a possible misclassification of this monitoring site, which is supposed to be more prone to urban traffic. Conversely, the traffic station “Montesilvano” being on the edge of the urban area is especially exposed to road traffic emissions.

### 3 Theoretical framework

In functional data analysis, the data are curves or more general functions that evolve over time, space or other continuous argument. However, the available data are usually vectors associated with the observation of a variable at a finite set of time points (longitudinal data). In fact, functional data can be seen as a particular case of high-

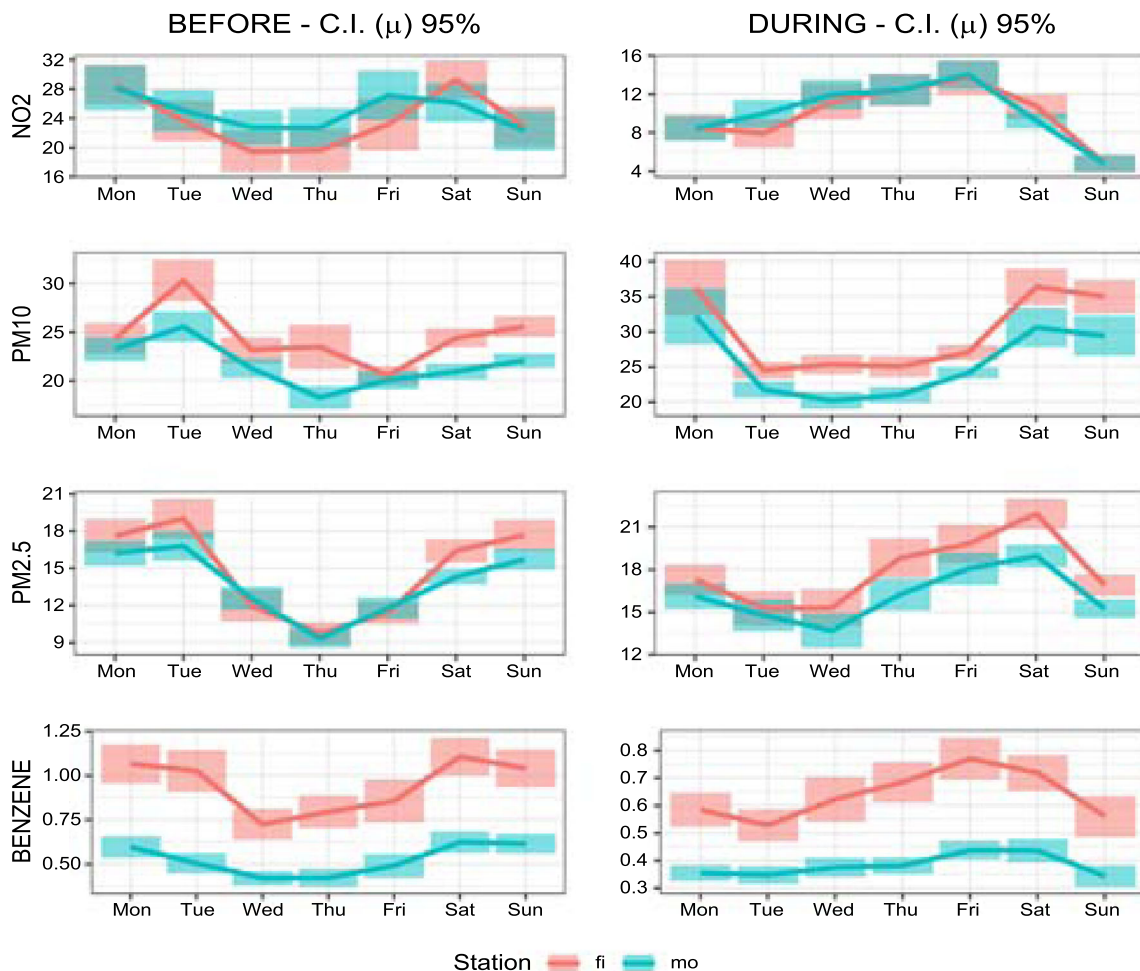
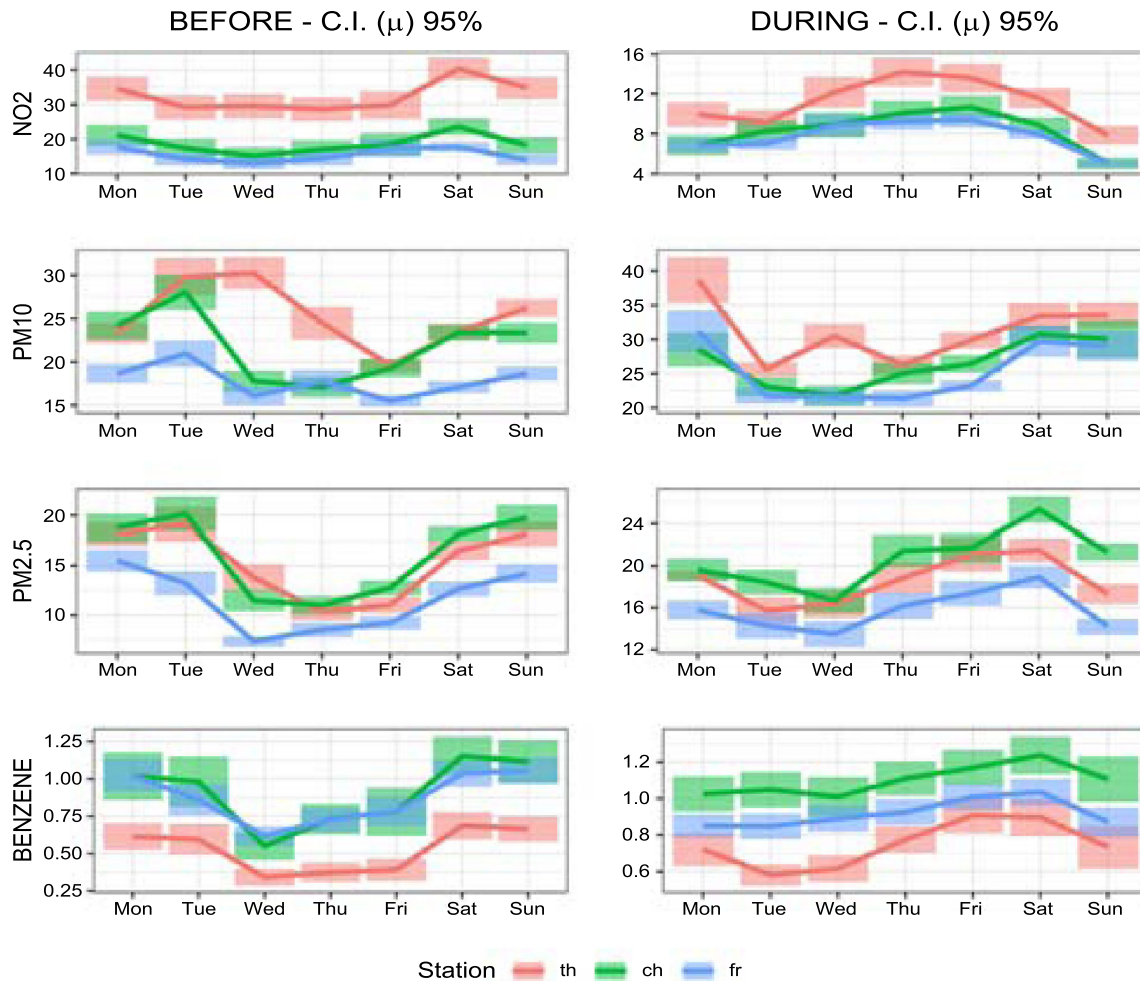


Fig. 3 Weekly concentrations of each pollutant at traffic stations before (left panel) and during (right panel) lockdown



**Fig. 4** Weekly concentrations of each pollutant at background stations before (*left panel*) and during (*right panel*) lockdown

dimensional data with a number of highly correlated variables that is usually much larger than the sample size. Because of this, the classical multivariate statistical methods are not usually efficient for functional data on account of problems related to the sample size and overfitting. In particular, it is well known that the multivariate parametric (ANOVA) and nonparametric approaches for testing homogeneity of both, independent and non independent (repeated measures) samples of vectors, do not work properly when the number of variables is large and many of them are not applicable when the dimension of the data exceeds the sample size (see, for example, Biswas and Ghosh 2014). The functional data analysis approaches solve these problems by introducing in the statistical models the true form of curves and considering the metric associated with the functional space they belong to. A detailed comparison of eleven existing functional tests for the one-way ANOVA problem for independent functional data was developed in an exhaustive simulation study from which guidelines for using the tests in practice were

provided (Gorecki and Smaga 2015). Different FANOVA approaches are proposed in this section for univariate and multivariate samples of dependent and independent curves, respectively.

Let  $X_{ijr} = (X_{ijr1}, \dots, X_{ijrH}), i = 1, \dots, g, j = 1, \dots, n_i, r = 1, \dots, R$  be a sample of curves. Note that  $g$  represents the number of independent groups,  $H$  is the number of observed response variables,  $R$  denotes the number of different periods of time (or conditions) where the response variable is observed (repeated measures) and  $n = \sum_{i=1}^g n_i$  is the sample size. It is considered that these curves are realizations of a  $H$ -dimensional stochastic process  $X = (X_1, \dots, X_H)$ , whose components are second order and continuous in quadratic mean stochastic processes with sample paths belonging to the Hilbert space  $L^2[\mathcal{T}]$  of squared integrable functions on  $\mathcal{T}$ , with the natural inner product

$$\langle f, g \rangle = \int_{\mathcal{T}} f(t)g(t)dt, \quad \forall f, g \in L^2[\mathcal{T}].$$



### 3.1 FANOVA for repeated measures

The goal is to test the equality of mean functions associated with the observation of a functional variable in two different conditions or periods of time for the same subjects. For instance, the problem laid out in this paper about the evolution of the quality of the air before and during the lockdown. That is, whether the level of each pollutant has changed during the lockdown. The theoretical framework involves the use of tools for repeated measures, and in particular, the analysis of variance. In the literature there are not many works related to this matter for the field of FDA. Martinez-Cambor and Corral (2011) introduced the first testing procedure for this problem by keeping in mind the between group variability. They proposed three different approaches in order to approximate the null distribution. The first technique consisted of applying a bootstrap parametric method through re-sampling some Gaussian process involved. The second and third methods were based on non-parametric approaches via bootstrap and permutation tests. Later, Smaga (2019) proposed another perspective focused on the Box-Type approximation. In that study, the four existent methods were compared, turning out to be the Box-Type approximation the quickest option from the computational viewpoint. In relation to size control and power all of them gave similar results, and its finite sample behaviour was very satisfactory. However, both works agree that for very small sample size, the bootstrap tests are lightly nonconservative. Smaga (2020) adapted two new statistics from the classical paired  $t$ -test to the functional data framework. This new approach is more powerful than the testing procedures aforementioned because takes the within group variability into account as well. Here, the distributions of the statistics were also approximated by parametric methods based on derived asymptotic distributions as well as non-parametric bootstrap and permutation approaches. The simulation study proved that the asymptotic and Box-Type tests are not recommended because of their liberality. Smaga (2020) suggested the permutation tests, although the non-parametric bootstrap methods also work correctly. Nevertheless, it was emphasized that there are evidences that the procedures proposed tend to be nonconservative for small sample size.

In what follows, a single functional variable is considered because they are going to be dealt separately. In this context, it is assumed that the sample functions can be represented as  $X_{jr}(t)$  with  $t \in \mathcal{T} = [a, b]$ ,  $j = 1, \dots, n$  and  $r = 1, \dots, R$ , such that  $E[X_{jr}(t)] = \mu_r(t)$ . Only two different conditions or periods of time are evaluated in the current work ( $R = 2$ ). Besides, each trajectory can be expressed as  $X_{jr}(t) = \mu_r(t) + e_{jr}(t)$  where  $e_{jr}(t)$  are random functions

centered in mean. In this kind of problem the pursued goal is to test the hypothesis

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

Martinez-Cambor and Corral (2011) proposed the following statistics in order to solve the statistical hypothesis testing

$$C_n = n \int_T (\bar{X}_1(t) - \bar{X}_2(t))^2 dt,$$

where  $\bar{X}_r(t) = n^{-1} \sum_{j=1}^n X_{jr}(t)$  is the mean function for each condition or period of time. This statistics avoid the homoscedasticity assumption.

Due to  $C_n$  only takes the between group variability, Smaga (2020) proposed the following two statistics in order to consider both the between and within group variabilities

$$\begin{aligned} \mathcal{D}_n &= n \int_T \frac{(\bar{X}_1(t) - \bar{X}_2(t))^2}{\hat{K}(t, t)} dt, \\ \mathcal{E}_n &= \sup_{t \in [a, b]} \left\{ \frac{n(\bar{X}_1(t) - \bar{X}_2(t))^2}{\hat{K}(t, t)} \right\}, \end{aligned}$$

with

$$\hat{K}(t, t) = \frac{\sum_{j=1}^n [(X_{j1}(t) - \bar{X}_1(t)) - (X_{j2}(t) - \bar{X}_2(t))]^2}{n - 1}.$$

One of the biggest problems in the practice is that curves are observed in discrete time because it is impossible to observe a set of functions continuously in time. Thus, the first step would be to reconstruct the functional form of the curves approximately. Ferraty and Vieu (2006) proposed to use non-parametric techniques for this purpose, meanwhile Ramsay and Silverman (2002, 2005) suggested an approach based on basis expansion of each sample curve. This last strategy consists of assuming that curves belong to a finite-dimension space spanned by a basis  $\{\phi_1(t), \dots, \phi_p(t)\}$ , so that they can be expressed as

$$X_{jr}(t) = \sum_{k=1}^p a_{jrk} \phi_k(t) = a'_{jr} \boldsymbol{\phi}(t), \quad j = 1, \dots, n; r = 1, 2,$$

where  $a_{jrk}$  represent the basis coefficients of the reconstruction for the corresponding sample curve with  $a_{jr} = (a_{jr1}, \dots, a_{jrp})'$  and  $\boldsymbol{\phi}(t) = (\phi_1(t), \dots, \phi_p(t))'$ . Note that  $p$  must be sufficiently large to guarantee an accurate precision. Besides, it is necessary to choose properly the dimension and the type of the basis by keeping in mind the nature of the curves. There are numerous basis systems but the most employed ones are Fourier functions (for periodic data), B-spline (for non-periodic and smooth data) and

wavelets (for curves with strong local behaviour). Finally, sample trajectories can be observed with error or without error. For the first case, least squares approximation is usually used in order to estimate the basis coefficients, whereas for the second scene some interpolation method could be applied. For more details about these methodologies, Ramsay and Silverman (2005) carried through an exhaustive study and Ramsay et al. (2009) extended these aspects to the software R.

$C_n$ ,  $\mathcal{D}_n$  and  $\mathcal{E}_n$  can be computed by considering the basis expansion. In fact, it is direct to prove that

$$\begin{aligned} (\bar{X}_1(t) - \bar{X}_2(t))^2 &= (\bar{a}'_1\phi(t) - \bar{a}'_2\phi(t))^2 \\ &= (\phi(t)'\bar{d})^2 = \phi(t)'\bar{d}\bar{d}'\phi(t), \end{aligned}$$

and

$$\begin{aligned} \hat{K}(t, t) &= \frac{\sum_{j=1}^n [(X_{j1}(t) - \bar{X}_1(t)) - (X_{j2}(t) - \bar{X}_2(t))]^2}{n - 1} \\ &= \text{Var}(X_1(t)) - 2\text{Cov}(X_1(t), X_2(t)) + \text{Var}(X_2(t)) \\ &= \hat{C}_1(t, t) - 2\hat{C}_{12}(t, t) + \hat{C}_2(t, t) \\ &= \phi(t)'(\hat{\Sigma}_1 - 2\hat{\Sigma}_{12} + \hat{\Sigma}_2)\phi(t), \end{aligned}$$

with

$$\bar{d} = (\bar{d}_1, \dots, \bar{d}_p)' = \bar{a}_1 - \bar{a}_2 = (\bar{a}_{11}, \dots, \bar{a}_{1p})' - (\bar{a}_{21}, \dots, \bar{a}_{2p})'$$

where  $\bar{a}_{rk} = n^{-1} \sum_{j=1}^n a_{jr k}$   $r = 1, 2; k = 1, \dots, p$ . Besides,  $\hat{\Sigma}_r$  is the sample covariance matrix of the matrix  $A_r$  of basis coefficients in the group  $r$ , whose elements are  $A_r = (a_{jr k})$ , and  $\hat{\Sigma}_{12}$  is the sample cross-covariance matrix between  $A_1$  and  $A_2$ . Note for major clarity that  $\bar{X}_r = n^{-1} \sum_{j=1}^n a'_{jr} \phi(t) = \bar{a}'_r \phi(t)$ .

### 3.2 Multivariate FANOVA for independent measures

The following idea in this kind of analysis is a little bit different to the case of repeated measures. Now, the aim is to test the equality of the mean functions coming from independent groups. For example, the evolution of level of benzene in the air in two different regions. If there is a response variable (e.g. level of benzene), the problem is known as Univariate FANOVA. Likewise, another fundamental aspect in these studies is the number of factors that determine the different groups. If it only exists one factor (e.g. regions) the problem is called one-way FANOVA. There are several existing methods for testing the one-way FANOVA problem (Faraway 1997; Cuevas et al. 2004; Zhang et al. 2013; Zhang 2014). On the other hand, Gorecki and Smaga (2015) made a detailed comparison of tests for the one-way ANOVA problem for functional data and presented tests based on a basis function representation.

These tests were inspired by the idea of the B-Spline method of Shen and Faraway (2004). In this line, Aguilera et al. (2021) suggested a novel approach by using Functional Principal Component Analysis (FPCA). This method consists of testing multivariate homogeneity on a vector of principal components scores. However, although there are available many works for the univariate case, the natural extension for the multivariate case (more than one functional response variable) is not a theme that it had been studied deeply. Some references are Jacques and Preda (2014), and Gorecki and Smaga (2017). Here, a novel approach based on multivariate FPCA is considered for dealing with the multivariate FANOVA problem. This new methodology can be seen as the extension of the parametric and nonparametric approaches proposed by Aguilera et al. (2021).

Suppose that  $X_{ijh}(t)$  with  $i = 1, \dots, g, j = 1, \dots, n_i$  and  $h = 1, \dots, H$  are a set of curves. Then, the information for each subject is a vector of curves denoted by  $\mathbf{X}_{ij}(t) = (X_{ij1}(t), \dots, X_{ijH}(t))'$ . Besides, it is assumed that  $\mathbf{X}_{ij}(t)$  are i.i.d. multivariate functional variables with mean vector  $\boldsymbol{\mu}_i = (\mu_{i1}(t), \dots, \mu_{iH}(t))'$  and matrix covariance function  $\mathbf{C}$  such that  $\mathbf{C}(t, s) = (C_{h,h'}(t, s))$ ,  $t, s \in \mathcal{T}$  and  $h, h' = 1, \dots, H$ . Note that if  $h = h'$ , then  $C_{h,h}$  is the covariance function and otherwise, that is  $h \neq h'$ ,  $C_{h,h'}$  represents the cross-covariance function. Now, the aim is to test

$$H_0 : \boldsymbol{\mu}_1(t) = \dots = \boldsymbol{\mu}_g(t) \quad \forall t \in [a, b],$$

against the alternative that its negation holds.

In the field of FDA, it is very common to deal with high dimension data. These type of data are defined as data associated to a great number of highly correlated variables where the sample size is too much small. For this reason, one of the most important technique in FDA is FPCA. This tool reduces the dimension of the problem and explains the main characteristics and modes of variation of the curves in terms of a reduce set of uncorrelated variables call functional principal components (PC's). Ramsay and Silverman (2002) presented the univariate approach and discussed the extension of FPCA to the case of bivariate functional data. This theory can be adapted for more than two response variables. PC's are obtained as some generalized linear combinations of the process variables with maximum variance. Formally, the  $m$ -th principal component scores are determined by

$$\begin{aligned} \zeta_{ijm} &= \int_{\mathcal{T}} (\mathbf{X}_{ij}(t) - \boldsymbol{\mu}(t))' \mathbf{f}_m(t) dt \\ &= \sum_{h=1}^H \int_{\mathcal{T}} (X_{ijh}(t) - \mu_h(t)) f_{mh}(t) dt, \end{aligned}$$

where  $\boldsymbol{\mu}(t) = (\mu_1(t), \dots, \mu_H(t))$  is the overall mean function and  $\mathbf{f}_m(t) = (f_{m1}(t), \dots, f_{mH}(t))'$  are the vector of weight

functions (or loadings) that maximizes the variance subject to  $\sum_{h=1}^H \int_{\mathcal{T}} f_{mh}(t)f_{m'h}(t)dt = 1$  if  $m = m'$  and 0 otherwise. These functions are obtained as the eigenequation system  $Cf_m = \lambda_m f_m$ , with  $C$  being the covariance operator and the sequence  $\{\lambda_m\}_{m \geq 1}$  of positive real eigenvalues decreasing to zero indicate the amount of variance attributable to each component. The aforementioned system can be written in detail as follows:

$$\begin{aligned} \int_{\mathcal{T}} C_{11}(s,t)f_{m1}(t)dt + \dots + \int_{\mathcal{T}} C_{1H}(s,t)f_{mH}(t)dt &= \lambda_m f_{m1}(s) \\ \int_{\mathcal{T}} C_{21}(s,t)f_{m1}(t)dt + \dots + \int_{\mathcal{T}} C_{2H}(s,t)f_{mH}(t)dt &= \lambda_m f_{m2}(s) \\ &\vdots \\ \int_{\mathcal{T}} C_{H1}(s,t)f_{m1}(t)dt + \dots + \int_{\mathcal{T}} C_{HH}(s,t)f_{mH}(t)dt &= \lambda_m f_{mH}(s). \end{aligned}$$

Highlight that each PC is a zero-mean random variable with maximum variance and uncorrelated with the remainder of PC's. Hence, in multidimensional context and similar to the univariate setting, this process admits the following orthogonal decomposition known as Karhunen-Loève expansion

$$X_{ij}(t) = \boldsymbol{\mu}(t) + \sum_{m=1}^{\infty} \xi_{ijm} f_m(t).$$

The principal advantage of this decomposition is that curves can be approximated by means of a principal reconstruction in terms of the first  $q$  PC's, that is  $X_{ij}^q(t) = \boldsymbol{\mu}(t) + \sum_{m=1}^q \xi_{ijm} f_m(t)$ . Normally,  $q$  is chosen so that the explained cumulative variability is as close as possible to 100%. With this approach, the dimension of the problem is considerably reduced.

Jacques and Preda (2014) explain the multivariate FPCA by means of the basis expansions and it is summarized in Schmutz et al. (2020). This approach is briefly explained hereafter. If the basis expansion is considered,  $X_{ij}(t)$  can be expressed as

$$X_{ij}(t) = \boldsymbol{\Phi}(t) \mathbf{a}'_{ij},$$

where the basis coefficients are gathered as  $\mathbf{a}_{ij} = (a_{ij11},$

$$\dots, a_{ij1p_1}, a_{ij21}, \dots, a_{ij2p_2}, \dots, a_{ijH1}, \dots, a_{ijHp_H})$$

with  $p_h$  being the number of basis functions for the  $h$ -th response variable and

$$\boldsymbol{\Phi}(t) = \begin{pmatrix} \phi_{11}(t) & \dots & \phi_{1p_1}(t) & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & \phi_{21}(t) & \dots & \phi_{2p_2}(t) & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & \phi_{H1}(t) & \dots & \phi_{Hp_H}(t) \end{pmatrix}.$$

In general  $\mathbf{X}(t) = \mathbf{A}\boldsymbol{\Phi}'(t)$ , where  $\mathbf{A}$  is the resultant matrix after joining by row all  $\mathbf{a}_{ij}$ . Thus, whether the mean vector is subtracted to each row of  $\mathbf{X}(t)$ , the spectral decomposition of the covariance operator  $C$  becomes

$$\boldsymbol{\Phi}(s)\Sigma_A \mathbf{W} \mathbf{b}'_m = \lambda_m \boldsymbol{\Phi}(s) \mathbf{b}'_m,$$

with  $\Sigma_A$  being the covariance matrix of  $\mathbf{A}$ ,  $\mathbf{b}_m$  being a row-vector that contains the basis coefficients of  $f_m(t) = \boldsymbol{\Phi}(t) \mathbf{b}'_m$  and  $\mathbf{W} = \int_{\mathcal{T}} \boldsymbol{\Phi}(t) \boldsymbol{\Phi}'(t) dt$  being the matrix of inner products between basis functions with dimension  $\sum_{h=1}^H p_h \times \sum_{h=1}^H p_h$ . Since the showed spectral decomposition is true for all  $s$ , the expression can be reduced as  $\Sigma_A \mathbf{W} \mathbf{b}'_m = \lambda_m \mathbf{b}'_m$ . Now, by considering  $\mathbf{u}_m = \mathbf{b}_m \mathbf{W}^{1/2}$ , the multivariate FPCA is equivalent to the multivariate PCA of the matrix  $\mathbf{A} \mathbf{W}^{1/2}$ , whose covariance matrix can be diagonalized as follows:

$$\mathbf{W}^{1/2'} \Sigma_A \mathbf{W}^{1/2} \mathbf{u}'_m = \lambda_m \mathbf{u}'_m.$$

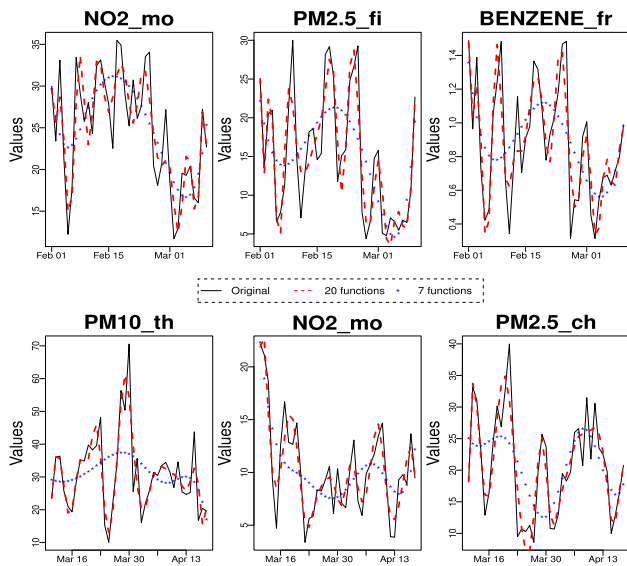
Therefore, the PC's are given by

$$\xi_{ijm} = \mathbf{a}'_{ij} \mathbf{W} \mathbf{b}_m.$$

In order to obtain the principal component scores for the multivariate case, Schmutz et al. (2019) implemented in the software R (R Core Team 2020) the package called "funHDDC". Once they are computed, there are suggested two different ways to solve the problem based on testing homogeneity on the vector of the first  $q$  principal components scores in the  $g$  groups. The first one is to perform univariate ANOVA on each principal component correcting the level of significance for the normality case. It is well-known whether the multivariate normality is suitable, the uncorrelatedness implies independence and then, it does not make sense to consider a multivariate approach. Otherwise, when the multivariate normality is not satisfied, the option is to apply non-parametric multivariate tests such as the extensions of the univariate Kruskal Wallis's test and Mood's's test. Normally, it is recommended to use the permutation version of the tests when the sample size is small (Oja 2010).

## 4 Results

We now illustrate the use of testing procedures previously described to ascertain whether the level of each pollutant has changed during the lockdown period. As pointed out in Sect. 2, to reveal the impact of restriction measures due to the COVID-19 on the air quality, the obtained environmental datasets were divided in two time frames, of the same length (39 days): i) pre-lockdown (February 1, 2020 - March 10, 2020) and ii) during-lockdown (March 11, 2020 - April 18, 2020). From a theoretical viewpoint, we have longitudinal functional data corresponding with the observation of the same functional variables in two different periods of time.



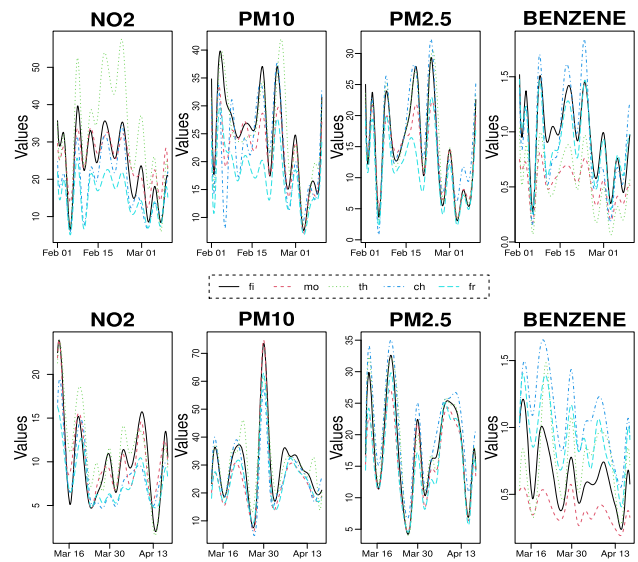
**Fig. 5** Functional approximation, using 7 and 20 basis functions, of some pollutants for stations before lockdown (*upper panel*) and during lockdown (*lower panel*)

### 4.1 Functional reconstruction of pollutant curves

As a first step of our data analysis, we reconstructed the functional form of curves from the initial points that come from the discrete values measured in the study. To convert the discretely observed data to smooth functions, the reconstruction of curves is made by using a cubic B-spline smoothing. The B-spline functions are one of the most prominent spline basis, used for non-periodic functions, which is proven to be numerically stable and flexible (Ramsay and Silverman 2005). Initially, in tailoring a basis system to fit our data, we used 7 basis functions. This option is conservative: it allows to capture the trend of curves but not their local behaviour. To recover the underlying functions of the observed data, we were increasing the number of basis functions up to 20. This choice preserves important information about the real form of the curves. Figure 5 illustrates the shape the data would take after smoothing them into these basis systems. It is clear that the increase of the number of basis functions produces smaller differences between the smoothed sample curves and the observed data (see Fig. 6). Hereinafter, an approximation of each sample curve in terms of a basis of cubic B-splines of dimension 20 is considered.

### 4.2 FANOVA results for repeated measures

Before moving on more complex studies, we carried out a univariate analysis to evaluate the behaviour of each pollutant before and during lockdown. To statistically confirm the effect of lockdown on the mean of each pollutant, we



**Fig. 6** Functional approximation, using 20 basis functions, of pollutants for stations before lockdown (*upper panel*) and during lockdown (*lower panel*)

**Table 2** FANOVA for repeated measures on the test statistics  $\mathcal{D}_n$  and  $\mathcal{E}_n$

<i>p-value</i>	$\mathcal{D}_n$	$\mathcal{E}_n$
NO <sub>2</sub>	0.034	0.035
PM <sub>10</sub>	0.000	0.034
PM <sub>2.5</sub>	0.028	0.030
Benzene	0.049	0.070

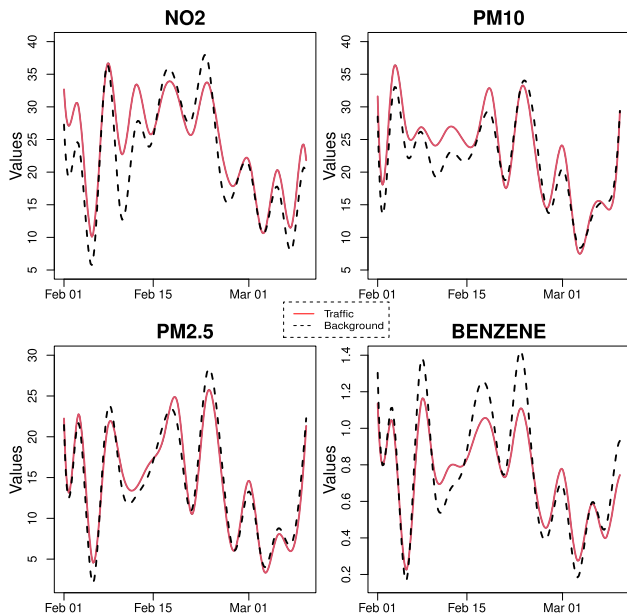
first implemented FANOVA for repeated measures, as defined in Section 3.2. Specifically, we apply the statistics  $\mathcal{D}_n$  and  $\mathcal{E}_n$  which are the best to control the between and within group variability that there are behind the repeated measures design. In order to construct the tests based on these statistics, a permutation method is used to approximate their null distributions. This technique consists of a random permutation of each sample unit. Let us denote the original data by  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  where  $X_j = (X_{j,1}, X_{j,2})$  ( $j = 1, \dots, n$ ), and the resampling vectors by  $\mathbf{X}_* = (X_{1*}, \dots, X_{n*})$  with  $X_{j*} = (X_{j,1*}, X_{j,2*})$  being a random permutation of the sample unit  $X_j$ . This process is repeated  $\Delta$  times, with  $\Delta$  a number sufficiently large, so that  $\mathcal{D}_n^{\delta*}$  and  $\mathcal{E}_n^{\delta*}$  are calculated for each replication, being  $\delta = 1, \dots, \Delta$ . Later, p-values are obtained as the proportion of times that  $\mathcal{D}_n^{\delta*}$  and  $\mathcal{E}_n^{\delta*}$  overcome  $\mathcal{D}_n$  and  $\mathcal{E}_n$ , respectively. Here, the p-values were obtained from 2000 replications. The results of the proposed testing procedures are shown in Table 2. The p-values of all tests are less than the significance level  $\alpha = 0.05$  for NO<sub>2</sub>, PM<sub>10</sub> and PM<sub>2.5</sub>. For benzene,  $\mathcal{E}_n$  shows no differences between both periods,



**Table 3** Multivariate FANOVA for independent measures

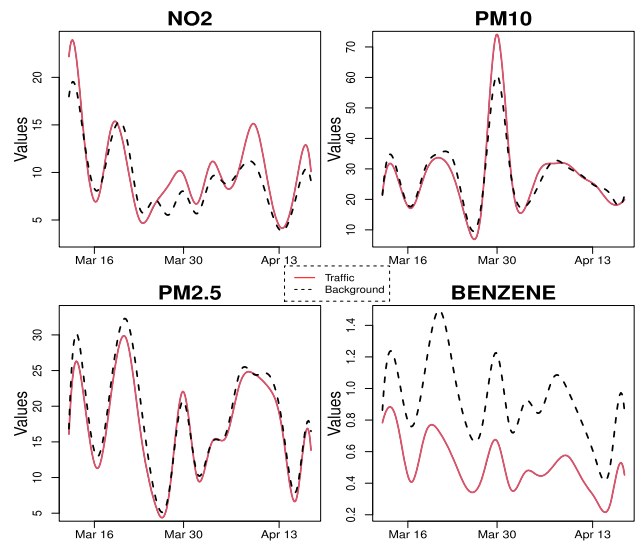
<i>p</i> -value	BL	DL
All pollutants	0.000	0.000
NO <sub>2</sub>	0.562	0.302
PM <sub>10</sub>	0.000	0.306
PM <sub>2.5</sub>	0.889	0.685
Benzene	0.186	0.000

Acronyms: *BL* Before Lockdown, *DL* During Lockdown



**Fig. 7** Mean function per pollutant of each group before lockdown

but it is very close to the limit region. Therefore, and taking into account the sample size, we have evidence to reject the null hypothesis for benzene and we state that there are also differences in the mean curves of this pollutant in the pre and during lockdown phases. These results statistically confirm the evidences already reported in Table 1 and discussed in Section 2.3. For some pollutants (PM<sub>10</sub> and PM<sub>2.5</sub>), we recorded an increase during the lockdown tenure probably due to a greater employment of domestic heating systems and the key roles of particular meteorological conditions that govern the formation and transport of particulate matter. For benzene, we pointed out a differentiate behaviour according to the type of monitoring stations, influenced by different emission sources (traffic vs domestic heating). Finally, the collapse of vehicular traffic during the quarantine days justifies the steep decline of NO<sub>2</sub> in both types of measuring sites.



**Fig. 8** Mean function per pollutant of each group during lockdown

### 4.3 Multivariate FANOVA results for independent measures

Once the impact of the lockdown has been studied, a further step of our data analysis has involved the assessment of equality of mean functions of individual groups. In our context, the groups have been individuated according to the location of the monitoring sites. In more detail, our interest lies in investigating if the mean function of all the pollutants measured in the background stations is equal to that of the urban traffic ones. The multivariate analysis of variance is carried out both before and during lockdown tenure to detect differences attributable to the government restrictions. This comparison has been evaluated firstly globally, considering all the pollutants together, and then for each variable separately. In Table 3, the results for multivariate and univariate FANOVA based on FPCA are displayed. On this matter, four principal components are chosen for both cases (multivariate and univariate analysis), since more than a 99% of total variability is explained with four components in all situations. Besides, due to the fact that the normality is in question and the sample size is really small, the extension of the univariate Kruskal-Wallis’s test with the permutation version is conducted by means of “*MNM*” R package.

Looking at the results of multivariate FANOVA, we found that, in the pre-lockdown phase, the groups are different from each other and the main discrimination is ascribable to the PM<sub>10</sub> concentrations. Furthermore, it seems that there could be indications of significance as well regarding the benzene because the *p*-value is 0.186 and by increasing the sample size we could reject the homogeneity in this pollutant. Conversely, the multivariate test is not able to distinguish the two groups in the lockdown period.

In fact, the *p*-value for the multivariate test is equal to 0.30. However, when we carry out the univariate tests, we record significant differences between the groups in relation to the benzene. This appreciation is also corroborated by the visual inspections of Figs. 7 and 8. A possible explanation for these results can be found in the simulation study performed by Aguilera et al. (2021) where it was shown that these approaches tend to be very conservative for small sample sizes.

As for fine particulate ( $PM_{2.5}$ ), the results of univariate test confirm, as already known in literature, that this pollutant is ubiquitous in urban environment due to the widespread diffusion of the sources (traffic, domestic heating etc.) and the important role played by meteorological variables in the processes of its formation and transport.

## 5 Conclusions

Recent studies suggested that lockdown measures, adopted by the most hard-hit countries around the world in the Spring of 2020 to prevent the spread of COVID-19, have had a positive significant impact on air quality. In this paper, novel approaches for functional analysis of variance with univariate repeated measures and multivariate independent measures are presented, as a new methodology for a more effective understanding of the impact of lockdown on four critical air pollutants, measured in five monitoring sites in the urban area of Chieti-Pescara (Central Italy). Being a powerful approach to modelling temporal observations, which is complementary to the usual time series techniques, FDA allowed us to reconstruct the temporal profiles of the studied pollutants for the lockdown and unlock phases in each measuring station. We have found significant reduction in  $NO_2$  levels during the lockdown period albeit some differences in magnitude are recorded according to the monitoring station. These results are in line with the findings of other published studies on this topic (Mahato et al. 2020; Berman and Ebisu 2020; Gautam and Trivedi 2020; Kerimray et al. 2020). Unlike the  $NO_2$  pollutant, for particulate matter, that is for  $PM_{10}$  and  $PM_{2.5}$ , the monitoring stations experienced an increase during the quarantine weeks. Besides, less clear was the impact of lockdown on benzene levels: the concentrations of this pollutant were smaller in the traffic stations while an increasing trend was observed in the background measuring sites. Equally important was to determine if these

differences were statistically significant. In this respect, the functional analysis of variance has proven to be beneficial to monitoring the evolution of air quality before and during the lockdown tenure and to assessing the equality of mean functions of individual groups, individuated according to the location of measuring sites. The considered FANOVA approaches based on basis expansion of sample curves, dimension reduction by using FPCA of pollutants curves and testing homogeneity on the vector of the most explicative principal component scores have made this analysis feasible providing contrasted evidence to reject the null hypothesis of equality in the mean functions of all pollutants, both in the time frame considered and the localization of monitoring stations. It should be noted that the results of multivariate FANOVA for independent measures, shown in Table 3, suggest a possible misclassification of the monitoring stations as concern the  $NO_2$ , because the proposed technique failed to discriminate between UB and UT measuring sites, despite the fact that  $NO_2$ , in urban areas, is a pollutant mostly produced by traffic emissions. The acknowledge of the presence of some redundant or misclassified monitoring stations will provide better support for managers to formulate a more adequate air pollution control strategy. For all air quality monitoring networks the identification of misclassified measurements is, in fact, an important task, not only for determining the cost of pollution monitoring, but also for determining the integrity of pollution information monitoring and the accuracy of air quality assessment. In general, the FDA framework has provided a valid understanding and knowledge of the temporal behaviour of air pollutants in a kind of controlled experiment such that offered by the lockdown. The COVID-19 restrictions reduced the anthropogenic emissions and created an “unprecedented scenario” in which the source of road traffic has been drastically dropped out.

We believe that the results of this study are of interest for environmental protection agencies involved in developing policies to achieve air quality improvements, encouraging them to establish mechanisms to reduce pollution emissions and properly redesign local monitoring network.

## Appendix

See Table 4.

**Table 4** Overview of selected studies on air pollution and COVID-19

Author (year)	Study area	Pollutant Types	Key observations
Lal et al. (2020)	Global	$NO_2$ , $CO$ , $AOD$	Reductions in $NO_2$ , $CO$ were observed in the major hotspots of COVID-19 outbreak during Feb–Mar 2020. Besides, the authors paid attention to investigate the AOD level variation during COVID-19 situation.
Dutheil et al. (2020)	Global and China	$NO_2$	Exploiting data from the Tropospheric Monitoring Instrument (TROPOMI) sensor on board ESA's Sentinel-5 satellite, massive reductions in $NO_2$ due to quarantine were observed near Wuhan, China (~ 30%) and worldwide.
Venter et al. (2020)	Global (27 countries, China, India and Europe)	$NO_2$ , $O_3$ , $PM_{2.5}$	Using satellite data and a network of more than 10,000 air quality stations, the authors find remarkable declines in ground-level nitrogen dioxide ( $NO_2$ : -29 with 95% confidence interval -44 to -13%), Ozone ( $O_3$ : -11; -20 to -2%) and fine particulate matter ( $PM_{2.5}$ : -9; -28 to 10%) during the first two weeks of lockdown.
Gope et al. (2021)	Global (most polluted cities worldwide)	$PM_{2.5}$ , $PM_{10}$ , $O_3$ , $NO_2$ , $CO$ , $SO_2$	It has been detected that the air quality of all the places has improved significantly. $NO_2$ concentration has decreased in all the cities around the world. $PM_{2.5}$ and $PM_{10}$ are the most affecting air concentration which control the air quality of all the selected places during and after lockdown.
Wang et al. (2020)	China	$PM_{2.5}$	Findings from this study supported that the suspension of anthropogenic activities (transportation and industry) contributed to the decrease of $PM_{2.5}$ concentrations. In this study, it is also shown that the benefits of emission reductions were overwhelmed by adverse meteorology.
Berman & Ebisu (2020)	USA	$NO_2$ , $PM_{10}$	The changes in levels of air pollutants across USA during COVID-19 pandemic have been investigated. The authors reported a significant reduction on $NO_2$ (up to -25.5%) and an overall decline in $PM_{2.5}$ , compared with pre-lockdown phase.
Bao & Zhang (2020)	44 cities in Northern China	$SO_2$ , $PM_{2.5}$ , $PM_{10}$ , $NO_2$ , $CO$	The vehicular restrictions during the lockdown period have led to a significant reduction of the concentrations of $SO_2$ , $PM_{2.5}$ , $PM_{10}$ , $NO_2$ and $CO$ , decreased significantly by 6.76, 5.93, 13.66, 24.67 and 4.58%, respectively.
Li et al. (2020)	Yangtze River Delta Region (China)	$SO_2$ , $NO_2$ , $CO$ , $O_3$ , $PM_{2.5}$ , $PM_{10}$ , $VOCs$	This study investigated the impact of reduced human activity on air quality over the Yangtze River Delta Region. During the most stringent Level I response period, primary pollutants like $SO_2$ , $NO_2$ , $PM_{2.5}$ and $VOCs$ have been reduced by 26, 47, 46 and 57%.
Zambrano-Monserrate et al. (2020)	China and Europe (France, Germany, Spain, and Italy)	$NO_2$ , $PM_{2.5}$	This research was aimed to highlight the positive and negative indirect effects that the new coronavirus has had on the environment. Lockdown measures led to reduced $PM_{2.5}$ and $NO_2$ concentrations. Conversely, among indirect negative effects, the increase in domestic and medical waste were mentioned.
Sicard et al. (2020)	Southern European cities (Nice, Rome, Valencia and Turin) and Wuhan (China)	$NO_x$ , $NO_2$ , $PM_{2.5}$ , $PM_{10}$ , $O_3$	In comparison to 2017–19, the lockdown measures led to a decrease of $NO_2$ (~ 53% in Europe and 57% in Wuhan), $NO$ (~ 63% in Europe), and $PM_{2.5}$ and $PM_{10}$ (~ 8% in Europe and ~ 42% in Wuhan) at urban stations. $NO_2$ and $NO$ decreased by ~ 65 and ~ 78% respectively at traffic stations in Europe. Conversely, $O_3$ increased (24% in Nice, 14% in Rome, 27% in Turin, 2.4% in Valencia and 36% in Wuhan).
Collivignarelli et al. (2020)	Milan (Italy)	$PM_{10}$ , $PM_{2.5}$ , $BC$ , benzene, $CO$ , $SO_2$ , $NO_2$ , $NO_x$ , $O_3$ , $NH_3$	The lockdown led to significant reduction of pollutant concentrations mainly due to vehicular traffic ( $PM_{10}$ , $PM_{2.5}$ , $BC$ , benzene, $CO$ and $NO_x$ )

Table 4 (continued)

Author (year)	Study area	Pollutant Types	Key observations
Tobias et al. (2020)	Spain (Barcelona)	BC, $PM_{10}$ , $NO_2$ , $O_3$	This study investigated the changes in air pollution levels during the lockdown in terms of urban background and traffic air quality observed stations. After two weeks of lockdown, the most significant reduction was estimated for BC and $NO_2$ (−45 to −51%), pollutants mainly related to traffic emissions. A lower reduction was observed for $PM_{10}$ (−28 to −31.0%). By contrast, $O_3$ , levels increased (from +33 to +57%).
Otmani et al. (2020)	Sale City (Morocco)	$PM_{10}$ , $SO_2$ , $NO_2$	Analysing air pollutants before and during the lockdown period, in this study it was found that $PM_{10}$ , $SO_2$ and $NO_2$ concentrations were reduced respectively by 75, 49 and 96%.
Sharma et al (2020)	India	$PM_{10}$ , $PM_{2.5}$ , CO, $NO_2$ , $O_3$ , $SO_2$	During March 16 <sup>th</sup> to April 14 <sup>th</sup> from 2017 to 2020 in 22 cities covering different regions of India were analyzed. Overall, around 43, 31, 10 and 18% decreases in $PM_{2.5}$ , $PM_{10}$ , CO, and $NO_2$ in India were observed during lockdown period compared to previous years.
Mahato et al. (2020)	India (Delhi)	$PM_{10}$ , $PM_{2.5}$ , $SO_2$ , $NO_2$ , CO, $O_3$ and $NH_3$	Data on seven pollutants, collected over 34 monitoring stations in Delhi, were analyzed during pre-lockdown periods and during the lockdown. Empirical findings gave evidence that air quality significantly improved during lockdown, with reductions of 60% ( $PM_{10}$ ), 39% ( $PM_{2.5}$ ), 53% ( $NO_2$ ) and 30% (CO) compared to 2019.
Gautam (2020)	India	AOD	From this study, it was emerged an up-gradation of air quality in the Indian region just a week after lockdown restrictions took place, as a result of the sizeable reduction aerosol optical thickness concentration.
Agarwal et al (2020)	India and China	$NO_2$ , $PM_{2.5}$	It has been recorded a visible improvement in air quality parameters in some cities of India and China, selected on the basis of their availability of historical air pollution data, population density, monitoring station network, and the number of positive COVID–19 cases per million people.
Kanniah et al. (2020)	Malaysia and Southeast Asia	AOD, $PM_{10}$ , $PM_{2.5}$ , $NO_2$ , $SO_2$ , CO, $O_3$	Over the Southeast Asia region, the analysis of air pollutants before and during the lockdown period reported a significant drop in AOD, $PM_{2.5}$ , $PM_{10}$ , $NO_2$ , $SO_2$ , and CO. Besides, the authors reported ~40 and ~70 % (in industrial and urban sites, respectively) reduction in AOD level in Malaysia during March–April 2020 as compared to the same period in 2019 and 2018.
Kerimray et al. (2020)	Almaty (Kazakhstan)	$PM_{2.5}$ , CO, $NO_2$ , $O_3$ , benzene, toluene	This study reported that in Almaty (Kazakhstan) $PM_{2.5}$ concentration reduced by 21% with spatial variations of 6–34% compared to the average of the same days in 2018–2019. CO and $NO_2$ concentrations reduced by 49 and 35%, respectively while $O_3$ concentrations increased by 15% compared to the preceding 17 days before the lockdown. Finally, concentrations of benzene and toluene were 2–3 times higher than in the same seasons of 2015–2019.
Zambrano-Monserrate & Ruano (2020)	Ecuador (Quito)	$NO_2$ , $PM_{2.5}$ , $O_3$	The quarantine policies adopted by the government of Ecuador have led to a significant reduction of $NO_2$ and $PM_{2.5}$ concentrations. Specifically, it was found that the $NO_2$ concentrations of 2020 were, on average, 5.6 times less than the 2018 concentrations and 4.8 times less than those from 2019. Likewise, compared with these years, the $PM_{2.5}$ concentrations were 1.5 and 1.6 times lower, respectively. On the other hand, regarding $O_3$ concentrations, it was arisen that the ozone levels in 2020 were much higher than the levels in 2018 and 2019.



**Table 4** (continued)

Author (year)	Study area	Pollutant Types	Key observations
Dantas et al. (2020)	Rio de Janeiro (Brazil)	$NO_2$ , $CO$ , $O_3$ , $NMHC$	The authors showed how the reduction of road traffic and economic activities led to the decrease in $CO$ and $NO_2$ levels and, by contrast, to the increase in ozone concentrations.
Nakata & Urban (2020)	São Paulo state (Brazil)	$NO_x$ , $NO_2$ , $CO$ , $O_3$	Data from four air quality stations in São Paulo Brazil were analyzed to assess air pollutant concentrations variations during the partial lockdown. Overall, drastic reductions on $NO_x$ (up to $-77.3\%$ ), $NO_2$ (up to $-54.3\%$ ), and $CO$ (up to $-64.8\%$ ) concentrations were observed in the urban area during partial lockdown compared to the five-year monthly mean. By contrast, an increase of approximately 30% in ozone concentrations was observed.

**Acknowledgements** This research was funded by project PID2020-113961GB-I00 of the Spanish Ministry of Science and Innovation (also supported by the FEDER program), project FQM-307 of the Government of Andalusia (Spain) and the PhD grant (FPU18/01779) awarded to Christian Acal. The authors also thank the support of the University of Granada, Spain, under project for young researchers PPJIB2020-01.

**Funding** Open access funding provided by Università degli Studi G. D'Annunzio Chieti Pescara within the CRUI-CARE Agreement.

**Data availability** All data used during the study are available from the corresponding author by request.

## Declarations

**Conflict of interest** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Agarwal A, Kaushik A, Kumar S, Mishra RK (2020) Comparative study on air quality status in Indian and Chinese cities before and during the COVID-19 lockdown period. *Air Qual Atmos Health* 13:1167–11178. <https://doi.org/10.1007/s11869-020-00881-z>
- Aguilera AM, Escabias M, Valderrama MJ (2008) Forecasting binary longitudinal data by a functional PC-ARIMA model. *Comput Stat Data Anal* 52(6):3187–3197. <https://doi.org/10.1016/j.csda.2007.09.015>
- Aguilera AM, Acal C, Aguilera-Morillo MC, Jimenez-Molinos F, Roldán JB (2021) Homogeneity problem for basis expansion of functional data with applications to resistive memories. *Math Comput Simul* 186:41–54. <https://doi.org/10.1016/j.matcom.2020.05.018>
- Aguilera-Morillo MC, Aguilera AM, Durban M (2017) Prediction of functional data with spatial dependence: a penalized approach. *Stoch Env Res Risk A* 31:7–22. <https://doi.org/10.1007/s00477-016-1216-8>
- Ambade B, Sankar TK, Kumar A, Gautam AS, Gautam S (2021) COVID-19 lockdowns reduce the Black carbon and polycyclic aromatic hydrocarbons of the Asian atmosphere: source apportionment and health hazard evaluation. *Environ Dev Sustain* 3:1–20. <https://doi.org/10.1007/s10668-020-01167-1>
- Bao R, Zhang A (2020) Does lockdown reduce air pollution? Evidence from 44 cities in northern China. *Sci Total Environ* 731:139052. <https://doi.org/10.1016/j.scitotenv.2020.139052>
- Berman JD, Ebisu K (2020) Changes in U.S. air pollution during the COVID-19 pandemic. *Sci Total Environ* 739:139864. <https://doi.org/10.1016/j.scitotenv.2020.139864>
- Bherwani H, Nair MM, Musugu K, Gautam S, Gupta A, Kapley A, Kumar R (2020) Valuation of air pollution externalities: comparative assessment of economic damage and emission reduction under COVID-19 lockdown. *Air Qual Atmos Health* 13:683–694. <https://doi.org/10.1007/s11869-020-00845-3>
- Biswas M, Ghosh AK (2014) A nonparametric two-sample test applicable to high dimensional data. *J Multivariate Anal* 123:160–171. <https://doi.org/10.1016/j.jmva.2013.09.004>
- Carslaw DC, Ropkins K (2012) Openair—an R package for air quality data analysis. *Environ Model Softw* 27–28:52–61. <https://doi.org/10.1016/j.envsoft.2011.09.008>
- Collivignarelli MC, Abba A, Bertanza G, Pedrazzani R, Ricciardi P, Miino MC (2020) Lockdown for COVID-2019 in Milan: What are the effects on air quality? *Sci Total Environ* 732:139280. <https://doi.org/10.1016/j.scitotenv.2020.139280>
- Cuevas A, Febrero M, Fraiman R (2004) An anova test for functional data. *Comput Stat Data Anal* 47(1):111–122. <https://doi.org/10.1016/j.csda.2003.10.021>
- Dantas G, Siciliano B, Franca BB, da Silva CM, Arbilla G (2020) The impact of COVID-19 partial lockdown on the air quality of the city of Rio de Janeiro, Brazil. *Sci Total Environ* 729(139085). <https://doi.org/10.1016/j.scitotenv.2020.139085>
- Donzelli G, Cioni L, Cancellieri M, Llopis Morales A, Morales Suárez-Varela MM (2020) The effect of the Covid-19 lockdown on air quality in three Italian medium-sized cities. *Atmosphere* 11:1118. <https://doi.org/10.3390/atmos11101118>

- Dutheil F, Baker JS, Navel V (2020) COVID-19 as a factor influencing air pollution? *Environ Pollut* 263:114466. <https://doi.org/10.1016/j.envpol.2020.114466>
- Escabias M, Aguilera AM, Valderrama MJ (2005) Modeling environmental data by functional principal component logistic regression. *Environmetrics* 16(1):95–107. <https://doi.org/10.1002/env.696>
- Escabias M, Valderrama MJ, Aguilera AM, Satofimia ME, Aguilera-Morillo MC (2013) Stepwise selection of functional covariates in forecasting peak levels of olive pollen. *Stoch Env Res Risk A* 27(2):367–376. <https://doi.org/10.1007/s00477-012-0655-0>
- Faraway J (1997) Regression analysis for a functional response. *Technometrics* 39(3):254–261. <https://doi.org/10.1080/00401706.1997.10485118>
- Ferraty F, Vieu P (2006) Nonparametric functional data analysis. Theory and practice. Springer Verlag, New York
- Galindo N, Varea M, Gil-Moltó J, Yubero E, Nicolás J (2011) The influence of meteorology on particulate matter concentrations at an urban mediterranean location. *Water Air Soil Pollut* 215:365–372. <https://doi.org/10.1007/s11270-010-0484-z>
- Gao HO, Niemeier DA (2008) Using functional data analysis of diurnal ozone and NOx cycles to inform transportation emissions control. *Transp Res D Transp Environ* 13(4):221–238. <https://doi.org/10.1016/j.trd.2008.02.003>
- Gautam S (2020a) The influence of COVID-19 on air quality in India: a boon or inutile. *Bull Environ Contam Toxicol* 104(6):724–726. <https://doi.org/10.1007/s00128-020-02877-y>
- Gautam S (2020b) COVID-19: air pollution remains low as people stay at home. *Air Qual Atmos Health*, pp 1–5. <https://doi.org/10.1007/s11869-020-00842-6>
- Gautam S, Trivedi U (2020) Global implications of bio aerosol in pandemic. *Environ Dev Sustain* 22:3861–3865. <https://doi.org/10.1007/s10668-020-00704-2>
- Gautam SA, Dilwaliya NK, Srivastava A, Kumar S, Baudh K, Singh D, Shah MA, Singh K, Gautam S (2021) Temporary reduction in air pollution due to anthropogenic activity switch-off during COVID-19 lockdown in northern parts of India. *Environ Dev Sustain* 23:8774–8797. <https://doi.org/10.1007/s10668-020-00994-6>
- Gope S, Dawn S, Das SS (2021) Effect of COVID-19 pandemic on air quality: a study based on Air Quality Index. *Environ Sci Pollut Res*. <https://doi.org/10.1007/s11356-021-14462-9>
- Gorecki T, Smaga L (2015) Comparison of tests for the one-way anova problem for functional data. *Comput Stat* 30(4):987–1010. <https://doi.org/10.1007/s00180-015-0555-0>
- Gorecki T, Smaga L (2017) Multivariate analysis of variance for functional data. *J Appl Stat* 44(12):2172–2189. <https://doi.org/10.1080/02664763.2016.1247791>
- Grigoratos T, Martini G (2014) Non-exhaust traffic related emissions. Brake and tyre wear pm. JRC Science and Policy Reports
- Gupta A, Bherwani H, Gautam S, Anjum S, Musugu K, Kumar N, Anshul A, Kumar R (2021) Air pollution aggravating COVID-19 lethality? *Environ Dev Sustain* 23:6408–6417. <https://doi.org/10.1007/s10668-020-00878-9>
- Hörmann S, Kidzinski L, Hallin M (2015) Dynamic functional principal components. *J R Stat Soc B* 77(2):319–348. <https://doi.org/10.1111/rssb.12076>
- Hosseini SE (2020) An outlook on the global development of renewable and sustainable energy at the time of COVID-19. *Energy Res Soc Sci* 68:101633. <https://doi.org/10.1016/j.erss.2020.101633>
- Jacques J, Preda C (2014) Model-based clustering for multivariate functional data. *Comput Stat Data Anal* 71:92–106. <https://doi.org/10.1016/j.csda.2012.12.004>
- Kanniah KD, Kamarul Zaman NA, Kaskaoutis DG, Latif MT (2020) COVID-19's impact on the atmospheric environment in the Southeast Asia region. *Sci Total Environ* 736:139658. <https://doi.org/10.1016/j.scitotenv.2020.139658>
- Kerimray A, Baimatova N, Ibragimova OP, Bukenov B, Pavel P, Karaca F (2020) Assessing air quality changes in large cities during COVID-19 lockdowns: The impacts of traffic-free urban conditions in Kazakhstan Almaty. *Sci Total Environ* 730:139179. <https://doi.org/10.1016/j.scitotenv.2020.139179>
- Lal P, Kumar A, Kumar S, Kumari S, Saikia P, Dayanandan A, Adhikari D, Khan ML (2020) The dark cloud with a silver lining: assessing the impact of the SARS COVID-19 pandemic on the global environment. *Sci Total Environ* 732:139297. <https://doi.org/10.1016/j.scitotenv.2020.139297>
- Li L, Li Q, Huang L, Wang Q, Zhu A, Xu J, Liu Z, Li H, Shi L, Li R, Azari M, Wang Y, Zhang X, Liu Z, Zhu Y, Zhang K, Xue S, Ooi MCG, Zhang D, Chan A (2020) Air quality changes during the COVID-19 lockdown over the Yangtze River Delta Region: An insight into the impact of human activity pattern changes on air pollution variation. *Sci Total Environ* 732:139282. <https://doi.org/10.1016/j.scitotenv.2020.139282>
- Mahato S, Pal S, Ghosh KG (2020) Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India. *Sci Total Environ* 730:139086. <https://doi.org/10.1016/j.scitotenv.2020.139086>
- Martínez J, Saavedra A, García-Nieto PJ, Piñeiro JI, Iglesias C, Taboada J, Sancho J, Pastor J (2014) Air quality parameters outliers detection using functional data analysis in the Langreo urban area (Northern Spain). *App Math Comput* 241:1–10. <https://doi.org/10.1016/j.amc.2014.05.004>
- Martínez-Cambor P, Corral N (2011) Repeated measures analysis for functional data. *Comput Stat Data Anal* 55(12):3244–3256. <https://doi.org/10.1016/j.csda.2011.06.007>
- Martínez Torres J, Pastor Pérez J, Sancho Val J, McNabola A, Comesaña M, Martínez, Gallagher J (2020) A functional data analysis approach for the detection of air pollution episodes and outliers: a case study in Dublin, Ireland. *Mathematics*, 8(225). <https://doi.org/10.3390/math8020225>
- Mogaji E (2020) Impact of COVID-19 on transportation in Lagos, Nigeria. *Transp Res Interdiscip Perspect* 6:100154. <https://doi.org/10.1016/j.trip.2020.100154>
- Nakada LYK, Urban RC (2020) COVID-19 pandemic: impacts on the air quality during the partial lockdown in Sao Paulo state Brazil. *Sci Total Environ* 730:139087. <https://doi.org/10.1016/j.scitotenv.2020.139087>
- Ocaña-Peinado FA, Valderrama MJ, Aguilera AM (2008) A dynamic regression model for air pollen concentration. *Stoch Env Res Risk A* 22:59–63. <https://doi.org/10.1007/s00477-007-0153-y>
- Oja H (2010) Multivariate nonparametric methods with R. Springer Science & Business Media, New York
- Otmani A, Benchrif A, Tahri M, Bounakha M, El Bouch M, Krombi MH (2020) Impact of COVID-19 lockdown on PM<sub>10</sub>, SO<sub>2</sub>, and NO<sub>2</sub>, concentrations in Sale City (Morocco). *Sci Total Environ* 735:139541. <https://doi.org/10.1016/j.scitotenv.2020.139541>
- Park A, Guillas S, Petropavlovskikh I (2013) Trends in stratospheric ozone profiles using functional mixed models. *Atmos Chem Phys* 13(22):11473–11501. <https://doi.org/10.5194/acp-13-11473-2013>
- Querol X, Alastuey A, Ruiz CR, Artiñano B, Hansson HC, Harrison RM, Buringh E, ten Brink HM, Lutz M, Bruckmann P, Straehl P, Schneider J (2004) Speciation and origin of PM<sub>10</sub> and PM<sub>2.5</sub>, in selected European cities. *Atmos Environ* 38(38):6547–6555. <https://doi.org/10.1016/j.atmosenv.2004.08.037>
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria [www.R-project.org/](http://www.R-project.org/)
- Rajput H, Changotra R, Rajput P, Gautam S, Gollakota ARK, Arora AS (2021) A shock like no other: coronavirus rattles commodity

- markets. *Environ Dev Sustain* 23:6564–6575. <https://doi.org/10.1007/s10668-020-00934-4>
- Ramsay JO, Silverman BW (2002) *Applied functional data analysis: methods and case studies*. Springer-Verlag, New York
- Ramsay JO, Silverman BW (2005) *functional data analysis*, 2nd edn. Springer-Verlag, New York
- Ramsay JO, Hooker G, Graves S (2009) *Functional data analysis with R and MATLAB*. Springer-Verlag, New York
- Sancho J, Martínez J, Pastor J, Taboada J, Piñeiro JJ, García-Nieto PJ (2014) New methodology to determine air quality in urban areas based on runs rules for functional data. *Atmos Environ* 83:185–192. <https://doi.org/10.1016/j.atmosenv.2013.11.010>
- Schmutz A, Jacques J, Bouveyron C (2019) *funHDDC: Univariate and multivariate model-based clustering in group-specific functional subspaces* R package version 2.3.0
- Schmutz A, Jacques J, Bouveyron C, Chéze L, Martin P (2020) Clustering multivariate functional data in group-specific functional subspaces. *Comput Stat* 35:1101–1131. <https://doi.org/10.1007/s00180-020-00958-4>
- Shaadan N, Deni SM, Jemain AA (2013) Assessing and comparing  $PM_{10}$  pollutant behaviour using functional data approach. *Sains Malays* 13(22):11473–11501
- Sharma S, Zhang M, Gao AJ, Zhang H, Kota SH (2020) Effect of restricted emissions during COVID-19 on air quality in India. *Sci Total Environ* 728:138878. <https://doi.org/10.1016/j.scitotenv.2020.138878>
- Shen Q, Faraway J (2004) An F test for linear models with functional responses. *Stat Sin* 14(4):1239–1257
- Sicard P, De Marco A, Agathokleous E, Feng Z, Xu X, Paoletti E, Rodriguez JJD, Calatayud V (2020) Amplified ozone pollution in cities during the COVID-19 lockdown. *Sci Total Environ* 735:139542. <https://doi.org/10.1016/j.scitotenv.2020.139542>
- Sigala M (2020) Tourism and COVID-19: impacts and implications for advancing and resetting industry and research. *J Bus Res* 117(3):312–321. <https://doi.org/10.1016/j.jbusres.2020.06.015>
- Smaga L (2019) Repeated measures analysis for functional data using Box-type approximation with applications. *REVSTAT* 17(4):523–549. [www.ine.pt/revstat/pdf/REPEATEDMEASURE\\_SANALYSISFORFUNC.pdf\(2017\)](http://www.ine.pt/revstat/pdf/REPEATEDMEASURE_SANALYSISFORFUNC.pdf(2017))
- Smaga L (2020) A note on repeated measures analysis for functional data. *AStA Adv Stat Anal* 104(1):117–139. <https://doi.org/10.1007/s10182-018-00348-8>
- Tobías A, Carnerero C, Reche C, Massagué J, Via M, Minguillón MC, Alastuey A, Querol X (2020) Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. *Sci Total Environ* 726:138540. <https://doi.org/10.1016/j.scitotenv.2020.138540>
- Valderrama MJ, Ocaña FA, Aguilera AM, Peinado FM Ocaña (2010) Forecasting pollen concentration by a two-step functional model. *Biometrics*, 66
- Venter ZS, Aunan K, Chowdhury S, Lelieveld J (2020) COVID-19 lockdowns cause global air pollution declines. *Proc Natl Acad Sci* 117(32):18984–18990. <https://doi.org/10.1073/pnas.2006853117>
- Wang P, Chen K, Zhu S, Wang P, Zhang H (2020) Severe air pollution events not avoided by reduced anthropogenic activities during COVID-19 outbreak. *Resour Conserv Recy* 158:104814. <https://doi.org/10.1016/j.resconrec.2020.104814>
- WHO (2020) Coronavirus disease (COVID-19) pandemic. <https://www.who.int/emergencies/diseases/novel-corona-virus-2019>
- Zabrano-Monserrate MA, Ruano MA (2020) Has air quality improved in Ecuador during the COVID-19 pandemic? A parametric analysis. *Air Qual Atmos Health* 13:929–938. <https://doi.org/10.1007/s11869-020-00866-y>
- Zambrano-Monserrate MA, Ruano MA, Sanchez-Alcalde L (2020) Indirect effects of COVID-19 on the environment. *Sci Total Environ* 728:138813. <https://doi.org/10.1016/j.scitotenv.2020.138813>
- Zhang JT (2014) *Analysis of variance for functional data*. CRC Press, Boca Raton
- Zhang JT, Cheng MY, Tseng CJ, Wu HT (2013) A new test for one-way ANOVA with functional data and application to ischemic heart screening. *Comput Stat Data Anal* 132:3–17. <https://doi.org/10.1016/j.csda.2018.05.004>
- Zoran MA, Savastru RS, Savastru DM, Tautan MN (2020) Assessing the relationship between surface levels of  $PM_{2.5}$  and  $PM_{10}$  particulate matter impact on COVID-19 in Milan, Italy. *Sci Total Environ* 738:139825. <https://doi.org/10.1016/j.scitotenv.2020.139825>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.