SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of

STATISTICS AND QUANTITATIVE METHODS

PhD program in **STATISTICS AND MATHEMATICAL FINANCE** Cycle **XXXIII**

Curriculum in **STATISTICS**

ESSAYS ON AIR QUALITY STATISTICS: TIME SERIES METHODS FOR EVALUATING ENVIRONMENTAL PROTECTION POLICIES IN LOMBARDY, ITALY

Surname: **MARANZANO** Name: **PAOLO**

Registration number: **800603**

Tutor: Prof. MATTEO **PELAGATTI**

Supervisor: Prof. MATTEO **PELAGATTI**

Co-Supervisor: Prof. ALESSANDRO **FASSÒ**

Coordinator: Prof.ssa EMANUELA **ROSAZZA GIANIN**

ACADEMIC YEAR 2019/2020

UNIVERSITY OF MILANO BICOCCA

DOCTORAL THESIS

Essays on air quality statistics: time series methods for evaluating environmental protection policies in Lombardy, Italy

Author:
Paolo MARANZANO

Supervisor:
Prof. Matteo PELAGATTI
Prof. Alessandro FASSÒ

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Department of Statistics and Quantitative Methods

21th December 2020

Abstract

The present manuscript attempts to contribute to the *Citizen science* philosophy, based on the collaboration between public institutions, research and active citizenship, by offering advanced statistical tools for air quality monitoring and the analysis of environmental protection actions. The main focus will be on the temporal modelling, hence involving the use of statistical methodologies for time series analysis, of air quality in Lombardy, the economic-financial heart of northern Italy, which has suffered from high pollution for decades and enjoys poor air quality despite the efforts made by the institutions. In this context, we will introduce and discuss new tools to estimate the impact on the air quality we breath of extraordinary events and unexpected shocks. In particular, we will consider both artificial events, such as the introduction of new restricted traffic zones in urban areas, and natural, such as the forced closure of all productive activities and restrictions on individual mobility following the events of the COVID-19 pandemic in 2020.
Overall, the manuscript has various objectives: on the one hand, it intends to offer a technical-scientific evaluation of the above events in order to validate environmental policy choices in the region and to open a discussion on public policies based on scientific evidence and data-driven approaches; on the other hand, it attempts to import from other realities, such as the financial world, statistical techniques useful for the analysis of data on air quality, opening up new research fronts; eventually, it wants to make data on environmental conditions in Lombardy region available to international users on a large scale through the development of specific software for downloading, visualising and managing this information.
Some of the new methodological proposals discussed in the paper are:

- the development of an algorithm, presented as a pipeline chain of distinct statistical methodologies, covering all the steps of the impact analysis: from the selection of key environmental factors to explain pollutant concentrations (model selection) to the estimation of the effects;
- the application of methodologies born in the world of financial statistics to air quality data and able to identify a level shift in multiple time series due to unexpected events;
- the application of filtering techniques to multiple time series in order to check their properties of stationarity and cointegration in the context of highly-noisy and non-Gaussian time series, such as pollution concentrations data;

The thesis is organised as follows. Chapter 2 proposes a robust overview of the European and Italian air quality monitoring systems, namely EIONET, and goes into rigorous detail on the current merger situation in Europe. In particular, the chapter analyses the case of Lombardy explaining its main geographical and physical characteristics, the composition of the economic and social network and the consequent effects on air pollution. The source of the data is also described in detail and will be examined in the following chapters. Air quality and weather data come from the Regional Agency for the Protection of the Environment (ARPA Lombardia), which manages and maintains the monitoring network and analyses the data collected. The chapter also discusses some issues related to advanced techniques to use weather and air quality data properly and lays the foundations for possible extensions.

Chapter 3 proposes the analysis of air quality time series using a purely-time series approach based on Unobservable Component Models (UCM) in State Space form and bootstrap resampling algorithms. The bootstrap use in a time series context makes it possible to correctly identify any variations in the average levels of the concentrations, taking into account the characteristics of positive asymmetry (right skewness) typical of pollutant concentrations. The use of state space models is well known and recognised in the statistical and econometric literature, which consider them adequate tools for the contemporary study of time series characteristics and the evaluation of the impact of policies and shocks on the variables of interest. The case study presented concerns the introduction of a new restricted traffic zone in the city of Milan, called Area B, and its possible effect on concentrations of pollutants. One of the analysis's main contributions is the introduction of a multi-stage model selection algorithm based on both fitting and forecasting performances criteria. The algorithm allows selecting the most relevant observable and non-observable features to explain the pollution concentrations before estimating the policy's unbiased effect.
In chapter 4, the analysis of a recent case study of great social and economic importance is proposed. We investigate the short-term impact of the restrictions on productive activities and mobility imposed by the Italian government, called in jargon lockdown, on the observed concentrations in Lombardy during the COVID-19 pandemic in 2020. Pollution concentrations are modelled using a class of time series techniques named autoregressive with exogenous covariate models or ARX models. This flexible and simple approach to time series modelling allows estimating the impact of covariates on a response variable, which is described by an autoregressive process. Model tuning and variable selection are performed by a LASSO algorithm which accounts for both seasonal and non-seasonal components. The lockdown is defined by a dummy variable which expresses the change in average levels of pollution due to the restrictive measures after being depurated by the effect of local meteorology. This study is part of the recent literature investigating the relationship between virus containment policies and environmental and social effects in Europe and actively contributes to highlighting the positive effects generated by this critical situation.
Chapter 5 presents and discusses the application of the so-called event studies methods to the time series of airborne concentrations. Event studies are analytical tools typically applied in financial and economic contexts and can assess the effects of exogenous shocks on one ore more variables of interest. Event studies methods test changes in observed levels of a series comparing a pre-event situation with the post-event situation. The comparison is performed using both parametric and non-parametric statistics. In particular, we will focus on the case of cross-sectionally correlated multivariate time series. Indeed, this feature can severely compromise classical statistics' properties even for moderate values of cross-sectional correlation among the data. The contribution to the literature is twofold: on the one hand, it broadens the scope of the above methodology to non-typical areas, such as air quality statistics; on the other hand, it presents and discusses two novel rank-based non-parametric statistics able to deal with strongly correlated series. The main statistical properties and performances of both statistics are compared to a set of other competing statistics available in the literature. Finally, all the considered test statistics are applied to air quality data for Lombardy using as the event of interest the introduction of mobility restrictions (lockdown) imposed to contrast the spread of COVID-19 virus, previously presented in Chapter 4.
Chapter 6 contributes to the literature investigating the stationarity of univariate time series and cointegration in multivariate contexts by analysing the effect of linear filters on the respective tests' critical values and the subsequent decisions to reject or accept the hypotheses. Under the null hypothesis, it is assumed that the series are non-stationary and buried by a high noise level. This immediately leads to the following implications: regarding the stationarity,

the classic ADF test will reject the non-stationarity hypothesis much more often than is tolerated by the theoretical values (i.e. large increase in test's size) and it will accept the stationarity hypothesis excessively (loss of test's power); for what concerns the cointegration, since the existence of high noise overstates the stationarity of the series, the Johansen's cointegration test will overestimate the cointegration rank, concluding that the number of cointegration relations is higher than the effective number (large increase in test's size). The linear filters considered are the moving average, the Kalman filter and the Kalman smoother. Given their noisy and volatile nature, we propose to test stationarity and cointegration on filtered and unfiltered data of observed nitrogen oxides concentrations at Milan between 2011 and 2019. The largest part of the statistical literature on air quality agrees about the airborne pollutants series's stationarity without producing sensitivity analyses or tests robust at extreme values. Using a 5-year moving window for parameter estimation and filtering, it is possible to identify possible modifications in the structure of NO_2 and NO_x concentrations and test whether linear filters' application leads to different conclusions from tests applied to raw data.
Chapter 7 concludes the manuscript, summarising the main contributions and emphasising future research developments.

Contents

List of Figures

List of Tables

"Quando le generazioni future giudicheranno coloro che sono venuti prima di loro sulle questioni ambientali, potranno arrivare alla conclusione che questi 'non sapevano': accertiamoci di non passare alla storia come la generazione che sapeva, ma non si è preoccupata."
— **Mikhail Sergeevich Gorbachev**

"Noi abbiamo bisogno di una nuova coscienza ambientale su basi globali. Per fare questo, abbiamo bisogno di educare le persone."
— **Mikhail Sergeevich Gorbachev**

Chapter 1

Introduction

1.1 Background and rationale for statistical air quality assessment: a human well-being perspective

Air is a gas mixture consisting of nitrogen (about 78%), oxygen (about 21%) and other gases, such as carbon dioxide or methane, present in much lower concentrations. We talk about atmospheric pollution whenever a substance is present in the air in a different concentration than the natural one. These alterations can be caused both by phenomena of natural origin, volcanic eruptions are an example, and by human activities, in this case we speak of anthropogenic origin. Typical artificial sources of atmospheric pollution are emissions of industrial pollutants, thermal power plants, heating systems, transport and road traffic. The substances contained in the emissions can be found directly in the environment (primary pollutants) or undergo transformation processes in the atmosphere, giving rise to new pollutants (secondary pollutants). When pollution is detected in open environments, we commonly speak of outdoor pollution, while the term indoor refers to indoor pollution such as homes, offices or public places.
People's health is affected by many environmental problems, some of them due to human activity and others simply natural. Among the latter, air pollution is of course one of the worst and most serious, particularly in urban areas of Europe. According to the special Eurobarometer survey 2017 on environmental attitudes of European people (European Commission, 2017) air pollution is currently the most important environmental risk to human health, and it is perceived as the second biggest environmental concern for Europeans, after climate change. Air pollution also damages vegetation and ecosystems. It leads to several important environmental impacts, which affect vegetation and fauna directly, as well as the quality of water and soil and the ecosystem services they support. The most harmful air pollutants in terms of damage to ecosystems are ozone (O_3), ammonia (NH_3) and nitrogen oxides (NO_X) (European Environmental Agency, 2019a).
Air quality assessment and environmental governance are two of the most important challenges that modern societies are facing for several decades and constitute a constantly-changing theme within the social, political and economic debate. Between 2001 and 2005 the United Nation (UN) developed a project named *Millennium Ecosystem Assessment* (Ecosystem, 2013) with the aim to assess the consequences of ecosystem changes for human well-being and to establish the scientific basis for actions needed to enhance the conservation and sustainable use of ecosystems and their contributions to human well-being. According to their view air quality is one of the key points conditioning human well-being, which is assumed to have multiple constituents, including:

- the basic material for a good life, such as secure and adequate livelihoods, enough food at all times, shelter, clothing, and access to goods;

- health, including feeling well and having a healthy physical environment, such as clean air and access to clean water; good social relations, including social cohesion, mutual respect, and the ability to help others and provide for children;
- security, including secure access to natural and other resources, personal safety, and security from natural and human-made disasters; and freedom of choice and action, including the opportunity to achieve what an individual values doing and being.

The environment is closely related, in both positive and negative ways, to key economic issues such as poverty, prosperity, jobs, production patterns, innovation, and resource availability or scarcity. On one hand, the economy is a major source of environmental problems, while environmental problems are increasingly causing economic losses. On the other hand, protecting the environment, as well as preventing and mitigating the impacts of pollution, are also major sources of economic opportunity, providing jobs, reducing poverty, driving innovation and addressing resource scarcity and depletion (Environment, 2019). This bilateral relationship is perfectly summarized by the concept of *Sustainable development*, firstly stated by the World Commission on Environment and Development (WCED), also known as the Brundtland Commission, in 1987 at the World Conference on Environment and Development in its annual report: "Our Common Future". Then sustainable development was defined as (Brundtland, 1987):

> *"Sustainable Development is development that meets the needs of the present (generations) without compromising the ability of future generations to meet their own needs."*

The UN effort for spreading the sustainability concept within the development framework has taken concrete form in the Sustainable Development Goals or SDGs (https://sdgs.un.org/goals). In 2015 the SDGs have been renewed and adopted by all UN Member States in 2015 as part of the 2030 Agenda for Sustainable Development of United Nations. They address 17 global challenges including poverty, inequality, economic growth, climate change, environmental degradation, peace, and justice. The development goals do not directly include air quality among the objectives, but deal with the topic in three sub-points: *Good Health and Well-Being* (Goal 3), *Sustainable Cities and Communities* (Goal 11) and *Responsible Consumption and Production* (Goal 12). Specifically, air quality is addressed in 3 targets to be achieved by 2030: substantially reduce the number of deaths and illnesses from hazardous chemicals and air, water and soil pollution and contamination (Target 3.9); reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management (Target 11.6); achieve the environmentally sound management of chemicals and all wastes throughout their lifecycle, in accordance with agreed international frameworks, and significantly reduce their release to air, water and soil in order to minimize their adverse impacts on human health and the environment (Target 12.4). The first target is measured by the mortality rate attributed to household and ambient air pollution, while the second is identified through the annual mean levels of fine particulate matter (i.e. $PM_{2.5}$ and PM_{10}) in urban areas. The third target does not have a specific associated indicator.

1.1.1 The issue of air quality statistics

As will be developed in the following sections of this manuscript, almost all the Governments around the world have set up international institutions at various levels for managing, monitoring and protecting the environment. Many of these institutions are able to issue legally binding regulations and promote awareness campaigns towards citizens. See for example the directives issued by the EU European Environmental Agency or by the US Environmental Protection Agency.

As a result of the growing demand from national and international institutions to increase knowledge and information database on environmental topics, in recent years various agencies have emerged to monitor the ecosystem and its condition. Environmental agencies work in a connected network in which the exchange of information between partners and to the external community is the main task. These institutions provide real-time data on many subjects: marine and submarine ecosystem, atmospheric pollution and air quality, climate change, health, energy, land use and so forth. Among others, the air quality monitoring function is one of the most important and requires constant methodological developments and case studies. Air quality statistics has now consolidated its position as hot-topic within the academic and applied research context. Just think, for example, that the study of environmental and spatial statistics has become a discipline taught in various international university courses. Some of the classic tasks entrusted to this research field are: estimation and forecasting of air quality long-run trends using measurements from monitors located across the territory; spatio-temporal modelling of the effect of meteorology on concentrations; dissemination of reports and analyses on the environmental situation; dissemination of public databases to other institutions and citizens to raise users' awareness of environmental issues;
As reported by the UN Environmental Outlook 2019 (Environment, 2019), environmental governance is increasingly important at all levels, including global, regional, national and sub-national governments, as well as business and civil society stakeholders. There is a need for greater synergies between governments and civil society organizations. Many efforts have been made to develop more effective facilitation methods to enable this collaboration, first of all, extensions to the citizen science, which engages citizens in scientific research (Kobori et al., 2016) to help address some of ecology's most pressing issues, such as global environmental change. Through the use of software, social media and online tools people can find projects that match their interests and learn the skills and protocols required to develop questions, collect data, submit data, and help process and analyze data collaboratively. In this framework, citizen science can be viewed as a producer of *big data* for spatial ecology. A simple example of this trans-disciplinary collaboration is the identification of butterflies in Tokyo, which is particularly difficult because of the large number of species. Volunteers submit butterfly photographs online along with their proposed species identifications and experts check each image to ensure correct identification. More than 18,000 butterfly records were uploaded over 4 years. In air quality research field citizen science can be a useful approach to collect specific data or model special observed situations. As an example, air pollution does not necessarily spread itself evenly throughout a territory. People may live next to a busy highway or some other pollution source, or live in areas with land topography that generates airborne pollutant pockets not identified by aggregate analysis. Citizen scientists can contribute more information to fill in the natural gaps in official statistics databases.

1.1.2 Air quality assesment models and weather normalisation

Air quality time series have some peculiar statistical characteristics that are more or less evident depending on the chosen time-frequency, often working with hourly or daily observations and the length of the period of interest. Time-series analysis of airborne pollution concentrations involves identifying long-term variation in the mean, namely the trend, and of cyclical or seasonal components (Salcedo et al., 1999). Given the link with other environmental drivers, such as meteorological seasons and weather conditions, infra-daily and annual seasonality and trend are certainly the key features to be considered. However, human activities must also be considered. For example, the infra-weekly seasonality generated by the working week and human activities, the more persistent traffic on certain working days or during holiday periods. Furthermore, again, consider the effects of climate mitigation policies, the closure of ever-larger

areas to traffic or environmental protection policies. All these factors influence the already complex Physico-chemical phenomena of the atmosphere and, consequently, affect identifying the underlying and data generating processes (DGPs) of the observed series.
In a recent review paper, Bai et al. (2018) tried to summarise the main statistical and artificial intelligence tools used for air quality forecasting by comparing the statistical performances presented by several studies using the main error assessment criteria available in the literature, such as the root mean square error (RMSE), mean absolute error (MAE) or mean absolute percentage error (MAPE). The review concludes by grouping the methods of analysis into categories based on the methodological area of origin. In particular, they propose some macro-categories of interest, highlighting their pros and cons:

- Statistical models: which have a wide application and require less time to build models, but they require a large amount of historical data and have a high dependence on the time series approach;
- Artificial Intelligence methods: such as the neural networks approach, have good performance and can solve nonlinear data, but the models are unstable and have a high dependence on data. Moreover, they easily fall into local optima;
- Hybrid models: which combines both statistical and machine learning algorithms.

According to their findings, AI methods applied to air quality forecasting show forecasting performances better than those of statistical modelling, but worse than those of hybrid models. Moreover, they state that forecast performance is better when considering the meteorological variables and geographic factors. Meteorological normalisation is a technique which controls for meteorology and weather overtime in an air quality time series analysis. The main aim is to assess the effect of exploratory intervention analysis or trend analysis for air quality robustly. At present, meteorological normalisation techniques range from the inclusion of weather control variables in regression designs up to machine learning algorithm, such as meteorological normalisation technique based on random forest algorithms (Grange et al., 2018; Grange and Carslaw, 2019) or gradient boosting machine (Petetin et al., 2020).
This present manuscript deals with air quality analysis and intervention assessment based on robust statistical approaches and principles, while attempting to integrate methodologies widely used in artificial intelligence and machine learning contexts, such as variable selection based on optimisation and penalisation criteria (e.g. LASSO) or meteorological normalisation. Meteorological normalisation occupies a key role in all the methodologies and their applications addressed in this manuscript. Often dealing with regression design contexts, meteorological variables will be used as control variables to cleanse the relationships of interest from any spurious influences in the data. In the case of intervention analyses, the presence of variables such as temperature or wind speed, allows to correctly identify the effect of policies and shocks net of possible variations in meteorology that could hide the real impact.

1.2 Outline and main contributions of the manuscript

The present manuscript attempts to contribute to the *Citizen science* philosophy, based on the collaboration between public institutions, research and active citizenship, by offering advanced statistical tools for air quality monitoring and the analysis of environmental protection actions. In this context, we will introduce and discuss new tools to estimate the impact on the air quality we breath of extraordinary events and unexpected shocks. In particular, we will consider both artificial events, such as the introduction of new restricted traffic zones in urban areas,

and natural events, such as the forced closure of all productive activities and restrictions on individual mobility following the events of the COVID-19 pandemic in 2020.
The main focus will be on the temporal modelling of airborne pollutant concentrations involving the use of statistical methodologies for time series analysis. The applications that will be presented concern the air quality conditions of Lombardy, the economic and financial core of northern Italy. The choice of this specific case study is motivated by the fact that northern Italy, and particularly the Lombardy, has suffered from high pollution for decades and enjoys poor air quality despite the efforts made by the institutions. Throughout the manuscript, especially in Chapter 2, we aim at describing in detail and with scientific evidence the specific situation of Northern Italy in terms of pollution, geography, orography and meteorology. In the following chapters, these characteristics will often be recalled in order to justify the choices of the proposed models or to contextualise the results obtained. For example, meteorology and orography play a key role in the discussion of the positive effects of the lockdown for COVID-19, as the effects will be more or less intense depending on the areas where the stations are located. Not surprisingly, the greatest reductions in oxides were found in urban metropolitan areas or along major roads, while agricultural areas were less affected by the restrictions.
Overall, the manuscript attempts to achieve various objectives. Regarding policy implications and socio-economic repercussions, the manuscript intends to offer a statistically based analytical assessment of the effects generated by the events being mentioned earlier. Data-driven analytical techniques for event impact evaluation free the discussion from subjectivism that can distort reality and introduce objective criteria based on scientific evidence to validate or refute public policy choices in the environmental, social and economic spheres.
On a methodological and theoretical plane, the thesis presents several time series methodologies that can be used together, or as alternatives, to estimate the impact of events on air quality in an undistorted fashion.
A key point of this thesis is that robust results with respect to the quality of the data (e.g. missing data, validated sources), with respect to the intrinsic issues of the environmental phenomenon (e.g. identifying the exogenous phenomena driving the physicochemical evolution of concentrations) and with respect to the methodologies employed (e.g. which effects estimation technique is selected) can be achieved using a chain or pipeline approach. In fact, by construction, each statistical technique solves only one piece of the problem, and therefore the integration of multiple approaches is necessary to achieve a common goal. At each step, it is necessary to use appropriate techniques. In the case of environmental policy impact assessment, for example, it is first necessary to identify the intrinsic characteristics of the phenomenon, i.e. to understand whether the phenomenon evolves following some specific pattern (e.g. seasonality or stationarity); then it is necessary to understand whether there are exogenous phenomena that significantly explain its evolution (e.g. concentrations vary as temperatures vary), and finally, the effects of the policy are estimated. The most striking example of this philosophy is presented in Chapter 3, in which the new Milan limited traffic zone, i.e. Area B, is presented. We discuss and develop an algorithm, presented as a pipeline chain of distinct statistical methodologies, covering all the steps of the impact analysis: from the selection of key environmental factors to explain pollutant concentrations (model selection) to the estimation of the policy effects.
To represent and analyse the time series of airborne pollutant concentrations, we propose three different classes of models: regression models with ARIMA errors (regARIMA), autoregressive moving average with exogenous covariates models (ARMAX) and unobserved (or latent) components models (UCM). These time series methodologies are based on radically different, and often competing, statistical backgrounds. We want to highlight that the same phenomenon can be addressed from various perspectives and that the use of different methodologies can lead to heterogeneous results and conclusions.

The different underlying background mainly lies in the fact that each technique represents the data, the time series in this situation, in somewhat different ways. The regression models with ARIMA errors used in the case study on the stationarity of oxide concentrations (Chapter 6) or for the events-studies (Chapter 5), assume that the concentrations can be statistically fully explained by exogenous covariates correlated with the response and that only the residual part evolves as a stochastic process to be estimated. Similarly, the ARMAX models introduced to estimate the effect of lockdown for COVID-19 (Chapter 4) model the dependent variable as a linear function of both some exogenous covariates (X component) and its intrinsic temporal dynamics (ARMA component). Hence the name autoregressive moving average models with exogenous covariates. The latent component models introduced to study the policy impact evaluation of Area B in Milan (Chapter 3) represent the time series by decomposing it into the sum of observable components (regressors) and unobservable but estimable components. The real power of this class of techniques is that, by analysing pollutant data, the latent components can model the physical characteristics of atmospheric concentrations, such as daily or annual seasonality, natural trends and estimate the effect of exogenous socio-economic and environmental factors.
The computational issue also represents a breakpoint between methodologies. Techniques based on autoregressive representations of time series, such as ARMAX or regARIMA, allow the estimation of unknown parameters using closed forms and theoretical results widely established in the literature. This reduces the computation time required to obtain the results and stabilises the output regardless of the software or algorithms used. In contrast, techniques such as the Kalman filter for UCM models rely on iterative estimation algorithms that depend on both the initial values and optimisation techniques. Fortunately, the data available to us do not suffer from any particular anomalies or sparsity and are of a volume that medium-power workstations can handle. Although it is not explicitely carried out in this manuscript, we recognise the importance of a computational comparison between the proposed algorithms to compare the numerical stability of the solutions, the sensitivity of the outputs and their time-consumption.
A further point of discussion covered in this thesis concerns the intrinsic characteristics of the time series of concentrations and influences the statistical techniques involved. In particular, we are interested in analysing the stationarity property of pollutant time series. While presenting the environmental data for Lombardy (Chapter 2), we show how the pollutant time series are strongly positive asymmetric and heavily fluctuate, generating high disturbance noise. In the presence of data asymmetry, classical inferential techniques fail by construction. Various examples of constructing confidence intervals for non-Gaussian data based on re-sampling techniques, such as bootstrapping, or methods based on robust indicators, can be found in the literature to counteract these issues. Non-stationarity, on the other hand, has a direct effect on the definition of time series models used to represent the data. The ARMAX and regARIMA models require by construction that the time structure of the data be established a priori, the autoregressive component of which is based on the very concept of stationarity. The incorrect identification of the stationarity (or non-stationarity) of the series would distort the estimation of the model parameters and compromise the results' validity and interpretation. From this discussion emerges the need to validate with robust techniques, such as linear filtering methods, to multiple time series in order to check their properties of stationarity and cointegration in the context of highly-noisy and non-Gaussian time series. The main theoretical and empirical results of these issues are discussed in Chapter 6.
Finally, we aim at importing and applying to air quality data some statistical and econometric methodologies originated in the world of financial statistics and having the purpose of identifying a level shift in multiple time series due to unexpected events. The key contribution of this part of the project is to open up new streams of research in environmental statistics that

take their inspiration from other fields and fit into the framework of interest. In particular, in Chapter 5, we show that methods known as event-studies methods can support statistical impact assessment by constructing indicators that can determine whether an event has generated a significant or non-significant effect on multiple time series of interest. This effect may concern both shifts in the average value of the series and changes in their variability. In the environmental sphere, event studies techniques can become a powerful complementary and preliminary tool for policy impact assessments, providing valuable suggestions on which hypotheses to test and which analytical tools to use to obtain robust estimates. Moreover, the proposed indicators explicitly consider the natural cross-sectional (spatial) dependence that characterises historical air quality series obtained through monitoring networks. The theoretical and empirical results presented here clearly show how even the slightest cross-correlation between the series can severely distort the results obtained, leading to misleading interpretations and thus to harmful policies for society and the environment.

1.2.1 Organisation of the manuscript

The thesis is organised as follows. Chapter 2 proposes a robust overview of the European and Italian air quality monitoring systems, namely EIONET, and goes into rigorous detail on the current merger situation in Europe. In particular, the chapter analyses the case of Lombardy explaining its main geographical and physical characteristics, the composition of the economic and social network and the consequent effects on air pollution. The source of the data is also described in detail and will be examined in the following chapters. Air quality and weather data come from the Regional Agency for the Protection of the Environment (ARPA Lombardia), which manages and maintains the monitoring network and analyses the data collected. The chapter also discusses some issues related to advanced techniques to use weather and air quality data properly and lays the foundations for possible extensions.
Chapter 3 concerns the analysis of air quality time series using a purely-time series approach based on Unobservable Component Models (UCM) in State Space form and bootstrap resampling algorithms. The bootstrap use in a time series context makes it possible to correctly identify any variations in the average levels of the concentrations, taking into account the characteristics of positive asymmetry (right skewness) typical of pollutant concentrations. The use of state space models is well known and recognised in the statistical and econometric literature, which consider them adequate tools for the contemporary study of time series characteristics and the evaluation of the impact of policies and shocks on the variables of interest. The case study presented concerns the introduction of a new restricted traffic zone in the city of Milan, called Area B, and its possible effect on concentrations of pollutants. One of the analysis's main contributions is the introduction of a multi-stage model selection algorithm based on both fitting and forecasting performances criteria. The algorithm allows selecting the most relevant observable and non-observable features to explain the pollution concentrations before estimating the policy's unbiased effect.
In chapter 4, the analysis of a recent case study of great social and economic importance is proposed. We investigate the short-term impact of the restrictions on productive activities and mobility imposed by the Italian government, called in jargon lockdown, on the observed concentrations in Lombardy during the COVID-19 pandemic in 2020. Pollution concentrations are modelled using a class of time series techniques named autoregressive with exogenous covariate models or ARX models. This flexible and simple approach to time series modelling allows estimating the impact of covariates on a response variable, which is described by an autoregressive process. Model tuning and variable selection are performed by a LASSO algorithm which accounts for both seasonal and non-seasonal components. The lockdown is defined by

a dummy variable which expresses the change in average levels of pollution due to the restrictive measures after being depurated by the effect of local meteorology. This study is part of the recent literature investigating the relationship between virus containment policies and environmental and social effects in Europe and actively contributes to highlighting the positive effects generated by this critical situation.
Chapter 5 presents and discusses the application of the so-called event studies methods to the time series of airborne concentrations. Event studies are analytical tools typically applied in financial and economic contexts and can assess the effects of exogenous shocks on one ore more variables of interest. Event studies methods test changes in observed levels of a series comparing a pre-event situation with the post-event situation. The comparison is performed using both parametric and non-parametric statistics. In particular, we will focus on the case of cross-sectionally correlated multivariate time series. Indeed, this feature can severely compromise classical statistics' properties even for moderate values of cross-sectional correlation among the data. The contribution to the literature is twofold: on the one hand, it broadens the scope of the above methodology to non-typical areas, such as air quality statistics; on the other hand, it presents and discusses two novel rank-based non-parametric statistics able to deal with strongly correlated series. The main statistical properties and performances of both statistics are compared to a set of other competing statistics available in the literature. Finally, all the considered test statistics are applied to air quality data for Lombardy using as the event of interest the introduction of mobility restrictions (lockdown) imposed to contrast the spread of COVID-19 virus, previously presented in Chapter 4.
Chapter 6 contributes to the literature on the stationarity of univariate time series and cointegration in multivariate contexts by investigating the effect of linear filters on the respective tests' critical values and the subsequent decisions to reject or accept the hypotheses. Under the null hypothesis, it is assumed that the series are non-stationary and buried by a high noise level. This immediately leads to the following implications: regarding the stationarity, the classic ADF test will reject the non-stationarity hypothesis much more often than is tolerated by the theoretical values (i.e. large increase in test's size) and it will accept the stationarity hypothesis excessively (loss of test's power); for what concerns the cointegration, since the existence of high noise overstates the stationarity of the series, the Johansen's cointegration test will overestimate the cointegration rank, concluding that the number of cointegration relations is higher than the effective number (large increase in test's size). The linear filters considered are the moving average, the Kalman filter and the Kalman smoother. Given their noisy and volatile nature, we propose to test stationarity and cointegration on filtered and unfiltered data of observed nitrogen oxides concentrations at Milan between 2011 and 2019. The largest part of the statistical literature on air quality agrees about the airborne pollutants series's stationarity without producing sensitivity analyses or tests robust at extreme values. Using a 5-year moving window for parameter estimation and filtering, it is possible to identify possible modifications in the structure of NO_2 and NO_x concentrations and test whether linear filters' application leads to different conclusions from tests applied to raw data.
Chapter 7 concludes the manuscript, summarising the main contributions and emphasising future research developments.

Chapter 2

The Italian environmental monitoring systems and the case of ARPA Lombardia

To face the issue of environmental governance, worldwide countries have adopted organizations and institutions at supra-national, national and local levels for air quality monitoring and analysis, environmental protection and planning of public policies toward sustainable development. The Italian institutional order belongs to a broad and complex international framework, in which the coordination of hierarchical institutions is fundamental. Since Italy is an EU member, it is primarily involved in the European system of environmental protection.

2.1 The Italian and European environmental protection systems

The Italian environmental protection system, translated as *Sistema Nazionale per la Protezione Ambientale (SNPA)*, is a national institution whose competences and objectives range from the protection and monitoring of environmental quality to local supervision and controls up to the scientific research. It is composed of the *Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA)* and a local environmental protection agency for each region Agenzie Regionali Protezione Ambiente or ARPA. In the Italian case, the system counts 19 ordinary agencies, an agency for the Bolzano autonomous province and another for Trento autonomous province.
SNPA activities are strictly connected to the political and administrative system and provide to decision-makers useful tools for policy actions, both in the short-run and in the long-run, and technical-scientific support. The system was established in 2016 through the law n. 132/2016 (Government, 2016), complementing the constitutive process of the individual ARPAs, begun in the 90s, and inserting Italy in the European Union context environmental protection.
The main activities implemented by the *SNPA* concern (https://www.snpambiente.it/chi-siamo/):

- inspection activities and environmental control;
- monitoring of the state of the national environment;
- control of water sources and pollution emissions;
- technical-scientific support to the activities of state, regional and local bodies that have active administration tasks in the environmental field;
- collection, organisation and dissemination of environmental data, reports, research and statistical analysis which constitute an official technical reference to be used for the purposes of the public administration activities.

Moreover, it expresses its binding opinion on the Government and local authorities' provisions regarding the environment. It indicates the opportunity for interventions, including legislative ones, to pursue sustainable development objectives, the reduction of land consumption, safeguarding and promoting the quality of the environment and the protection of natural resources.
The last listed activity consists in collecting and providing data on the Italian environmental situation, both at local and national levels, through the environmental informative system (*Sistema Informativo Nazionale Ambientale or SINA*), locally managed by each territorial ARPA. The Italian *SINA* is part of the *European Environment Information and Observation Network (EIONet)*, a European partnership network managed by the *European Environmental Agency (EEA)* aiming at uniforming, coordinating and spreading the national informative systems. *EIONet* consists of the EEA itself, eight European Topic Centres (*ETCs*) and a network of around 1500 experts from over 400 national environment agencies and other bodies dealing with environmental information (European Environmental Agency, 2019b).
In the *EIONet* context, each national informative system is referred as *National Focal Point (NFP)*, and it cooperates mutually with all the other members and the central agency. The mutual collaboration is also carried out through national research centres and international consortia dealing with specific environmental topics and contracted by the EEA to perform specific tasks of its work program. ISPRA plays the role of *National Reference Centre* (NRC) for Italy. Figure 2.1 shows the subjects involved in the circular informative system.

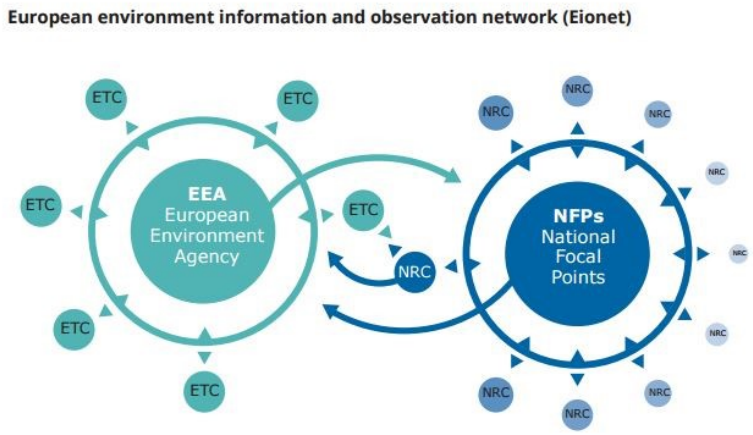


FIGURE 2.1: EIONet circular scheme. *ETCs* are the European Topic Centres, while *NRCs* are the National Reference Centres
Source: EEA-EIONet website

Figure 2.2 gives a representation of the *EEA* data flow process. Data are collected locally (Step 1), then they are aggregated at a regional level (Step 2) and national level (Step 3). Aggregated data are then transmitted to the European Environmental Agency, which harmonises the data and drafts reports and analyses to other international organisations.

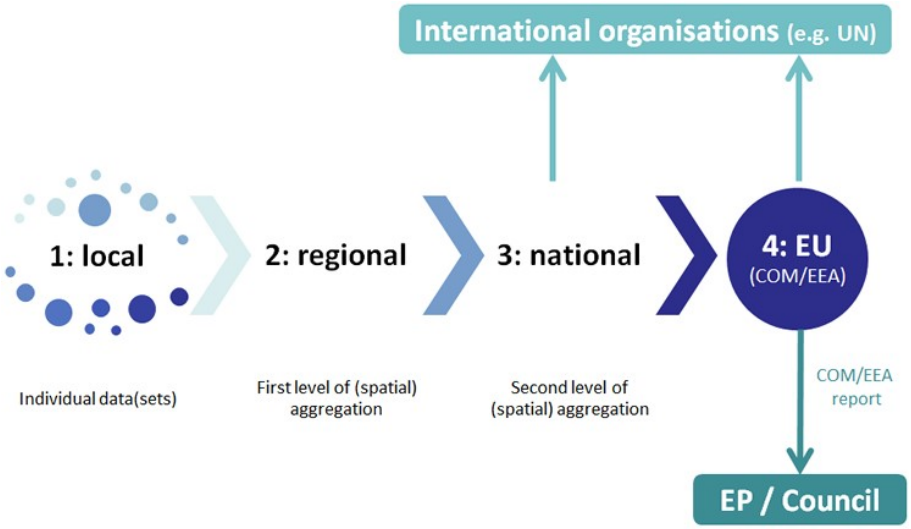


FIGURE 2.2: EEA data flow chart. *Source: European Commission website*

The *EEA* is an agency of the European Union whose task is to provide sound, independent information on the environment in Europe. It aims to support sustainable development by achieving significant and measurable improvement in Europe's environment by providing timely, targeted, relevant and reliable information to policymakers and the public (European Environmental Agency, 2019b).
The Agency was officially established with the Regulation (EC) n. 401/2009 of the European Parliament (European Union, 2009), jointly with the *EIONet*, and it is located in Copenhagen; its activities and operations came into force in 1993.
Being part of the *EEA* structure, the *EIONet* gathers information by all the EU member countries, Iceland, Norway, Balkans countries (i.e. Bosnia-Herzegovina, Serbia, Montenegro, Kosovo, North Macedonia and Albania, and Turkey. Following its withdrawal from the EU, the United Kingdom ceased on 31 January 2020 to be part of the EEA's institutional networks and governance. Country members and cooperating countries are reported in Figure 2.3.

FIGURE 2.3: EEA country members and cooperating countries. 1 February 2020.
Source: EEA-EIONet website

2.2 An overview on airborne pollution

2.2.1 Macro-and-micro pollutants

With the term macro-pollutant, we refer to airborne pollutants easily measurable with automatic systems, largely diffused in the atmosphere and present the highest concentrations. The monitoring of the quantities of macro-pollutants present in the atmosphere has a double meaning: from a *micro* or *local* perspective, it allows the characterisation of the state of atmospheric air quality and its composition at a local level, also taking into account most regulatory requirements; from another *macro* point of view, i.e. at global level, it allows the verification of the general state of the atmosphere and the control of climate change and global warming phenomena. About the former purpose, the pollutants of interest are sulphur dioxide (SO_2), nitrogen oxides (NO_X), carbon monoxide (CO) and ammonia (NH_3) and particulate matters (PM_{10} and $PM_{2.5}$), while about the latter can be interesting to monitor some greenhouse gases, such as carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O).
Other pollutants, which remain in the air in much lower concentrations and for this reason are called *micro-pollutants*, have a higher ascertained harmfulness, and their quantitative determination is very complex. Micro-pollutants important for harmfulness or diffusion are the formaldehyde, nitrous acid (HNO_2) and some hydrocarbons, such as benzene (C_6H_6) or benzo(a)-pyrene ($C_{20}H_{12}$). Also, the fine particulate matters, which present a considerable difficulty in providing a quantitative determination of the respirable fine fraction, belong to the latter category, whether they have an aerodynamic diameter of less than 10 μm (PM_{10}) or one less than 2.5 μm ($PM_{2.5}$).

Primary and secondary sources

The main airborne pollutants can be divided into two groups by source of emission: *primary* and *secondary* pollutants. The former are released into the atmosphere directly from sources, whether anthropogenic or natural, while the latter form in the atmosphere subsequently due to chemical or physical reactions involving other species of pollutants, both primary and secondary. Primary pollutants tend to persist in the form they were emitted into the atmosphere, although they will undergo chemical transformations over time. Secondary airborne pollutants can participate, shortly after their emission, in chemical reactions that lead to the formation in the atmosphere of new pollutants under certain conditions.
Table 2.1 reports for the most dangerous airborne pollutant the main emission sources and its classification as a primary or secondary pollutant.

TABLE 2.1: Main emission sources by type of pollutant

Pollutant	Classification	Anthropogenic sources
PM_{10}	Both	Combustion and mechanical actions (erosion, friction, etc.), chemical-physical processes that occur in the atmosphere starting from precursors even in the gaseous phase.
$PM_{2.5}$		
O_3	Secondary	There are no significant anthropogenic emission sources into the atmosphere.
NO_2	Both	Heating systems, motor vehicle traffic, power plants, industrial activities (combustion processes).
CO	Primary	Road traffic.
SO_2	Primary	Heating systems, power plants, combustion of organic products of fossil origin containing sulfur (diesel, coal, fuel oils).

Source: ARPA Lombardia, 2018

2.2.2 A brief description of the main airborne pollutants

We now provide a brief description of the main pollutants, highlighting the fundamental characteristics and sources.

Oxides

Nitrogen monoxide (NO) is a colourless, odourless and tasteless gas, also called nitric oxide, while nitrogen dioxide (NO_2) comes in the form of a reddish gas with a strong and pungent smell. NO is formed in all combustion processes in the presence of air by reaction of nitrogen with atmospheric oxygen, especially in conditions of high temperature. When the NO thus reacts with the atmospheric oxygen (O2) it transforms into nitrogen dioxide (NO_2). NO_2 is considered a secondary pollutant, even if small quantities of this gas are formed during the combustion process itself. The main artificial sources of NO, and therefore of NO_2, are heating systems, industrial processes, and motor vehicles' exhaust gases. Among the sources of natural origin, there is the action of lightning, fires and volcanic emissions. However, the decrease of these pollutants is extremely rapid as soon as the cause of their production disappears.

Particulate matters

By atmospheric particulate, we mean all the particles present in the atmosphere, which remain suspended in the air for very long periods thanks to their tiny size. The particles are classified

according to the aerodynamic diameter, and the unit of measurement conventionally used is the micrometre (μm), i.e. the fine fraction of the particulate. The conventional classification divides them into PM_{10}, which includes all particles of molecular size up to 10 micrometres in diameter and $PM_{2.5}$, which includes all particles up to 2.5 micrometres in diameter.
The origin of these particles is both natural (sands, dust from volcanic eruptions, pollen and spores) and anthropogenic, generally produced by combustion and industrial processes. Those produced by combustion (car engines, cigarettes, candles, heating, fireplaces, etc.) are mainly below the micrometre and therefore are the most dangerous. The particles can also carry heavy metals, i.e. lead, cadmium, zinc, and hydrocarbons. The main sources of particulate matter in indoor environments are combustion systems and tobacco smoke. Other secondary sources are the sprays, fumes from cooked food, bacteria, spores and pollen.

Ozone

When oxygen occurs in the form of a triatomic molecule, it is called ozone (O_3). In the contemporary ecological debate, ozone appears in a dual role: as "good" ozone, naturally present in the upper layers of the troposphere (10-15 km of altitude), acting as a filter for the ultraviolet component of solar radiation, which is highly harmful to live cells. This is the ozone we are talking about about the problem of the thinning of the ozone layer in the polar regions (ozone hole); on the other hand, the ozone in the breathable air (lower layers of the atmosphere, in contact with the ground) is an index of pollution of the air itself ("bad" ozone), and we refer to it when speaking of ozone as a pollutant. The "bad" ozone is generated, starting from solar radiation's action on the nitrogen dioxide molecules present in the atmosphere. Ozone is a typical secondary pollutant, characteristic of the spring and summer months with the highest irradiance. It reaches the highest concentrations on the outskirts of urban polluted areas where O_3 molecules, produced during sunny hours in urban environments polluted and transported by local breezes, they encounter purer air and can accumulate, as no pollutant reacts with them, destroying them. Indoor pollution sources of ozone can derive from air purifiers, ultraviolet lamps, photocopiers or laser printers. In the absence of specific sources, the normal conditions of ventilation of buildings the main source of indoor ozone remains the outside air.
Ozone is also a greenhouse gas; that is, it can significantly change the earth-atmosphere system's balance, producing global warming of the atmosphere. Its percentage contribution to global warming was estimated at 8%, compared to 50% of carbon dioxide, 20% of chlorofluorocarbons, 16% of methane and 6% of nitrous oxide.

2.2.3 Main effect of air pollutant exposure to human health

Many factors can influence outdoor air pollutant concentrations. These include the quantity of air pollutants released by sources, the distance from the sources, and meteorological conditions such as air temperature, the stability of the air, wind speed and direction. Some air pollutants can be carried by the wind and affect the air quality in locations hundreds to thousands of kilometres away from the sources.
Atmospheric particulate matter and oxides have a significant environmental impact on climate, visibility, contamination of water and soil, buildings, and all living beings' health. Both the short-term and long-term health effects of exposure to air pollution are well established in the literature.
In fact, the finer the particle, the more it can penetrate deeply into the respiratory system and settle. The main health effects of particulate exposure are:

- increases in premature mortality due to cardio-respiratory diseases and lung cancer;

- increases in hospital admissions and urgent visits for respiratory problems;
- chronic diseases related to the respiratory system.

For example, oxide and ozone emissions are responsible for irritating the eyes' mucous membranes and damage to the respiratory tract and lung function such as chronic bronchitis, asthma, and pulmonary emphysema. Even at low concentrations, long exposures cause a drastic decrease in lung defences with a consequent increase in the risk of respiratory tract diseases. Instead, the harmfulness to human health of fine particles depends on both the chemical composition and the size of the particles: those with a diameter between 5-10 μm reach the trachea and bronchi, while those with a diameter of less than 5 μm can penetrate as far as the pulmonary alveoli, causing carcinogenic effects for long-run exposures. Finally, the atmospheric particulates produce the exposed surfaces' degradation and reduce visibility (haze that we often see in the air, and we generally call smog). On a large scale, they produce changes in the climate.
The subjects most at risk are children and people already suffering from respiratory diseases (asthmatics) and those residing near roads with high traffic density due to long-term exposure. According to the 2013 Global Burden of Disease study (Factors, 2015), outdoor air pollution is placed at the fifth place among the deadliest risks for human health, and it kills about 5.5 million worldwide annually. This fact is further confirmed by studies using the Environmental Burden of Disease (EBD), i.e. measures that allow quantifying the effect of environmental factors on human health expressing diverging health effects in disability-adjusted life years unit. Hänninen et al., 2014 discussed the results of a study developed in 6 European countries to rank the most dangerous environmental risk factors for health. Their findings report that particulate matter exposure is the worst factor, followed by second-hand smoke and radon exposure.

The case of Lombardy

The literature concerning the evaluation of the impacts of exposure to atmospheric pollutants has given considerable space to the case of Lombardy and, more in general, to Northern Italy. Many epidemiological studies conducted using Lombardy ARPA data provided strong evidence of the association between poor air quality exposure and short-and-log term effects on health, particularly the association with mortality. Recent studies performed meta-analyses to assess how much particulate matter and oxides affect people's health in many Lombardy locations, including Milan's urban areas, starting from the early 2000s. In Baccini et al., 2011 effects have been evaluated in terms of numbers of attributable deaths under some counterfactual scenarios of air pollution reduction based on WHO guidelines and European Union limits. The authors found that current annual PM_{10} levels are responsible for over than 13% attributable death rates per year using WHO limits. By reducing existing concentrations by a fifth, it could be possible to reduce by more than 30% the burden of short-term deaths linked to ambient air pollution.
Other further factors to be considered studying the Lombardy case are the transport and commuting among municipalities. In fact, mobility in Lombardy is predominant, with the highest percentage in Italy of residents commuting daily to workplaces or schools, about 5 million incoming and out-going in 2013, half of them out of their residence municipality. Baccini et al., 2015 considered commuting data to estimate the very-short increase in mortality, i.e. within 2 days from the exposure, due to PM_{10} in Lombardy. Their findings suggest that the largest estimated impacts are placed along the Po River basin and in the largest cities. Commuting contributes to the spatial distribution of the estimated impact.
In Carugno et al., 2016 the authors estimated the number of attributable deaths and hospital admission at the areal level through Poisson regression, showing that both NO_2 and PM_{10}

affect people differently based on the season and age of the patient and that to an increase of $10\mu g/m^3$ in the average airborne levels is associated an increase in death rate up to 1.6%. Moreover, Carugno et al., 2017 showed that although the average levels of particulate matter in Lombardy are progressively reducing, in all the most affected areas the population-weighted exposure levels decreased but never met the WHO threshold.

2.2.4 International air quality standards

As the main legal instrument for air quality monitoring and guiding countries' behaviour towards environmental protection, international institutions have established some quantitative limits on pollutant concentrations. The limits in force on the European territory, which take the name of air quality standards, have been formalised by the European Union through two EU Ambient Air Quality Directives (European Union, 2004; European Union, 2008) and by the World Health Organization through the air quality guidelines (WHO, 2000; WHO, 2006). In particular, the 50/2008 European Air Quality Directive (AQD) (European Union, 2008) requires EU Member States to design appropriate air quality plans for zones where the air quality does not comply with the AQD limit values. The EU and WHO limits for the main airborne pollutant are here reported in Tables 2.2 and 2.3 respectively.

TABLE 2.2: Air quality standards given in the EU Ambient Air Quality Directives

Pollutant	Averaging period	Limit value	Comments
PM_{10}	daily mean yearly mean	50 $\mu g/m^3$ 40 $\mu g/m^3$	Not to be exceeded on more than 35 days per year
$PM_{2.5}$	yearly mean	25 $\mu g/m^3$	Maximum exposure concentration* 20 $\mu g/m^3$
O_3	max daily 8-hour mean hourly mean	120 $\mu g/m^3$ 180 $\mu g/m^3$	Not to be exceeded on more than 25 times per year
NO_2	hourly mean yearly mean	200 $\mu g/m^3$ 40 $\mu g/m^3$	Not to be exceeded on more than 18 times per year
SO_2	hourly mean daily mean	350 $\mu g/m^3$ 125 $\mu g/m^3$	Not to be exceeded on more than 24 times per year (hourly) and 3 times per year (daily)
CO	max daily 8-hour mean	10 mg/m^3	

Source: Air Quality in Europe 2019, EEA report
Note on exposure to $PM_{2.5}$: the values are determined at national level and are based on the average exposure indicator (AEI). The exposure is determined as a 3-year running annual mean $PM_{2.5}$ concentration averaged over selected monitoring stations in urban background locations to best assess the $PM_{2.5}$ exposure of the general population.

TABLE 2.3: World Health Organization air quality guidelines

Pollutant	Averaging period	Limit value
PM_{10}	daily mean	50 $\mu g/m^3$
	yearly mean	20 $\mu g/m^3$
$PM_{2.5}$	daily mean	25 $\mu g/m^3$
	yearly mean	10 $\mu g/m^3$
O_3	max daily 8-hour mean	100 $\mu g/m^3$
NO_2	hourly mean	200 $\mu g/m^3$
	yearly mean	40 $\mu g/m^3$

Source: Air Quality in Europe 2019, EEA report

The term *averaging period* refers to the period of time during which the collected data is used to calculate the reported value. In contrast, *limit value* is the atmospheric concentrations established based on scientific knowledge to avoid, prevent or reduce harmful effects on human health and the environment. Limit values can be computed using different averaging periods: *daily mean* when the reported value is the average of the hourly values between 01.00 and 23.00 for which at least 75% of the values are not-missing, *annual mean* when it is the average of the daily values between 1 January and 31 December of the calendar year and *max daily 8-hours mean* if the value is determined as the maximum concerning 8-hour running averages, calculated based on hourly data and updated every hour.

2.3 Air quality in Europe

Air quality situation for Europe is constantly updated by the European Environmental Agency (EEA), which present a yearly report reviewing the progress made towards meeting the air quality standards established in the EU Ambient Air Quality Directives and towards the World Health Organization (WHO) air quality guidelines (AQGs). The report also assesses progress towards the long-term objectives of achieving air pollution levels that do not lead to unacceptable harm to human health and the environment, as presented in the European Union, 2002 and European Union, 2013 environment action programmes.
Data presented in the present section have been extracted by the latest available report, edited in 2020 and concerning the air quality in Europe from 2000 to 2017, namely, European Environmental Agency, 2019a.
Over the past decade, air quality has slowly improved in many of Europe's cities due to more robust air quality policies across various governance levels, the introduction of targeted measures and actions, and technological improvements that have reduced emissions from various sources. These reductions were achieved through various means, including implementing regulations, non-regulatory instruments, and technological improvements for transportation vehicles and industrial processes. The adoption of more environmentally sustainable practices by consumers and industry, such as using public transit and carpooling, and optimising production processes to reduce energy use, have also contributed to the decrease. Nevertheless, many cities and regions still experience excesses of the regulated limits for air pollutants (European Environmental Agency, 2018). Figure 2.4 shows the temporal evolution of the main macro-pollutants within EEA country members and clearly states the decreasing trends for all of them. Extreme are the reductions associated with sulphur dioxide (SO_2) and nitrogen oxides (NO_X), which halve emissions compared to 1990.

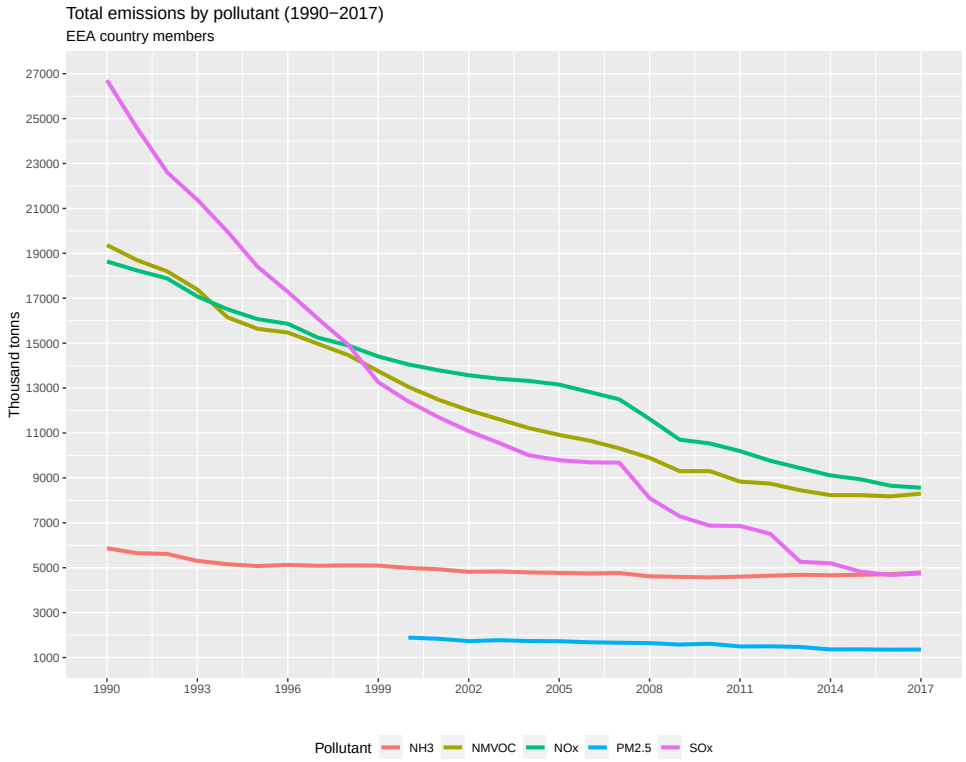


FIGURE 2.4: Total emissions in EEA country members since 1990
Source: EEA website

As mentioned above, pollutants have very different origins and derive from many sources of anthropogenic emissions. As shown in Figure 2.5, the air of European cities is strongly negatively affected by road traffic emissions and industrial production processes. These sectors significantly influence CO, particulate matters, and oxides emissions, accounting for more than half of total emissions.
Figure 2.6 reports the historical evolution of the urban population's share exposed to excessive pollutant limits computed by the EEA since 2000. The percentage is calculated as the share of urban population exposed to airborne pollutant concentrations exceeding the limit value set by EU legislation. All the pollutants considered show a constant decreasing trend, making it possible to reach minimum levels between 2015 and 2017.
The percentage is calculated as the share of urban population exposed to airborne pollutant concentrations exceeding the limit value set by EU legislation. For $PM_{2.5}$ the EEA considers the population exposed to annual concentrations above $25\mu g/m^3$, for PM_{10} the population exposed to daily concentrations exceeding $50\mu g/m^3$ for more than 35 days a year, for ozone the population exposed to maximum daily 8-hour mean O3 concentrations exceeding $120\mu g/m^3$ for more than 25 days a year and for NO_2 the population exposed to annual concentrations above $40\mu g/m^3$.

2.3.1 Brief insight on particular matter (PM) dynamic in Europe

Concentrations of particulate matter (PM) continued to exceed the EU limit values, and the WHO AQGs in large parts of Europe in 2017. In particular, for PM with a diameter of 10 μm or less (PM_{10}), concentrations above the EU daily limit value were registered at above 20% of the reporting stations in 23 over of the EIONet country members. For $PM_{2.5}$, concentrations

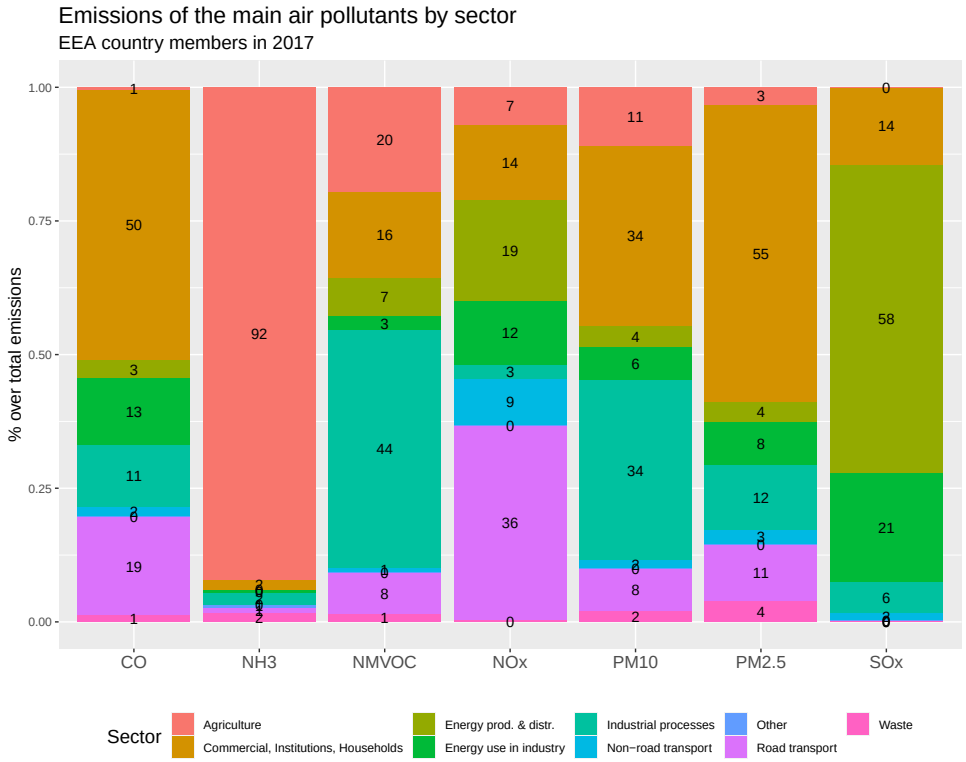


FIGURE 2.5: Distribution of total emissions by sectors in EEA country members in 2017
Source: EEA website

above the annual limit value were registered at 7% of the reporting stations in 10 reporting countries. The long-term WHO AQG for PM10 was exceeded at 51% of the stations, whereas the long-term WHO AQG for PM2.5 was exceeded at 69% of the stations located in almost all the reporting countries.
As presented in Figure 2.5, emissions from the commercial, institutional and households sectors account for over half of the current primary $PM_{2.5}$ emissions for the EEA region. Within this sector, emissions are almost exclusively from households, over 95%, and current emissions are 7% lower than those in 1990. Emissions from road transport account for approximately 11% of total $PM_{2.5}$ emissions in EEA country members, but account for about 20% of the reduction in total emissions since 1990. This reflects the improved emission control technologies that have been introduced, particularly for diesel vehicles.
A total of 17 % of the EU-28 urban population was exposed to PM_{10} levels above the daily limit value and 44 % was exposed to concentrations exceeding.

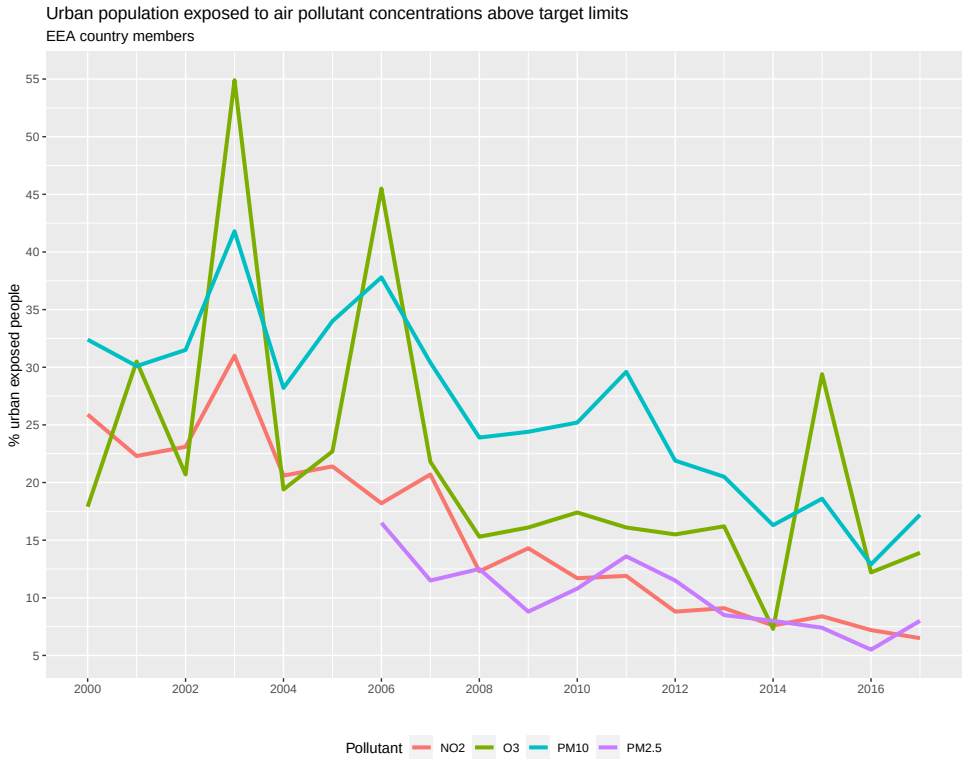


FIGURE 2.6: Urban population exposed to air pollutant concentrations above target limits from 2000 to 2017.
Source: EEA webiste in 2020

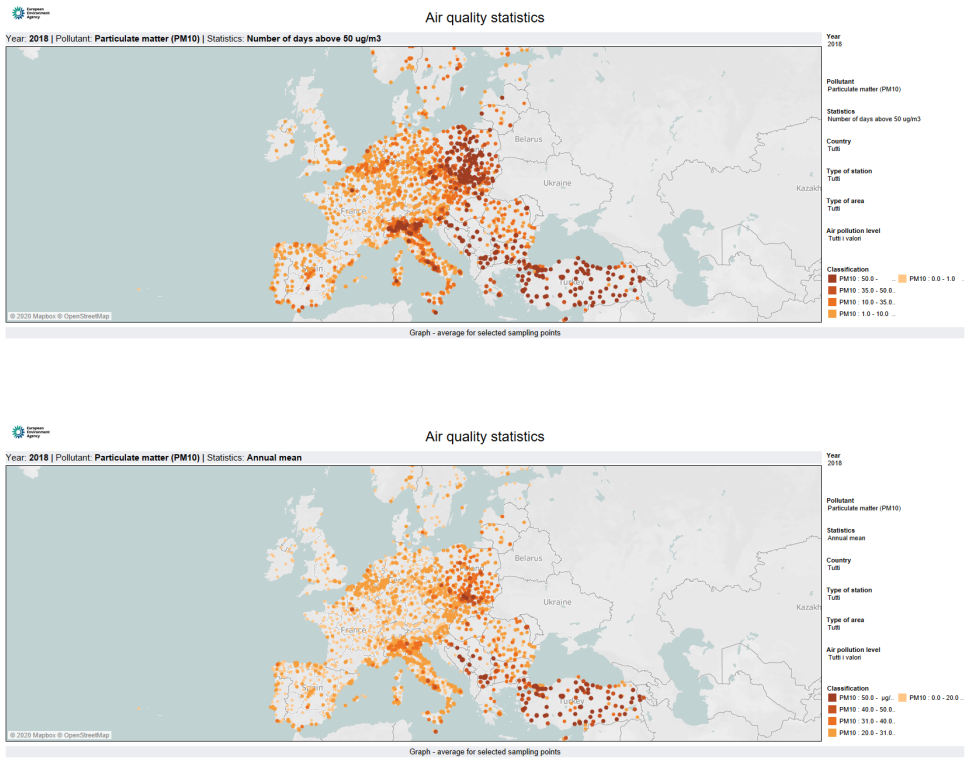


FIGURE 2.7: Number of days in which PM_{10} excedeed the the limit value of 50 $\mu g/m^3$ (top panel) and average concentrations of PM_{10} ($\mu g/m^3$) in 2018.
Source: EEA website

Figure 2.7 gives a graphical representation of the above situation: the concentrations of particular matters remain in the atmosphere for long periods and maintain much higher levels, both in terms of time and quantity, than the limits set. The Italian Po Valley situation and many areas of Eastern European countries and Turkey seem to be particularly serious; in 2018 the concentrations in those areas exceeded the limits for over 50 days, with daily average values close to 50 $\mu g/m^3$.

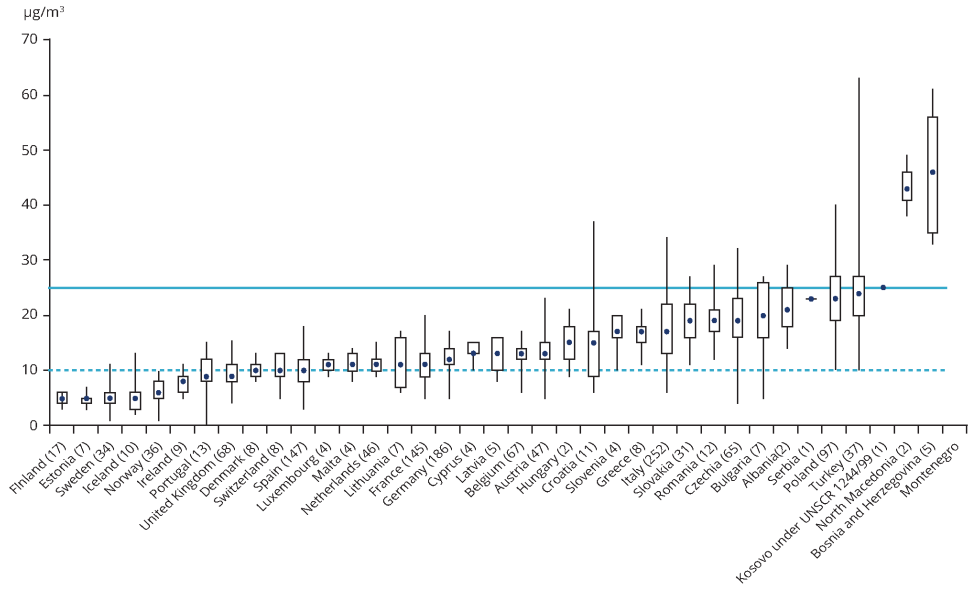


FIGURE 2.8: Average concentrations ($\mu g/m^3$) of $PM_{2.5}$ in 2017 for all the EEA country members. Rectangles are box-plots with upper limit and lower limit corresponding to the maximum and minimum concentration observed, respectively. Values in parenthesis are the number of stations considered by country. The limit value set by EU legislation is marked by the horizontal line.
Source: EEA website

Figure 2.8 shows the distribution of $PM_{2.5}$ concentrations by country in 2017. The plot highlight how EEA country member has very different and heterogeneous pollution levels. Some countries, in particular the Balkans and Eastern European countries, reach concentrations greater than 2 or 3 times those of the Scandinavian or Central European countries. Italy is characterised by a median level lower than the EEA limit (fixed at 25 $\mu g/m^3$ for $PM_{2.5}$), but with high variability that leads the country to exceed frequently the maximum limit value set by EU legislation.

2.3.2 Brief insight on nitrogen oxides (NO_X) dynamic in Europe

Concentrations above the annual limit value for nitrogen dioxide (NO2) are still widely registered across Europe, even if concentrations and exposures continue to decrease. In 2017, around 10% of all the reporting stations recorded concentrations above both EU and WHO standards. These stations were located in 20 reporting countries, 16 belonging to EU-28 and 4 non-EU members. In total, 86% of concentrations above this limit value was observed at traffic stations.

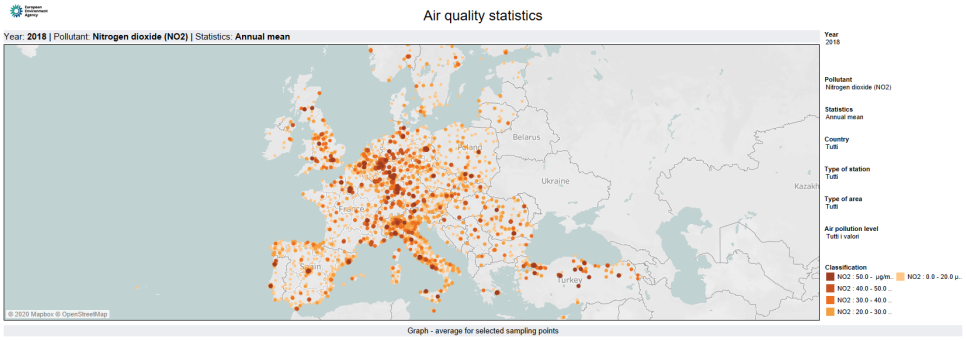


FIGURE 2.9: Average concentrations of NO_2 ($\mu g/m^3$) in 2018.
Source: EEA website

The NO_2 dynamics just described tells a very similar story to particular matters and summarised by Figure 2.9. In fact, the oxides have very high values in vast areas of Europe, particularly Italy, in the industrial belt between Germany, France and Benelux and at specific points near major cities.
Emissions from road transport and non-road transport combine to account for around half of the current NO_X emissions in the EEA members. Since 1990, there have been considerable reductions in NO_X in the road transport sector, despite the general increase in transport activity within this sector over the period. This sector alone has contributed to over 40% of the total reduction in nitrogen oxides emissions. Emissions of NO_X have also declined in the energy production and distribution sector, and current emissions are at approximately 40% of the emission level of 1990. This dynamic is well illustrated in the two previous graphs.
In 2017 around 7% of the EU-28 urban population was exposed to concentrations above the annual EU limit value for NO_2. This fact is a successful milestone and represents the lowest value since 2000. Table 2.4 reports the EU and WHO reference concentrations and the percentage of the urban population in the EU-28 exposed to air pollutant concentrations in 2015-2017 by type of air pollutant.

TABLE 2.4: Percentage of the urban population in the EU-28 exposed to air pollutant concentrations above EU and WHO reference concentrations (minimum and maximum observed between 2015 and 2017)

Pollutant	EU reference value	WHO reference value	Urban population exposure
PM_{10}	50 $\mu g/m^3$ per day	20 $\mu g/m^3$ per year	13%-19%
$PM_{2.5}$	25 $\mu g/m^3$ per day	10 $\mu g/m^3$ per year	6%-8%
O_3	120 $\mu g/m^3$ each 8 hours	100 $\mu g/m^3$ each 8 hours	12%-29%
NO_2	40 $\mu g/m^3$ per year	40 $\mu g/m^3$ per year	7%-8%

Source: Air Quality in Europe 2019, EEA report

Figure 2.10 shows the distribution of NO_2 concentrations for all the EEA country members. As other important countries, such as France and Spain, Italy reports values in line with the largest part of European countries but sometimes presents very high and abnormal values in terms of nitrogen dioxide concentrations. The upper tail of the distribution (over the 75th percentile), in fact, is characterised by extreme values well above the median and the rest of the distribution.

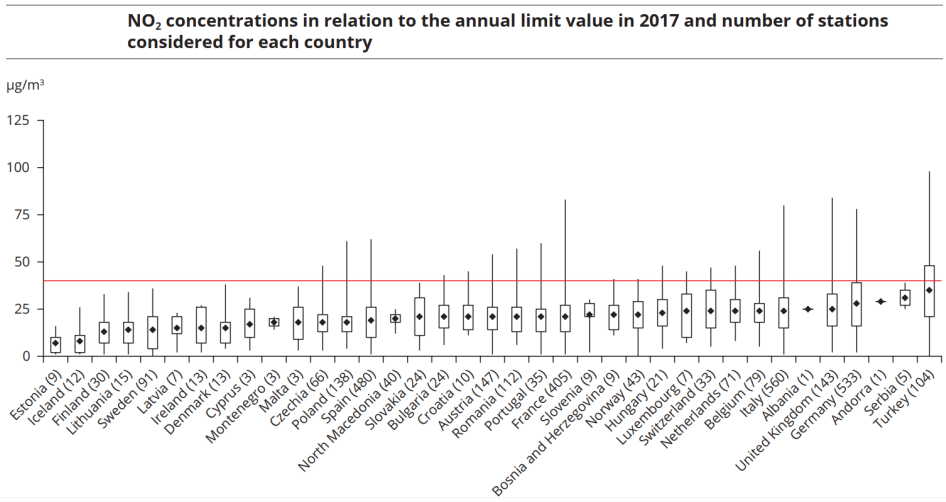


FIGURE 2.10: Average concentrations ($\mu g/m^3$) of NO_2 in 2017 for all the EEA country members. Rectangles are box-plots with an upper limit and lower limit corresponding to the maximum and minimum concentration observed, respectively. Values in parenthesis are the number of stations considered by country. The horizontal line marks the limit value set by EU legislation.
Source: EEA Air quality in Europe 2019

All the information just discussed, and the graphs raise some important critical issues in the civil society and among European institutions regarding the quality of the air we breathe and its management. The concentration and emission levels of the main pollutants harmful to health, first of all, particulate matters, are very critical and require interventions for drastic reductions.

2.3.3 The new European Air Quality Index

In November 2017 the European Environmental Agency launched a new overall air quality index for EEA member and reporting countries in order to provide to the civil society an easy and intuitive tool to evaluate and compare the air we breathe. The new *European Air Quality Index* (EAQI) reflects the potential impact of air quality on health, driven by the pollutant for which concentrations are poorest due to associated health impacts. The Index indicates the short-term air quality situation. Hence it does not reflect the long-term (annual) air quality situation, which may differ significantly. EAQI substituted the previously existing *Common Air Quality Index* (CAQI) established in 2006.
The index is accessible through an interactive map that shows the local air quality situation at station level, based on five key pollutants that harm people's health and the environment: particulate matter ($PM_{2.5}$ and PM_{10}), ground-level ozone (O_3), nitrogen dioxide (NO_2) and sulphur dioxide (SO_2). The index is computed hourly assembling information gathered from more than 2000 air quality monitoring stations across Europe and accounting for special procedures according to the type of monitoring stations, i.e. traffic or other types.

Pollutant	Index level (based on pollutant concentrations in µg/m3)					
	Good	Fair	Moderate	Poor	Very poor	Extremely poor
Particles less than 2.5 µm ($PM_{2.5}$)	0-10	10-20	20-25	25-50	50-75	75-800
Particles less than 10 µm (PM_{10})	0-20	20-40	40-50	50-100	100-150	150-1200
Nitrogen dioxide (NO_2)	0-40	40-90	90-120	120-230	230-340	340-1000
Ozone (O_3)	0-50	50-100	100-130	130-240	240-380	380-800
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-750	750-1250

FIGURE 2.11: EAQI composition by type of pollutant
Source: EEA AQI website

The EAQI returns the short-term air quality near each station using a 6-levels Likert scale, here reported in Figure 2.11, which depends on the type of registered airborne pollutant and their mixing in terms of concentrations. In general, when EAQI increases, public health risks increase. As shown in Table 2.5 to each level of the scale is associated with a different message to the European population and providing particular attention to the population most at risk.

TABLE 2.5: Percentage of the urban population in the EU-28 exposed to air pollutant concentrations above EU and WHO reference concentrations (minimum and maximum observed between 2015 and 2017)

AQ Index	General population	Sensitive population
Good	The air quality is good. Enjoy your usual outdoor activities.	The air quality is good. Enjoy your usual outdoor activities.
Fair	Enjoy your usual outdoor activities	Enjoy your usual outdoor activities
Moderate	Enjoy your usual outdoor activities.	Consider reducing outdoor activities, if you experience symptoms.
Poor	Consider reducing activities outdoors, if you experience symptoms such as sore eyes, a cough or sore throat	Consider reducing outdoor activities, especially if you experience symptoms.
Very poor	Consider reducing activities outdoors, if you experience symptoms such as sore eyes, a cough or sore throat	Reduce outdoor activities, especially if you experience symptoms.
Extremely poor	Reduce physical activities outdoors.	Avoid physical activities outdoors.

Source: EEA Air Quality Index website

European Air Quality Index is the European response to other air quality indices developed

by international government agencies to communicate how polluted the air currently is or how polluted it is forecast to become. Other famous indices have been developed by Canada (Air Quality Health Index), US (US Air Quality Index) and Malaysia (Air Pollution Index). In particular, the air quality index developed by the U.S. Environmental Protection Agency (Agency, 2018) is used as the main reference for always developing more sophisticated index adapted to specific situations, such as indoor (Kang and Hwang, 2016) or thermal comfort (Saad et al., 2017) monitoring.

2.4 ARPA Lombardia monitoring system

According to the *EEA-EIONet* system, national data on air pollution concentration are collected at the local level and directly managed by the regional agencies for environmental protection. Recall from Section 2.1 that the Italian monitoring network is composed by twenty-one regional agencies, named ARPAs, and by a central institution called *ISPRA*, which coordinates the regional bodies.
Being the body responsible for local environmental protection and air quality, the ARPA monitoring system controls the main meteorological and hydrogeological measurements. Each regional agency has its own territorial monitoring system consisting of fixed and mobile ground stations. Stations can gather information on air quality, on meteorology or both. Installation, management, maintenance, replacement and calibration of the stations are under the agencies' responsibility, which must also provide IT and analytical tools, e.g. databases and dashboards, necessary to spread knowledge about collected information to the public. In the following, we will also refer to *stations* with the term *control units*.
The results that will be discussed in the next chapters derive from applying statistical methodologies to data relating to the ARPA Lombardy terrestrial monitoring network exclusively collected by fixed stations. Given the intrinsic dynamism generated by the displacement of the control units, the data deriving from the mobile stations would not allow analysing uniquely, from a spatial perspective, the dynamics of air quality in Lombardy.

2.4.1 The role of ARPA Lombardia

ARPA Lombardia plays a key role in all the matters relating to the environment and climate within the region; therefore, it can carry out many tasks.
Collection of air quality data and emission inventories arrangement constitutes basic information, available to the civil society, particularly for public administration, for the private sector and scientific research. ARPA, in particular, carries out the following actions in terms of air quality:

- Design, construction and management of the air quality monitoring network;
- Application of statistical and mathematical models of dispersion of pollutants in the atmosphere for near real-time and forecast evaluations;
- Scenario studies on the effects of policy actions;
- Development and updating of the inventory of INEMAR (AIR Emissions Inventory) emissions, which makes it possible to know the sources of the main air pollutants in each Lombard municipality, broken down by combustion and activity;
- Collection, processing and dissemination of data and report;

- Participation in technical-scientific discussions on sector issues set up by the European Commission, the Ministry of the Environment, the Higher Institute for Environmental Protection and Research, the Lombardy Region and local authorities.

ARPA Lombardia also manages the hydro-meteorological service, which provides a public service of meteorological forecast and real-time hydro-meteorological monitoring: this knowledge base has numerous applications, for example, protection from natural risks, management of water resources and environmental protection. All data acquired by the monitoring networks are validated, archived and made available to the public, both in raw format and processed in analysis products (daily, weekly, monthly, yearly bulletin). Meteorologists process weather forecasts based on numerical weather modelling and observational data, and they are constantly updated to meet the user's specific needs.
Concerning the meteorology, among the others, ARPA Lombardia provides the following hydrogeological services:

- Weather monitoring and forecasting;
- Processing, archiving and dissemination of meteorological, hydrological and climatological data;
- Reporting of meteorological, hydrological and climatological analysis;
- Research and development in the context of national and international projects

Regione Lombardia open data

As in the case of the INEMAR emissions inventory, the whole database concerning atmospheric concentrations and meteorology collected by ARPA is also freely available and usable by the public. The data can be consulted and downloaded through two channels: the first is the specific dashboard prepared by ARPA Lombardia (https://www.arpalombardia.it/Pages/Aria/qualita-aria.aspx), which also provides real-time mapping of air and water quality in addition to periodic reporting; the second channel is the open database owned by Lombardy regional administration (https://www.dati.lombardia.it/), an open-source portal which contains all the territorial statistics regarding the environment, energy, local government, health, culture, productive activities and others.
The open-access database contains information about air quality and weather measures collected through the ground sensors and descriptive information on the stations: unique ID code within the ARPA monitoring system, geographical coordinates (latitude and longitude), altitude, administrative membership (province or other political aggregations) and municipality, the service start date and the eventual disposal date. This last case implies that the sensor, or the entire station, has been deactivated and no longer produces information. Full description of air quality station registry is available at the link https://www.dati.lombardia.it/Ambiente/Stazioni-qualit-dell-aria/ib47-atvt and at https://www.dati.lombardia.it/Ambiente/Stazioni-Meteorologiche/nf78-nj6b for the weather stations.
All the data used in this manuscript were directly downloaded via regional open data. In this regard, one of the innovative proposals of this paper concerns the development of a new software tool, written using the R language, aiming at downloading and managing air quality and weather data from the regional open database easily. The reason for this proposal relies on the fact that currently, open data allows the user to download only one year at a time in *xlsx* or *csv* format with dimensions that are often not easily manageable. This lead to a reduction in the public diffusion of the data to non-technical users, such as journalists or policy-makers.

Emission inventory for Lombardy

In Lombardy, the INEMAR system has been used to estimate and update emissions' regional inventory into the atmosphere. INEMAR inventory is an essential tool for air quality planning for regional, provincial and municipal authorities in Lombardy and a useful data source largely used by public and private environmental organisations. The inventory is directly managed by the Air Quality Modeling and Inventories Organizational Unit of ARPA Lombardia and is updated every 2-4 years. The latest available version refers to 2017, currently under public review.
Emission inventory for Lombardy can be freely consulted through the project website (https://www.inemar.eu/xwiki/bin/view/Inemar/WebHome), which allows an open-source distribution of raw data for researchers and all the subjects interested in the air quality topics.
ARPA makes available emission inventories up to the municipal level and classifies the emission sources according to the SNAP (Selected Nomenclature for sources of Air Pollution) activity nomenclature, which includes, as an example, non-industrial combustion plants, combustion in manufacturing industry or road transport. INEMAR is developed respecting the European guidelines drawn up in EMEP-Corinair inventory guidebook, in particular, ARPA Lombardia adopted the Corinair SNAP-97 nomenclature integrated over the years the most significant and updated activities added to the basic nomenclature. A detailed report on the structure of the atmospheric emission inventory for the Lombardy Region and on the algorithms used to compute estimations has been realised by Caserini et al., 2004.
According to the EMEP-Corinair Emission Inventory approach, The SNAP-97 system (European Environmental Agency, 2019b) is set up according to three levels:

- the upper level (11 source categories) which features sources' grouping as commonly performed;
- the intermediate level (75 source sub-categories) comprehending technological and social-economic criteria;
- the lower level (485 source activities) aiming at an exhaustive enumeration of sources and sinks to spot homogeneous sections in generating emissions.

The upper level of INEMAR classification is composed of 11 macro-sectors:

- Combustion in energy and transformation industries
- Non-industrial combustion plants
- Combustion in the manufacturing industry
- Production processes
- Extraction and distribution of fossil fuels
- Solvent and other product use
- Road Transport
- Other mobile sources and machinery
- Waste treatment and disposal
- Agriculture
- Other sources and sinks

Sectorial classification is available at aggregate, i.e. regional, and local levels, such as provinces or municipalities.
Available data refer both to the common airborne macro-pollutants (SO_2, NO_X, CO, NH_3), micro-pollutants (particulates) and the main greenhouse gases (CO_2, CH_4, N_2O).
The inventory also provides information about the emissions classified by fuel and estimates the average traffic emission factors to estimate road transport emissions. The emission factors are available for different levels of aggregation:

- by type of vehicle (cars, light vehicles, heavy vehicles and buses, mopeds and motorcycles)
- by type of road (motorways, suburban roads, urban roads)
- by fuel (petrol, diesel, LPG, methane)
- by legislative type, i.e. Euro category (from Euro 0 to Euro VI).

2.4.2 Air quality in Lombardia

For many years, Northern Italy has been at the top of the ranking of the most polluted areas in Europe (European Environmental Agency, 2018; European Environmental Agency, 2019a). The Lombardy region is the geographic and economic epicentre of this area: it is organised in 11 provinces, and it counts on more than 10 million inhabitants and the highest gross domestic product per inhabitant of the country, approximately 35 thousand Euros per capita at current prices (Regional Statistical Yearbook, 2017a; Regional Statistical Yearbook, 2017b). The region's geographical position within the European context is clearly visible in Figure 2.12.
The region can be geographically and economically divided into three zones: the mountain range of the Alps at North and the Apennine at South, the sloping foothills at the Middle-North, and the highly industrialised and populated basin of the Po River in the middle. Most of its major cities are located in the Po River basin, which crosses the entire region and acts as the Southern natural border for many kilometres. As shown in the left panel of Figure 4.2b, the Lombard basin is bordered on two sides by mountains, i.e. the southern and northern borders, which makes the atmospheric dispersion poor and which render air mass exchange very low. Wind speed measured in the Po River plain is among the lowest in Europe, about 1.5 m/s on average, causing smog and pollution trapping close to the ground. According to a recent simulation study by Raffaelli et al., 2020, if Po Valley had the same meteorological conditions typical of central–northern Europe and kept the same emission levels, average monthly concentrations of PM_{10} and NO_2 would be lowered by 60 to 70% and 60% compared to concentration levels of 2013. Peaks of 70-80% reduction would be reached in the western Po Valley. Consequently, it is clear that it is more difficult for Po Valley regions to comply with international air quality standards compared to other EU and non-EU member states.

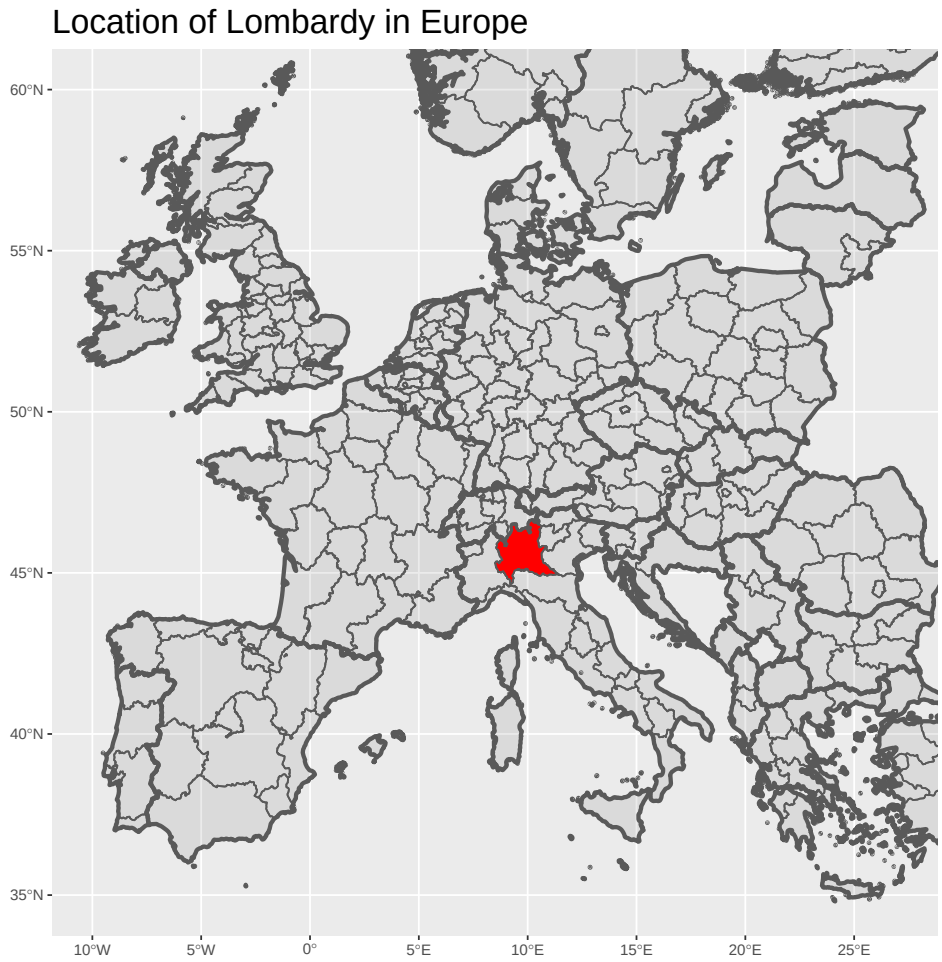
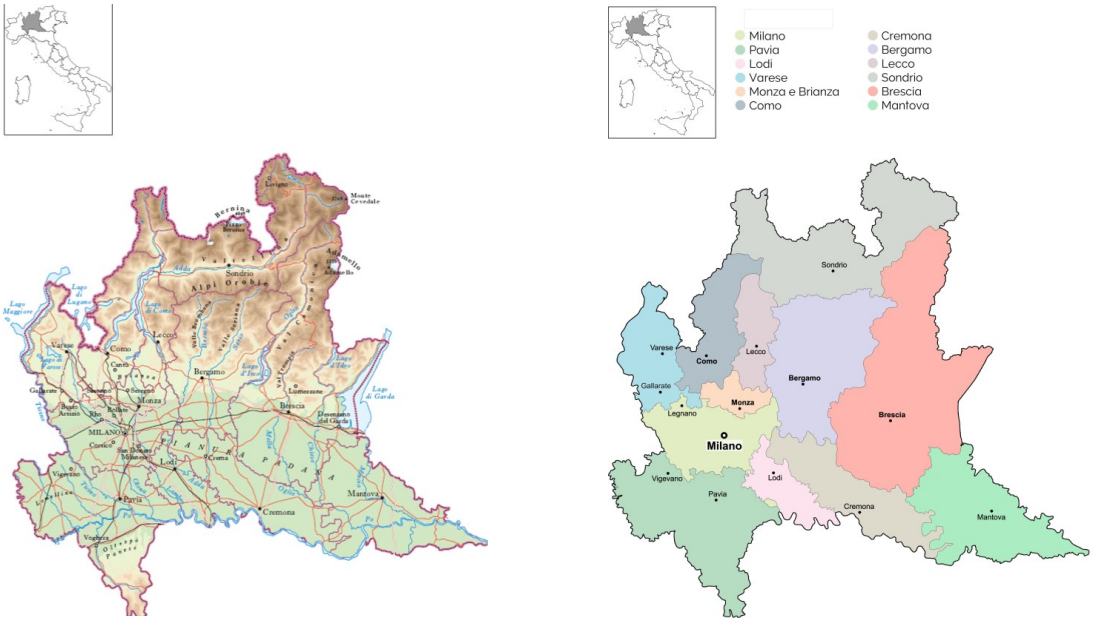


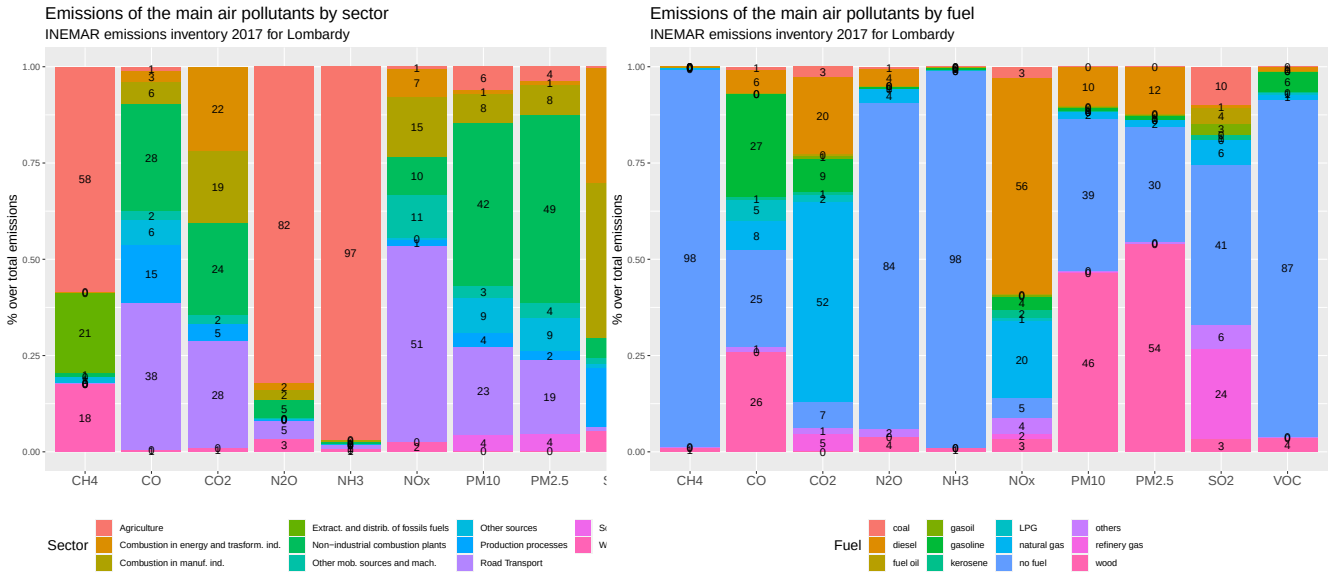
FIGURE 2.12: Location of Lombardy in Europe



(A) Physical map of Lombardy (B) Political map of Lombardy

FIGURE 2.13: Physical map (left panel) and political map (right panel) of Lombardy, Italy.
Elaboration of the author on web images.

Besides, Lombardy counts many industrial facilities and small and medium enterprises for which road transport is an essential component for economic viability. According to INEMAR Emission Inventory of Lombardy 2017 (INEMAR ARPA Lombardia Settore Aria, 2020), industrial and non-industrial combustion plants and road transport represent more than 73% of particulate matter emission sources and more than 76% of nitrogen oxides emissions in the region. Comparing Figures 2.14a and 2.5, regarding at oxides, carbon dioxide and particulate matters, road transport emissions diverge significantly from the European levels with values much higher than the EU average. In particular, in the metropolitan area of Milan, which will be the subject of a specific focus in the next chapters, road traffic alone is responsible for the emission of 65% and 69% of the total emissions of NO_X and CO respectively.



(A) Emissions by macrosectors

(B) Emissions by fuel

FIGURE 2.14: INEMAR Lombardia 2017 emissions by macrosectors and by fuel type
Source: INEMAR emission inventory for Lombardy

Coherently with their anthropogenic emissions sources, fine particulate matters are mainly generated by wood-fired heating systems (i.e. wood), while oxides are generated by chemically treated fossil fuels and petroleum derivatives (i.e. diesel and oil).
Figure 2.14b also reports that the large part of other airborne pollutant emissions, particularly for ammonia, volatile organic compounds and methane, are due to non-combustible materials. The two figures also highlight that emissions from agricultural activities in Lombardy are the major source of pollution regarding methane (CH4), nitrous oxide (N2O) and ammonia (NH3), for which they amount respectively to 58%, 87% and 97%. Regional agriculture land makes up around 69% of the total area. Lombardy is the first Italian region for agriculture production and contains 1.5 million bovines and 4 million swine, respectively, about 48% and 25% of the national headcount. The INEMAR 2017 inventory estimates that about 95% of total ammonia (NH_3) emissions in Lombardy are caused by livestock and fertilisers. Indirectly, atmospheric ammonia reacts with atmospheric nitric and sulphuric acids and turns to particulate matters, generating a consequent increase in $PM_{2.5}$ concentrations (Hristov, 2011).

The Lombardy region is also the most densely populated country, with large and very dense urban agglomerations. The average population density in Lombardy is around 419.9 inhabitants/km^2, while only 200 inhabitants/km^2 at national level (Regional Statistical Yearbook, 2017a). This fact also reflects on the spatial distribution of airborne emissions. In fact, the four largest and populated provinces, i.e. Milano (MI), Monza (MB), Bergamo (BG) and Brescia (BS), generated the 52% of total emissions of NO_x and the 51% of particulate matters in 2017.

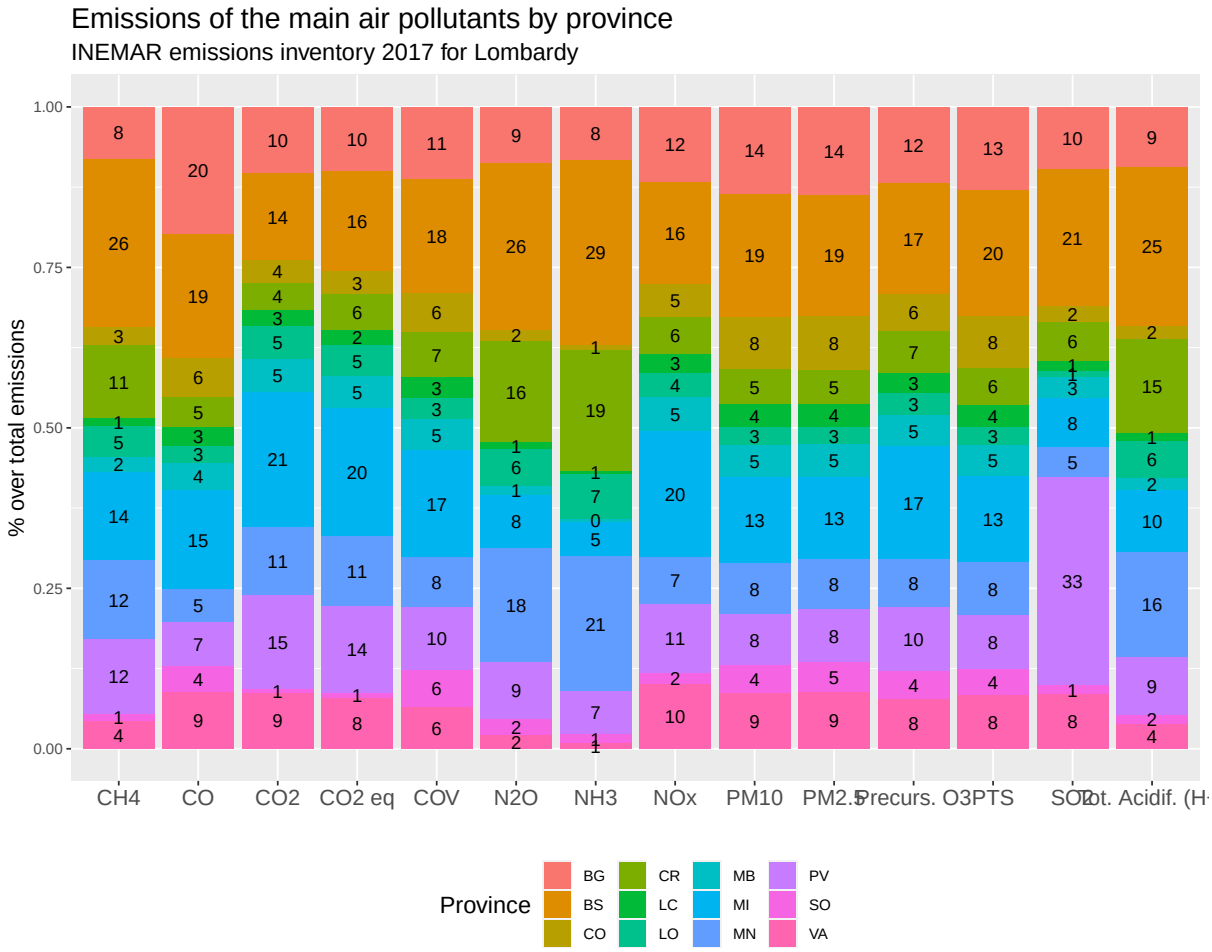


FIGURE 2.15: INEMAR Lombardia 2017 emissions by province
Source: INEMAR emission inventory for Lombardy

Overall, its unfavourable geographical context, aggressive land use, climate characteristics and emission sources create a high level of air pollution, therefore determining long favourable periods for accumulating pollutants.
Figure 2.16 shows the average weekly concentrations, reinforced by the 95% Gaussian confidence bounds, registered in Lombardy from January 2014 to November 2020. The regional average and the standard errors have been obtained by aggregating the weekly observations from all the available stations located in Lombardy during the period. The seasonal behaviour of the series is well-defined: oxide and particulate matters concentrations reach maximum peaks during the winter, while their minimum is while the minimum is achieved in the summer months; at the opposite, ozone is a summer pollutant that reaches serious levels in August and minimum levels during winters. Regarding oxide concentrations, both NO_2 and NO_x in the last 2 or 3 years showed reductions in the maximum values observed, especially at the apex of winter, supporting the hypothesis of a generalised decreasing trend. Despite this, during the

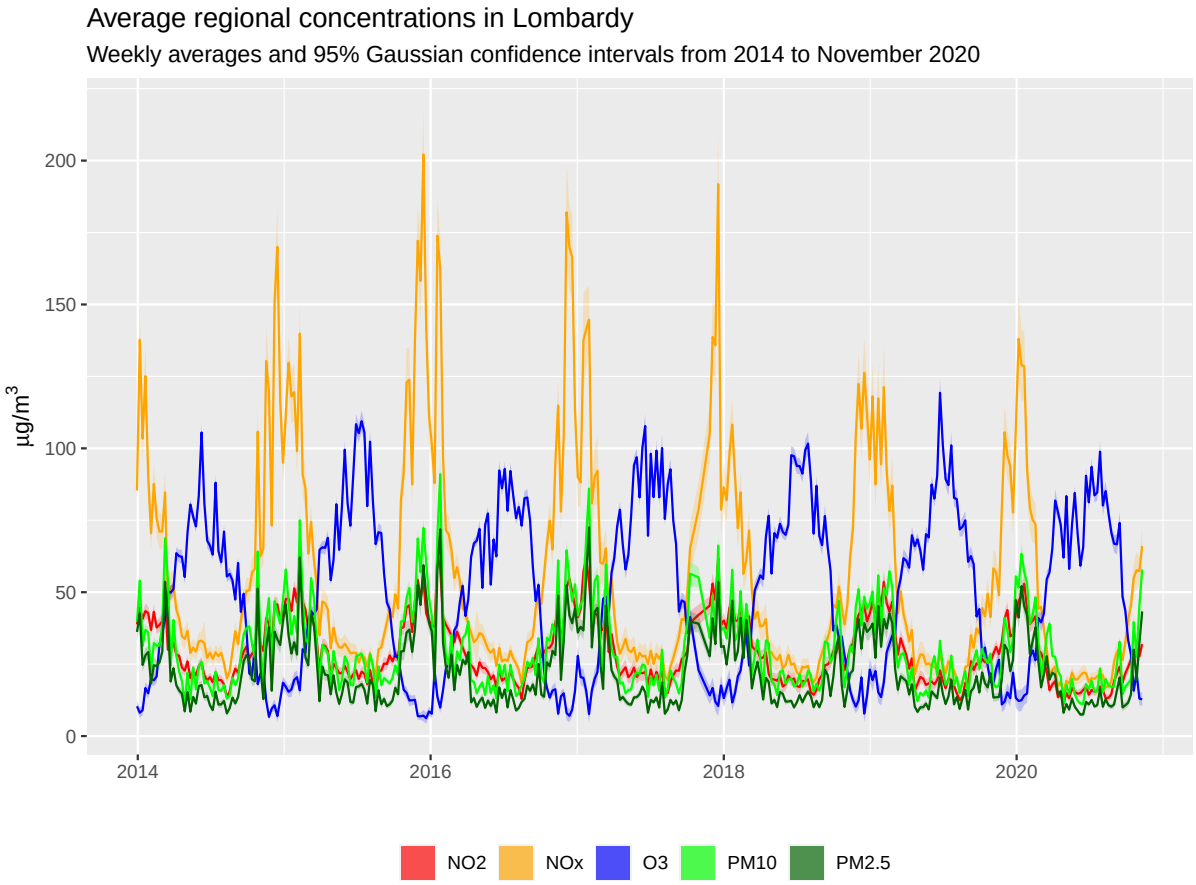


FIGURE 2.16: Average weekly concentrations of NO_2, NO_x, PM_{10}, $PM_{2.5}$, and ozone in Lombardy. Matt ribbons are the 95% Gaussian confidence bounds for the average.

winter period, nitrogen dioxide concentrations show average values above the EU and the WHO's annual limit, which amounts to 40 $\mu g/m^3$. The graph also reveals a similar pattern for particulate matters, defined by a period of very high concentrations between 2015 and 2018, and strong reductions in 2019 and 2020. The historical evolution of average concentrations is now represented according to the provinces in which they are recorded. The patterns are shown in Figure 3, which reports in four panels the weekly average concentrations for NO_2, NO_x, PM_{10} and $PM_{2.5}$. For what concerns the oxide, the provinces of Monza-Brianza (MB), Milan (MI), Sondrio (SO) and Mantova (MN) are clearly identifiable and opposite. The first two dominate the high values of the concentrations placing almost always above all the other provinces, while the last two always have the lowest values recorded. Each timestamp their observed differences of NO_2 is near 40 to 50 $\mu g/m^3$. This fact can be explained by considering the physicochemical process generating the oxides and already discussed in Section 2.2.1. As stated before oxides are emitted primarily by anthropogenic sources, such as combustion processes. Milan and Monza areas are rich in manufacturing and industrial activities, as well as being densely urbanised. Sondrio is an alpine province with only mountains, while Mantova is a wide and flat area with various rivers and agricultural lands. Both areas are sparsely inhabited. As far as PM_{10} and $PM_{2.5}$ are concerned, the graph is more complex but still provides some important suggestions. None of the provinces seems to overcome the others, but Sondrio shows the lowest concentrations again. Cremona (CR), Mantova (MN) and Pavia (PV) recorded very high quantities of airborne $PM_{2.5}$ during the winter months, and in some cases, such as

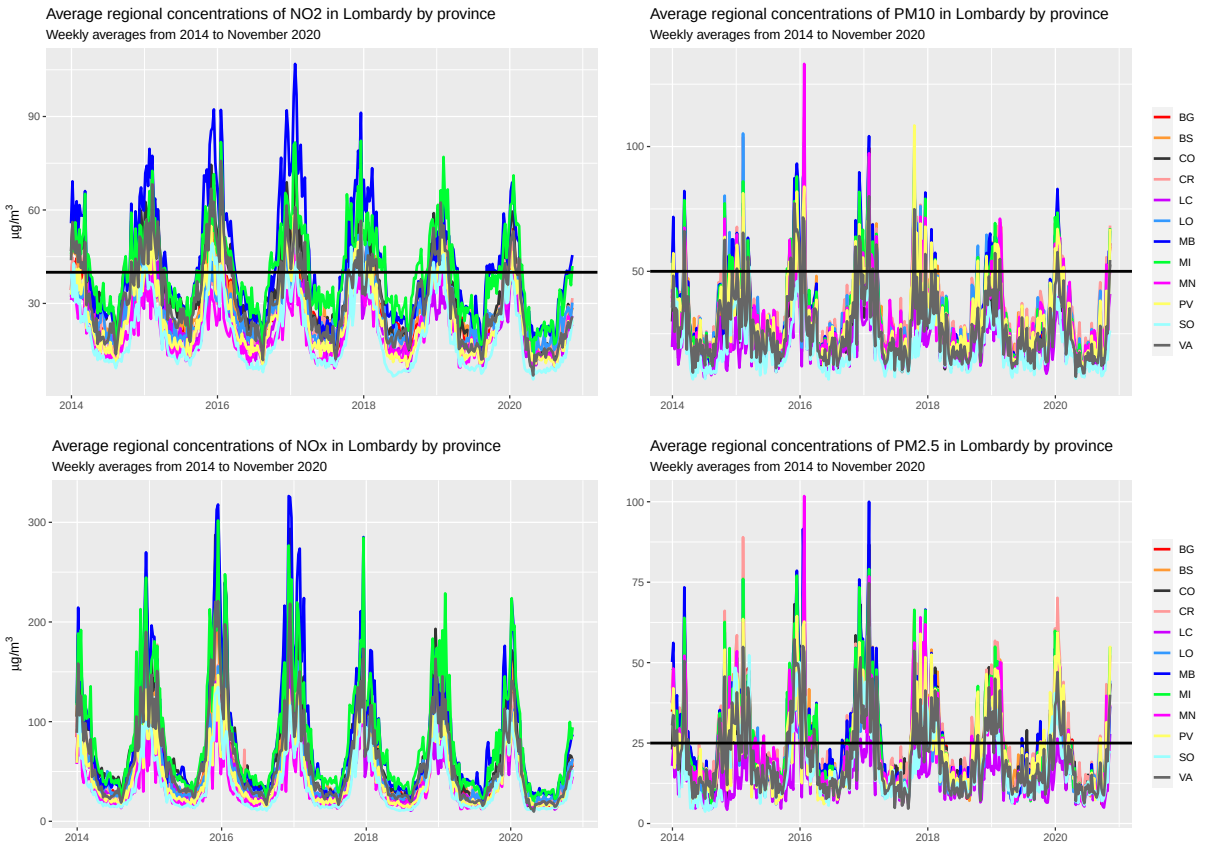


FIGURE 2.17: Average weekly concentrations of NO_2, NO_x, PM_{10}, $PM_{2.5}$ in Lombardy by province. Horizontal line for NO_2 is the WHO yearly average limit, while the horizontal lines for PM_{10} and $PM_{2.5}$ are the daily WHO limits. Upper left panel: average weekly concentrations of NO_2 by province; lower left panel: average weekly concentrations of NO_x by province; upper right panel: average weekly concentrations of PM_{10} by province; lower right panel: average weekly concentrations of $PM_{2.5}$ by province.

in 2016, they reached or overcame Milano (MI) or Monza-Brianza (MB). This can be due to the triangular relationship between agricultural activities, ammonia emissions and $PM_{2.5}$. Intensive agriculture generates high levels of ammonia, which converts into ultra-fine particulates, increasingly its concentrations.
In Section 2.4.3 these average concentrations will be further analysed considering various types of stations showing how the oxide concentrations vary greatly depending on the monitored area. The presence of a decreasing trend of oxide concentrations in Lombardy will be further investigated in the following chapters, which will also discuss the specific case about the metropolitan area of Milan (Sections 3.4.1 and 3.5.1).
As preliminary observational results, it is possible to state, in line with what reported by the European Environmental Agency, that Lombardy suffers from concentrations well above the allowed limits for almost all the pollutants considered, especially in winter periods, and needs immediate interventions on the production structure and mobility.

2.4.3 ARPA Lombardia air pollution emission monitoring system

Air quality network

At the present date, air quality detection network in Lombardy consists of 84 ground stations that use automatic analyser tools and provide continuous data at regular time intervals, generally on an hourly or daily basis, depending on the pollutant of interest. On the monitoring stations are installed 495 sensors, each of them measuring a single type of pollutant among those available. The number of sensors for each station ranges from 1 to 14, with a median value of 5 pollutants.

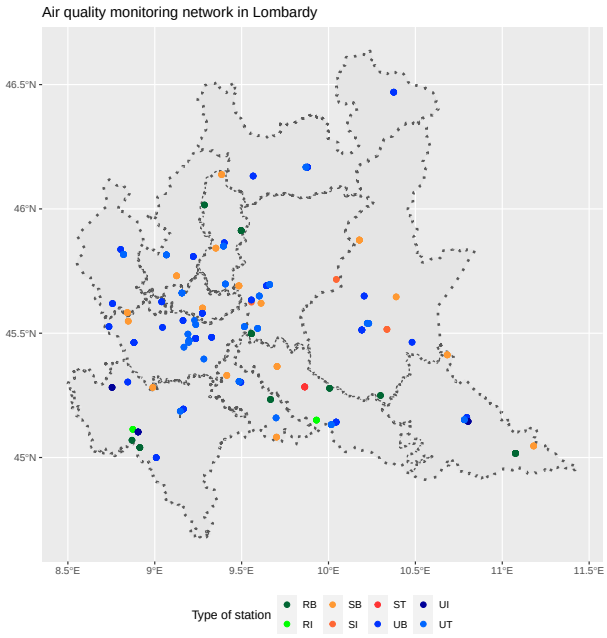


FIGURE 2.18: Regional distribution of air quality stations by typology.
Source: ARPA Lombardy

The stations are distributed throughout the regional territory according to the population density and the type of environmental context. As represented in Figure 2.21, the territorial distribution covers all the provinces and pays particular attention to the central belt, i.e. Milan, Monza, Bergamo and Brescia provinces, which corresponds to the industrial and most dense area. Stations are classified according to the type described in detail in Section 2.4.3.
With 8 mobile laboratories, ARPA also investigates local situations and specificities, following the municipal administrations' notification.

Air quality equipments

Depending on the environmental context (urban, industrial, traffic or rural) in which the monitoring is active, the type of pollutants that must be detected is different. Therefore, not all stations are equipped with the same technical instrumentation. Since the network is constantly updated, thanks to improvements in technical equipment or the replacement of obsolete components, the control units can be installed in different years or interrupt their life cycle prematurely. This is an important factor generating heterogeneity of the monitoring network and consequently of the data available. In other words, stations can measure different types of pollutants with different periods of functioning.

The monitoring stations consist of a basic structure, instruments for measuring atmospheric pollutants and equipment for displaying, processing and transmitting the measured parameters' values. All the air quality control units are equipped with a storage unit that allows to acquire the data measured by the instrumentation and provide for their transmission to the processing centre. The atmospheric pollutant analysers' main feature is to determine, automatically and continuously over 24 hours, the measurement of the substance under examination with high sensitivity, even when present in low concentrations. Technical details on the physical-chemical methodologies for detecting pollutants and on instrumentations can be retrieved from the ARPA Lombardia site, section detection criteria.

Data quality

The instrumentation installed in the air quality monitoring stations, operating 24 hours a day 365 days a year, is periodically subjected to checks and maintenance, aimed at ensuring the proper functioning of the equipment over time and the given product's reliability.
ARPA Lombardy activated the so-called "quality assurance procedures" to guarantee the data's quality and accuracy. An independent team, comparable to an audit team of the financial and corporate world, carries out specific checks of quality control and quality assurance activities to ensure that the entire network functions adequately and that the data produced throughout the stations are high quality and comparable. For example, in the *field audits*, the audit office's analysers, which represent the regional reference for a specific air pollutant, operate for a certain period in parallel with the network instrumentation. Any discrepancies found are investigated to identify the reason and intervene from a technical perspective to restore a correct measurement. The more two data are aligned, the more reliable the instrument is.

Stations classification

EEA and ARPA classify air quality control units according to the environmental context in which they are active. We can distinguish two classes of classifications: the first depending on the distribution (density) of buildings and the type of area surrounding the installation site and a second based on the predominant emission sources recorded by the monitoring station (European Environmental Agency, 2019a).
The first classification clusters the stations in urban (U), suburban (S) and rural (R) stations. Urban stations are units installed in continuously built-up urban areas, suburban stations are installed in largely built-up urban areas, while rural stations provide information on sparsely populated areas.
The second classification divides the stations in the background (B), i.e. pollution levels are representative of the average exposure of the general population or vegetation, traffic (T), i.e. stations located close to a single major road, and industrial (I), i.e. stations located near an industrial area or an industrial source.
The combination of the two classifications allows identifying the functioning of any station properly. For example, a station labelled as *UT* describes a control unit installed in a large urban centre near a highly-traffic zone; a station labelled as *RB* describes a station capturing background pollutants in a rural area.

TABLE 2.6: Classification of air quality stations by type of area and by predominant emission sources

By type of area			**By emission source**		
Classes		# Stats	Classes		# Stats
Urban	U	53	Traffic	T	25
Suburban	S	20	Background	B	49
Rural	R	11	Industrial	I	8

Figures 2.19 and 2.20 report the average weekly regional concentrations of NO_x, NO_2 and particulates classified by type of area and by emission souces. From the examination of these

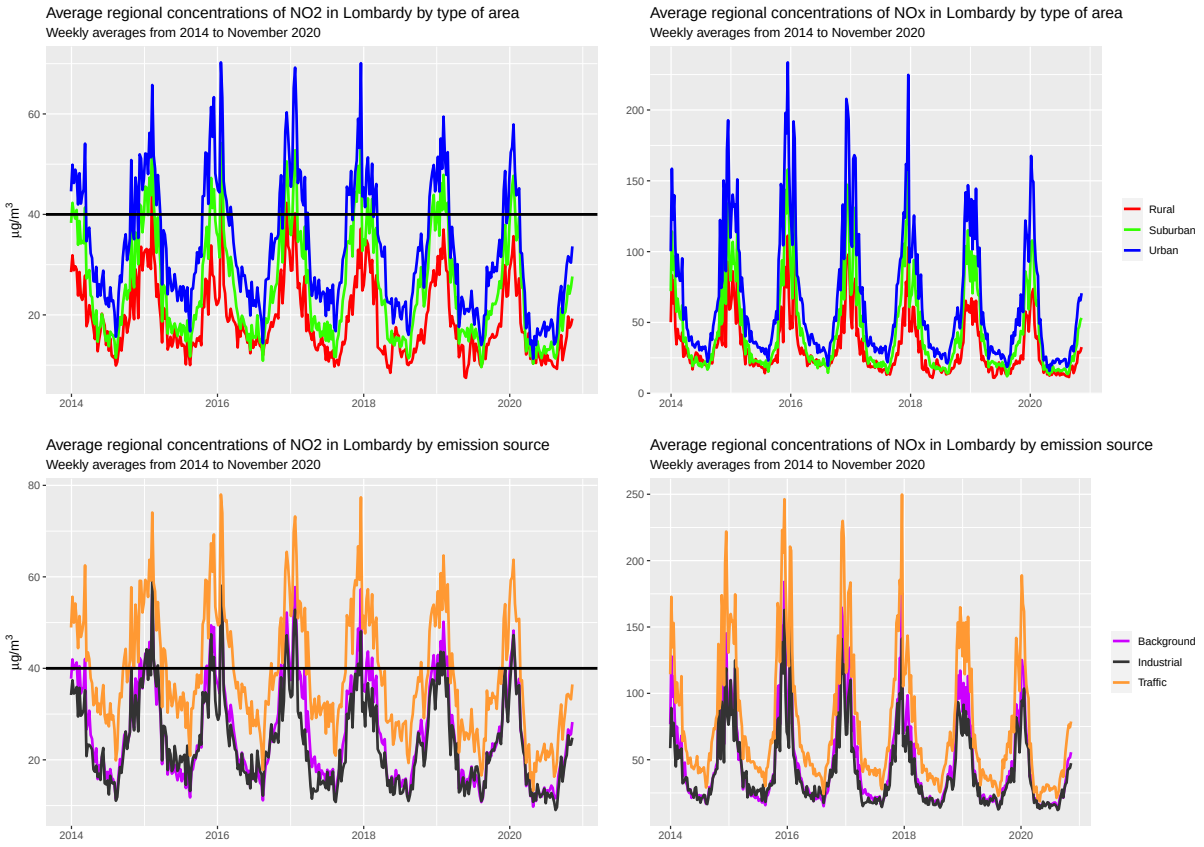


FIGURE 2.19: Average weekly concentrations of NO_2 and NO_x by type of area (rural, urban or suburban) and by type of emission source (traffic, background or industrial) in Lombardy. Horizontal lines for NO_2 are the WHO yearly average limit. Upper left panel: average weekly concentrations of NO_2 by type of area; lower left panel: average weekly concentrations of NO_2 by emission source; upper right panel: average weekly concentrations of NO_x by type of area; lower right panel: average weekly concentrations of NO_x by emission source.

figures it is clear that the average concentrations observed depend strongly on the type of station, and therefore on the territory it monitors. This is particularly true for oxides. First of all, both total oxides and nitrogen dioxide have distinct average concentrations considering the classification by area and source: urban and traffic monitoring units, which often are associated, record values well above the other typologies for the whole period considered. Their values are spring or summer they settle around 25 to 30 $\mu g/m^3$, while in the winter months

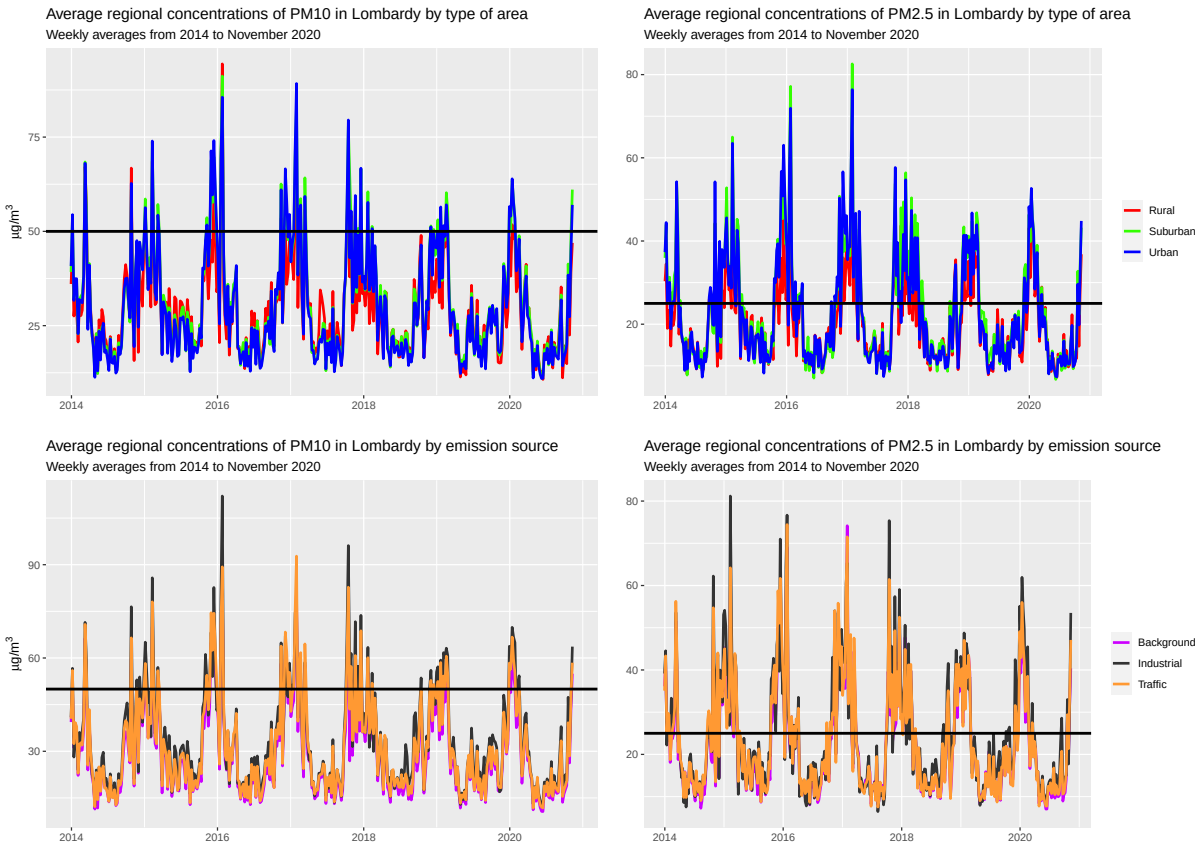


FIGURE 2.20: Average weekly concentrations of particulate matters by type of area (rural, urban or suburban) and by type of emission source (traffic, background or industrial) in Lombardy. Horizontal lines are the WHO daily average limit. Upper left panel: average weekly concentrations of PM_{10} by type of area; lower left panel: average weekly concentrations of PM_{10} by emission source; upper right panel: average weekly concentrations of $PM_{2.5}$ by type of area; lower right panel: average weekly concentrations of $PM_{2.5}$ by emission source.

they reach heights of 65 to 70 $\mu g/m^3$ (beyond the above mentioned limits). The rural stations have permanently lower values under 40 $\mu g/m^3$. Background stations are reasonably in the middle, given their function of monitoring sites not directly close to emission sources.
Instead, both fine and ultra-fine matters do not seem to be affected by station classifications and show homogeneous concentrations. Taking as reference the daily WHO limits, i.e. 50 $\mu g/m^3$ for PM_{10} and 25 $\mu g/m^3$ for $PM_{2.5}$, the concentrations seem out of control during the winter and under control during the rest of the year. The data show that from 2019, particulate matter has suffered strong reductions compared to the period 2016-2018, but there is still a large potential for improvement.

Air pollutant

The pollutant elements monitored by the ARPA stations are reported in Table 2.8. The table also contains information about the time coverage, collection frequency and the number of sensors installed for each airborne pollutants.

TABLE 2.7: Synthesis of the airborne pollutants monitored by ARPA Lombardia

Pollutant	Symbol	Frequency	Sensors	Time coverage
Total nitrogen oxides	NO_X	Hourly	88	2011-2020 (current)
Nitrogen dioxide	NO_2	Hourly	88	
Particulate matters 2.5	$PM_{2.5}$	Daily	10	
Particulate matters 10	PM_{10}	Daily	64	
Black carbon	BC	Daily	2	
Ozone	O_3	Hourly	53	
Carbon Monoxide	CO	Hourly	46	
Sulphur dioxide	SO_2	Hourly	35	
Benzene	C_6H_6	Daily	24	
Benzo(a)-pyrene	$C_{20}H_{12}$	Daily	14	
Ammonia	NH_3	Daily	10	
Cadmium	Cd	Daily	14	
Nickel	Ni	Daily	14	
Arsenic	As	Daily	14	
Lead	Pb	Daily	14	
Black carbon	BC	Daily	2	2017-2020 (current)

In the following chapters, we will focus on the statistical analysis of nitrogen oxides and particulate matters. Two reasons mainly guide the choice: the first, of a statistical nature, concerns the amount of data available, the second is instead linked to the effects that the two elements have on human health.
About the first reason, particulates and oxides data are gathered over the 52% of the total number of available sensors, which involve 83 over 84 stations. Moreover, these sensors are often those installed for the longest time, guaranteeing high temporal coverage.

2.4.4 ARPA Lombardia weather monitoring system

Parallel to the environmental and air quality monitoring network, ARPA Lombardia also manages the automatic meteorological and hydrogeological monitoring network. The weather monitoring network counts on 984 sensors installed on 249 active stations. As for the air quality stations, the weather control units are located homogeneously on the regional territory to return the most extensive spatial coverage possible. Figure 2.21 shows the spatial distribution of weather stations within the provinces.
The sensors have been installed since the late 1980s and share the maintenance and installation process and the data quality standards, described above for the air quality stations in Sections 2.4.3 and 2.4.3.
The next Table summarises weather stations' information; in particular, it provides the number of sensors for each variable and their unit of measures. All the weather variables cover a time.

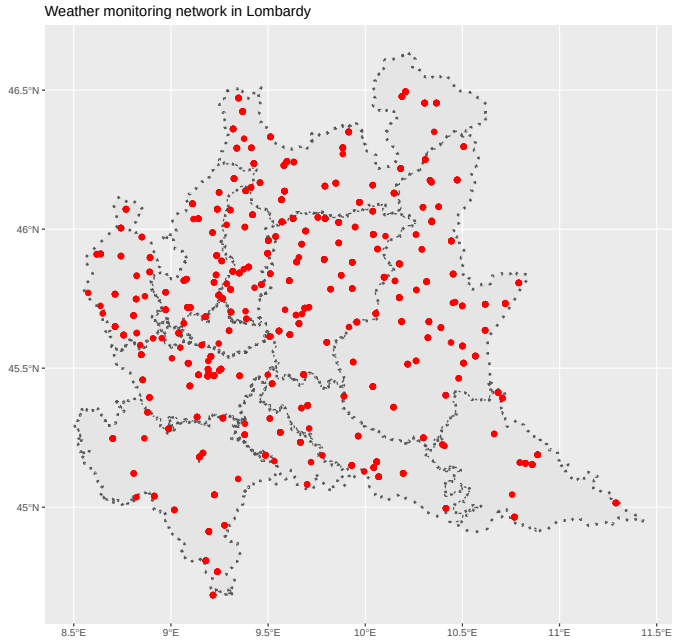


FIGURE 2.21: Regional distribution of meteorological stations in Lombardy.
Source: ARPA Lombardia

TABLE 2.8: Synthesis of the airborne pollutants monitored by ARPA Lombardia

Variable	Unit measure	Frequency	Sensors	Time coverage
Temperature	Celsius degrees (^{o})	One observation each 10 minutes	205	2011 – 2020 (current)
Rainfall	millimeters (mm)		223	
Solar global irradiance	watt per square metre (W/m^2),		94	
Relative humidity	percentage (%)		160	
Wind speed	meter per second (m/s)		126	
Wind direction	clockwise degrees (^{o}) where 90^o means wind from East		126	
Snow height	centimeters (cm)		33	
Water height	centimeters (cm)		65	

Observed values of temperature, solar irradiance, relative humidity and wind speed and direction represent the average value recorded during the 10 minutes of reference, while rainfall and snow and water height are cumulated over time.
Wind direction represents the direction of origin of blowing wind, and it is conventionally stored using North-degrees, i.e. ranging from 0^o to 360^o, where 0^o and 360^o coincide and represent a wind blowing from North. In the following chapters, wind direction and wind speed will be treated jointly summarising the two measures into a set of four new variables, each representing the average wind speed in the four quadrants of the Cartesian plane. For each timestamp, i.e. observations at a temporal resolution of 10 minutes, we have recorded the prevalent wind direction over the ten minutes and the average speed. Wind direction is divided into four blocks, each of them associated to one of the four quadrants of the Cartesian plane: the north-eastern (NE) quadrant is represented by wind directions lying between 0 and 90 degrees, the south-eastern (SE) quadrant collects directions lying 90 and 180, the south-west

(SW) quadrant from 180 to 270 and the north-western (NW) quadrant from 270 to 360. Aggregating the data into hourly or daily observations, we can construct the new variables averaging the observed wind speed in each of the four quadrants. The results are the average wind speed in NE, NW, SE and SW directions for each hour or day.
Given this manuscript's purpose, focused on air quality, we will not consider information about snow and water height, since they are not directly involved in the air pollution spread mechanism. The literature provides dozens of examples of research papers involving the cited weather variables in air quality analysis. Obviously, given the nature of the problem, temperature, rainfall, and wind are the main candidate factors for explaining the phenomenon of atmospheric concentrations and emissions.

Time coverage

All the air quality and meteorological variables presented above cover a time interval that starts from 2011 and goes up to the current days through periodic updates of the database. Years before 2011 have significant information gaps and missing data, so they will not be considered. This reasoning is obviously conditioned by the installation date of the station.

2.4.5 Meteorology in Lombardy

Lombardy is subject to a continental climate, characterised by humid hot summers and cold winters. During anticyclonic episodes, the region experiences higher atmospheric aerosol concentrations, persistent fog, and haze (Lolli et al., 2012). Meteorology is considered a key factor in monitoring air quality data, especially because ignoring the relationship between weather and pollutants would inevitably lead to biased estimates of concentrations and even to underestimates (Pernigotti et al., 2012). As discussed in Section 2.4.2, the Po river basin, particularly the Lombardy region, has an unfavourable microclimate that generates large accumulations and pockets of pollutants. The presence of the Alps at North and the Appennini at South inhibits the wind circulation heavily.
Using the available data from ARPA Lombardy open data, it is possible to obtain a general overview of the region's meteorological situation. The wind speed is deficient and does not allow air recycling and cleaning. This situation is well described in Figure 2.22, which distinguishes the region's wind.
As confirmed by other studies and reports, the average hourly wind speed registered is minimal, most of the records are well below 2 m/s or even below 1 m/s. The distribution in circular degrees of direction shows that the wind blows mainly from the North (high concentration of cases around 0 and 360) with some reduced peaks from the East (values in the 80 to 100 range). During the year it is possible to notice some heterogeneity between the seasons. For example, in the winter months (December-February) the main wind comes from the north and has a higher speed than the rest of the year, while between spring and summer the wind increases from the south and east.
Another relevant factor for air cleaning is the rainfall, here reported by Figure 2.23. The Lombardy region registers a modest daily precipitation of water, almost always below 5 mm overall. In May, August and November the highest falls are recorded.
Finally, it is possible to show also the regional trend of the temperatures. In particular, we report the average, minimum and maximum weekly temperatures recorded all over the region from 2016 to November 2020 (Figure 2.24) and the average weekly temperature by province (Figure 2.25). Another relevant factor for air cleaning is the rainfall, here reported by Figure . Another relevant factor for air cleaning is the rainfall, here reported by Figure .

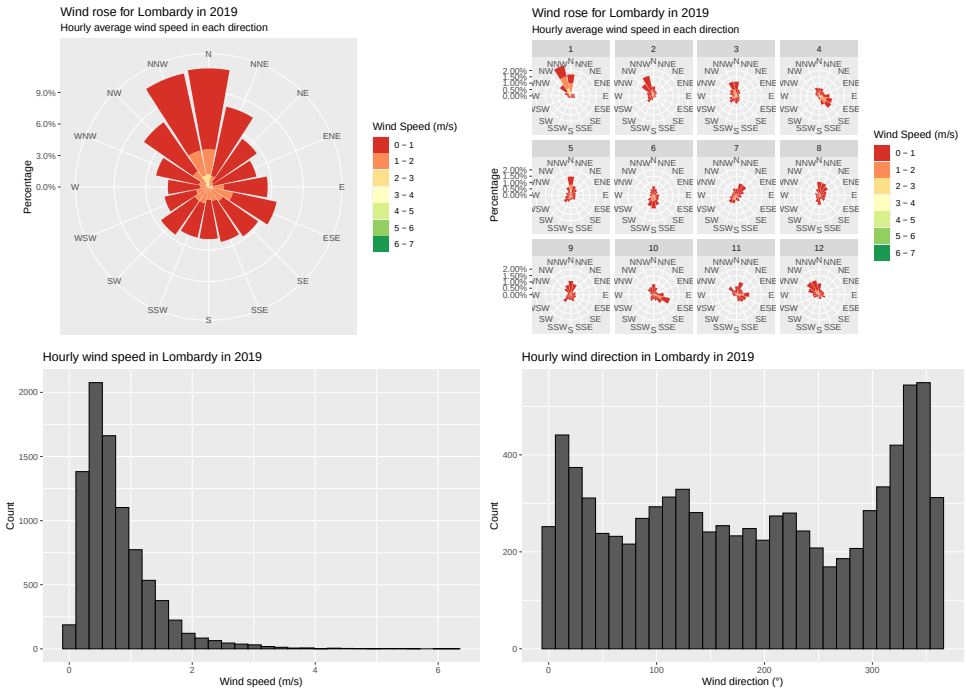


FIGURE 2.22: Main characteristics of the wind in Lombardy in 2019. Wind compass with average wind speed in each direction at regional level (upper-left panel); Wind compass with average wind speed by month (upper-right panel); histogram of the hourly wind speed (lower-left panel); histogram of the hourly wind direction (right-panel);

Both figures also highlight that in recent years, Lombardy's average temperature is increasing. This is evidenced by the evident linear growing trend overlapped with the observed temperatures. In particular, the regional average minimum temperature (blue line in Figure 2.24) is rapidly increasing. At the end of 2019, the minimum temperature stabilised around zero Celsius degrees, recording the highest minimum temperature since 2016. The increases in average observed temperature are also visible at the provincial level. All the provinces recorded higher average temperatures in 2019 and 2020 compared to the previous years. It is interesting to note that the province of Sondrio (light-blue line in Figure 2.25), located in the extreme north of the region and bordering with Switzerland, registered values around zero at the beginning of 2019, while in the same period of 2016 or 2017 the temperature was widely below zero. Many macro-scale factors, such as global warming, are clearly affecting local climatology and could lead to dramatic changes with ever faster and more serious events if not corrected in the short-term.

2.5 Side issues concerning air quality and weather data

2.5.1 Solar irradiance in regression

The use of solar irradiance in statistical models is more complex due to its physical characterisation and bi-direction relationship with pollution. In fact, it is often used as the response variable in regression models where airborne pollutant or air quality indices are used as covariates to study their statistical relationship (Yao et al., 2017; Zhang et al., 2020).

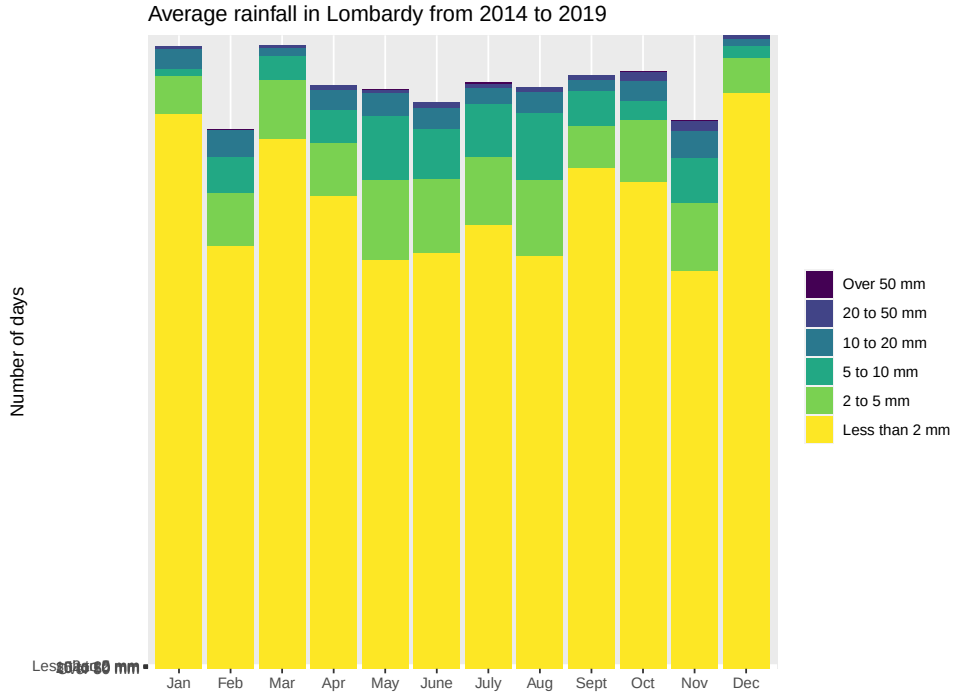


FIGURE 2.23: Number of rainy days during the month classified according to the precipitation quantity.

2.5.2 Humidity and temperature

Also, relative humidity can present challenging issue inside research designs due to its physical relationship with temperature. Humidity tells us the moisture content of the atmosphere, or how much water vapour there is in the air. The relative humidity is given as a percentage and tells us how close the air is to being saturated. If the relative humidity is 50%, the air contains half the water vapour required for it to be saturated. Relative humidity depends on air temperature: if the water vapour content stays the same and the temperature drops, the relative humidity increases. Suppose the water vapour content stays the same, and the temperature rises, the relative humidity decreases. This inversely-proportional relationship could then generate redundant information in the models, translated into statistical multicollinearity. This fact leads to consider a different specification of temperature and humidity using the concept of *humidex*.
Humidex is an index designed to quantify human exposure to the joint effect of heat and humidity during the day and which gives a measure of climatic well-being and thermal comfort of humans combining temperature and relative humidity.
As represented in Figure 2.26, Humidex scores are positive values lying mainly within the range [20,60]. According to its values, the thermal comfort can be classified using an ordinal Likert-like scale that for high Humidex values associates dangerous exposure to heat (red area) and for low values of Humidex associates comfortable human exposure (blue area). Temperature and relative humidity data can be combined to construct humidex based on the standard formula proposed by Masterton and Richardson (Masterton and Richardson, 1979) of Canada's Atmospheric Environment Service:

$$H = T + \frac{5}{9}\left(6.112 \cdot 10^{\frac{7.5T}{237.7+T}} \cdot \frac{R}{100} - 10\right) \tag{2.1}$$

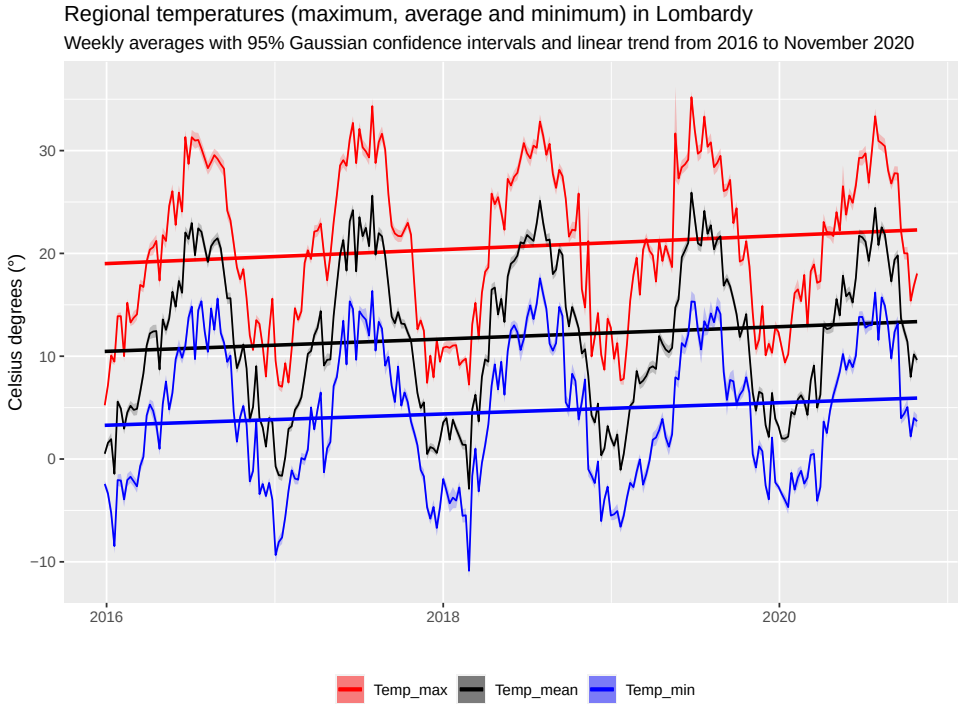


FIGURE 2.24: Average, minimum and maximum weekly temperature at regional level from 2016 to November 2020.

where H is the resulting humidex, T is the air temperature in Celsius degrees, and R is the relative humidity. The statistical relationship between humidex and air quality, both in terms of concentrations and emissions, is gaining ground in the literature, especially the engineering one related to the production of real-time monitoring systems. Ma'rufatin and Kusnoputranto, 2018, based on Indonesian data, suggest the presence of an inverse (negative) relationship between the variables and that humidex can explain differences in greenhouse gases emissions. Belyakhina et al., 2017 developed an indoor air quality monitoring and prediction system using humidex to determine the level of user's comfort.

The circular nature of wind direction measure

As stated above, wind direction measures the direction of origin of blowing, and it is conventionally represented through the so-called *Compass rose*, a representation system that assigns each degree a direction and a name to the wind.
The degree representation of wind direction opens the statistic theme linked to the observed data's circular nature. A more sophisticated approach, with respect to that one presented in section 2.4.4 to deal with circular data, uses data transformations and models that account properly for this particular shape of data. Details and references on circular data can be found in Mardia and Jupp, 2009 and Jammalamadaka and Sengupta, 2001, which give a deep overview of the main analytical tools for circular statistics.
Within a regression framework, when circular variables are used as covariates, we refer to the regression techniques as *circular-linear regression models*. These kinds of models are mainly based on non-parametric approaches grounding on the concepts of circular kernels (Di Marzio et al., 2009) and circular distribution (Von Mises, 1918). The literature offers many application of non-parametric regression techniques to the air quality-wind direction-wind speed regression

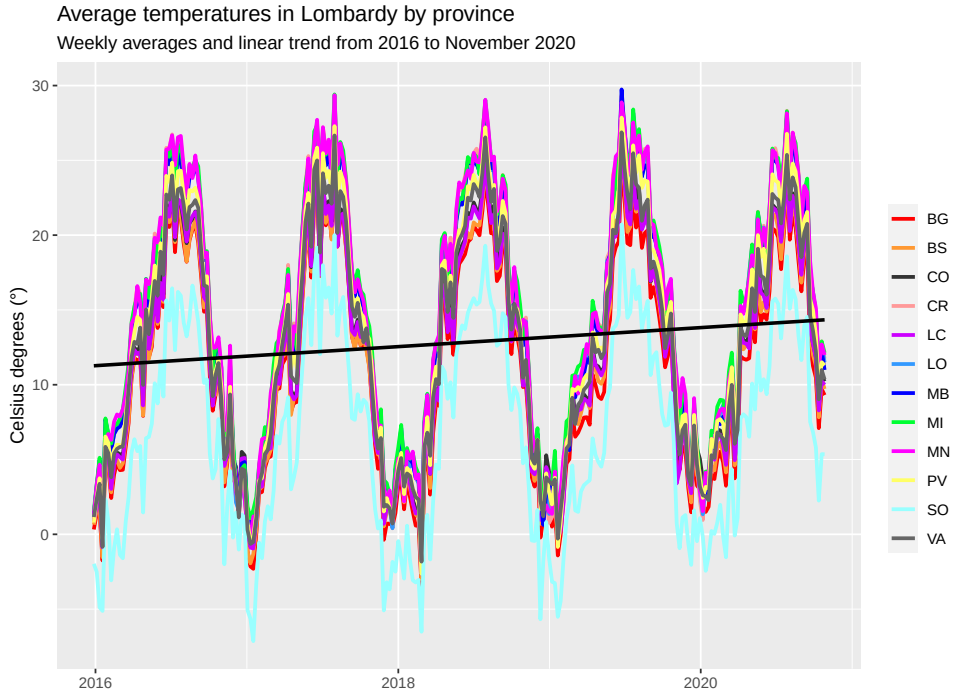


FIGURE 2.25: Average weekly temperature at provincial level from 2016 to November 2020.

case, as in Donnelly et al., 2011, Yu et al., 2004, Jammalamadaka and Lund, 2006 and Erdem and Shi, 2011.

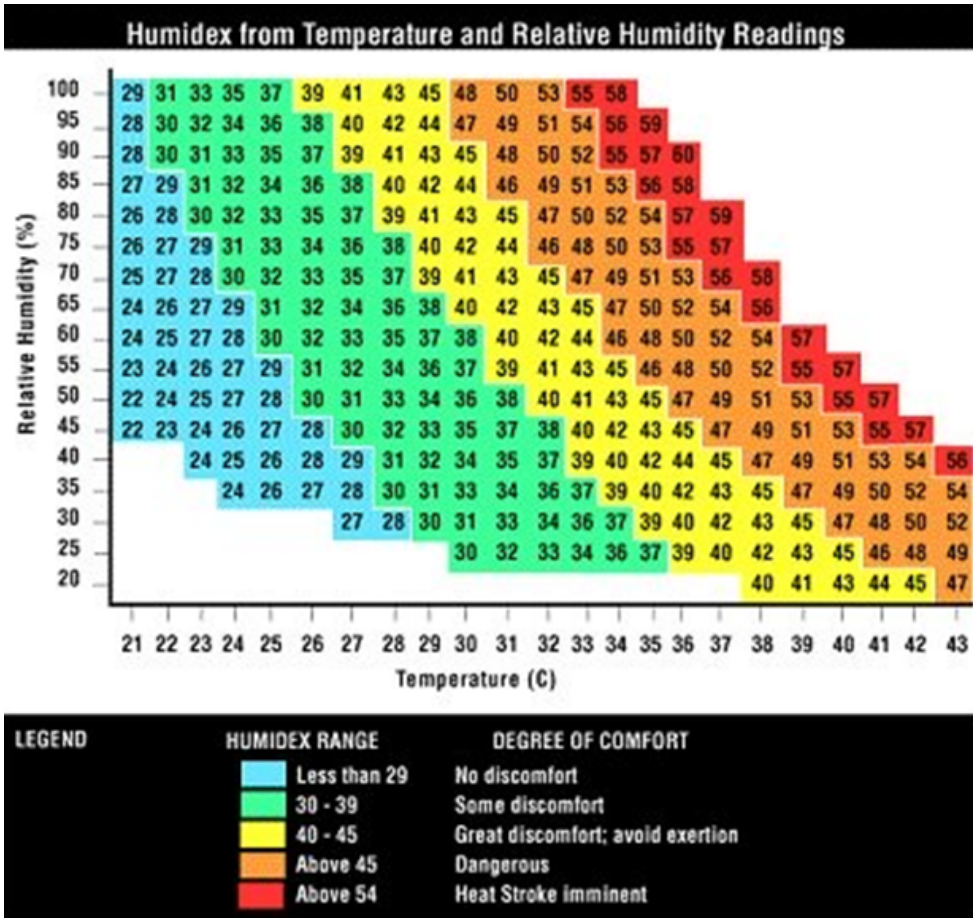


FIGURE 2.26: Humidex scale as a function of relative humidity (vertical axis) and temperature (horizontal axis)
Source; web image.

Chapter 3

Statistical Modeling of the Early-Stage Impact of a New Traffic Policy in Milan, Italy

Based on:

Maranzano, P., Pelagatti, M., Fassò, A. & Mudelsee, M.
"Statistical Modeling of the Early-Stage Impact of a New Traffic Policy in Milan, Italy"
International Journal of Environmental Research and Public Health, 2020, 17(3), 1088)
DOI: https://doi.org/10.3390/ijerph17031088

3.1 Urban air quality management and political issues

In the last few decades, European countries implemented various modeling methods to assess the effects of local and regional emission abatement policy options on air quality and human health (Thunis et al., 2016). On the one side, they include scenario approaches, in which running a chemical-physical simulation model with and without a specific emission source allows for quantifying the impact on air quality levels (Cuvelier et al., 2007; Vautard et al., 2007). On the other side, they also include more comprehensive and multidisciplinary approaches, such as Integrated Assessment Models (IAM), which combine simultaneously many features of the economy, society, and scientific findings. These models are based on the combination of multiple mathematical tools and allow for assessing the impact of environmental policies or to improve the air quality control system. Typical tools are the full cost-benefit analyses (Gao et al., 2016), in which abatement measures, costs, and benefits are expressed in monetary units, optimization, and spatial analysis (Carnevale et al., 2012; Vlachokostas et al., 2009).
According to the above EU rules, governments adopted standards and quantitative limits for pollutant emissions to make economic agents responsible and implement abatement policies. In particular, the maximum concentration for NO_x and NO_2 is set to 40 $\mu g/m^3$ annual average and 200 $\mu g/m^3$ hourly not to be exceeded more than 18 times in a single year. As stated in Section 2.4.2 The Northern Italy, and in particular the Lombardy region, stands out as a heavily polluted area with difficulties in pollution management. The negative impact on society is not limited to health only. There is increasing evidence showing that bad air quality in general, and high NO_2 concentrations in particular, impact the economy, including finance (Liu et al., 2019) and tourism (Xu et al., 2019).

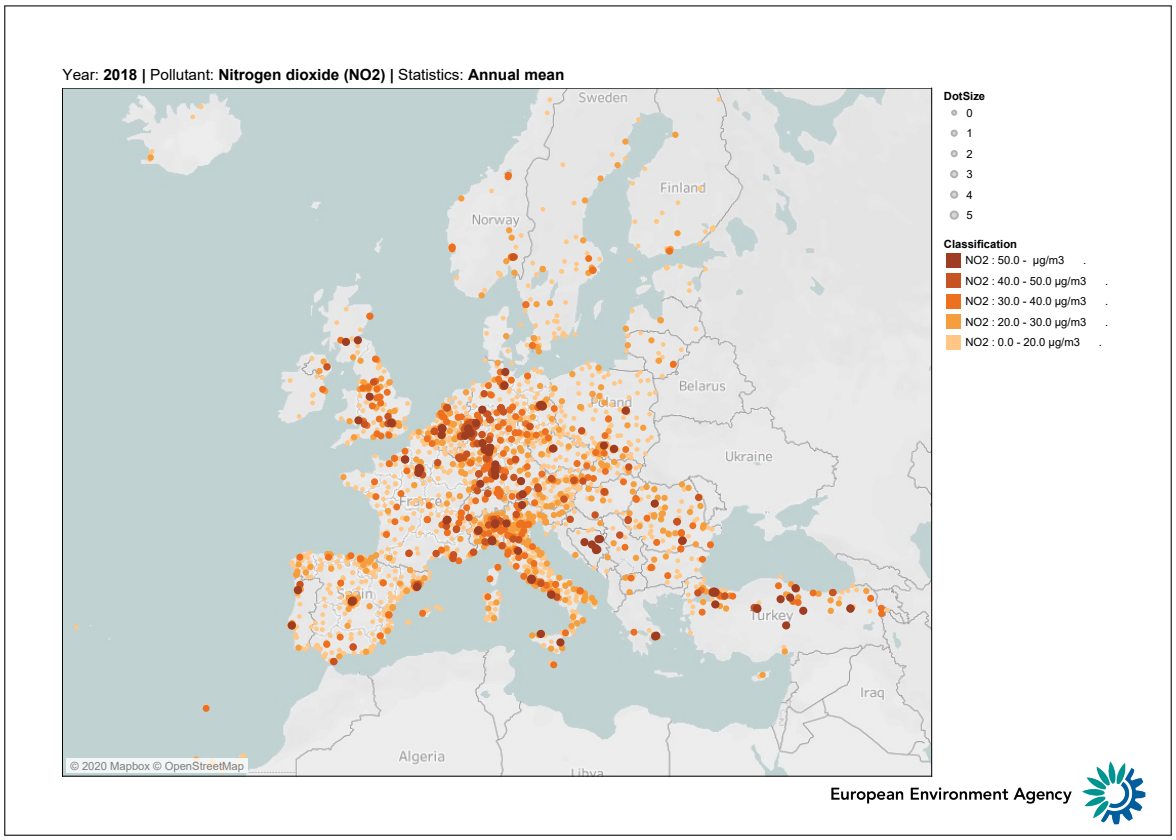


FIGURE 3.1: Annual average NO_2 concentrations ($\mu g/m^3$) in Europe during 2018. Levels are expressed in $\mu g/m^3$. Source: European Environmental Agency

3.2 Framework and reserch hypothesis

3.2.1 The new limited traffic zone in Milan: Area B

In this chapter we present the analysis of the case study concerning the introduction of the an air quality control policy in the municipality of Milan, which started on 25 February 2019 and directly acts on traffic rules. The administration defined an extended limited traffic zone, named Area B (https://www.comune.milano.it/aree-tematiche/mobilita/area-b), where the access and circulation for the most polluting vehicles, as well as those longer than 12 meters, have some partial restrictions, enforced by a monitoring system of entrance gates controlling each license plate and imposing a fine on unauthorized vehicles. The access prohibition concerns Euro 0 petrol vehicles and a large part of Euro 0, 1, 2, and 3 diesel vehicles, with specific exemptions for public transport, itinerant traders, and residents, and it is active from Monday to Friday during business hours (from 7:30 a.m. to 7:30 p.m.), except holidays. According to the municipality of Milan, the share of cars registered in the Milan metropolitan area and involved in the restrictions in 2019 is close to 17%, while the share of freight transport vehicles is around 53%(Milano, 2018). Area B is a progressive policy divided into various phases, which will concern an increasing number of vehicle classes. In terms of NOx emissions, the administration expects a reduction of 4-5% per year until 2022 and a reduction of 11% between 2023 and 2026. The policy will be fully operative within October 2030.
Area B extends the previously existing limited traffic zone, named Area C, which covers just the historical city centre, renamed in Italian as *Cerchia dei Bastioni*. The physical coverage of the two restriction zones is represented in Figure 3.2, which highlights the arrangement of both

within the city borders. Area B covers almost the entire area of the city, excluding extreme peripheral districts.
Statistical literature on air quality grew up sharply in the last decades. Two main statistical modeling directions have been developed. One has a focus on pollutants concentration and the other on human exposure. Regarding the latter, recent advances are based on crowdsourced data, such as smartphone data modeling (Finazzi and Paci, 2019). Regarding pollutants concentrations, increasing attention is being given to latent component models; see, as an example Fassò and Porcu, 2015 and for the problem of misalignment. In particular, the use of the INLA-SPDE approach for misalignment between pollutant concentration and epidemiological data (Cameletti et al., 2019) and PCA based methods with missing data (Vu et al., 2019).
When the territory under study is large and spatial correlation is important, spatio-temporal models are appropriate. See, for example, the multivariate state space approach of Calculli et al. (Calculli et al., 2015), which is capable of handling jointly PM_{10}, NO_2 and weather variables, the approach of Menezes et al. (Menezes Piairo et al., 2015) for modeling daily NO_2 trends in Portugal. Moreover, the land-use regression model (LUR) under a state space approach has been used for modeling air pollution in Tehran Taghavi-Shahri et al., 2019. Despite this growing spatial literature, time series analysis methods have been recently developed to understand the effect of meteorology on pollutant concentration Chen et al., 2018, which will be the main focus of this paper.
The previous Milan limited traffic zone, known as Area C, has already been treated in literature by Fassò Fassò, 2013, who analyzed its introduction through spatio-temporal models, by Invernizzi et al. Invernizzi et al., 2011, who considered its impact on black carbon, and by Percoco (Percoco, 2014) who considered its effect on traffic. Moreover, similar problems have been studied for London "sulphur-free zone" (Jones et al., 2012) and the "low emission zone" in Munich Qadir et al., 2013. In Fassò Fassò, 2013, the author considered both particulates and nitrogen oxides and observed the presence of a more pronounced permanent reduction of the latter within the restricted area, despite the data showing a strong spatial variability depending on the type of pollutant. This is consistent with the known emissions pattern of particulate matters and nitrogen oxides. The choice of NO_2 and NO_x as the main pollutants of interest is given by physical-chemical reasons well known in the literature that classify oxides as both primary and secondary gaseous pollutants emitted mainly by anthropogenic activities, such as heating systems, motor vehicle traffic, power plants, industrial activities and combustion). Moreover, recall that the emission inventory for Lombardy (INEMAR ARPA Lombardia Settore Aria, 2020) estimates that road traffic is responsible for 51% of the annual NO_x emissions for the Lombardy region and 65% for the metropolitan city of Milan. Moreoverm, another reason is due the fact that oxides are airborne pollutant that responds immediately to emission shocks, while atmospheric particulate matter has more complex and slower reactions. Therefore, for this case study about the short-term effects of Area B on air quality, we will take into account NO_x and NO_2 and postpone the analysis of PM_{10} and $PM_{2.5}$ to further research.
To adjust for confounding factors, we will consider weather conditions in Milan, the main calendar events, and the concentration levels of oxides observed in neighbouring towns, as in a pseudo-treatment-control approach.

3.2.2 Research hypothesis

The study aims to identify and quantify variations in pollutant levels due to the above described Area B. Hence, the present paper will try to investigate and test the following two scientific hypotheses:

- **Hypothesis 1**. The introduction of Area B achieved significant changes in pollution concentration for the city of Milan;

- **Hypothesis 2**. The variation occurred homogeneously on the territory and the stations do not show spatial variability of the effects.

The first hypothesis aims at quantifying the impact of the policy on pollution levels measured by several air quality stations scattered around the city and to assess whether this evidence is significantly supported by the data. The impact is evaluated both regarding the statistical significance of the estimates, the absolute magnitudes of the coefficients, and their signs. From the policy maker perspective, the expected coefficients should be negative, indicating a reduction effect on concentrations due to the car traffic restrictions. However, given the complexity of the phenomenon, a change of opposite sign cannot be ruled out either. The second research hypothesis is dedicated to the comparison of the estimates for the considered stations: the effect can be considered homogeneous when the sign and the magnitude of the coefficients for all the stations are similar.

3.2.3 Chapter structure

The rest of the chapter is structured as follows. Section 3.3 describes the dataset. In particular, we briefly explain the composition of both weather and air quality monitoring systems in Milan, available data sources, and metadata information. Then, in Section 3.4, we present the methodologies implemented for the preliminary analysis and the state space approach to time series analysis for air quality data. Section 3.5 reports and discusses the empirical results of the estimated models and their implications. Section 3.6 concludes the paper discussing the two research hypotheses in light of what emerged from the data analysis and gives some hints for future research developments.

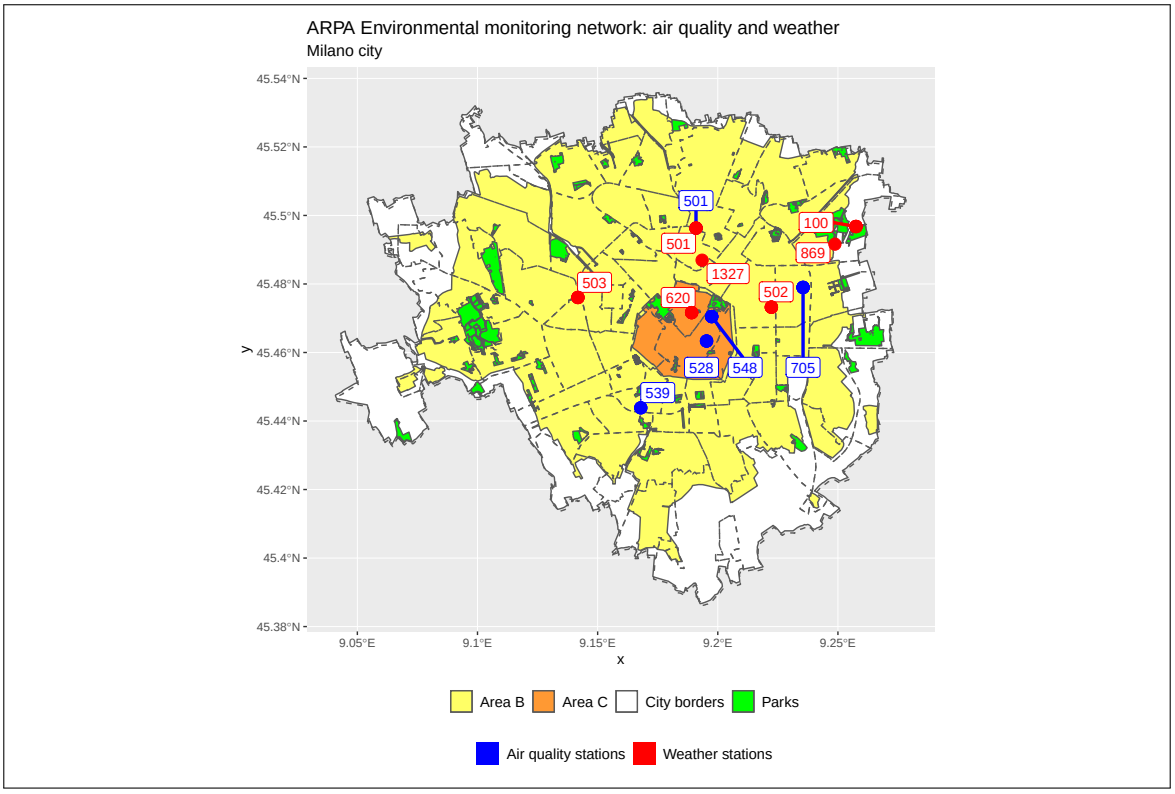


FIGURE 3.2: Monitoring system in Milan. Air quality stations (blue points): Marche (501), Verziere (528), Senato (548), Liguria (539) and Città Studi (705). Weather stations (red points): Lambrate (100), Zavattari (503), Brera (620), Feltre (869), Rosellini (1327), Juvara (502), and Marche (501).

3.3 Available data on air quality in Milan

In this section, we present the structure of the ARPA dataset and briefly introduce the methodologies implemented for the analyses. Section 3.3.1 introduces the data source for the Milan case study and the spatio-temporal structure of the data and provides a brief description of the variables taken into consideration. Section 3.4.2 designs the preliminary analysis, which introduces two temporal treatment-control experiments based on bootstrap resampling to highlight the differences in concentration levels before and after the policy intervention. Section 3.4.3 gives a detailed overview of the use of state space models in time series analysis for the study of air quality data, including also a specific subsection for model selection and policy intervention.

3.3.1 Air Quality and Weather Monitoring Network in Milan

Data on pollution and weather conditions of Lombardy are collected from the Lombardy Regional Agency for Environmental Protection (ARPA Lombardia), which makes a large open data portal fully available to users (`https://www.dati.lombardia.it/`). The agency manages a diffuse monitoring system distributed among the regional territory and counting on hundreds of monitoring stations collecting intra-daily information on climate and pollution through sensors.
Installed within the borders of Milan are seven weather monitoring stations and five air quality monitoring stations. Air quality stations are classified according to a taxonomy system that identifies the type and function in the network. The stations Liguria (ARPA code 539), Marche (ARPA code 501), Senato (ARPA code 548), and Verziere (ARPA code 528) are urban traffic control units: sensors installed near important roads and intersections in order to accurately quantify the pollution generated by traffic. The station Città Studi (ARPA code 705) is instead of type urban background, that is, the station is located in such a position that the level of pollution is not mainly influenced by specific sources but by the integrated contribution of all the upwind sources at the station with respect to the predominant directions of the winds on the site (European Environmental Agency, 2019a). The seven weather stations are Marche (ARPA code 501), Lambrate (ARPA code 100), Zavattari (ARPA code 503), Brera (ARPA code 620), Feltre (ARPA code 869), Rosellini (ARPA code 1327), and Juvara (ARPA code 502).
Figure 3.2 georeferences on the map the exact position of each station and allows for identifying the position with respect to Area B and Area C. Air quality stations are represented as blue points, while weather stations are the red points. Marche station (ARPA code 501), in the upper side of the map, is the only one to collect both weather and pollution data and is represented with a double label, the first one blue and the second red.
The spatial distribution of the stations is not uniform: air quality stations cover northern, eastern, central, and southern parts of the city, leaving the western districts uncovered; climate stations cover in detail the city centre and all the northern neighbours but are not installed in the south.

3.3.2 Temporal Coverage, Pollutants, and Weather Measures

The analysis presented in this paper takes into account daily measures from 1 January 2014 to 30 September 2019, generating an overall sample of 2099 daily observations.
The whole, the monitoring system provides information about many urban pollutants, such as carbon dioxide, particulates, and oxides. All the pollutants are measured as $\mu g/m^3$. As already stated in the Introduction, we focus our attention on concentrations of total nitrogen

oxides (NO_x) and nitrogen dioxide (NO_2), which are mainly primary gaseous pollutants, hence considered as proxies of pollution emissions due to human activities, first of all car traffic.
Weather stations provide measures of local temperature ($^\circ C$), rainfall (cumulated mm), humidity (%), global radiation (W/m^2), wind speed (m/s), and wind direction. The wind direction is expressed in clockwise degrees from 0° to 360°; for example, 90° identifies winds going from east to west. To make results easier to interpret, we decide to aggregate the measurements on wind direction and speed by constructing a set of new variables that describe the average speed in the four quadrants of the compass rose. The Northeast quadrant (Q_{NE}) corresponds to degrees between 0 and 90, the Southeast quadrant (Q_{SE}) to degrees from 90 to 180, the Southwest quadrant (Q_{SW}) to degrees from 180 to 270 and the Northwest quadrant (Q_{NW}) to the remaining values lying between 270 and 360 degrees.
These measures will be used in the modeling part to capture local weather conditions specific to the city of Milan. Instead of using the data referring to the weather station closest to each air quality station, we preferred to aggregate each of the climate variables through a daily city average valid for each pollution station. In this way, the subsequent models will be fully comparable guaranteeing the maximum possible spatial coverage.

3.3.3 Anthropogenic Activities

Human activities, and therefore the quality of the air we breathe, are often affected by calendar events that are recorded based on national, local, and religious holidays and weekends. Calendar effects are captured by a set of covariates, which identify the weekends and the main Italian holidays, both religious and secular. Holidays are collected in a dummy variable named *Holidays*, while the weekends are contained in a dummy variable named *WeekEnd*. Specific effects related to the behavior of people can be observed when holidays coincide with the weekend; therefore, we considered two terms of interaction between the two dummies. The interaction terms include those holidays that fall on Saturday, denoted as *Saturday:Holiday*, and those on which they fall on Sunday, which is *Sunday:Holiday*.
For a correct assessment of the effects of the traffic policy on pollutants concentrations in Milan, it is necessary to purify the estimates from any external weather or socio-economic effects overlapping with the policy and which may hence alter policy effects. This operation is accomplished by introducing a counter-factual term into the models represented by the pollution levels observed in other cities surrounding Milan. We considered seven important urban centres located in the Lombardy Po Valley area, which show socio-demographic and economic characteristics and weather conditions similar to Milan, but which cannot be directly affected by the limited traffic zone. These urban centres are Bergamo (East), Brescia (far East), Cremona (far Southeast), Lodi (Southeast), Pavia (South), Saronno (North) and Treviglio (East). As reported in Figure 3.3, the considered candidates cover a large territory surrounding Milan in all the directions while maintaining a sufficient distance to be considered independent in terms of traffic.

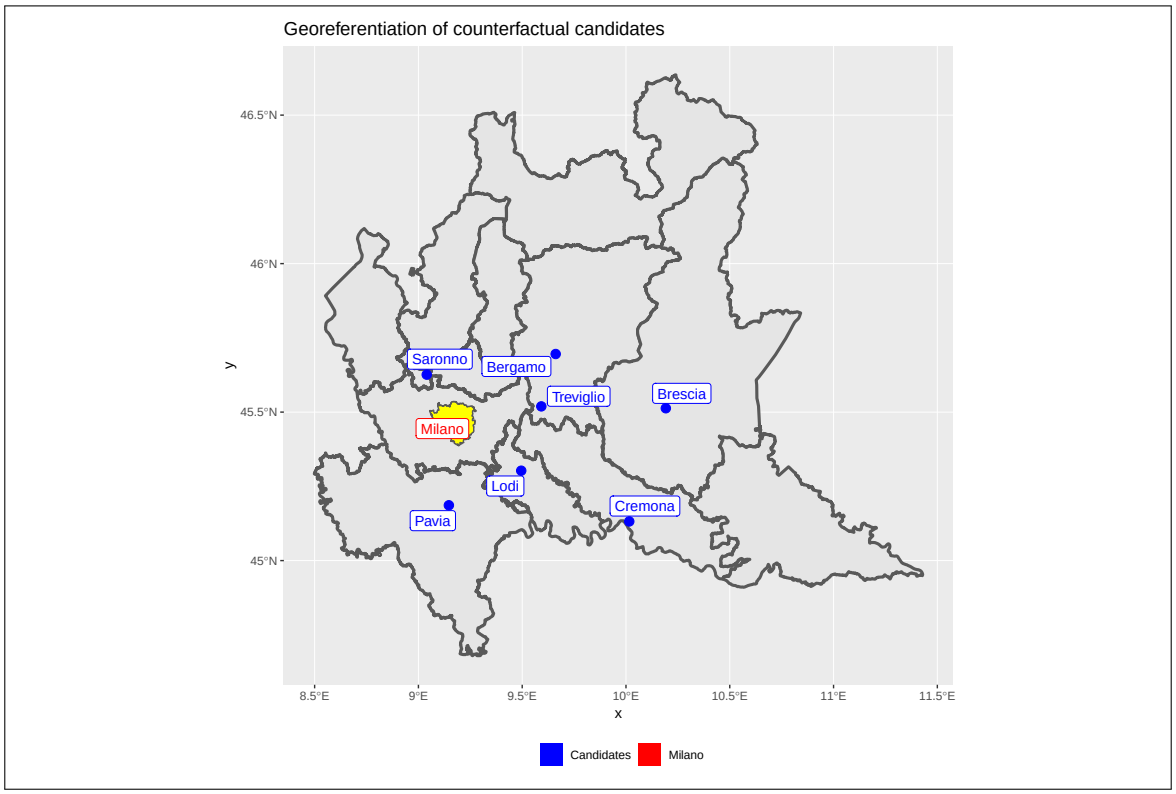


FIGURE 3.3: Georeferentiation of counter-factual candidates. Geographical positioning of the counter- factual candidates with respect to Milan.

3.4 Time series methods for level shifts detection in air quality series

3.4.1 Some facts about air quality in Milan during 2019

Figure 3.4 shows the temporal evolution of yearly average and median concentrations in the period preceding and following the entry into force of the policy for each control units located in Milan. According to the figure, starting from 2015, the city of Milan recorded a generalized reduction of concentration levels especially in peripheral areas, such as Marche and Liguria. Observed mean values for 2019 present a further reduction of concentrations rather apparently anomalous and significant. The comparison between the levels of NO_x and NO_2 pairs for each station shows obvious common trends between the two pollutants both considering the annual average and median values. Averages and medians follow similar temporal patterns, but focusing on nitrogen oxides sensors, it is possible to note that the medians are significantly smaller than the averages, highlighting the heavy-tailed characteristic of the distribution (positive asymmetry) and the presence of extreme values. Following these facts, an interesting question to investigate is if, and how much, the greater difference observed in 2019 can be attributed to traffic restrictions, or if it is due to a general de-carbonization trend that the city is experiencing, or to weather variations not considered yet.

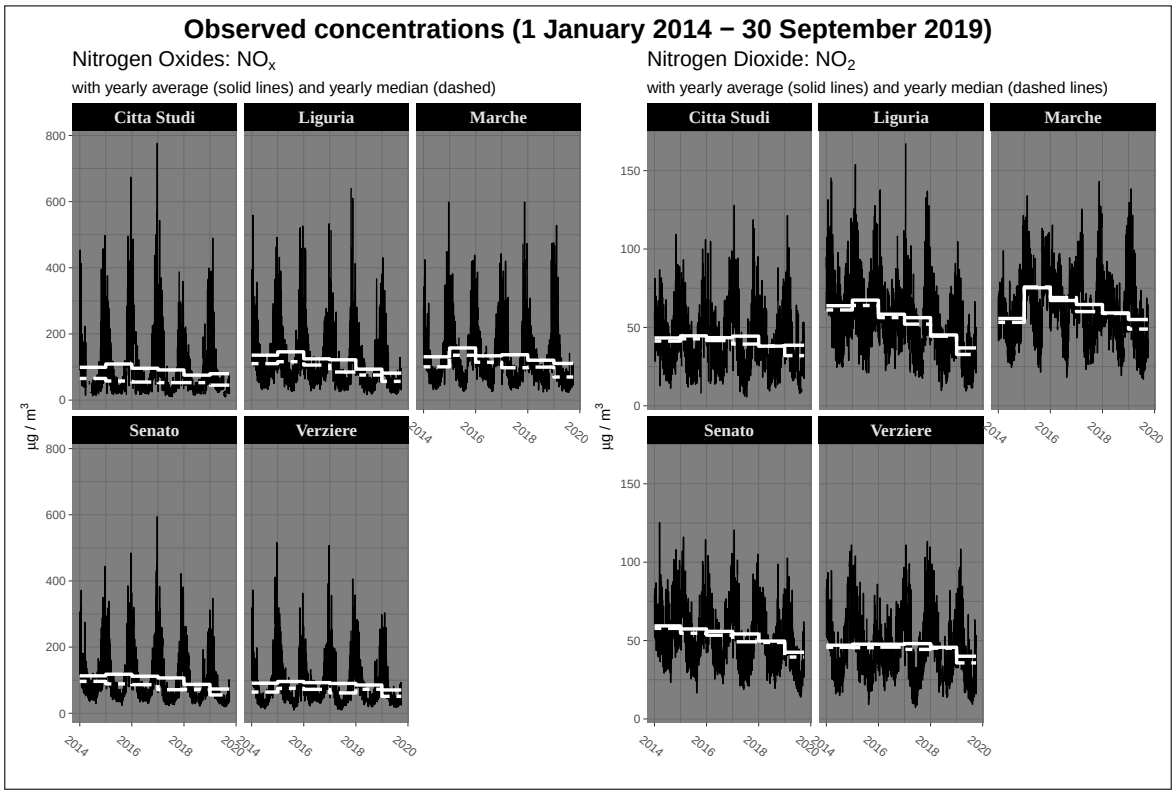


FIGURE 3.4: Pollutant levels in Milan ($\mu g/m^3$). Observed concentrations levels of NO_x and NO_2 between 2014 and 2019 with yearly average and median values. Values are expressed as $\mu g/m^3$.

3.4.2 Comparison of the average and median concentrations pre-and-post policy intervention using bootstrap methods

Before investigating the factors and causes that may have generated these sharp reductions, we perform a preliminary analysis of the concentration levels pre-and-post policy, in order to quantify the changes observed in 2019 both in Milan and in the other centres.
The comparison is performed through the design of two experiments, the first based on the difference of central tendency indicators and another based on the average and median differences among observed concentrations. Both experiments use the observations collected between 25 February and 30 September of each year, with a total length of 214 days. Approaches of this type can be framed in a context of treatment-control analysis, in which the data referred to the year 2019 constitute the treatment group, while the observations collected between 2014 and 2018 compose the control group. Control data refer to a 5-year-period; therefore, the concentrations measured on the same calendar day are aggregated into a single representative value calculated as the daily average concentration of the period 2014 to 2018. Denoting as c_{ij} the observed pollutant concentration during the day j, where $j = 25\ February, ..., 30\ September$, of the year i, where $i = 2014, ..., 2018$, the average for a generic calendar day j is computed as $c_j = \frac{\sum_{i=2014}^{2018} c_{ij}}{5}$.
Let $U = \{u_j, \quad j = 1, 2, ..., 214\}$ be the treatment observations and $V = \{v_j, \quad j = 1, 2, ..., 214\}$ the control observations.
The first experiment includes two statistics and considers control and treatment groups as unpaired independent samples; the two calculated statistics are respectively the difference of average and median between the registered concentration levels in 2019 and the mean levels

within the control group, denoted respectively byd_{AVG} andd_{MED}. The first statistic computes the difference between the average of the observations gathered after the policy intervention and the average of observations referring to the sub-period 2014-2018. The second statistics consists of computing the difference between median concentration levels observed in 2019 and before that year. The difference in average concentrations is denoted by d_{AVG}, whereas the difference in median concentrations is denoted as d_{MED}. Since both sub-periods are treated as independent of each other, from the statistical perspective, the statistics are assimilable to unpaired samples statistics.
The second experiment is designed as a two-sample paired experiment which evaluates the mean and median difference between the control and treatment pollutant levels observed in the same paired days. After computing the paired-day difference, i.e. $u_j - v_j \quad \forall j = 1, 2, ..., 214$, the paired-day average distance (AVG_d) and the paired-day median distance (MED_d) are computed respectively as the sample mean and the sample median of the differences.
Both experiments will be applied to the 5 stations installed in Milan city, three of them in peripheral areas and two within the city center, and to all the counterfactual cadidates presented in 3.3.3. This allows us to have an immediate comparison between the possible variations occurred in Milan and the air quality situation recorded in other urban areas of Lombardy.
According to the literature in climatology and environmental sciences, pollution data often show not-Gaussian and heavy-tailed distributions and persistence over time which compromise the accuracy of classical procedures for estimation variability calculus and could lead to distorted results. Therefore, we compare the central values obtained both considering the sample mean and the sample median, which is notoriously a more robust indicator if outliers occur (Mudelsee and Alkio, 2007; Mudelsee, 2014). To fix this issue we purpose to estimate the variability of all previous indicators through bootstrap confidence intervals techniques (Mudelsee and Alkio, 2007). Mudelsee (Mudelsee, 2014) suggests to implement different bootstrap schemes accordingly to the type of time series structure we are considering. The first experiment (difference of averages and medians) uses the data as time series, while the second experiment (average and median differences) compute the paired-day difference but lose the original time series structure. For these reasons we computed the variability of the estimators, and confidence intervals, implementing a moving-block-bootstrap (MBB) scheme for unpaired samples experiment and a non-parametric bootstrap with I.ID. resampling (Davison and Hinkley, 1997) for the paired data experiment. Moving-block-bootstrap (Lahiri, 2003) represents a natural extension of classic bootstrap methods which takes into account the presence of dependency among the data. The algorithm resamples blocks of consecutive observations at a time, preserving the underlying dependence structure of the original data generating process without any parametric model assumptions.
Table 3.1 sums up the previously defined indicators and gives the analytical definition of their computation.

Experiment	Bootstrap type	Indicator	Symbol	Formula
Experiment 1	MBB	Difference of average	d_{AVG}	$AVG(U) - AVG(V)$
Experiment 1	MBB	Difference of medians	d_{MED}	$MED(U) - MED(V)$
Experiment 2	IID	Average difference	AVG_d	$AVG(U - V)$
Experiment 2	IID	Median difference	MED_d	$MED(U - V)$

TABLE 3.1: **: Treatment effect indicators for two-sample experiments**. Note: $AVG(.)$ is the temporal sample mean operator and $MED(.)$ is the temporal sample median operator

3.4.3 Level shift detection using time series models in State Space form

In this section, we discuss the time series models used to identify the policy intervention effect on pollutant levels, the estimation algorithms, and the related inference. Firstly, we will introduce a brief description of State Space approach to time series analysis and the estimation algorithm using Kalman's filter. Then, we introduce a brief description of the basic structural model (BSM) using a state space approach for time series analysis and the estimation algorithm based on the Kalman filter (Durbin and Koopman, 2012; Pelagatti, 2015a). Following, we present a three-step procedure used to select the most representative model in terms of predictive power and quality of fit. As a last step, we explain how the policy intervention is included in the models and how it should be interpreted.

State Space analysis and Kalman Filter-Smoother estimation

State space modelling, hereafter SSM, is a generalized and unified framework for time series analysis which allows to study wide classes of models, including ARIMA, linear models and unobservable ccomponents (UCM). With the term *unobservable* we mean that the component cannot be directly observed and measured.
In the UCM model framework, each time series can be seen as the sum of directly observable and non-observable components. The latter can in turn be seen as both deterministic processes and stochastic latent processes, thus with their own source of randomness. When a component is said to be *stochastic*, it means that it evolves stochastically over time. The underlying idea is that the evolution over time of a series $Y_t = y_1, y_2, ..., y_n$ is guided by vectors of unobserved series $\alpha_1, \alpha_2, ..., \alpha_n$, whose relationship is specified by the state space (Durbin and Koopman, 2012). A natural and easy way to design such situations, is to image the observed Y_n series as the sum of trend, seasonality, cycle, deterministic covariates and white noise terms:

$$Y_t = \mu_t + \gamma_t + \psi_t + \theta X_t + \varepsilon_t \tag{3.1}$$

where μ_t is the trend, γ_t is the seasonality, ψ_t is the cyclical component, X_t is a set of deterministic covariates and ε_t is the white noise term.
Each SSM can be represented as a system of two equations: the first is named *measurement equation* (eq. 3.2) and identifies the evolution of the observable time series Y_t according to a set of unoberservable components α_t, whose evolution is described by the *transition equation* (eq. 3.3). The system is expressed as follows:

$$Y_t = c_t + Z_t \alpha_t + \varepsilon_t \tag{3.2}$$

$$\alpha_{t+1} = d_t + T_t \alpha_t + \nu_t \tag{3.3}$$

where $\varepsilon_t \sim N(0, H_t)$ and $\nu_t \sim N(0, Q_t)$ are uncorrelated Gaussian processes, Z_t is the *measurement matrix* and T_t is the *state matrix*. In the case of univariate process, both T_t and Z_t are scalars.
In the present chapter all the results refer to linear models which assume error term components as Gaussian distributed processes and named Gaussian linear state space models. The parameters estimation procedure involved the use of maximum likelihood (MLE) techniques, while the state estimation was performed using the *Kalman Filter* algorithm (KF). When dealing with Gaussian linear state space models, the parameters estimated using a maximum likelihood (ML) approach inherit the asymptotic properties of ML estimators (Pelagatti, 2015a). The distribution of the MLE is asymptotically approximated using a Gaussian distribution, which

allows deriving the usual asymptotic confidence intervals and t-tests for significance. Assuming a significance level of 5%, the estimates are statistically significant if the standardized value lies outside of the interval [−1.96,1.96], obtained using the quantiles of a Standard Normal distribution.
Kalman Filter is an iterative procedure which aim is to update the knowledge of sytem's states each time a new information y_t is provided; at each iteration t the filter uses just the previously observed information $y_1, ..., y_{t-1}$ to predict the current observation y_t and estimates new values for the model parameters via MLE.
The algorithm is composed by three ordered phases: 1) *Prediction step* in which the states α_t and the corresponding variance-covariance matrices P_t are updated using just the observations up to time $t-1$; 2) *One-step-ahead forecast* during which the predictions of the observed values and the forecast errors are computed; 3) *Update step* where the states are updated using both the prediction of step 1 and the current observation y_t.
The Kalman Filter requires as input the initial values of the states, i.e. $\alpha_{1|0}$, and their var-covar matrix, i.e. $P_{1|0}$, which will sensibly influence the final estimations. To fix this issue, after having run the filter, the states are smoothed through the *Kalman Smoother* algorithm (KS). This algorithm computes the values of the states using the complete time series Y_n and proceeding recursively backward. In this way, the smoothed states and parameters will take into account twice the information contained in the data, becoming more precise and less tied to the arbitrarily fixed initial values.

Basic Structural Model (BSM) for air quality time series

According to their physical characteristics, air pollution concentrations time series are often characterized by seasonality, high persistence (Windsor and Toumi, 2001; Chelani, 2013), strong right skewness with uni-modal distribution, and scale invariance (Lee, 2002). Therefore, we analyze the concentrations using the basic structural model (Harvey and Todd, 1983; Harvey and Koopman, 2014) augmented by deterministic regressors for weather conditions and socio-economic features.
BSM is defined as a simple unobservable components model composed by local linear trend (LLT), stochastic seasonality, and irregular (white noise) component. LLT describes both the temporal evolution of the series level and its slope, while the seasonal component aims to capture cyclical behaviors given by natural and anthropogenic phenomena. We modeled the seasonal component using a trigonometric form for daily data, hence with period $s = 365$, and considering only a few harmonics given the very regular and almost deterministic behavior of the series. This fact avoids the risk of a model over-parametrization.
Let $\{y_1, y_2, ..., y_n\}$ be the time series of the observed pollution concentrations in logarithmic scale, the state space form of BSM without regressors is composed by the following equations:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \,, \tag{3.4}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is the measurement error and

$$LLT\ (Level): \quad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \,, \qquad \eta_t \sim WN(0, \sigma_\eta) \,, \tag{3.5}$$

$$LLT\ (Slope): \quad \beta_t = \beta_{t-1} + \zeta_t \,, \qquad \zeta_t \sim WN(0, \sigma_\zeta) \,, \tag{3.6}$$

$$Stochastic\ seasonality: \quad \gamma_t = \sum_{j=1}^{k} \gamma_{j,t} \,, \tag{3.7}$$

where $k \leq \lfloor \frac{s}{2} \rfloor$ is the number of included harmonics and $\gamma_{j,t}$ is the non-stationary stochastic cycle

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma^*_{j,t} \end{bmatrix} = \begin{bmatrix} cos(2\pi j/s) & sin(2\pi j/s) \\ -sin(2\pi j/s) & cos(2\pi j/s) \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma^*_{j,t-1} \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega^*_{j,t} \end{bmatrix} , \tag{3.8}$$

$\omega_t \sim WN(0, \sigma^2_\omega)$ and $\omega^*_t \sim WN(0, \sigma^2_\omega)$ are white-noise processes with mean zero and variance σ^2_ω.
Equation (3.4), which is the measurement equation, describes the evolution of the observed series as the sum of the underlying components, while Equations (3.5), (3.6), and (3.7) are the transition equations. Equations (3.5) and (3.6) compose the LLT and describe respectively the unobservable processes of the level and the slope, whereas Equation (3.7) describes the trigonometric seasonality evolution. Weather, socio-economic factors, and policy intervention will be included in the models adding a set of deterministic components to the measurement Equation (3.4). Since the BSM with Gaussian errors belongs to the class of Gaussian linear models, the estimation step has been performed using the Kalman Filter algorithm and the model's parameters by maximizing the Gaussian likelihood function. Moreover, since the dependent variable is expressed in logarithmic scale, the coefficients have to be interpreted as relative increases or decreases in concentration levels due to a unitary increase in the explanatory variable.

3.4.4 Three-Step model selection algorithm

Selection algorithm

We now propose a three-step procedure for model selection, which considers multiple rules based on cross-validation, information criteria, and stepwise regression. To avoid estimation bias due to the policy introduction, all the steps are computed using only the observations before the introduction of Area B that is, from 1 January 2014 to 24 February 2019.
The purposed three-steps model selection algorithm can be summarised as follows:

- **Step 1. Trigonometric seasonal component selection:** choice of the optimal number of harmonics able to describe the natural seasonality of air quality series;
- **Step 2. Counterfactual selection:** choice of the optimal counterfactual term able to model socio-economic features common to LOmbardy cities;
- **Step 3. Selection of the most relevant calendar and weather covariates:** choice of the most relevant features able to model local weather conditions and calendar effects;

Step 1 is designed for selecting the most predictive seasonal component, defined in Equation (3.7), comparing different model specifications, which consider a varying number of harmonics k for the trigonometric function. Specifically, we fit 10 alternative models for each station: in each of them, the trigonometric seasonality is modeled by an increasing number of harmonics ranging from $k = 1$ to $k = 10$. The use of an increasing number of sinusoids, in our case up to 10, allows the modeling of complex seasonality with strong variations within short periods, but at the same time increases the model complexity.
Once the seasonal component has been selected, *Step 2* introduces in Equation (3.4) a counterfactual term x_t able to capture weather and socio-economic factors common to Lombardy and affecting the air quality of Milan. In our approach, the counter-factual candidates are the time series introduced in Section 3.3.3 and which refer to the measurements of pollutant concentrations in seven important cities around Milan. The new measurement Equation can be written as follows:

$$y_t = \mu_t + \gamma_t + \theta x_t + \varepsilon_t \, , \tag{3.9}$$

where x_t is the logarithm of the counter-factual time series and θ is its coefficient, μ_t follows Equations (3.5) and (3.6), and γ_t follows the specification obtained by step 1. The expected sign of θ is positive: higher levels of air pollution should correspond to high values in nearby cities due to similar conditions.
Eventually, in *Step 3*, we identify the best subset of calendar events and weather covariates, capturing residual variations not yet covered by the counter-factual or by the latent components. These residual variations are estimated by the smoothed observation disturbances from Equation (3.9) that is $\hat{\varepsilon}_t$, and describe residual patterns that have not been explained by the persistence of series, the seasonality or characteristics common to nearby territories of the region. Relevant weather and calendar covariates are selected through a backward-forward stepwise regression algorithm, which uses as a starting model the auxiliary linear regression expressed in Equation (3.10). The equation represents the full model which sets up the smoothed observation errors $\hat{\varepsilon}_t$ as dependent variable and the weather conditions and calendar events as covariates:

$$\begin{aligned}\hat{\varepsilon}_t = {} & \tau_1 Holidays + \tau_2 WeekEnd + \tau_3 Saturday : Holidays + \tau_4 Sunday : Holidays \\ & + \tau_5 Temperature + \tau_6 Rainfall + \tau_7 Radiation + \tau_8 Humidity \\ & + \tau_9 WindSpeedQ_{\text{NE}} + \tau_{10} WindSpeedQ_{\text{NW}} + \tau_{11} WindSpeedQ_{\text{SW}} + \tau_{12} WindSpeedQ_{\text{SE}} \\ & + e_t \ ,\end{aligned} \tag{3.10}$$

The stepwise regression is set up twice for each station: in one case, it selects the model according to the Akaike's Information Criterion (AIC), while, in the other, it uses the Bayesian Information Criterion (BIC). The algorithm starts estimating the full model and computes the AIC or the BIC. Iteratively, it drops out the predictors one at a time; at each step, it computes the new information criterion and considers whether the criterion is improved by adding back in a variable removed at a previous step. The procedure ends when the reintroduction of each omitted variable does not improve the information criteria.

Rolling origin cross-validation

In the first two steps, we select the seasonal component and the counter-factual term by fitting and comparing alternative models based on Equations from (3.4) to (3.7) according to their predictive power and their ability to adapt adequately to the observed data. The first principle, which tests the predictive power of the models, relies on the minimization of the cross-validated mean square forecasting error (MSFE) evaluated for up to 10-step-ahead forecast horizon, that is, $\hat{y}_{t+h} \quad \forall\ h = 1, 2, ..., 10$, while the second compares the models in terms of estimation quality. The latter computes both corrected Akaike's Information Criteria (AICc) and BIC intending to select the model that minimizes both. To identify a unique model for all the stations located in Milan, we proceed to a global comparison, both graphical and analytical, of the two blocks of indicators, giving attention to the overall performances and not focusing only on individual outputs.
According to the cross-validation principle for time series (Bergmeir et al., 2015; Bergmeir et al., 2018), we split the full time series into two subsets: a training set for model estimation and a test set for evaluating the out-of-sample forecast performances. The training set includes all the measurements until 24 February 2018, while the test set contains observations relative to the sub-period 25 February 2018–24 February 2019, for a total count of 365 out-of-sample observations. The exclusion of observations after the start of the traffic restrictions makes it possible to obtain unbiased estimates of the policy effects avoiding overlapping with other unidentified factors. Before starting the iterative loop, the model to evaluate is estimated just on time

using the observations included in the original training set. At the end of the estimation, the cross-validation algorithm is iteratively implemented as follows. For each iteration, the algorithm extracts the first ten observations available in the test set, generating a forecasting set, and computes three quantities: the 1-to-10 step-ahead forecasts that is $\hat{y}_{t+h} \quad \forall\, h = 1, ..., 10$, the forecast errors $\hat{y}_{t+h} - y_{t+h}$ and the quadratic forecast errors $(\hat{y}_{t+h} - y_{t+h})^2$. The first out-of-sample observation is discarded and the set of forecasting observations is updated right-shifting the forecast horizon by 1 unit and adding the new observation. These operations are repeated for a number of times equal to the length of the test set, in our cases 365 times. The algorithm returns the output of 365 different sequences of 1–10 step-ahead forecasts; for each step-ahead $h = 1, 2, ..., 10$, the MSFE is calculated as $MSFE(h) = \frac{\sum_{j=1}^{365} (\hat{y}_{t+h} - y_{t+h})^2}{365}$. Current statistical literature refers to this procedure as *rolling-origin forecast CV* because the origin at which the forecast is based rolls forward in time. Figure 3.5 shows an example of the rolling-origin mechanism for 1-step-ahead (left panel) and 4-step-ahead (right panel) forecasts. On the horizontal axis are reported the time stamps (time), while on the vertical axis are reported the iterations of the algorithm. Blue points represent the training set, while the red bullets are the observations of the test set.

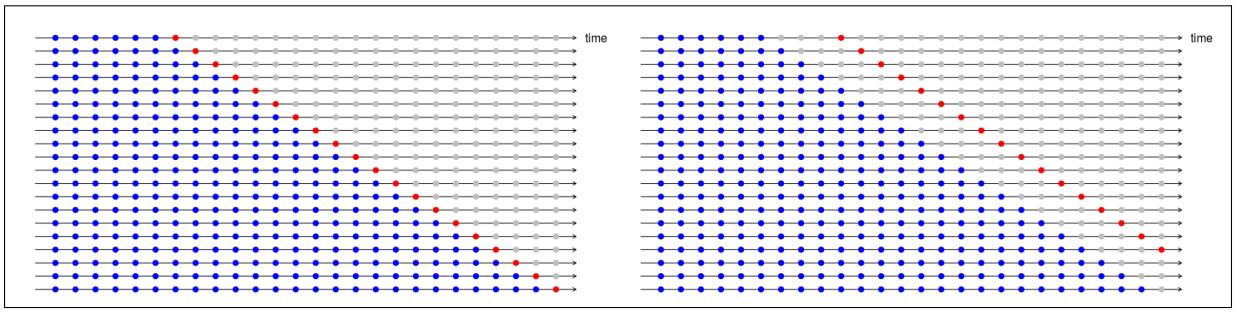


FIGURE 3.5: Rolling origin Cross-Validation mechanism. Left panel: 1-step-ahead forecasts. Right panel: 4-step-ahead forecasts.
Image by Rob J. Hyndman blog.

3.4.5 Modelling the policy intervention using ARMA-like transfer functions

The introduction of new rules or limitations to individual behaviors can lead to the co-existence of multiple effects with different structure, such as simultaneous immediate changes and adaptive changes that take a long time before visible effects occur. Take into consideration that this fact leads to implement intervention analysis, which includes both permanent and transitory effects. Further details and examples of ARMA-like transfer function applied to intervention analysis are available in Pelagatti (Pelagatti, 2015a).
The policy intervention is modeled through the combination of two individual effects: (1) a permanent effect, estimated by δ_1 that measures the level shift of pollutant concentrations given by the treatment and modeled as a step dummy, which is D_{1t}, which assumes a value equal to 1 starting from 25 February 2019; (2) a transitory effect, estimated by δ_0 and evolving according to a first-order difference dynamics of the type

$$w_t = \lambda w_{t-1} + \delta_0 D_{2,t}, \tag{3.11}$$

where $D_{2,t}$ is a impulse dummy, which assumes value equal 1 for 25 February 2019 and 0 otherwise and λ measures the persistence of the transitory effect. The sum of the two effects returns the total effect, which expresses the estimated overall reduction or increase in air pollutant levels generated by the policy. The measurement equation after the three-step model selection and

augmented by the policy intervention is eventually expressed as follows:

$$y_t = \mu_t + \gamma_t + \theta x_t + Z_t\Phi + \delta_1 D_{1t} + w_t + \varepsilon_t \,, \tag{3.12}$$

where y_t is the logarithm of pollution concentrations in one of the stations in Milan, x_t is the logarithm of pollution concentrations in the optimal counter-factual station, μ_t is the LLT evolving according to Equations (3.5) and (3.6), γ_t is the optimal seasonal component selected in step one, Z_t is a matrix containing the set of optimal subset of weather and calendar covariates selected in step 3, and Φ is the associate vector of coefficients.

Software

All the statistical computations and figures have been carried out using the statistical software R (R Core Team, 2020). For state space models estimation, the *KFAS* package (Helske, 2016) was used. Cross-validation, forecasting, and model selection codes have been developed by the authors. The graphic elaborations were obtained by using the packages *ggplot2* (Wickham, 2016) and *sf* (Pebesma, 2018).

3.5 Results

In this section, we present and comment on the empirical results relating both to the preliminary analysis performed comparing pre-and-post policy concentrations and to the policy intervention analysis described in section 3.4.3. In Sections 3.5.1 and 3.5.2 we show the empirical results of the bootstrap experiments for NO_x and NO_2 for all the stations installed in Milan and for the other seven cities around it. We will focus also on the statistical significance of estimated variations using the bootstrap confidence intervals and the related hypothsis test theory. Section 3.5.3 presents the model selection results, the values of the selection criteria, and the final model specifications. Section 3.5.4 reports the empirical estimates of the policy effects obtained through the basic structural model augmented by the policy intervention.

3.5.1 Average and median differences in NO_2 and NO_x concentrations

Estimations obtained by MBB bootstrap resampling on the first experiment are presented in Figures 3.8 and 3.9. The plots show respectively the resamplings histograms for nitrogen oxides and nitrogen dioxide for treatment stations, i.e. Milano city, and for control stations, i.e. the seven stations located in other large urban centers of Lombardy presented in 3.3.3.
The plots represent in left panels the bootstrap estimations for the difference of mean values (d_{AVG}) of treatment and control periods and in the right panel the bootstrap results for the difference of medians (d_{MED}). All the variations are measured in $\mu g/m^3$.
The estimates highlight large negative differences in oxides concentrations between 2019 and the period 2014-2018, both in the metropolitan area of Milan and in almost all the surrounding towns. The simultaneous abatement inside and outside Milan confirms the presence of a general decreasing trend in the aggregate levels of pollutants for the Lombardy as already indicated by the previous analyses.
Regarding the Milan city stations, almost all the estimates highlight negative differences both for the average and the median of NO_x and NO_2, but some of them are not statistically significant. The differences registered are relevant both in suburban districts, such as the stations Liguria and Marche, and in the historical centre at the Senato station. Città Studi (East Milan) station includes within the 95% confidence bands the null value for the difference in average and median concentrations for NO_x. Accordingly to the classic inference theory, these facts

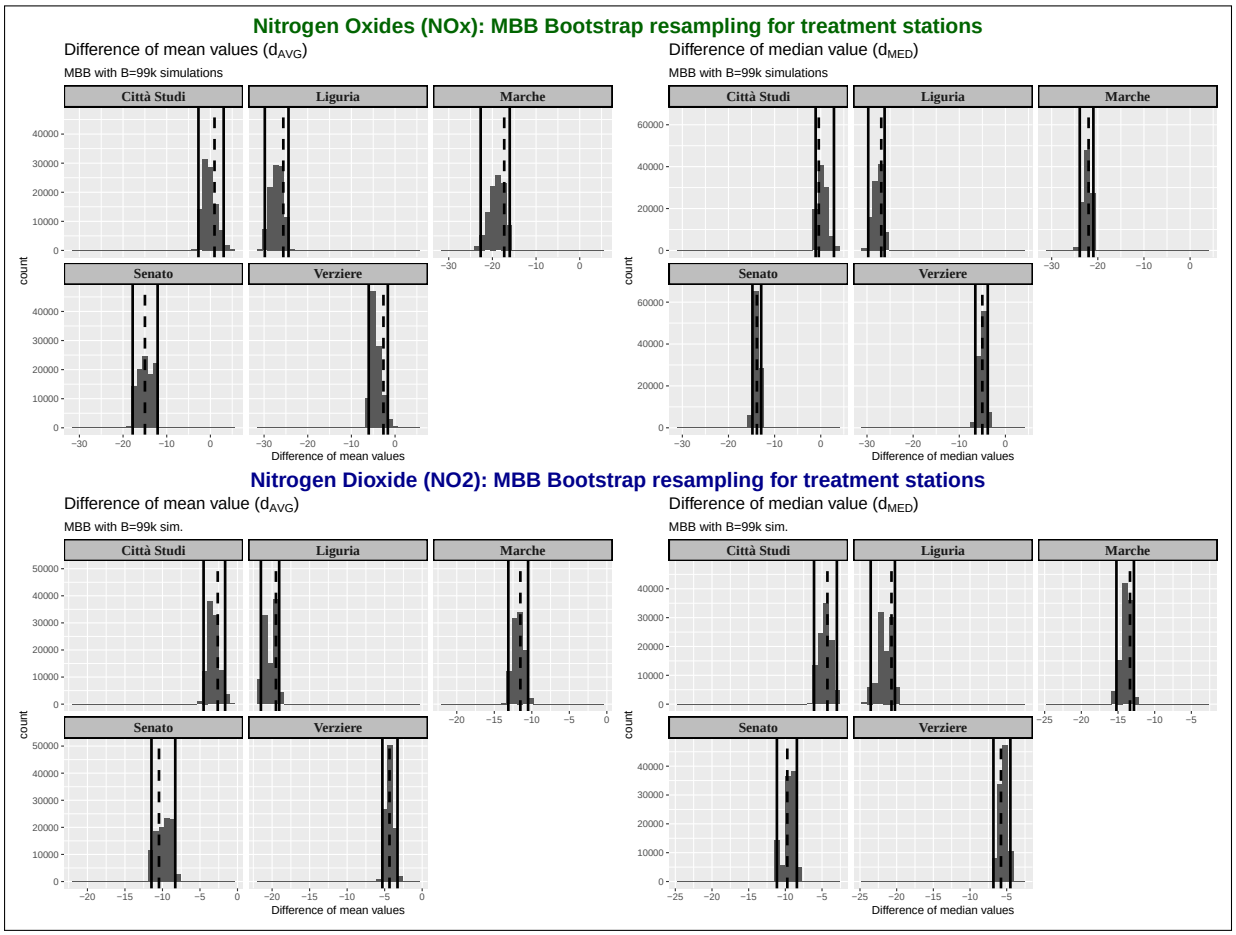


FIGURE 3.6: Difference of average and median differences for NO_x and NO_2 in Milan (treatment stations). Upper-left panel: difference of average concentrations for NO_x. Upper-right panel: difference of median concentrations for NO_x. Lower-left panel: difference of average concentrations for NO_2. Lower-right panel: difference of median concentrations for NO_2. Confidence bands correspond to the 2.5% and 97.5% percentiles of the MBB bootstrap simulations, whereas the vertical dashed lines are the sample averages. Values are measured in $\mu g/m^3$.

can be considered empirical evidences of statistically not-significant reductions in average observed concentrations of NO_x and NO_2 in Milan. Also Verziere (City center) registered low differences near the zero, but never including the null value within its bounds. At the opposite, Marche station (North Milan), Senato (Milan center) and Liguria (West) provide strong evidences of negative variations in both nitrogen oxides and nitrogen dioxide concentrations. Turning now to the Lombard cities near Milan we can see that the situation is quite heterogeneous and different from the metropolis. In general the estimated difference in mean and median levels are more contained and less significant. Only the cities of Bergamo (East), Brescia (far East), Pavia (South-West) and Saronno (Nort-West) recorded statistically signficant reductions comprised between -10 $\mu g/m^3$ and -5$\mu g/m^3$. Moreover, in Treviglio (South-East) and Cremona (far South-East) positive average and median variations, that is increase in concentrations, have been estimated for both NO_x and NO_2 levels. Cremona's increases are even statistically significant.

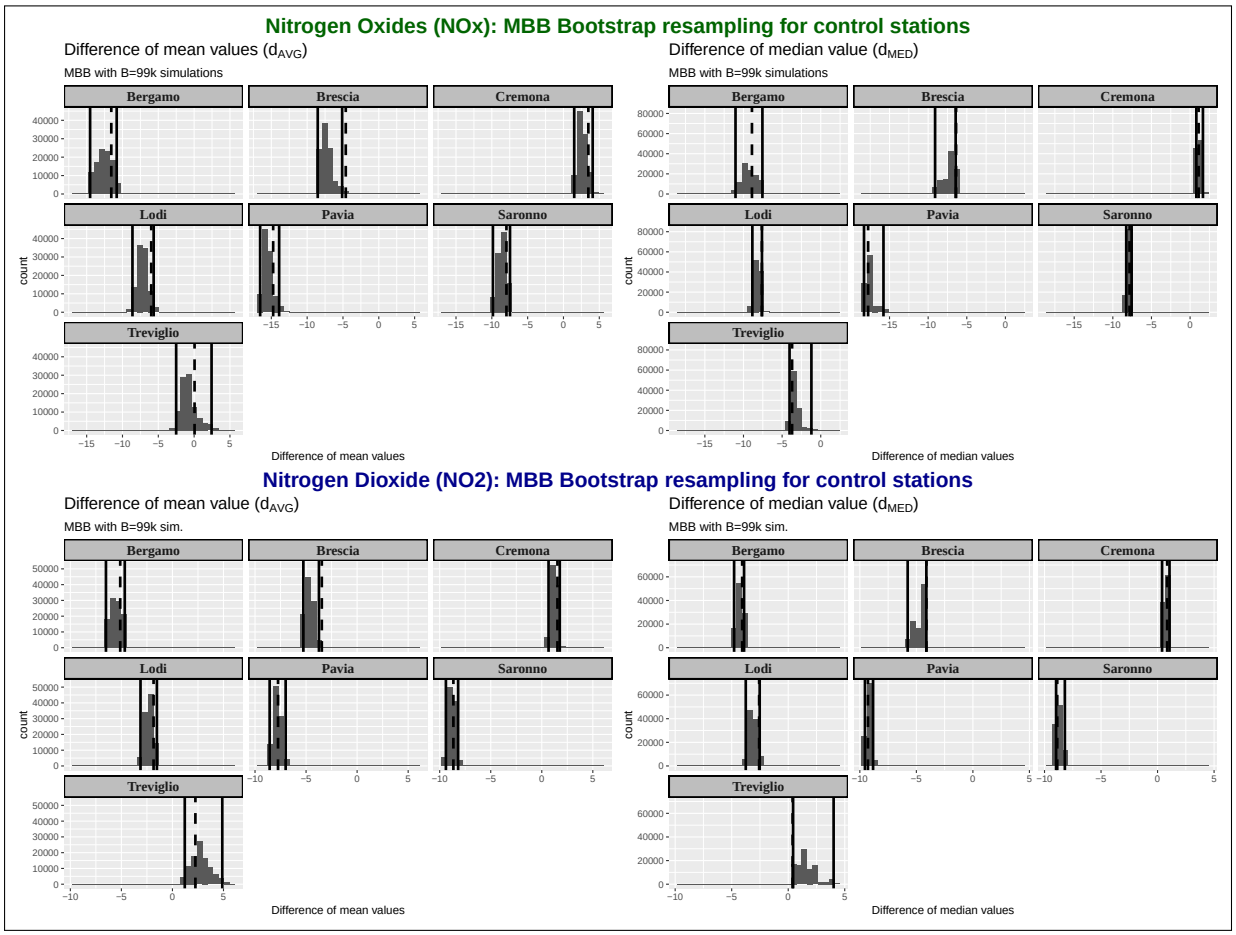


FIGURE 3.7: Difference of average and median differences for NO_x and NO_2 in other large uban centers (control stations). Upper-left panel: difference of average concentrations for NO_x. Upper-right panel: difference of median concentrations for NO_x. Lower-left panel: difference of average concentrations for NO_2. Lower-right panel: difference of median concentrations for NO_2. Confidence bands correspond to the 2.5% and 97.5% percentiles of the MBB bootstrap simulations, whereas the vertical dashed lines are the sample averages. Values are measured in $\mu g/m^3$.

3.5.2 Average and median differences in observed NO_2 and NO_x concentrations

The results of the second bootstrap experiments, carried on using a IID bootstrap approach, are summarized in Figures 3.8 e 3.9.
AVG_d and MED_d bootstrap estimations for Milan are very similar to the results of experiment 1 and lead to the same considerations: some stations show strong average and median difference among 2019 and the period 2014-2018, while the others are uncertain but in any case negative. As for the first experiment, the most relevant and significant reductions have been estimated for Senato, Liguria and Marche stations; while uncertain variations are estimated for Verziere and Città Studi. The first stations show statistically significant reductions lying between -30$\mu g/m^3$ and -10$\mu g/m^3$ for NO_x and lying between -20$\mu g/m^3$ and -10$\mu g/m^3$ for NO_x. On the contrary, in the last two cases the confidence bounds include the null value considering the average difference in NO_x levels among the periods, while the medians appears to be weakly statistically significant. This fact is related to the skewed and non-symmetric characteristics of the distribution involved, as shown also in Figure 3.4. In the same stations, the estimated variations

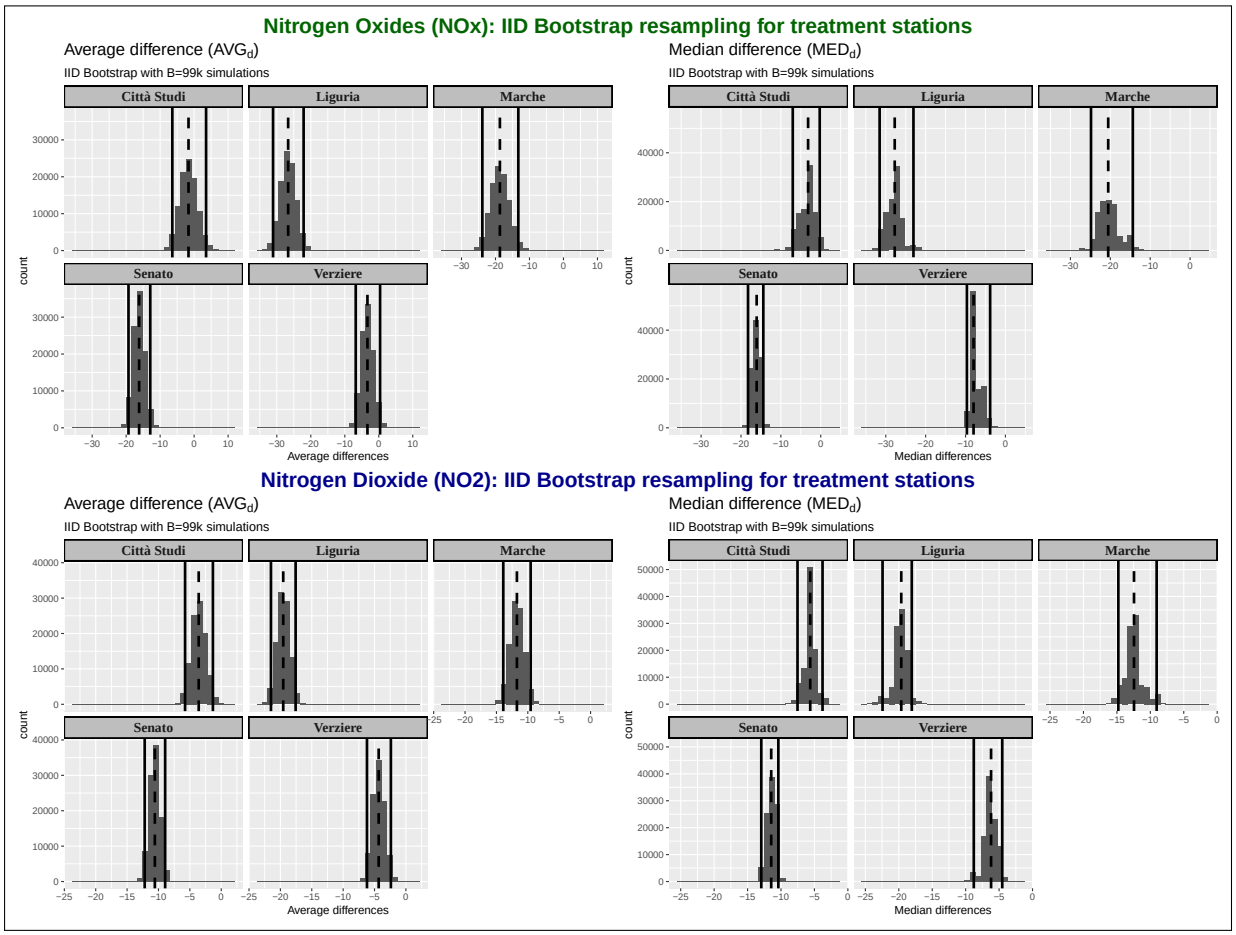


FIGURE 3.8: Average and median differences for NO_x and NO_2 in Milan (treatment stations). Upper-left panel: Average difference for NO_x. Upper-right panel: Median difference for NO_x. Lower-left panel: Average difference for NO_2. Lower-right panel: Median difference for NO_2. Confidence bands correspond to the 2.5% and 97.5% percentiles of the IID bootstrap simulations, whereas the vertical dashed lines are the sample averages. Values are measured in $\mu g/m^3$.

in NO_2 concentrations seem to be more robust and significant; despite the fact that they are largely reduced compared to the other stations.
The second experiment applied to the concentrations observed in other large urban centers produced almost identical results to the first experiment. Here, too, a territorial heterogeneity is evident in the differences between 2019 and the previous four-year period. Bergamo, Brescia, Saronno and Pavia stations show reduced but significant reductions in the concentrations of both oxides, while the cities of Cremona and Treviglio are not statistically significant.
As a final indication of the estimated reductions, we report in Table 3.2 the empirical values of the four considered statistics for both NO_x and NO_2.
These prelimnary results are consistent with the decreasing trend observed for Milan and highlighted in Figure 3.4 and anticipated in section 2.4.2 when discussing air quality in Lombardy and the visible decrease in concentrations and emissions on the territory. In this sense we can also state that the city of Milan in 2019 has recorded a generalized reduction in oxide concentrations for almost the entire city, while just part of the other considered urban centers lived the same situation. These preliminary results do not allow for identifying the causes of the reductions and to state if they depend on common causes related to the environment and climatic factors or if they have been generated by the introduction of the new policy in Milan.

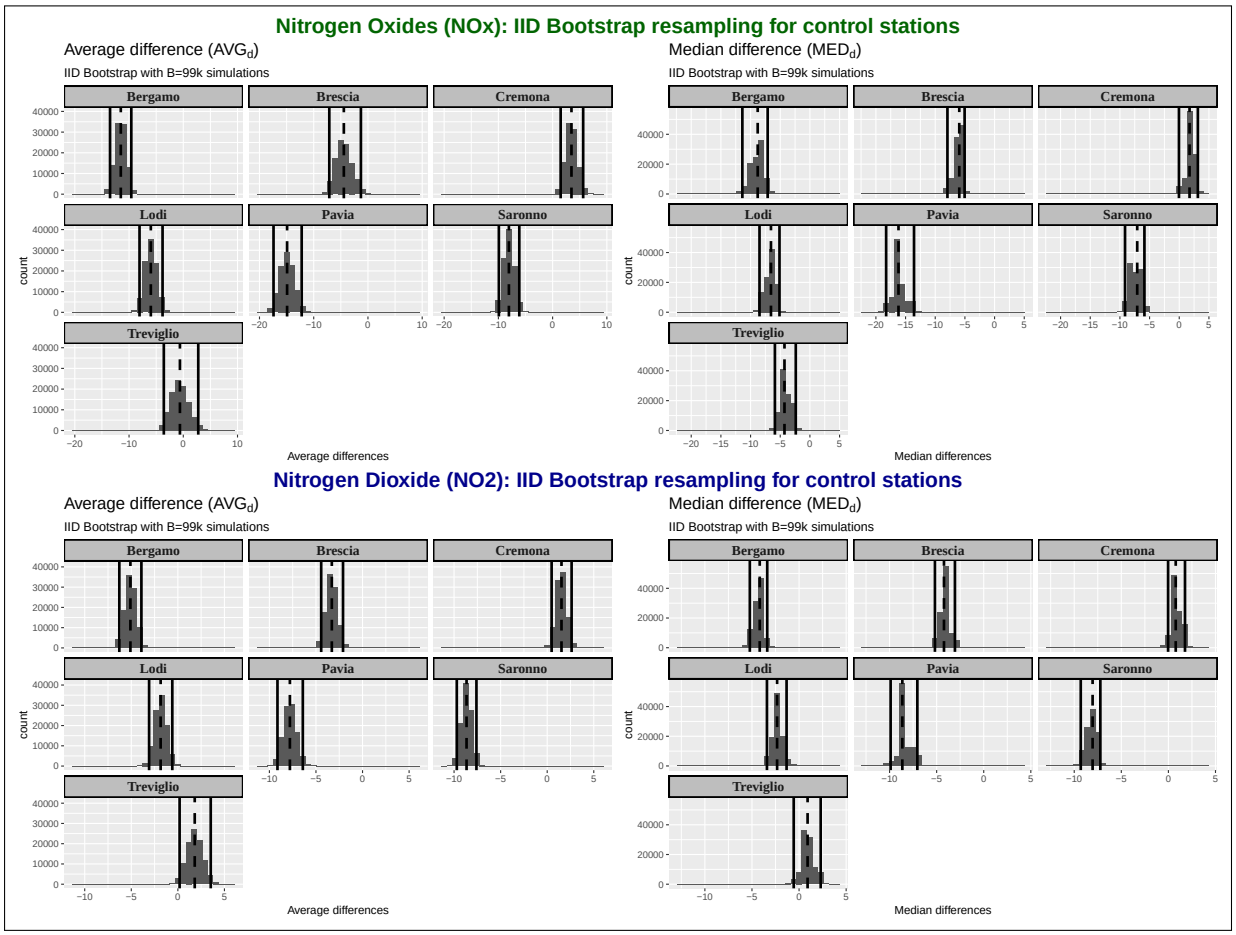


FIGURE 3.9: Average and median differences for NO_x and NO_2 in counterfactual candidates (control stations). Upper-left panel: Average difference for NO_x. Upper-right panel: Median difference for NO_x. Lower-left panel: Average difference for NO_2. Lower-right panel: Median difference for NO_2. Confidence bands correspond to the 2.5% and 97.5% percentiles of the IID bootstrap simulations, whereas the vertical dashed lines are the sample averages. Values are measured in $\mu g/m^3$.

The next section will attempt to investigate the variations through the modeling of possible environmental and anthropogenic factors able to influence the air quality of the city.

3.5.3 Three-steps model selection

Step 1: detection of the seasonal components

Results relative to the first step of model selection are summarized in Figures 3.10 and 3.11, which show the evaluation criteria for all the stations. For each station, the plots are organized in paired-panels; the left panel represents the 10-steps-ahead MSFE as a function of the forecast horizon and the number of harmonics modeling the seasonality (scale colour); the right panel shows the AICc–BIC pairs for each model. The optimal number of harmonics to model the seasonality is identified as the one that evaluates the minimum pair of AICc and BIC and returns the lowest MSFE curve. Both the estimates for NO_x and NO_2 for the city of Milan agree

TABLE 3.2: Bootstrap statistics. Differences between the concentration levels of the sub-period 2014-2018 and the treatment period 2019. Differences are expressed in $\mu g/m^3$.

Station name	Nitrogen Oxides				Nitrogen Dioxide			
	d_{AVG}	d_{MED}	AVG_d	MED_d	d_{AVG}	d_{MED}	AVG_d	MED_d
Milano city stations								
Città Studi	0.94	-0.43	-1.65	-3.17	-2.63	-4.26	-3.59	-5.63
Liguria	-25.59	-26.91	-26.67	-27.76	-19.49	-20.71	-19.55	-19.63
Marche	-17.29	-22.00	-18.71	-20.53	-11.53	-13.37	-11.74	-12.44
Senato	-14.99	-13.84	-16.18	-16.07	-10.46	-9.73	-10.58	-11.46
Verziere	-2.66	-5.00	-3.34	-7.99	-4.35	-5.78	-4.34	-6.20
Other urban centres in Lombardy (counterfactual candidates)								
Bergamo	-11.56	-8.90	-11.52	-8.80	-5.11	-4.07	-5.09	-4.18
Brescia	-4.61	-6.39	-4.45	-5.95	-3.46	-4.12	-3.30	-4.22
Cremona	3.47	1.10	3.47	1.74	1.53	0.85	1.53	0.78
Lodi	-5.99	-7.62	-5.99	-6.56	-1.85	-2.57	-1.85	-2.33
Pavia	-14.75	-17.77	-14.91	-16.19	-7.75	-9.28	-7.80	-8.64
Saronno	-7.96	-7.84	-8.06	-7.08	-8.66	-8.87	-8.66	-8.05
Treviglio	0.06	-3.71	-0.62	-7.99	2.25	0.39	1.81	0.91

unanimously in the selection of the model in which the seasonality is composed by a single harmonic ($k = 1$); therefore, it can be rewritten as

$$Optimal\ seasonal: \quad \gamma_t = \gamma_{1,t}\ , \tag{3.13}$$

where $\gamma_{1,t}$ is

$$\begin{bmatrix} \gamma_{1,t} \\ \gamma^*_{1,t} \end{bmatrix} = \begin{bmatrix} cos(2\pi/365) & sin(2\pi/365) \\ -sin(2\pi/365) & cos(2\pi/365) \end{bmatrix} \begin{bmatrix} \gamma_{1,t} \\ \gamma^*_{1,t} \end{bmatrix} + \begin{bmatrix} \omega_{1,t} \\ \omega^*_{1,t} \end{bmatrix}\ , \tag{3.14}$$

$\omega_t \sim WN(0, \sigma^2_\omega)$ and $\omega^*_t \sim WN(0, \sigma^2_\omega)$.

Step 2: detection of the counterfactual component

After selecting the seasonal component, we proceed to the selection of the counter-factual term. Estimates are summarized in Figures 3.12 and 3.13, which show the results for NO_x and NO_2. The plots are graphically organized like those related to step one, with the difference that the MSFEs and the AICc-BIC pairs are functions of one of the seven counter-factual candidates instead of the number of harmonics. The selection criteria follow the same rules used for the previous step.
The search for the optimal counter-factual term requires greater attention and detail than in the previous step as the minimizers are not unique. According to the plots, there is a restricted set of stations that are good candidates for the counter-factual role. The set includes the following cities: Treviglio, Pavia, Saronno, and Cremona. In particular, Pavia's station achieves one of the best forecast and fitting performances for almost all the stations in Area B for both NO_2 and NO_x. Based on this last consideration, we select as the counter-factual term for future models the air quality monitoring station of Pavia, located South to Milan. Therefore, the final specification of the basic structural model will include as counter-factual term the logarithm of the concentrations in Pavia, $x_t = log(Pavia_t)$.

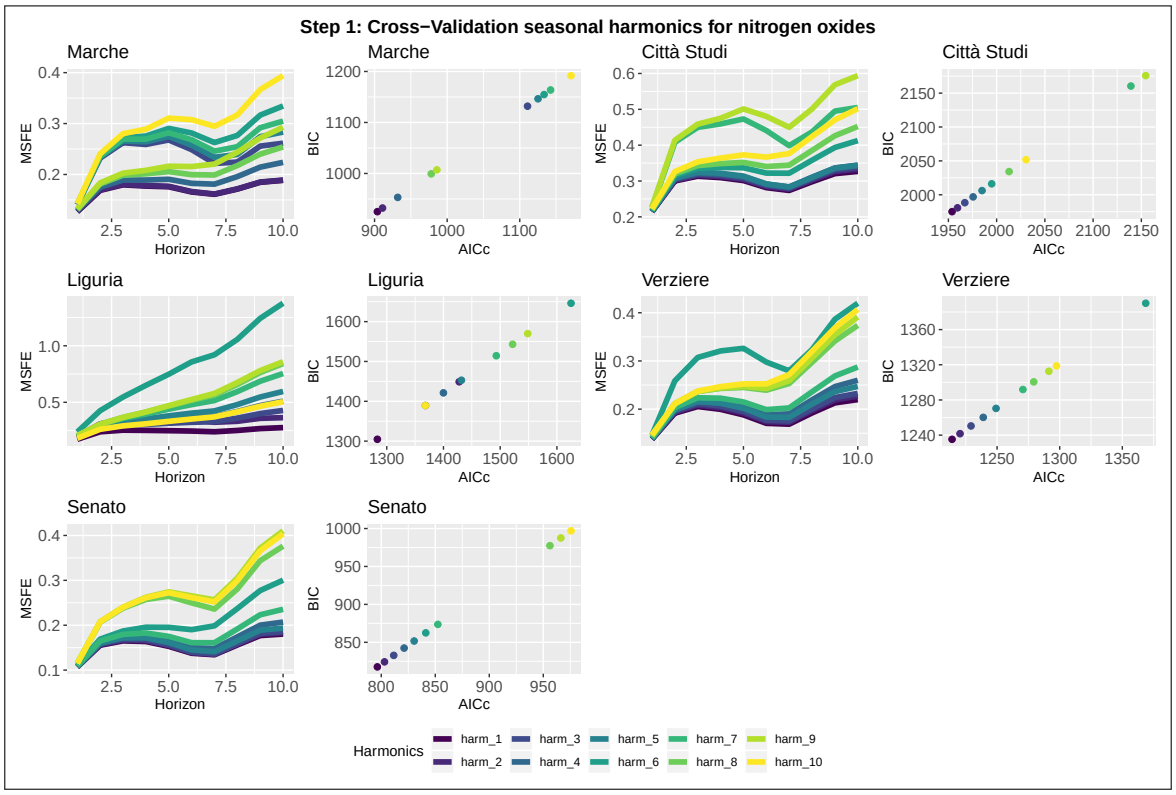


FIGURE 3.10: Model selection-Step 1-NO_x. Seasonal component selection for the five nitrogen oxides stations in Milan. Left panel: 10-steps-ahead MSFE in log-scale as function of the number of harmonics. Right panel: AICc and BIC pairs for each model.

Step 3: detection of the weather and calendar covariates

The last step of model selection aims to select the optimal subset of local weather and calendar regressors after having selected the optimal unobservable components, common weather, and socio-economic factors, captured by the counter-factual. For each station and pollutant, the best models are reported in Tables 3.3 and 3.4.

As expected, BIC-based models, being more parsimonious, retain fewer covariates than AIC-based models. Following this fact, we will use the BIC-selected models, but we now discuss some details about the selection process. Concerning the calendar, both criteria include in almost all cases the holidays, week-end, and Sunday holidays effects. AIC suggests adding also the interaction term between Sunday and holidays. Even if the interaction between Saturday and holidays is not always included, the final model will take into account the full set of calendar events and their interactions.

Regarding weather covariates, except for the wind speed, none of the others is included within the final models. Winds blowing from Southwest (Q_{SW}) are always selected and those coming from Northwest (Q_{NW}) are often included, and hence kept in the final model. Moreover, temperature, rainfall, solar radiation, and humidity are considered only by the AIC. This fact can be explained by the presence of the counter-factual, which captures not only common characteristics in terms of human behaviour and air quality conditions but also homogeneous climatic conditions common to all the areas considered.

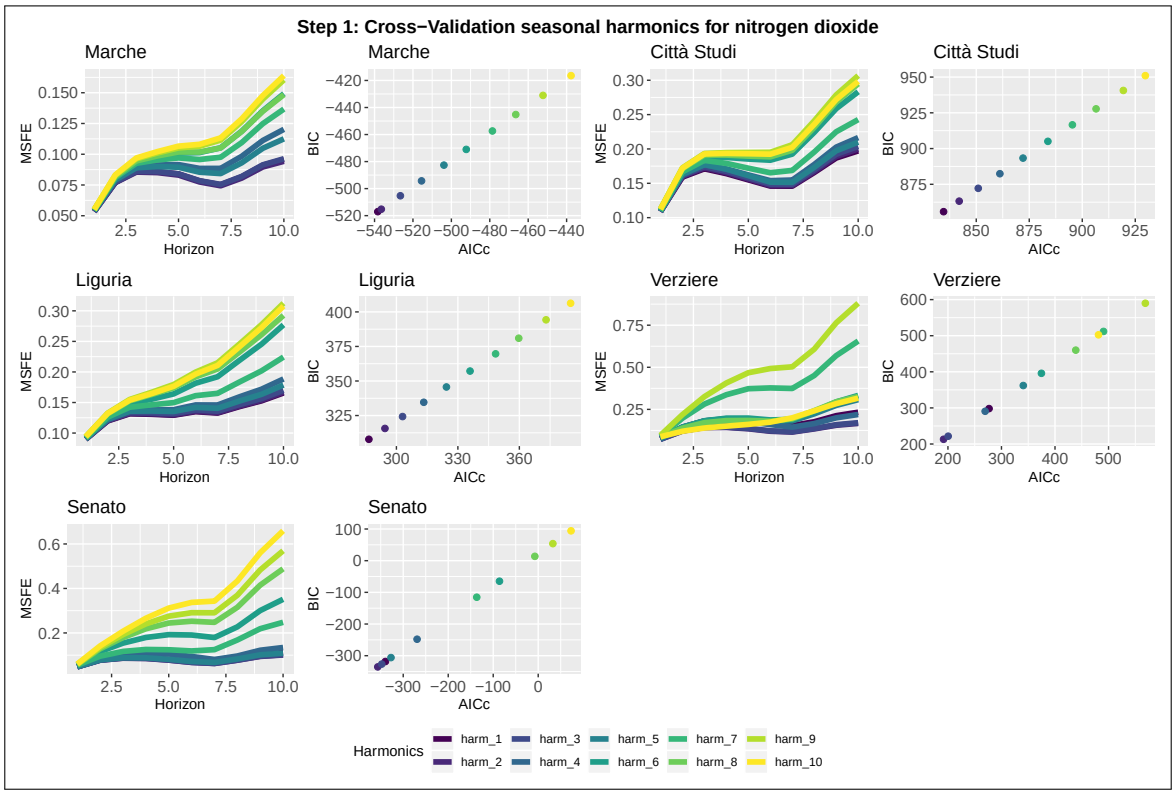


FIGURE 3.11: Model selection-Step 1-NO_2. Seasonal component selection for the 5 nitrogen dioxide stations in Milan. Left panel: 10-steps-ahead MSFE in log-scale as function of the number of harmonics. Right panel: AICc and BIC pairs for each model.

Final model specification

Based on the results of the three-step model selection procedure, the final specification of the BSM augmented by the policy intervention can be expressed using the following model:

$$\begin{aligned} \textit{Measurement}: \quad y_t = \mu_t + \gamma_t + \theta x_t + \phi_1 Holidays + \phi_2 WeekEnd \\ + \phi_3 Saturday : Holidays + \phi_4 Sunday : Holidays \\ + \phi_5 WindSpeedQ_{\text{SW}} + \phi_6 WindSpeedQ_{\text{NW}} \\ + \delta_1 D_{1t} + w_t + \varepsilon_t \,, \end{aligned} \tag{3.15}$$

where $\varepsilon_t \sim N(0, \sigma_\epsilon^2)$

$$\textit{Seasonal component}: \quad \gamma_t = \gamma_{1,t} \,, \tag{3.16}$$

$$\textit{Level}: \quad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \qquad \eta_t \sim WN(0, \sigma_\eta^2) \,, \tag{3.17}$$

$$\textit{Slope}: \quad \beta_t = \beta_{t-1} + \zeta_t \qquad \zeta_t \sim WN(0, \sigma_\zeta^2) \,, \tag{3.18}$$

$$\textit{Transitory policy}: \quad w_t = \lambda w_{t-1} + \delta_0 D_{2,t} \,, \tag{3.19}$$

where y_t is the logarithm of pollution concentrations in one of the stations located in Milan and $x_t = log(Pavia_t)$ is the logarithm of pollution concentrations in Pavia.

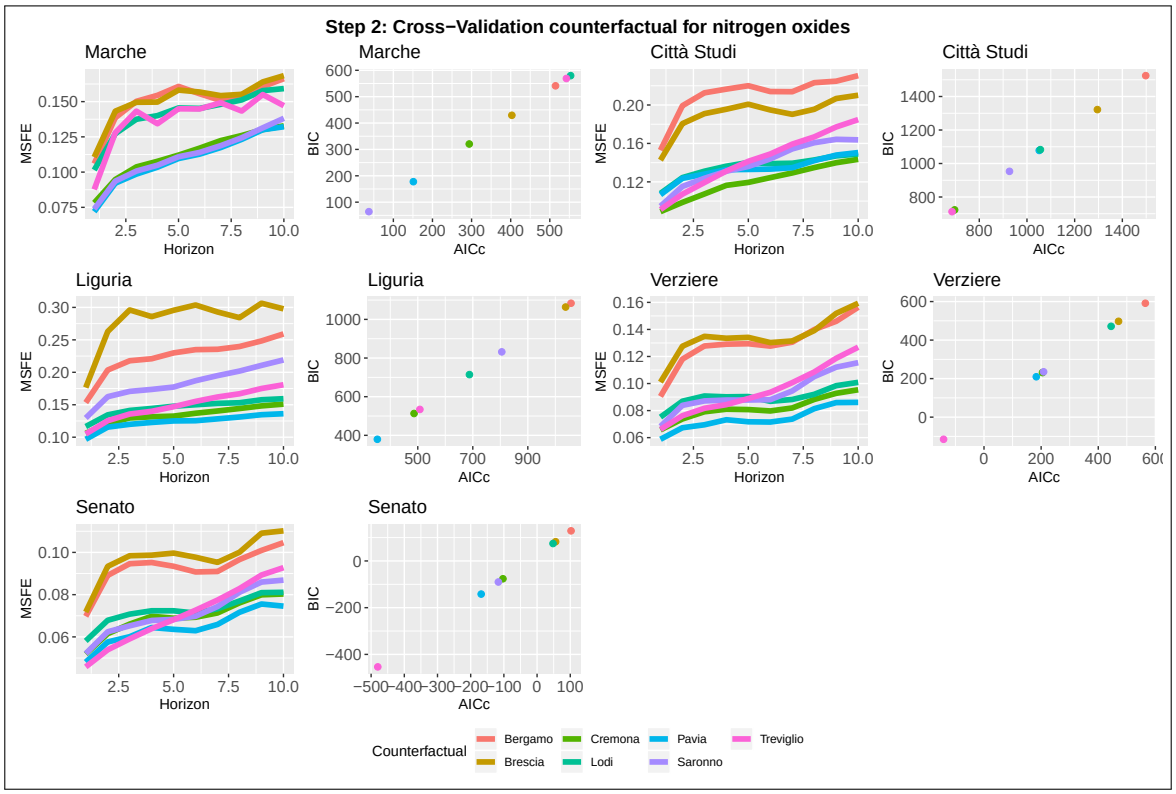


FIGURE 3.12: Model selection-Step 2-NO_x. Counter-factual term selection for the five nitrogen oxides stations in Milan. Left panel: 10-steps-ahead MSFE in log-scale as function of the candidate. Right panel: AICc and BIC pairs for each model.

3.5.4 Basic Structural Model and policy intervention

In this section, we show the numerical results obtained using state space modeling to estimate both permanent and transitory effects generated by the introduction of Area B controlling for local weather conditions, anthropogenic effects, and common areal trends.
The maximum likelihood estimates of the coefficients and the components' variances for the five air quality monitoring stations installed in Milan are reported in Tables 3.5 and 3.6. The results appear to be coherent both for nitrogen oxides and nitrogen dioxide. First, the models identify statistically significant and positive coefficients for the counter-factual term, highlighting its capability to capture socio-economic and climatic factors common to neighbouring areas and coherent with the expected sign. Second, weekends and holidays exert a negative effect on concentration levels probably due to the reduction in the movements and productive activities of the city in those days. Their interactions are almost everywhere not statistically significant but with a positive sign and always less than the sum of the individual effects of the weekend and holidays. This fact underlines how the holiday weekends enjoy more contained effects of emission reductions compared to generic weekends of the year. Third, as to be expected, winds blowing from the West (Q_{SW} and Q_{NW}) greatly reduce the amount of pollutants all over the city with peaks over 40% to 50%.
The short-term impacts adjusted for common anthropic and weather factors are summarized in Table 3.7, which shows the estimated permanent and transitory effects for each station in Milan expressed in logarithmic scale, hence interpretable as relative variations in concentrations levels. None of these two coefficients identifies an improvement of the considered pollutant

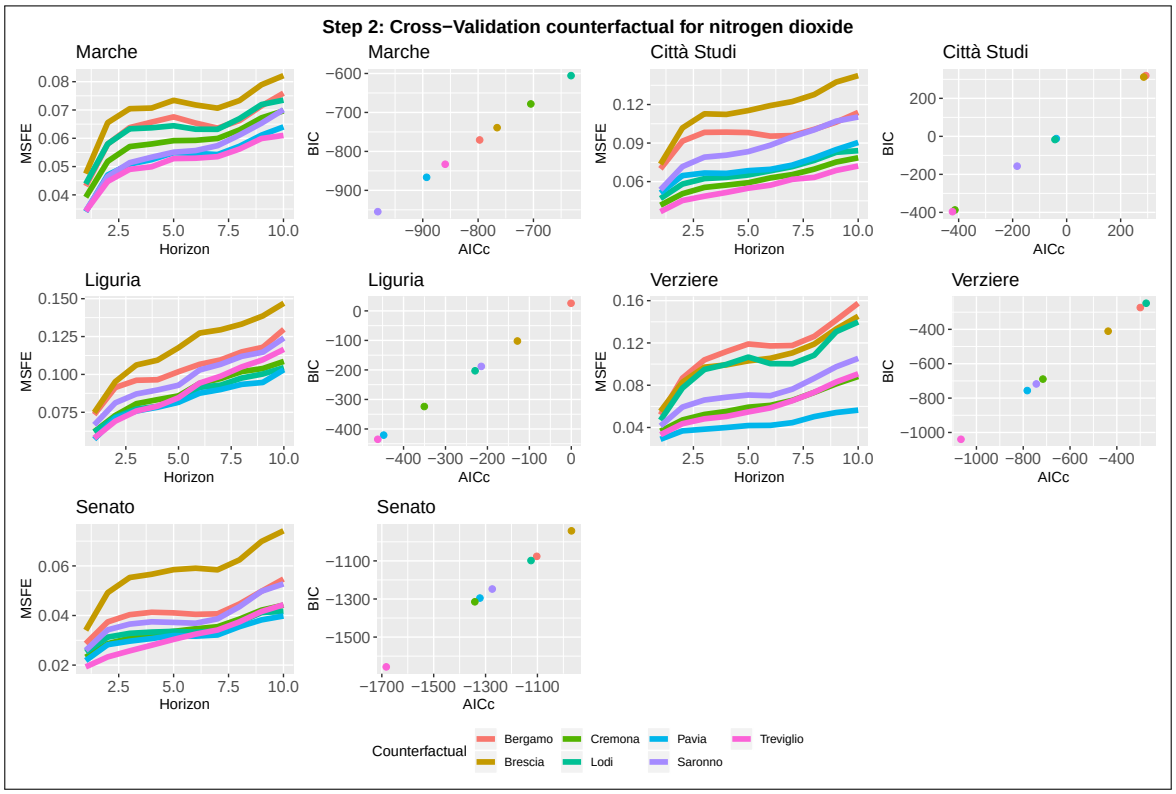


FIGURE 3.13: Model selection-Step 2-NO_2. Counter-factual term selection for the five nitrogen dioxide stations in Milan. Left panel: 10-steps-ahead MSFE in log-scale as function of the candidate. Right panel: AICc and BIC pairs for each model.

concentrations. Moreover, the permanent effect (δ_1) is always positive and in some cases moderately statistically significant. This means that, compared to the generally decreasing areal trend, Milan air quality went worst. It is worth observing that the most significant results are obtained at the Senato station, which is located in the already existing Area C, hence already subject to some car traffic restrictions.
Such a result could be justified by the presence of multiple causes. As a first justification, we are approaching the initial phase of a progressive policy and the time elapsed since its outset may be too short to assess any significant impacts on pollutant levels. This fact is consistent with the forecasts expected by the municipality of Milan about the reductions in nitrogen oxide levels; in fact, the first significant reductions should be observed starting from 2022 (Milano, 2018). Furthermore, since we are dealing with limitations to human behavior and social perception of new norms, it is not always clear how agents adapt to changes. The deterioration in the air quality of the centre could be linked to new traffic congestions in that area or to a panic shock of drivers, who need time to understand the functioning of the restrictions and adapt their behavior, exactly as in situations mismanagement of individual and organizational changes (Perlman and Takacs, 1990; Elrod and Tippett, 2002). Eventually, the recent climate changes and the extreme weather conditions that affected the Po valley, such as temperatures higher than the seasonal average and extreme atmospheric events, could increase the noise present in the data and thus mask the real repercussions of the limitations.

TABLE 3.3: Model selection-Step 3-NO_x: Best subset of covariates using backward-forward stepwise algorithms for NO_x.

	Marche		**Verziere**		**Senato**		**Liguria**		**Citta Studi**	
	AIC	***BIC***	***AIC***	***BIC***	***AIC***	***BIC***	***AIC***	***BIC***	***AIC***	***BIC***
Holidays	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Week-End	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Saturday:Holidays	✓		✓		✓		✓		✓	
Sunday:Holidays	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wind speed Q_{NE}										
Wind speed Q_{SE}							✓			
Wind speed Q_{SW}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wind speed Q_{NW}	✓	✓	✓				✓		✓	✓
Temperature										
Rainfall	✓		✓		✓				✓	
Global radiation			✓		✓					
Humidity										

Note: symbol ✓indicates that the regressor is selected within the best subset of covariates.

3.6 Conclusions about the Area B case study

In this chapter we analyzed the early-stage effects on air quality of the new traffic policy in Milan, the so-called Area B. The concentrations of nitrogen oxides (NO_x) and nitrogen dioxide (NO_2), which are mainly primary pollutants with high reactivity to changes in emission sources, have been considered as proxies of pollution emissions. Recalling that the policy started the 25th February 2019, we considered a temporal sample from the 1^{st} January 2014 to the 30^{th} September 2019.

The policy evaluation started with a preliminary analysis aiming at estimating the variations in observed concentrations of oxide levels using a temporal treatment-control. In this framework the concentrations of NO_x and NO_2 in 2019 were considered as treatment, while the average daily observation of the period 2014-2018 were the control data. To account for possible extreme values and the natural right skewness characterizing air quality data we supported the estimations by bootstrap resampling algorithm for uncertainty evaluation.

We then modelled the concentrations series for each air quality station in Milan using a Basic Structural Model (BSM) in State Space form and estimated through maximum likelihood. Model components and covariates were selected using an innovative three-stage model selection algorithm which evaluates both the forecasting performance of the models and their goodness-of-fit. In particular, our algorithm computes the cross-valited mean square forecasting error using a rolling-window cross-validation approach for time series. To adjust for possible unknown socio-economic effects we included ad covariate the observed oxide levels observed in a large urban center with similar urban conditions to Milan but not affected by the Area B intervention. Also this counterfactual have been selected using the cross-validation algorithm.

The first research hypothesis we posed investigated the presence of a significant effect of the

TABLE 3.4: Model selection-Step 3-NO_2: Best subset of covariates using backward-forward stepwise algorithms for NO_2.

	Marche		**Verziere**		**Senato**		**Liguria**		**Citta Studi**	
	AIC	***BIC***	AIC	***BIC***	***AIC***	***BIC***	***AIC***	***BIC***	***AIC***	***BIC***
Holidays	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Week-End	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Saturday:Holidays	✓		✓		✓		✓	✓	✓	
Sunday:Holidays	✓	✓	✓	✓	✓	✓	✓		✓	✓
Wind speed Q_{NE}										
Wind speed Q_{SE}										
Wind speed Q_{SW}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wind speed Q_{NW}	✓								✓	
Temperature										
Rainfall			✓		✓					
Global radiation	✓		✓		✓				✓	
Humidity	✓								✓	

Note: symbol ✓indicates that the regressor is selected within the best subset of covariates.

road traffic limitations on the air quality of the city. As a first point, the preliminary results show that concentrations during spring and summer 2019 are lower than during the same seasons in the previous five years, hinting for a reduction effect due to the policy. On the other side, a similar reducing trend has been observed in various neighbouring cities around Milan, which belong to a homogeneous meteorological, social, and economical cluster within the Po valley. Their similar behavior is used here as an areal common trend capturing both weather and anthropogenic components. Our approach, which adjusts for local weather conditions and the areal common trend, does not provide a further reduction effect for any station comparing to this trend. Instead, in Senato station, which is inside the historical city centre and was already covered by Area C, the estimates provide a strong, but moderately statistically significant, increase for both considered pollutants. This is coherent with the fact that the restriction introduced is very limited as it concerns just some classes of old vehicles, which are a small percentage of the entire vehicle pool, both in terms of number of cars and emissions.
Since the first research hypothesis is confirmed just to a minor extent and with an opposite sign with respect to what was expected, the second research hypothesis, concerning the territorial homogeneity of the estimated effects, assumes now only a technical scope. It is confirmed just for what concerns the positive direction of the changes, but not for their significance. In fact, among all the estimated permanent effects, only Senato station is significant at 5%. Moreover, the estimated transitory effects are always not significant at any confidence level.
The above facts hint that, compared to the common trend of the considered area, Milan air quality is improving slowly, and, in this sense, the first phase of Area B seems to have a negative effect on air quality. Due to the limited scope of this first phase and its progressiveness, it is not unexpected to find a limited or a zero effect. Nonetheless, the negative effect needs some more explanations.
Although finding the ultimate motivation for this is not the aim of this case study, a discussion follows. Firstly, the statistically significant increase found is limited in space and is located

TABLE 3.5: ML estimates of BSM parameters and variances for NO_x.

	Parameter	***Marche***	***Citta Studi***	***Liguria***	***Verziere***	***Senato***
log(Pavia)	θ	0.51 ***	0.93 ***	0.73 ***	0.66 ***	0.57 ***
		(0.01)	(0.02)	(0.02)	0.02	(0.01)
Holidays	ϕ_1	−0.06 ***	−0.02	−0.05 *	−0.10 ***	−0.09 ***
		(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
WeekEnd	ϕ_2	−0.11 ***	−0.09 ***	−0.09 ***	−0.17 ***	−0.15 ***
		(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Saturday:Holidays	ϕ_3	0.10	0.17	0.04	0.10	0.14 ***
		(0.08)	(0.13)	(0.10)	(0.14)	(0.05)
Sunday:Holidays	ϕ_4	0.09 **	0.09	0.12 *	0.08	0.10 ***
		(0.05)	(0.08)	(0.06)	(0.05)	(0.03)
WindSpeed Q_{SW}	ϕ_5	−0.44 ***	−0.20 ***	−0.56 ***	−0.30 ***	−0.27 ***
		(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
WindSpeed Q_{NW}	ϕ_6	−0.31v***	−0.27 ***	−0.12 ***	−0.20 ***	−0.12 ***
		(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Level variance	σ^2_η	0.0047	0.0065	0.0037	0.0038	0.0027
Slope variance	σ^2_ζ	0.0000	0.0000	0.0000	0.0000	0.0000
Seasonality variance	σ^2_ω	0.0000	0.0000	0.0000	0.0000	0.0000
Error variance	σ^2_ε	0.0298	0.0745	0.0437	0.0370	0.0308

Note 1: values in parenthesis are standard errors. *Note 2*: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

inside the previously introduced restricted Area C. It may be possible that this further restriction increased congestion of public transport buses, which are often very old vehicles, or to the aforementioned adaptation shocks. This could explain only a part of the results. In fact, this first point is also related to the other sources of nitrogen oxides. According to INEMAR INEMAR ARPA Lombardia Settore Aria, 2020, road traffic is about 68% of the total emissions. Hence, a transition to house heating green techniques slower in Milan comparing the other considered cities could have an influence on this result. Moreover, also the other stations experienced a comparative deterioration of air quality and the second-worst station is Città Studi, which is an urban background station, hence with limited relation to local traffic congestion. Second, the increase due to road traffic may have temporal dynamics. Since the traffic restriction is limited to business hours, there may be an increase in congestion early in the morning and in the late evening, affecting the daily average.
In conclusion, although environmental protection policies are in general a fundamental step for sustainability improvement, in some cases, they may not be sufficient or their implementation may be misleading. In our case, we considered only the early-stage of a policy, which is progressive in time. Hence, the results of this paper may be regarded as physiological, provided that they characterize only the initial part of the policy implementation and are improved soon. It follows a recommendation to the municipal government to develop the policy more strongly. Additional research could be developed in the future. In particular, the effect on traffic congestion inside Area C could be investigated further using historical data related to the vehicle movements crossing the access points. Moreover, the use of a multivariate approach, which includes other pollutants such as PM_{10} and $PM_{2.5}$, and spatio-temporal modeling could highlight hidden effects, which are not visible considering the single stations. Eventually, the extension to hourly data could consider both the presence of intra-daily effects and explaining the spatial

TABLE 3.6: ML estimates of BSM parameters and variances for NO_2.

	Parameter	***Marche***	***Citta Studi***	***Liguria***	***Verziere***	***Senato***
log(Pavia)	θ	0.36 ***	0.85 ***	0.69 ***	0.65 ***	0.55 ***
		(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Holidays	ϕ_1	−0.06 ***	−0.08 ***	−0.08 ***	−0.10 ***	−0.08 **
		(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Week-end	ϕ_2	−0.06 ***	−0.10 ***	−0.09 ***	−0.14 ***	−0.11 ***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Saturday:Holidays	ϕ_3	0.06	0.18 ***	0.06	0.14 **	−0.09 ***
		(0.05)	(0.08)	0.07	(0.06)	(0.02)
Sunday:Holidays	ϕ_4	0.08 ***	0.11 ***	0.11 ***	0.06	0.10 ***
		(0.03)	(0.05)	(0.04)	(0.03)	(0.03)
WindSpeed Q_{SW}	ϕ_5	−0.35 ***	−0.13 ***	−0.42 ***	−0.23 ***	−0.18 ***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
WindSpeed Q_{NW}	ϕ_6	−0.17 ***	−0.16 ***	−0.08 ***	−0.13 ***	−0.08 ***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Level variance	σ^2_η	0.0051	0.0065	0.0046	0.0044	0.0024
Slope variance	σ^2_ζ	0.0000	0.0000	0.0000	0.0000	0.0000
Seasonality variance	σ^2_ω	0.0000	0.0000	0.0000	0.0000	0.0000
Errors variance	σ^2_ε	0.0093	0.0282	0.0182	0.0154	0.0114

Note 1: values in parenthesis are standard errors. *Note 2*: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

dynamics related to traffic.

TABLE 3.7: Estimated permanent and transitory effects in log scale on NO_x and NO_2 for each station.

Stations	**Effect**	**Nitrogen Oxides**			**Nitrogen Dioxide**		
		Estimate	**S.E.**	***t*-Statistic**	**Estimate**	**S.E.**	***t*-Statistic**
Senato	Perm. eff. δ_1	0.38	0.19	2.03 **	0.29	0.12	2.40 **
	Trans. eff. δ_0	-0.12	0.16	-0.76	-0.14	0.11	-1.25
Verziere	Perm. eff. δ_1	0.26	0.21	1.27	0.22	0.15	1.50
	Trans. eff. δ_0	-0.01	0.18	-0.05	-0.06	0.14	-0.40
Liguria	Perm. eff. δ_1	0.12	0.22	0.54	0.20	0.16	1.25
	Trans. eff. δ_0	-0.02	0.19	-0.10	-0.10	0.15	-0.67
Marche	Perm. eff. δ_1	0.15	0.19	0.80	0.23	0.13	1.82 *
	Trans. eff. δ_0	-0.11	0.17	-0.63	-0.19	0.13	1.56
Citta Studi	Perm. eff. δ_1	0.35	0.29	1.19	0.25	0.19	1.30
	Trans. eff. δ_0	0.08	0.25	0.30	0.02	0.18	0.10

Note 1: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 4

The impact of the lockdown restrictions on air quality during COVID-19 pandemic in Lombardy

Based on:

Maranzano, P., Fassò, A.
'The impact of the lockdown restrictions on air quality during COVID-19 pandemic in Lombardy, Italy'

4.1 Introduction

Environmental agencies and scientists around Europe have reported that COVID-19 lockdown caused an extended environmental clean-up. Considering air quality, we focus on the Lombardy region (Northern Italy), which is at the same time the most populous region and the area most affected by COVID-19 in Italy. Lombardy is also one of the most polluted areas in the European Union. The central research hypothesis concerns if and how the first-wave restrictions imposed during the 2020 spring have improved the air quality in Lombardy and if the improvements are similar throughout the territory. To answer these questions, we use weekly data from January 2015 to mid-June 2020 for 74 ground monitoring stations and provided by the regional environmental protection agency (ARPA Lombardia). We estimate an autoregressive time series model with exogenous covariates (ARX) to assess the combined impact of meteorology, seasonality, trend, and lockdown on the NO_2 concentrations at each monitoring site. We also propose using the LASSO algorithm to select the set of relevant covariates to model the concentrations and then estimate the effect of lockdown restrictions with a maximum likelihood post-LASSO estimator. Statistical modelling confirms a generalised NO_2 reduction due to the lockdown throughout the whole region, despite considerable variability due to the morphological and geographical heterogeneity of Lombardy. Compared to the observed average variations, the estimated lockdown impacts are mitigated by meteorology and natural trends. Expectantly, the most significant and remarkable NO_2 reductions have been estimated near urban and congested areas and in the proximity of industrialised sites.

4.2 Introduction

With the global COVID-19 pandemic, the quality of the air we breathe every day has acquired a role of primary importance worldwide, both in scientific research and in daily media storytelling.
Air pollution is seen as having both an active and passive role in managing the epidemic. On the one hand, many research studies have provided evidence of strong correlations between poor air quality and COVID-19 spread, especially in large urban centres with severe air quality conditions. On the other hand, air quality has improved significantly worldwide due to partial and total lockdown measures and to human mobility restrictions imposed on citizens to avoid spreading the virus.
This paper analyses the case of Lombardy, the economic and financial centre of northern Italy, to assess how restrictive measures to control the spread of the COVID-19 virus have influenced the concentrations of oxides, in particular nitrogen dioxide (NO_2), in the atmosphere. We focus on the first-wave lockdown restrictions imposed on citizens and their productive activities between 9th March and 18th May 2020 all over the country. The choice of NO_2 as the main pollutant of interest is justified by physical-chemical reasons well known in the literature. Oxides are classified as both primary and secondary pollutants emitted mainly by anthropogenic activities, such as heating systems, motor vehicle traffic, power plants, industrial activities and combustion and therefore directly affected by the lockdown restrictions. In particular, the emission inventory for Lombardy (INEMAR ARPA Lombardia Settore Aria, 2020) estimates that road traffic is responsible for 51% of the annual NO_X emissions in the Lombardy region and the 65% in the metropolitan area of Milan.
Moreover, the choice to analyse NO_2 is also reasonable as it is a pollutant that responds immediately to emission shocks, while other airborne pollutants, such as the atmospheric particulate matter, have more complex and slower reactions. Therefore, nitrogen dioxide is a suitable candidate for a preliminary investigation into the impacts of a severe limitation to anthropogenic activities, especially vehicle traffic, on air quality. We pose some questions: have the restrictions improved the air quality in the region? Are there similar improvements throughout the region? Can we measure the variation of airborne pollutants related explicitly to car traffic?
Although the measures may have generated socio-economic issues in the population, they may have improved air quality in an area that suffers from longstanding environmental problems. Expectations were for substantial reductions in concentrations generalised to all areas of the region. We implemented a time series modelling approach for the available air quality monitoring stations, controlling for the effect of local meteorology, seasonality, and trends.
We propose to use a data-oriented approach based on the combination of time series and statistical learning methods to simultaneously address the statistical variable selection issue and the model estimation using a penalised least squares approach. We estimate the effects of local meteorology, long-term weather trends, and lockdown restrictive measures on NO_2 concentrations in Lombardy implementing an order-one autoregressive with exogenous covariate model, namely ARX(1) model, for each considered station. These models express the response time series as a linear combination of both autoregressive components and exogenous variables and are easily estimated using the maximum likelihood approach. The ARX form allows to estimate the effect of covariates on the response and at the same time allows modelling the temporal dynamics of the dependent variable explicitly. Each time series is modelled as a function of several regressors and one lag of the dependent variable, also known as the ARX(1) model. In our application, the dependent variable is the weekly concentration of NO_2 observed at each ground site, while the covariates include local meteorology, long-run trends and the lockdown. The lockdown's estimated effect will depend both on the estimated coefficient for the lockdown event and the autoregressive dynamics of the dependent variable. The most suitable model to

describe the temporal evolution of concentrations for each control unit is determined using the least absolute shrinkage and selection operator (LASSO) algorithm. To validate the correct specification assumptions, we will also provide several diagnostic checks on the residuals. Finally, the estimated effects are discussed and contextualised concerning their geographical location, the area surrounding the station and the type of station they are associated with.
The rest of the paper is organised as follows. In Section 4.3, we report and discuss the most recent contributions to the COVID-19 literature quantifying the impact that the lockdown restrictions on the levels of airborne pollutant concentrations. Section 4.4 reports the leading facts connected to the pandemic in Italy and Lombardy, and describes the environmental context of the region, focusing on the primary airborne pollution sources and the geographical features. In Section 4.5, we introduce and describe the data set. In Section 4.6, we present the statistical modelling approach by illustrating the main characteristics of the ARX(1) model, how the lockdown effect is computed and the proposed model selection algorithm. In Section 4.7, we discuss the LASSO algorithm's output, the fitting of the models and some diagnostic checks for the regression residuals. In Section 4.8, we discuss the estimated NO_2 variations due to the lockdown restrictions' main findings regarding the estimated lockdown variations, analysing the gaps between the observed and the estimated reductions and contextualising the estimates with respect the geography of Lombardy and its socio-economic structure. Finally, Section 4.9 sums-up the discussion, giving some concluding remarks and outlining some possible extensions to be developed in future works.

4.3 State-of-the-art research on the relationship between the COVID-19 pandemic and air quality in the world

4.3.1 The active role of air quality: how pollution has influenced the spread of the virus

Recent research about the relationship between air pollutants and the number of cases and mortality rates due to COVID-19 have focused their attention on those areas mostly affected by the epidemic, such as China (Jiang et al., 2020; Zhu et al., 2020a), Europe (Ogen, 2020), and Italy (Coccia, 2020; Fattorini and Regoli, 2020), with particular attention on northern Italy (Bontempi, 2020; Martelletti and Martelletti, 2020; Zoran et al., 2020a; Zoran et al., 2020b) and the United States (Adhikari and Yin, 2020; Wu et al., 2020). All these contributions confirm the fact that poor air quality, i.e. high air pollution levels, is associated to excess in mortality rates and higher cases due to COVID-19, especially when the polluting elements are fine particulates ($PM_{2.5}$ and PM_{10}), nitrogen oxides (NO_2 and NO_X), and carbon monoxide (CO). Zhu et al., 2020a explored the relationship by considering data on 120 Chinese cities from January to February 2020 and choosing an incubation period of 14 days for COVID-19. They used a Generalized Additive Model with Gaussian distribution family to estimate the associations between the moving average concentrations of air pollutants in a 0-14-day lag period and COVID-19 confirmed cases. Their findings suggest that an increase of 10-$\mu g/m^3$ in $PM_{2.5}$, PM_{10}, O_3, and NO_2 levels are significantly associated with an increase in daily confirmed new cases of 2.24%, 1.76%, 4.76%, and 6.94% respectively, and that a 1-mg/m^3 increase in CO levels is significantly associated with a 15.11% increase in new cases. Other researchers have investigated the correlations between pollutants and spread of the virus by using both contemporary data, i.e. the quality of air observed during the pandemic, as in the Chinese cases, and long-term exposure to severe air quality. In the latter case the tests were conducted by using the average air pollutant levels for years prior to the crisis (Fattorini and Regoli, 2020), or the number of days exceeding the limit set by the legislative standards (Martelletti and Martelletti, 2020; Coccia, 2020). These

approaches led to conclusions similar to previous ones, confirming that high spread of the contagious virus in some areas was likely to be linked to long-term poor air pollution conditions. For example, in the case study presented by Fattorini and Regoli, 2020, following the Chinese descriptive approach, the results showed positive correlations of more than 50% between the number of COVID-19 cases and long-term exposure to pollutants such as NO_2, particulates, and ozone. Moreover, in a study by Coker et al., 2020, the authors implemented an aggregate counting-process regression model for the period 1st January - 30th April 2020 in all the municipalities of northern Italy, i.e. 4393 cities across eight regions. Their findings suggest that a unitary increase in $PM_{2.5}$ concentration ($\mu g/m^3$) is associated with a 9% increase in COVID-19 related mortality.
Furthermore, environmental factors such as temperature, albeit with discordant signs, and relative humidity are also highly correlated with the number of confirmed cases (Jiang et al., 2020; Li et al., 2020). The correlation among meteorological parameters, COVID-19 spread and air quality in Lombardy has been investigated by (Lolli et al., 2012). They considered multiple weather measures, such as temperature, humidity, pressure and wind speed, and airborne pollutant concentrations in order to correlate the observations with the spread of the virus. According to their approach, a reliable variable for modelling the virus spread is the anomaly of the daily number of Intensive Care Unit (ICU) hospitalized patients. The authors considered both linear and non-linear correlation measures. Their findings suggest that a positive correlation exists among anomaly of the daily number of intensive care units and $PM_{2.5}$, while the correlation is not statistically significant for NO_2 levels. Moreover, temperature, absolute humidity, water vapor are negatively correlated with the virustransmission, whereas wind speed seems to be uncorrelated. Another case study has been presented by Fazzini et al., 2020, which analysed the spread of COVID-19 in Lombardy during the period 1th March to 20th April. The authors considered both air quality indicators, i.e. sulphur dioxide and nitrogen dioxide, and weather measure, such as minimum and maximum temperature, relative humidity, wind direction and speed, other than municipal epidemiological parameters. Their results show that there is no climate variable that is acceptably correlated with the temporal evolution of the pandemic, with the exception of relative humidity. However, the influence of climatic parameters on COVID-19 remains extremely controversial and discussed among researchers.
In a study by Blangiardo et al., 2020, the authors investigated the spatio-temporal differences in excess mortality rates during the pandemic in Italy at municipality level. The models included both demographic and climate variables, such as temperature. According to their findings Lombardia showed higher mortality rates than expected from the end of February, while North-West and North-East regions showed higher mortality rates from the beginning of March. Controlling for local temperature, differences in the spatio-temporal distribution of excess mortality rates were also observed at municipal levels.
In addition to environmental factors, anthropogenic factors, such as human mobility and networking, may also have allowed the virus to spread. According to Vasquez-Apestegui et al., 2020, the amount of intensive use of food markets by the citizens of Lima, Peru, could have affected the spread of the virus in the districts of the city, making the number of infections heterogeneous. Similar considerations were deployed in a study by Zhu et al., 2020b, which jointly analysed human mobility data, measured as intra-city migration, and airborne pollution in China. The paper reported that, at a lag of 14 days, a unit increase in the proposed human mobility index was associated with a 6.45% increase in daily confirmed COVID-19 cases and that air pollution significantly mediated around 19.47% of this association.
To further explore this specific topic, we suggest the detailed reviews of scientific papers published during the first half of 2020 by Copat et al., 2020 and Domingo et al., 2020. Both propose a systematic review of the literature on the relationship between COVID-19 disease spread and

air quality, describing the contextual effects of environmental conditions and chronic exposure to air pollutants.

4.3.2 The passive role of air quality: did we breathe better air quality during the COVID-19 lockdown?

The passive role played by the air quality refers to the positive effects on the air we breathe generated by lockdowns and restrictions imposed around the world. Scientists and researchers are compact in stating that, in general, air quality has significantly improved everywhere, reaching minimal levels of pollution. This fact holds particularly true in large urban centres and densely populated areas, which are often affected by poor air quality.
The largest improvements were registered in Europe, where oxides and particulate concentrations were reduced by around 40% to 70% as in Spain (Baldasano, 2020; Tobías et al., 2020), Italy (Collivignarelli et al., 2020), and France (Connerton et al., 2020), and in Brazil, where concentrations were reduced by 50% (Nakada and Urban, 2020). Smaller, but still significant, reductions were reported for the United States (Berman and Ebisu, 2020; Zangari et al., 2020) and other Asian countries, such as Kazakhstan (Kerimray et al., 2020) and India (Sharma et al., 2020; Singh and Chauhan, 2020), in which both the average levels of both oxides and particulates fell from 20% to 30%. All these papers measured the impact of lockdown restrictions on air quality controlling for weather conditions, possible counterfactual terms, and long-run trends in the concentrations.
In some cases, including in the models temporal trends as in Zangari et al., 2020 for New York City, or Cameletti, 2020 for the city of Brescia (northern Italy), no significant difference between the years was found, suggesting that the reduction in concentration levels in 2020 was similar to that measured in the previous five years.
However, in other studies which analysed the impact both on macro-areas, such as US counties (Berman and Ebisu, 2020), and on micro-areas, such as Sao Paulo in Brazil (Nakada and Urban, 2020), Los Angeles, New York and Paris (Connerton et al., 2020), or Barcelona (Baldasano, 2020; Tobías et al., 2020), the differences proved to be significant while including meteorological factors and time trends. According to all these studies, all the common air pollutants were affected by significant reductions: nitrogen dioxide (NO_2) in Barcelona, Sao Paulo, Los Angeles, New York, Paris, and Almaty (Kazakhstan) fell by 50%, 30%, 38%, 25%, 39%, and 35% respectively; particulate matter concentrations fell approximately by 31%, 12%, 37%, 36%, and 28% and carbon monoxide (CO) decreased by 40% in São Paulo, 49% in Almaty, 24% in Los Angeles, 19% in New York, and 67% in Paris. By contrast, ozone (O_3) levels increased in most parts of the World: it increased by 30% in São Paulo, 7% in New York and 12% in Paris. At a national level, prominent air quality improvements were detected in many countries. As example, in the United States NO_2 and $PM_{2.5}$ reduced at county level of 25.5% and 4.45% respectively (Berman and Ebisu, 2020) and in United Kingdom (Higham et al., 2020; Ropkins and Tate, 2020), which registered reductions in oxide concentrations of -32% (NO_2), -38% (NO_x) and -50% (NO). Or also in India, which registered a general reduction in $PM_{2.5}$, PM_{10}, CO, and NO_2 levels by 43%, 31%, 10%, and 18% during the lockdown period (Sharma et al., 2020).
The presence of strongly contrasting results and estimates of the impact of the lockdown on air quality, resulting from the inclusion of temporal and meteorological covariates or the use of oversimplified methodologies (such as correlation or simple linear models), opens at least a couple of discussions: the first linked to the role of statistical uncertainty in the models and the choice of appropriate instrument, and another on a reasoned selection of influential variables in the analysis. Given these premises, we propose a data-oriented approach based on time series methods to address simultaneously the tasks of statistical variable selection and estimation using penalized regression models, such as the LASSO algorithm. This approach

allows simultaneous identification of the meteorological factors relevant for the explanation of the observed concentrations and estimation of the effects of meteorology, long-term weather trends, and restrictive measures on NO_2 concentrations during the lockdown.

4.4 Air quality in Lombardy during the COVID-19 lockdown

For many years, the Po Valley in northern Italy has been ranked among the most polluted areas in Europe and the world (European Environmental Agency, 2019a). According to the report about air quality in 2018 by the European Environment Agency (EEA), in Europe around 3.9 million people live in areas where the limits of the air pollutants are frequently exceeded. Among these, the 95% live in the Po Valley. Moreover, Italy is at the first place in Europe as concerns the number of premature deaths attributable to exposure excess of nitrogen dioxide (around 14600 victims per year) and ozone (around 3000 victims per year) and at second place after Germany as regards deaths due to fine particulates ($PM_{2.5}$) pollutant. As represented by the upper panel of Figure 4.1, which shows the average NO_2 concentrations in Europe in 2018, the Po Valley area is easily identifiable since it reports dark spots (high concentrations) all over its surface. The main physical and geomorphological features of the Po Valley are depicted in the lower panel of 4.1. The area is surrounded by a C-shape mountain range, which acts as a barrier and prevents wind movement from the West. As a result, the wind speed on the Po Valley is among the lowest in Europe, about 1.5 m/s on average, causing a high accumulation of smog and pollution close to the ground. According to a recent simulation study by Raffaelli et al., 2020, if Po Valley had the same meteorological conditions typical of central-northern Europe and kept the same emission levels, average monthly concentrations of PM_{10} and NO_2 would be lowered by 60 to 80% compared to concentration levels of 2013. Consequently, it is more difficult for the Po Valley region to comply with international air quality standards than other EU and non-EU member states. The Lombardy region is the economic and financial centre of Po Valley. It is organised in eleven provinces and is home to more than ten million inhabitants. The region holds the highest gross domestic product per inhabitant of the country (Regional Statistical Yearbook, 2017a). In Lombardy are located many industrial facilities, as well as small and medium enterprises, and the road transport is an essential component of economic structure. Lombardy is also the most densely populated region of Italy, with large and very dense urban agglomerations. The average population density in Lombardy is around 419.9 inhabitants/km^2, whereas at national level it is 200 inhabitants/km^2 (Regional Statistical Yearbook, 2017a). This fact also reflects on the spatial distribution of airborne emissions. In fact, the four largest and populated provinces, i.e. Milano (MI), Monza (MB), Bergamo (BG) and Brescia (BS), generated the 52% of total emissions of NO_x and the 51% of particulate matters in 2017. Figure 4.2 shows that the region can be geographically divided into three zones. The mountain range of the Alps in the North, the sloping foothills in the mid-north, and the flat southern area. The strong heterogeneous physical conformation of the territory influences the socio-economic organisation of the society and, consequently, the airborne pollution levels in the atmosphere. The mountainous area is sparsely populated and less trafficked, the central area between the hills and the plain is densely urbanised and industrialised, while the southern rural area is less densely populated and oriented towards agriculture and farming. The major urban centres (the metropolitan area of Milan and the cities of Monza, Bergamo and Brescia) are located in the central foothills belt. The poor air quality in the proximity of large urban centres is mainly due to the emissions of oxides (NO_X and NO_2), carbon (CO and CO_2), and sulphur dioxide (SO_2), produced by industry, heating plants, and road transport. At the opposite, the southern rural provinces are dominated by emissions of pollutants produced by agricultural and breeding activities. In particular, ammonia (NH_3), methane (CH_4), and fine particulates

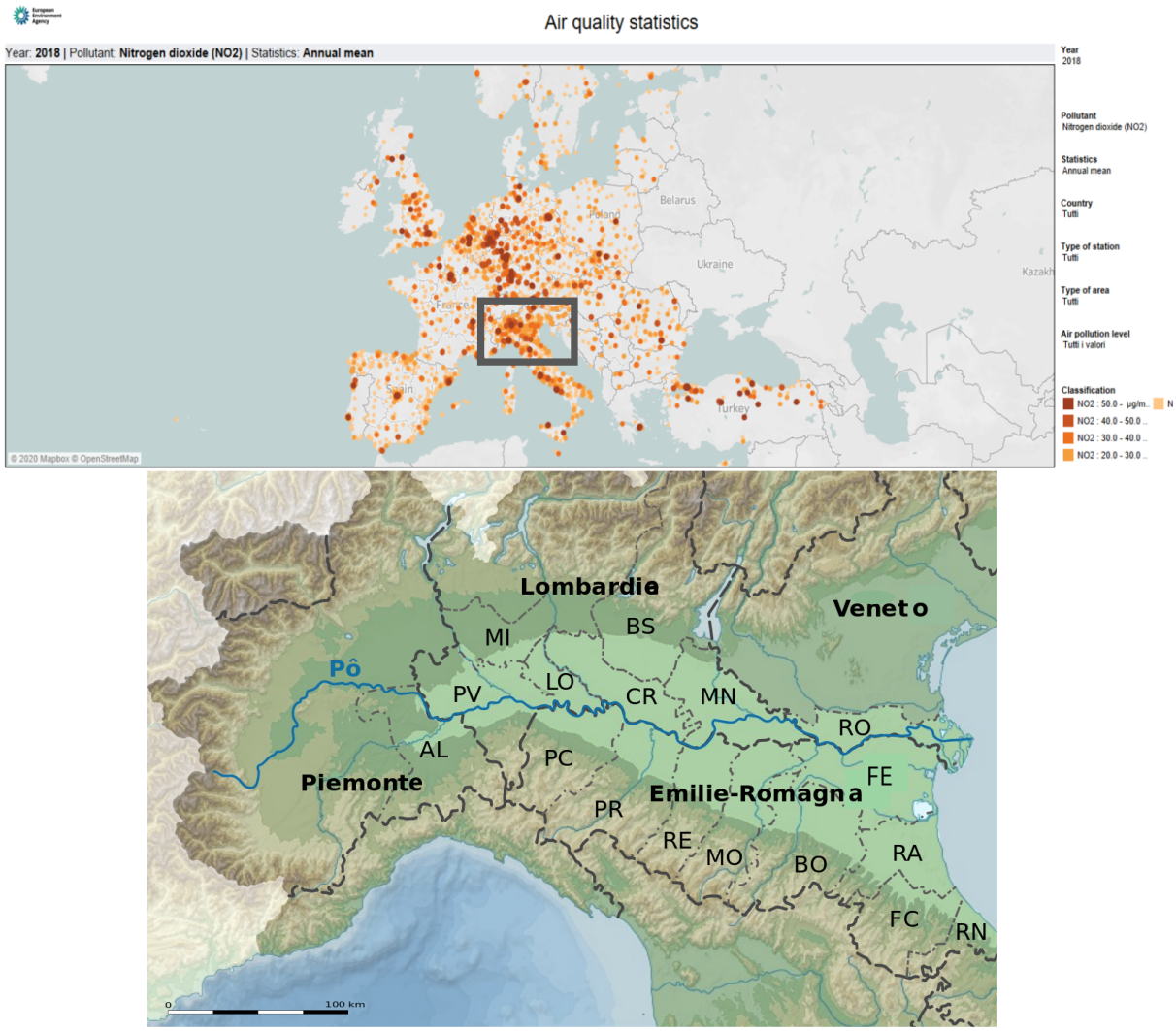


FIGURE 4.1: Average NO_2 concentrations in Europe in 2018 (upper panel) and physical map of Po Valley (lower panel).

($PM_{2.5}$) emissions. Furthermore, according to the emission inventory for Lombardy (INEMAR ARPA Lombardia Settore Aria, 2020), in 2017 the sum of industrial combustion plants, non-industrial combustion plants (i.e. house heating), and road transport represented more than 73% of particulate matter emission sources and more than 76% of total nitrogen oxide emissions in the region. In the metropolitan area of Milan, road traffic alone was responsible for the 65% and 69% of the total emissions of NO_X and CO respectively. Overall, its unfavourable geographical context, an aggressive land use, climate characterstics and high pollutant emissions turn in the accumulation of toxic elements in the atmosphere.
As will be shown in section 4.8.2, estimates of the effects of the restrictions on nitrogen oxide concentrations will depend strongly on the type of area surrounding the ground station and by the geomorphological structure of the territory and its reliefs. In particular, it will be shown that the proximity of the stations to congested or urban areas leads to substantial reductions in concentrations during the lockdown phase, while in agricultural or rural areas the reductions will be much smaller and, in several cases, zero.

(A) Location of Lombardy in Europe

(B) Physical map of Lombardy

FIGURE 4.2: Location of Lombardy in Europe (left panel) and physical map of Lombardy (right panel).

Eventually, seasonal and meteorological variations and human activities have a relevant influence on air quality in Lombardy, as they act as confounders when dealing with the assessment of the lockdown effect on airborne pollutant concentrations. Recent studies of the air quality in Lombardy, such as the case study by Maranzano et al., 2020 about the city of Milan or the analysis of Carugno et al., 2017 concerning the main urban centres of Lombardy, highlighted that the whole region has been experiencing a significant and constant reduction in airborne pollutant concentrations since the early 2000s. To show such decreasing pattern, we reported in Figure 4.3 the average NO_2 concentrations observed at a highly trafficked area of Milan which shows a reduction of approximately 4 $\mu g/m^3$ per year starting from 2016. In the next sections of this paper, we will devote much effort to the effect of meteorology on estimates of the impacts of restrictions. In particular, we will use several environmental parameters, such as air temperature and wind strength, to filter out spurious effects from the regressions and obtain adjusted estimates.

4.4.1 The COVID-19 lockdown in Italy

The Italian government imposed a total lockdown period between 9th March and 18th May 2020, totalling 71 days. It was characterised by the closure of all non-essential activities and enterprises and minimised individual mobility and social distancing. It resulted in a generalised

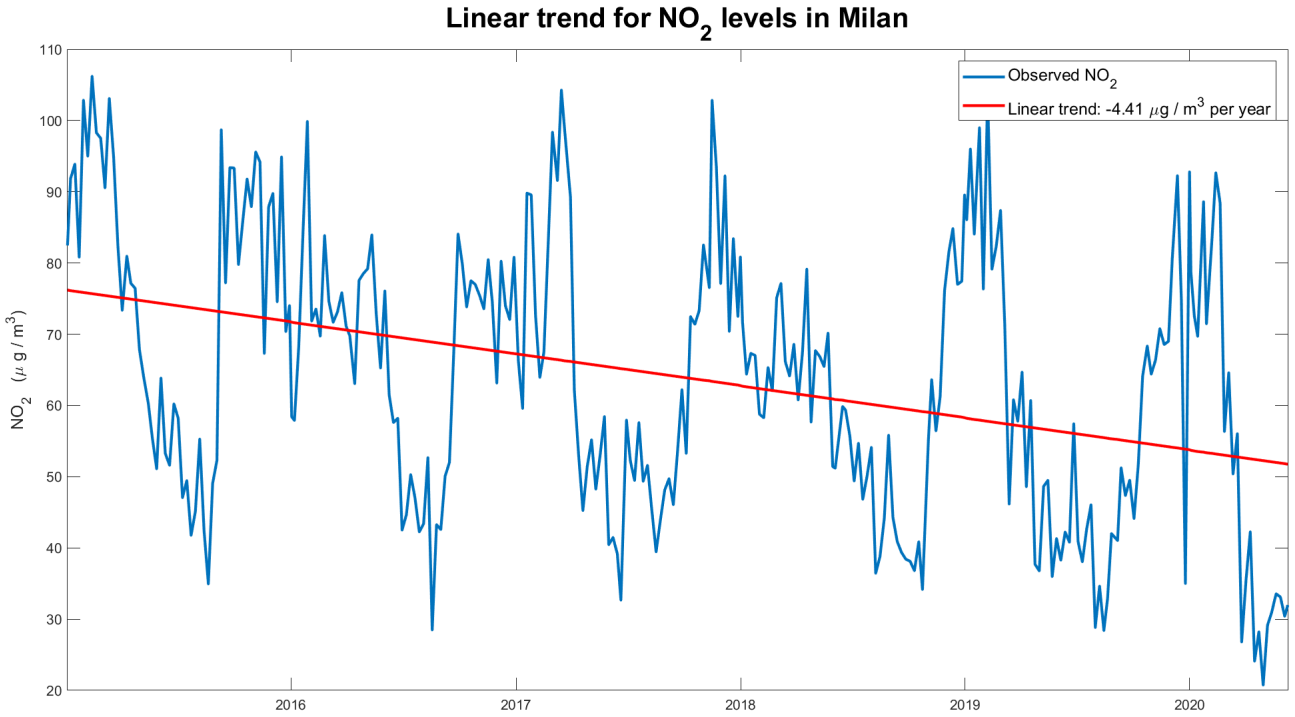


FIGURE 4.3: NO_2 concentration in the North of Milan (Marche traffic monitoring station), years 2015-2020. The plot shows seasonality, intra-seasonal variability and a decreasing trend (red line) of $4\mu g/m^3$.

reduction of car traffic and personal travel, as discussed by Finazzi and Fassò, 2020. According to their estimates, obtained with a sample of 20 thousands Italian users of an earthquake-tracker application, at the peak of the lockdown, the daily mean distance travelled decreased by approximately 50%, and the percentage of users who did not move within 24 hours reached 65%.

In principle, the variation due to lockdown can be measured by the difference between the average NO_2 concentration during the event and the corresponding average before. Hence, we could compare air quality during the lockdown and the average of the preceding 71 days. Since this means comparing spring and winter concentrations, it is a biased comparison.

To avoid the impact of seasonal variations, we compared the lockdown average concentration with the average of the same period (9th March – 18th May) in the preceding years. Figure 4.4 shows how the average concentrations fell in all provinces [1] during the lockdown, especially in large urban areas, such as Milano, Bergamo, Monza and Brescia.

At a glance, it is clear that the restrictions generated a positive effect on the whole territory, especially in the urbanised central area. The provinces of Monza-Brianza and Milano experienced reductions close to $20\mu g/m^3$, while Bergamo, Brescia and the northern mountain provinces recorded reductions of around $15\mu g/m^3$. In these latter cases, the reductions were around 50% compared to previous years. The southern areas, which are less inhabited and less industrialised, experienced smaller decreases.

[1] Provinces: Varese (VA), Como (CO), Lecco (LC), Sondrio (SO), Bergamo (BG), Brescia (BS), Milano (MI), Monza e Brianza (MB), Pavia (PV), Lodi (LO), Cremona (CR) and Mantova (MN)

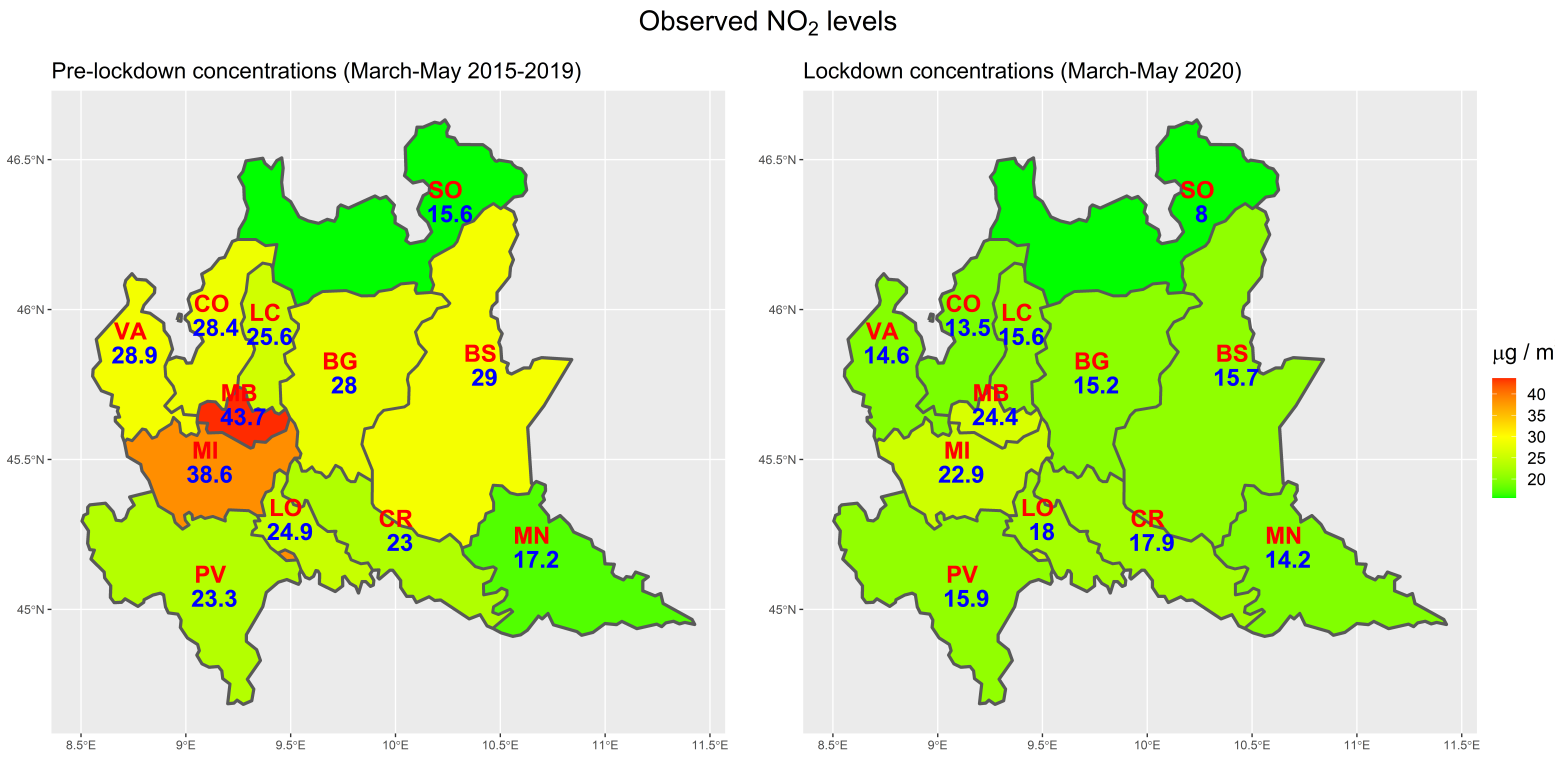


FIGURE 4.4: Observed pre-lockdown and in-lockdown air quality. Average NO_2 levels by province observed between 9th March and 18th May of years 2015-2019 (pre-lockdown) and 2020 (lockdown).

Provinces: Varese (VA), Como (CO), Lecco (LC), Sondrio (SO), Bergamo (BG), Brescia (BS), Milano (MI), Monza e Brianza (MB), Pavia (PV), Lodi (LO), Cremona (CR) and Mantova (MN)

Are these reductions due to the lockdown measures, or other factors occurred? In other words, uncertainty and biases may still be present due to variations between years. These may be related to both meteorological variations and anthropogenic variations, such as traffic policies or changes in vehicle emissions (Maranzano et al., 2020). Our purpose is to measure the variation of airborne pollutants such as nitrogen dioxide (NO_2), which is related to car traffic limitations and activity restrictions during the COVID-19 period, controlling for meteorology and trends characterising the phenomenon.

4.5 Air quality and weather data

We collected air quality and weather measurements from *Regione Lombardia Open Data portal* (https://www.dati.lombardia.it/), the regional open-source database which stores, among other things, environmental data acquired by the regional agency for environmental protection, ARPA Lombardia.

The full sample comprises weekly observations from 1st January 2015 to 8th June 2020, a total of 289 observations, from 74 monitoring stations. Stations are classified as background (B), traffic (T), or rural (R), according to the environmental context in which they are installed. Overall, the number of background, traffic, and rural stations are 42, 25, and 7, respectively. Monitoring stations are located heterogeneously over the provinces, ranging from three control units in Como to sixteen units in Milan while considering the population density of the areas. Moreover, given the geographical specificities of each province, the rural stations are predominantly

located in the flat southern strip of land, while traffic stations are placed in the large urban areas in the centres.
The lockdown period (9th March - 18th May 2020) is modelled through a dummy variable which assumes value 1 for the lockdown weeks, for a total of 10 observations, and value 0 for the other observations. Thus, the whole period can be divided into three parts: the pre-lockdown period composed of 275 weeks from 1st January 2015 to 4th March 2020, the lockdown period of 10 weeks, and the post-lockdown term from the 20th May 2020 to the latest available observation.
To obtain an unbiased estimate of the effect of lockdown restrictions on nitrogen dioxide concentrations, we considered a set of weather variables to control for possible confounding effects: temperature (measured in degrees celsius), rainfall (cumulative millimetres), relative humidity (%), and the average wind speed for each quadrant of the Cartesian plane (metres per second) [2]. Therefore, overall, we have considered the weekly NO_2 concentrations and eight meteorological variables for each station.

4.6 Statistical modelling

4.6.1 The autoregressive model with exogenous covariates

For each station, the lockdown effect has been estimated using an autoregressive model with covariates, namely ARX model. The autoregressive moving average models with exogenous covariates, or ARMAX(p_y,p_x,q), belongs to a wide class of models called *linear input-output models*, where the dependent variable, i.e. the output, can be described using a ARMA representation augmented by one or more control variables, i.e. the input, characterized by its own process (Bittanti, 2019). A autoregressive moving average model with a single exogenous covariate, or ARMAX(p_y,p_x,q) can be analitically expressed using the following equations system:

$$\begin{aligned} y_t =& \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_{p_y} y_{t-p_y} \\ &+ \beta_1 x_{t-1} + \beta_2 x_{t-2} + ... + \beta_{p_x} x_{t-p_x} + v_t, \end{aligned} \tag{4.1}$$

where

$$v_t = \varepsilon_t + \eta_1 \varepsilon_{t-1} + \eta_2 \varepsilon_{t-2} + ... + \eta_q \varepsilon_{t-q}, \tag{4.2}$$

where p_y represent the lag order for the dependent variable, p_x is the lag order for the covariate and q is the moving-average order for the residuals. The noise ε_t is a white-noise process with null mean and constant variance σ^2. Gaussian ARMAX models assumes that the white-noise term follows a Gaussian distribution, i.e. $\varepsilon_t \sim N(0, \sigma^2)$. In particular, we used a steady-state representation of the ARX(1) model. In this specification, the covariates and the residuals are both assumed as idiosyncratic, hence without lags, while the response variable presents a lag-one operator. A ARX(p_y) model can be considered as a special case of ARMAX(p_y,p_x,q) having fixed $p_x = 0$ and $q = 0$, i.e. the ARX(p_y) is equivalent to ARMAX(p_y,0,0). See for example the case study of Fassò, 2013 in which the author used a simple ARX(1) model to assess the effect on oxide concentrations due to a new limited traffic zone in Milan metropolitan area.
We set as dependent variable the weekly concentration of NO_2 observed for each single station. The list of covariates includes the above seven meteorological variables, a linear trend component, and the dummy for the lockdown. Given the natural seasonality of the phenomena under

[2] We computed four variables measuring the average wind speed blowing from North-East, South-East, South-West and North-West

analysis, we decided to consider the interactions between the meteorological variables, excluding the linear trend, and the dummies for the four climatic seasons. In this specific case, we considered spring to be the period between March and May, summer from June to August, autumn the weeks from September to November, and winter from December to February. Hence, the full model considers a constant term, the autoregressive term and 30 covariates, i.e. 28 seasonal weather variables, the linear trend and the lockdown dummy.
Let $y(s_i, t)$ be the weekly observation of NO_2 concentration from the station located in $s_i, i = 1, 2, \ldots, 74$ and time $t = 1, \ldots, 289$ (weeks) and let $\boldsymbol{X_t}$ be the vector of weekly seasonal weather covariates, T_t be the linear trend component, and let L_t be the lockdown dummy variable for station i at time t. The model equation for each location $i = 1, 2, ..., 74$ is the following:

$$y_t = \alpha + \beta y_{t-1} + \nu T_t + \gamma L_t + \boldsymbol{\theta X_t} + \varepsilon_t, \tag{4.3}$$

where ε_t is a Gaussian random noise, γ defines the NO_2 variation at the corresponding station i due to the lockdown, β represents the lag-1 autoregressive coefficient for pollution concentrations, and ν is the linear trend parameter.

4.6.2 Model selection and inference using LASSO algorithm

In order to consider the best set of covariates able to describe the local meteorology, we performed a variable selection by implementing the Least Absolute Shrinkage and Selection Operator, or LASSO estimator (**Tibshirani1996**), trained with a 20-fold cross-validation setup. For each station, the optimal model has been selected using the *one standard error empirical rule* (**ESL2017**; **SLS2015**), that is we selected the model with largest penalisation parameter (lambda) value such that the MSE is within one standard error of the minimum MSE. This approach ensures that the most parsimonious model whose error is no more than one standard error above the best model's error is selected while considering the randomness generated by the out-of-sample randomisation used to construct LASSO. All the weather covariates, the linear trend, the autoregressive term and the lockdown dummy have been included in the LASSO algorithm. The inclusion of the dummy among the LASSO inputs provides a handy tool to understand the effective impact of restriction measures on oxide concentrations. If the LASSO includes the dummy in the list of relevant variables, this would indicate a great relevance of lockdown restrictions in explaining the concentrations pattern in the weeks of interest; otherwise, its exclusion would be a signal of a modest and negligible variation during the shutdown period.
The model parameters identified by the LASSO algorithm are then re-estimated using a maximum likelihood approach under the hypothesis of Gaussian distribution of the errors. It is well-known that the LASSO introduces a bias in the estimates of the regression parameters to reduce the variance. This bias can be reabsorbed by applying the OLS or ML estimators to estimate the regression parameters of the variables selected by the algorithm. As discussed by **Belloni2013**, this approach, namely OLS post-LASSO estimator, performs at least similarly as the LASSO regression in terms of convergence rate and can achieve a smaller bias for the estimated parameters, even when the algorithm fails in selecting the actual variables. Finally, the post-selection statistical inference has been made by applying the delta method to provide the approximate lockdown impact standard errors.
As stated in Fassò, 2013, the AR(1) dynamics imply that the scalar steady state impact on weekly NO_2 concentrations is given by

$$\hat{\delta} = \frac{\hat{\gamma}}{1 - \hat{\beta}}. \tag{4.4}$$

Moreover, ignoring the uncertainty of the pre-intervention estimation of β and applying the delta method to the estimated parameters, it is possible to approximate the variance for δ as follows:

$$VAR(\hat{\delta}) \cong \frac{VAR(\hat{\gamma})}{(1-\hat{\beta})^2}. \tag{4.5}$$

The parameter standard errors were simply obtained by calculating the square root of the formula of the above variance.
To correctly assess the estimated coefficients, particularly the one associated with the impact of lockdown, it is necessary to check the whiteness of the regression residuals. In particular, we point out that the variance of the parameters depends on residuals autocorrelations. Indeed, serial correlation generates bias effects on the estimated variances of the coefficients and consequently on the respective p-values and confidence intervals (**Hamilton1994**). This necessarily leads to erroneous assessments of the significance of the estimates. Thus, we assessed the goodness of the estimated models by performing several misspecification diagnostics on the regression residuals. In particular, each residuals time series has been checked for serial correlation by analysing the sample ACF function and by using the Ljung-Box test (Ljung and Box, 1978). We also tested for non-normality through the Jarque-Bera test (Jarque and Bera, 1987).
Several practical and contextual reasons can be argued to justify the choice of the ARX model with a predefined number of lags to model a complex phenomenon such as atmospheric NO_2 concentrations. Firstly, recall that the considered observations have a weekly frequency, allowing the serial correlations to vanish after a few lags. Besides, we want to favour statistical modelling that is as simple but effective as possible, so that the empirical results are unbiased and meaningful. In particular, we recall that in an ARX setting, the effect of each covariate depends both on the estimated coefficient and on the autoregressive dynamics (see equation 4.3). Thus, fixing a lag order common to all the stations, we make it possible to quantify the local impact of lockdown limitations on NO_2 concentrations, while keeping a standard temporal dependency structure across the region and preserving simplicity and ease interpretation of the estimates. In the end, recall that airborne pollution is a natural phenomenon mainly driven by atmospheric conditions, such as the air temperature and the wind. Thus, by including a suitable set of weather covariates in the regression models, it is possible to reasonably model the seasonality and persistence of the series without requiring more sophisticated and complex models. The residuals diagnostics presented in Section 4.7.3 validate the previous statements. The residuals are in many of the stations non-autocorrelated, and at those stations where some remaining autocorrelation persists, it always shows very small values that do not bias the coefficient estimates and their significance.

4.7 Model fitting and selection

In this section, we discuss in detail the results regarding the model selection performed through the LASSO algorithm and the OLS post-LASSO estimated coefficients obtained by re-estimating the parameters of the optimal model for each station.

4.7.1 LASSO performances

To understand how the LASSO algorithm identified the optimal models is primarily necessary to assess which penalty values, i.e. the lambda parameter, minimise the cross-validated mean square errors. Recall that we identified the optimal lambda using the one standard error empirical rule that is we selected the model with the largest penalisation parameter such that the

MSE is within one standard error of the minimum MSE. Figure 4.5 shows the scatterplot of the one-SE lambdas and the corresponding cross-validated MSE. For each combination of lambda and MSE, it is also reported in an error bar the estimated uncertainty. Recall that the LASSO

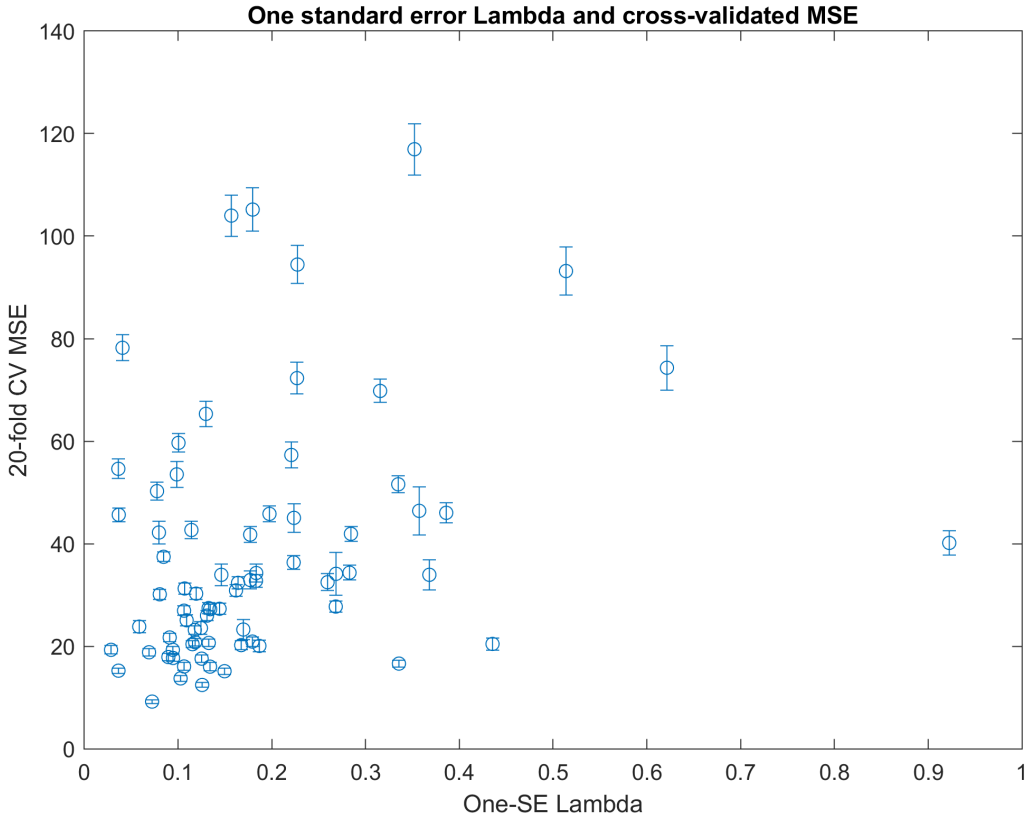


FIGURE 4.5: Selected 1-SE Lambda and corresponding MSE. Error bars represent the cross-validated standard errors of each penalisation parameter.

solution corresponding to a null value of the penalisation parameter is equivalent to the OLS estimator solution. Observing the horizontal axis one can notice that the optimal penalisation values are always lower than one and largely located around 0. Considering the full sample, the largest part reports a lambda value below 0.20. It means that the model selection procedure provided estimates that were not very penalised and relatively close to OLSE solutions. Hence, the models will contain a large number of relevant covariates. The cross-validated MSE are concentrated between 10 and 30, while the largest mean squared errors are more dispersed. Observing the variability, it can also be noticed that lower lambdas correspond lower uncertainty, while to greater penalisation value the uncertainty increases.

4.7.2 Meteorology and long-run trend

In aggregate, the estimated models show that some peculiar features can characterise air quality in the region in terms of long-term trends and meteorology. As reported by Table 4.1, both the linear trend and the autoregressive term play a key role in explaining the air quality variability.
The autoregressive coefficient has been selected by the LASSO algorithm as a relevant variable in the 85% of the stations. Moreover, the estimated coefficients are statistically significant and

TABLE 4.1: Summary statistics for the estimated linear trends and autoregresive coefficients

	Negative sign		**Positive sign**		**Null**				
	Signif.	Not signif.	Signif.	Not signif.	**coefficients**	**Min**	**Mean**	**Max**	**StdDev**
	%	%	%	%	%	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$
AR (β)	0	0	85.1	0	14.9	0.30	0.50	0.80	0.10
Trend (ν)	43.2	24.3	2.7	5.4	24.3	-1.9	-0.60	1.50	0.60

Note: we considered as statistically significant the coefficients showing p-values below 5%. Percentages are calculated with respect to the total number of involved stations, that is 74. The column *Null coefficients* reports the percentage of stations with null estimates, that is those stations in which the LASSO did not selected the covariate and set up to zero by the ML post-LASSO estimator.

with a positive sign, showing how concentrations are particularly persistent even after several days. The estimated variability is low (standard deviation is around 0.10 $\mu g/m^3$), meaning that the same serial dependence structure characterizes almost all the stations. Furthermore, such a large majority of significant values, combined with the low serial correlation identified in the residuals, indicate a successful specification of the model regarding NO_2 concentrations' temporal evolution. Also, the trend appears to be a relevant factor. In the 43% of the stations (32 out of 74), the linear trend coefficient has been estimated as statistically significant with a negative sign, whereas it is positive or not-significant in some cases. In average, the coefficient is -0.60 $\mu g/m^3$ per week. This fact confirms what we have observed in the city of Milan and in other parts of the region: in the long-run, the NO_2 concentrations are significantly decreasing almost all over the region, with some rare exceptions where the trend is opposite. As we will see below, stations with null or positive trends are often isolated from urban centres and monitor areas where oxides in the atmosphere are low and difficult to reverse.
Regarding the local meteorology, Table 4.2 summarises the estimated seasonal coefficients for each of the seven climate variables considered. Recall that the coefficients for the temperature have to be read as the change in NO_2 concentrations (in $\mu g/m^3$) associated with an increase in temperature of one degree Celsius; the coefficients for rainfall as the change due to an increase of one mm in rainfall fall; the coefficients for wind as the change due to one m/s increase in wind speed, and those for humidity as the change in NO_2 due to a 1% increase in relative humidity. The seasonality of the variables of interest is well defined by the alternation of the estimated signs and their respective significance. For example, during spring and summer the coefficients always assume negative and often significant values, indicating climatic factors' strength in reducing concentrations during warm periods. Temperature and wind remain throughout the year the most important factors in reducing the levels of oxide concentrations. Except for summer, the temperature is selected and estimated as negative and significant in 70% of the stations, while the wind is estimated as significant and negative in 30 or 40% of the models. Although the wind in Lombardy is generally very weak, winds blowing from the East and North are of particular importance. Humidity is often discarded, excepting during summer, where it contributes significantly to reducing the concentrations. On the other hand, cumulative rainfall is highlighted as crucial for reducing concentrations especially during the winter.

4.7.3 Models fitting and diagnostic checks

To assess the goodness of the estimates and model fitting, we performed an analysis of multiple goodness-of-fit indicators. For each station, we evaluated the in-sample adjusted R-squared (R2) index, the residuals root mean squared error (RMSE), and the corrected Akaike Information Criterion (AICc). Estimated values are summarised in Figure 4.6, which represents the scatterplot between adjusted-R^2 and residuals RMSE weighted by the model AICc. Consider-

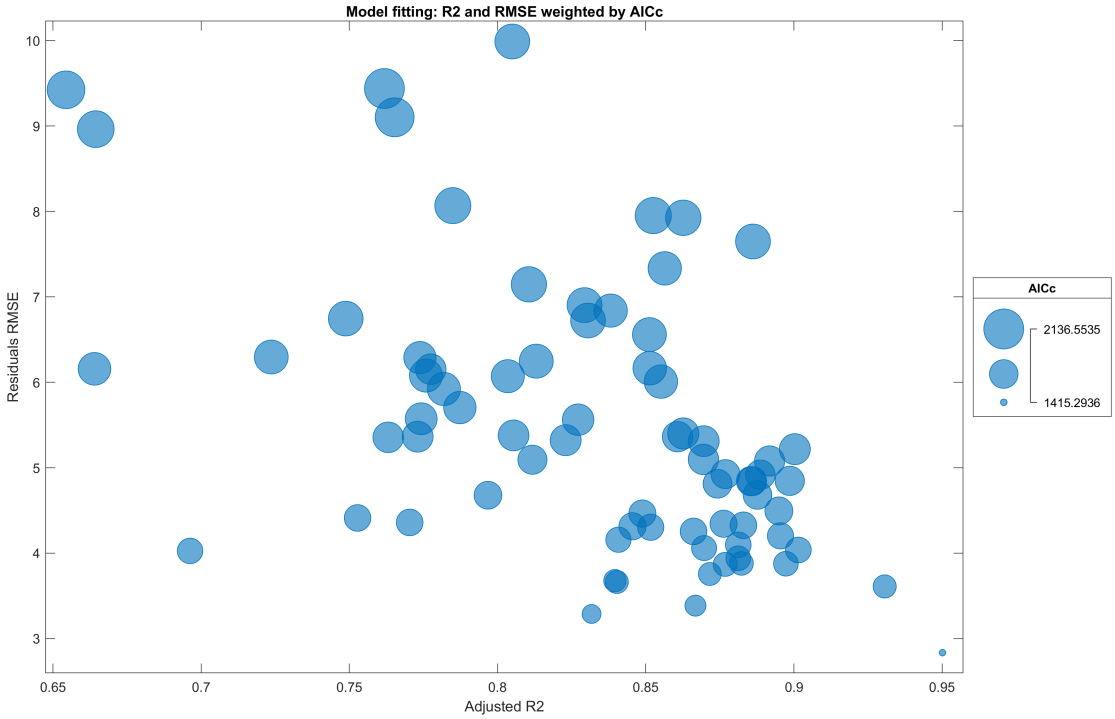


FIGURE 4.6: Scatterplot of stations residuals RMSE against adjusted R-squared index. Bubble size is given by the AICc.

ing the 74 stations, the models fitting can be considered satisfactory. Indeed, most of the models report an R-squared above 80% and an RMSE below 6 $\mu g/m^3$. The minimum goodness-of-fit value is above 65%, which can be considered an acceptable share of explained variability. The models that fit better provide also the lowest prediction error. The models with greater R^2 index and lower RMSE are also associated with the minimum AICc values. The smallest points are located on the lower-right side of the plot, while the largest marks are on the upper-left part. Given these characteristics, from a fitting point of view, it can be stated that the strategy of estimating through ML post-LASSO estimator is a correct tool for correctly modelling NO_2 concentrations for the chosen sample.
To validate the assumptions concerning the AR(1) structure of the model in equation 4.3, we implemented several diagnostic checks on the residuals. We firstly checked for the residual serial correlations both using graphical and analytical methods. The sample ACF and the estimated autocorrelation tests reported in Figure 4.7 provide favourable indications. In particular, looking at 20-lags Ljung-Box test, among the 74 stations, 55 of them had p-values above 5% and 64 report a p-value larger than 1%. Thus, at a 1% significance level, only 10 of 74 present significantly autocorrelated residuals. Moreover, the average sample ACF distribution is always below 10%, while the maximum estimated ACF is around 27%. In Section 4.6, we mentioned

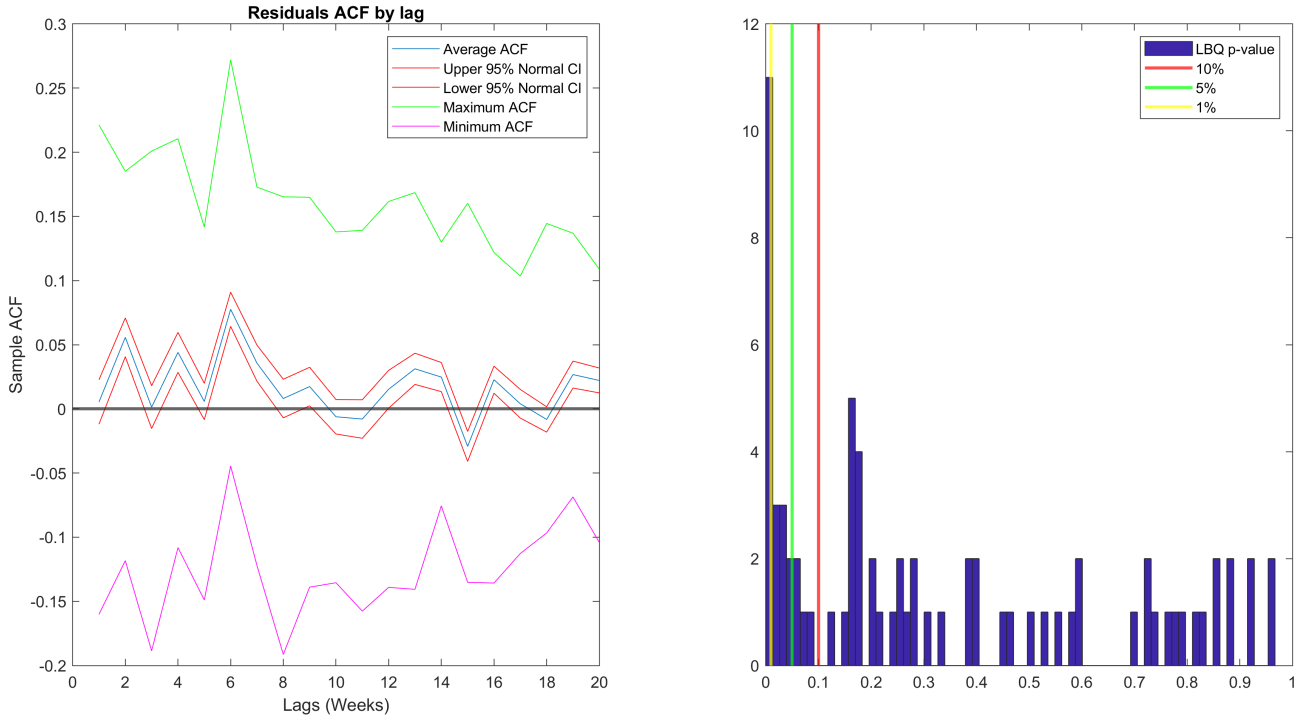


FIGURE 4.7: Residuals ACF by lag (left panel) and empirical distribution of the p-values for the Ljung-Box test with 20 lag applied to the regression residuals (right panel).

that the residual serial correlation induces a bias on the variances of the estimated parameters, and hence on their respective p-values, leading to misjudgments regarding the statistical significance of the coefficients. The above empirical results show that sample autocorrelations are very moderate in absolute value. Therefore, we can consider the bias as negligible and having a minimal impact on the estimated significance.
The Gaussianity assumption of the residuals, verified through the Bera-Jarque test, was confirmed for just five stations out of 74. It means that the stations with independent and normally distributed residuals are only five out of 74.

4.8 Lockdown results

This section discusses the main empirical results related to the estimated variations due to the COVID-19 lockdown. The discussion will focus on the geographical and landscape profile of the considered stations to characterise and justify the estimates. In addition to comments on the estimates by station type and province, we will also comment on some specific stations that have an interesting relevance for an extensive understanding of the lockdown phenomenon. In particular, we will describe the cases of Meda and Schivenoglia as examples of null lockdown effect and other cases in Monza, Bergamo and Brescia provinces regarding local meteorology and geography. The location of the specific stations is shown in Figure 4.8.
The proposed statistical modelling confirms a generalised reduction of NO_2 levels consequently to the lockdown. Compared to the observed average reduction, reported in Figure 4.4, the lockdown effect is mitigated by seasonal trends and meteorology. Moreover, a non-null lockdown effect was estimated for a subset of stations among those included in the sample. In fact, the

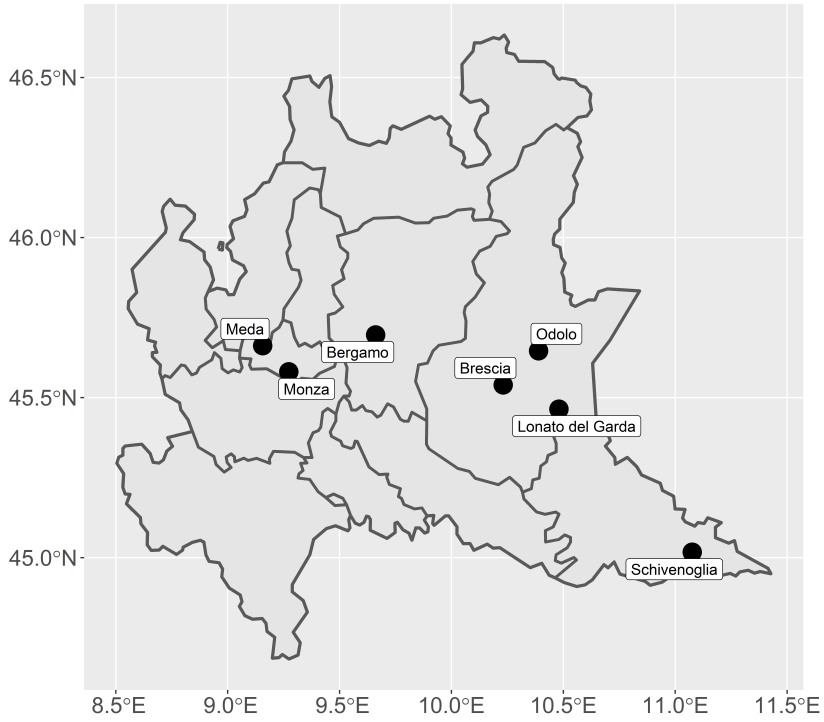


FIGURE 4.8: Geolocation of some sites of interest for the COVID-19 lockdown in Lombardy

LASSO algorithm selected the lockdown dummy among the covariates 63 times out of 74 stations. For these locations, the lockdown variable can be considered an important factor. For the remaining eleven stations where the dummy was discarded, the lockdown impact on NO_2 concentrations has to be considered null. The reasons for this result are many and will be explored more deeply in the following paragraphs discussing the different types of stations and their geographical distribution. From the statistical modelling perspective, a first intuitive explanation can be ascribed to the presence of external factors other than the lockdown, either embedded in the model or unknown, able to better explain the observed reductions. First of all, the specific local weather conditions. Table 4.3 reports both the pre-post lockdown observed average reduction and the model-based effects, aggregated by station type.
Regarding the stations where the lockdown dummy was estimated via MLE, the estimates of the coefficients were largely statistically significant. Indeed, considering the alternative hypothesis of a statistically significant negative variation, in 47 over 63 cases, the coefficients were significant at 5%, while 35 of them were statistically significant at a 1% significance level.
To investigate which factors may have led to some null lockdown coefficients, we now provide two illustrative examples. Both examples are represented in Figure 4.9, which depicts the time series of the NO_2 concentrations at the Meda station in Monza-Brianza province (upper panel) and Schivenoglia (lower panel), located in Mantova province. In both plots, the blue marks represent the weekly observed concentrations between March and May in 2015-2019, while the red marks are the weekly observed concentrations between March and May 2020.
Although the lockdown impact has been estimated as zero at both locations, the historical evolution of the concentrations and the surrounding anthropogenic context are widely different. Meda control unit is classified as a traffic sensor and is located in a residential area far from the major high traffic roads. Schivenoglia is a rural station surrounded by many kilometres of

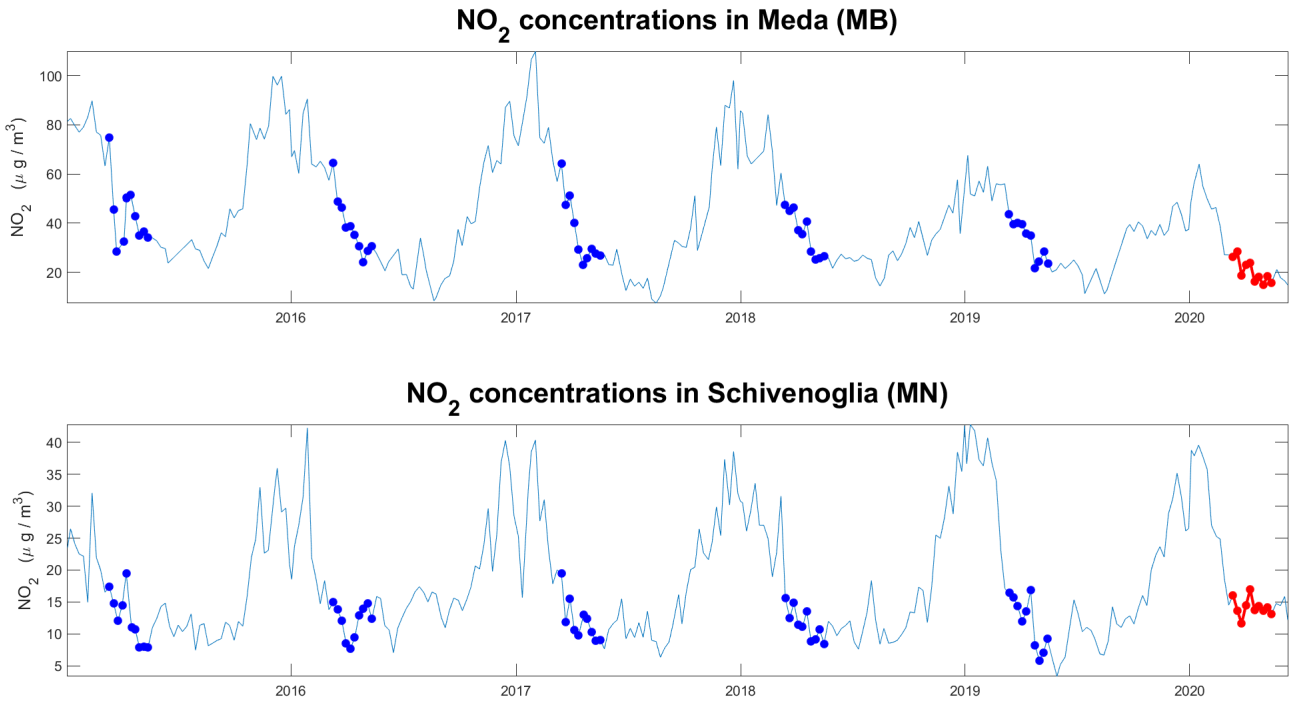


FIGURE 4.9: NO_2 concentrations from 2015 to 2020 at Meda (MB) ground station (upper panel) and NO_2 concentrations from 2015 to 2020 at Schivenoglia (MN) ground station (lower panel). Blue marks are the weekly observed concentrations between March and May in 2015-2019. Red marks are the weekly observed concentrations between March and May in 2020.

farmland and livestock. Several secondary rivers and canals flow near its location.
Regarding Meda station, the plot shows a decreasing trend in average NO_2 concentrations starting from 2016 and a considerable reduction of the concentrations during the winter season in 2019 and 2020. Between 2015 and 2018, at that location were recorded maximum winter peaks approximately from $80\mu/g^3$ to $110\mu/g^3$, while in the same period in 2019 and 2020, the observed values hovered around $60\mu/g^3$. The concentrations recorded in the 2020 lockdown period are lower and less volatile than those observed n the previous five-year period. However, the variations do not appear to be large-scale. The abrupt drop in NO_2 concentrations between 2018 and 2019 leads to thinking about possible changes in the viability around the station or at least a drastic change in the surrounding emission sources. The downward trend (common throughout the region) and a possible change in the road network around the station have reduced the concentrations considerably, mitigating the positive effects of the lockdown on air quality. The case of Meda can be considered as a suitable example of air quality improvement due to viability changes.
On the other hand, the second plot depicts a different story about the concentrations in rural areas. NO_2 concentrations in Schivenoglia are always very low but show a marked seasonality. In winter there are always maximum values around $40\mu/g^3$, while in summer, the concentrations reach approximately $5\mu/g^3$. The chart also shows that the concentrations are always decreasing between March and May and with a slight variability. Despite the lockdown restrictions, 2020 is no exception to the above.
Moreover, more compact values close to $15\mu/g^3$ are observed. The values observed during the

lockdown seem even higher than in the past. There is no downward trend, but the scale of values and the variability are much lower than in the previous case. The effect of the lockdown is negligible.

4.8.1 Evaluation of the lockdown effect based on area type

If we simultaneously analyse Table 4.3 and Figure 4.10, which reports the geo-location of the control units dividing the non-null and the null lockdown impact estimates, it is interesting to note that among the stations with null lockdown effect, five are classified as rural, and they are mainly located in the southern plain area. The non-null effects are evenly distributed on the map. This fact is crucial in understanding how the lockdown restrictions acted on pollution concentrations. There was a considerable difference between the variations of traffic, background, and rural stations in both observed and estimated reductions. For traffic control units, the observed and estimated average reductions were -17.5μ/g^3 and -14.7μ/g^3 (-38.2%) respectively. At rural sensors, the estimated variations amounted to -3.5μ/g^3 (-21.7%). Estimated variations at background stations stand in the middle, with an average reduction of -8.7$\mu g/m^3$ (-34.4%). Also, the average variability associated with the estimates, i.e. the estimated standard errors, follows the previous order. Indeed, the highest variability has been estimated for traffic sensors (5.1 μ/g^3), while the lowest is associated with the rural stations (2.8 μ/g^3). These facts can be primarily attributed to the large differences in the counting of stations for each type (39 background, 22 traffic, and two rural), and secondly to the spatial heterogeneity generated by the geographical location of the stations that could capture different traffic intensities, for traffic stations, for industrial concentrations, and background stations.

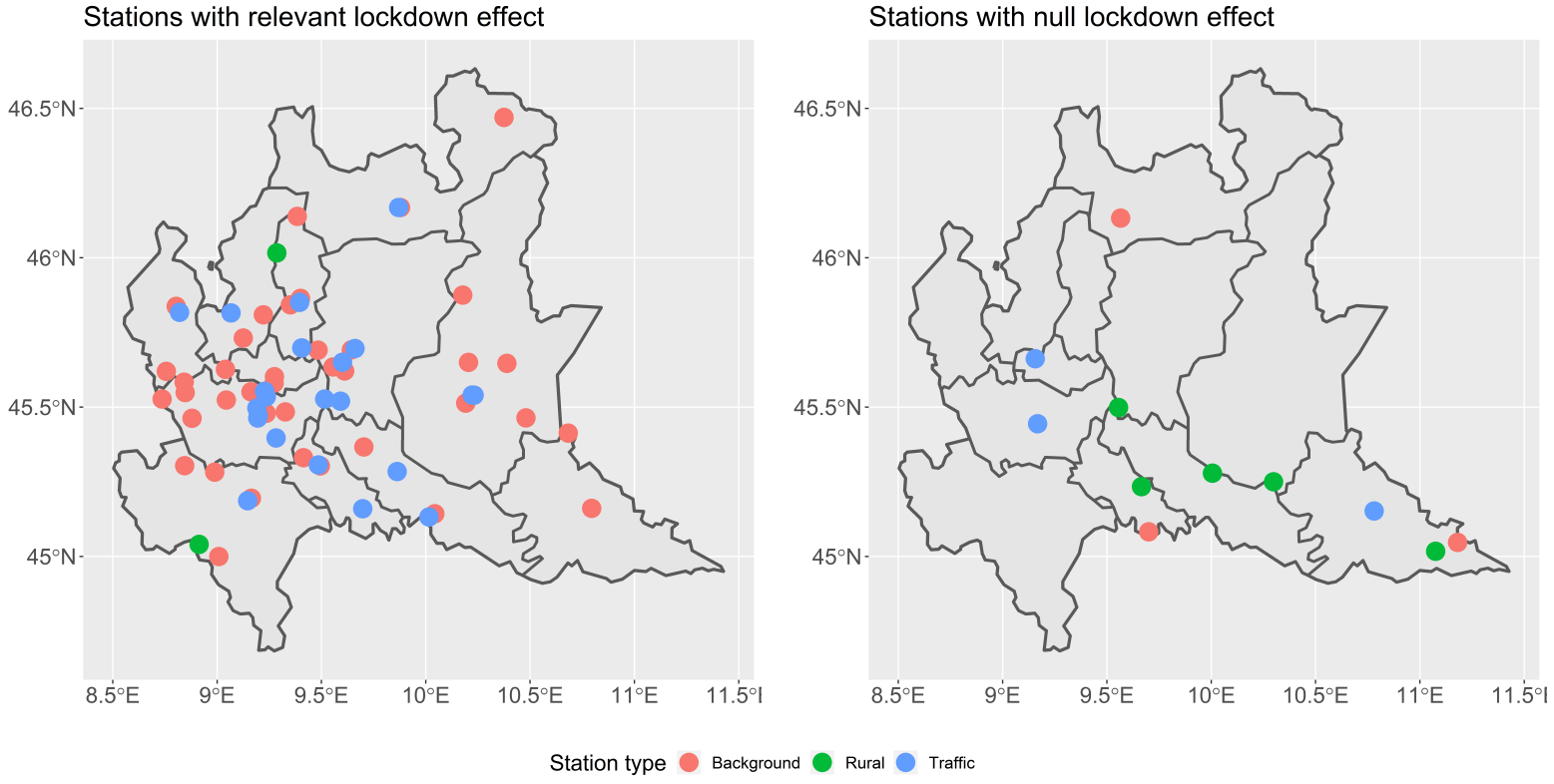


FIGURE 4.10: Geo-location of air quality stations with relevant lockdown effect (left panel) and geo-location of stations with null lockdown effect (right panel).

As already mentioned in Section 4.4, which describes the air quality conditions in Lombardy, geographical and morphological conformation plays a fundamental role in determining variations due to lockdown. The prevalence of rural stations with null lockdown impact, and the remarkable differences in estimated effects by station types, provided strong evidence of how the limitations to traffic, human mobility, and productive activities worked decisively and sharply in dense urban areas and more lightly in agricultural areas. The former are subject to high levels of oxides (NO_X and NO_2), carbon (CO and CO_2), and sulphur dioxide (SO_2), produced by industry and road transport. At the same time, the latter is characterized by emissions of pollutants produced mainly by agricultural and breeding activities, i.e. ammonia (NH_3), methane (CH_4), and fine particulates ($PM_{2.5}$). See, for example, the annual reports from INEMAR Lombardia (2020) and the European Environmental Agency (2019) on pollution sources and abatement policies in Lombardy. All these facts are consistent with the analyses on movements and mobility during COVID-19 in Lombardy by Finazzi and Fassò, 2020, which reported reductions of up to 65% in human movements all over Italy that led to a natural fall in road traffic.

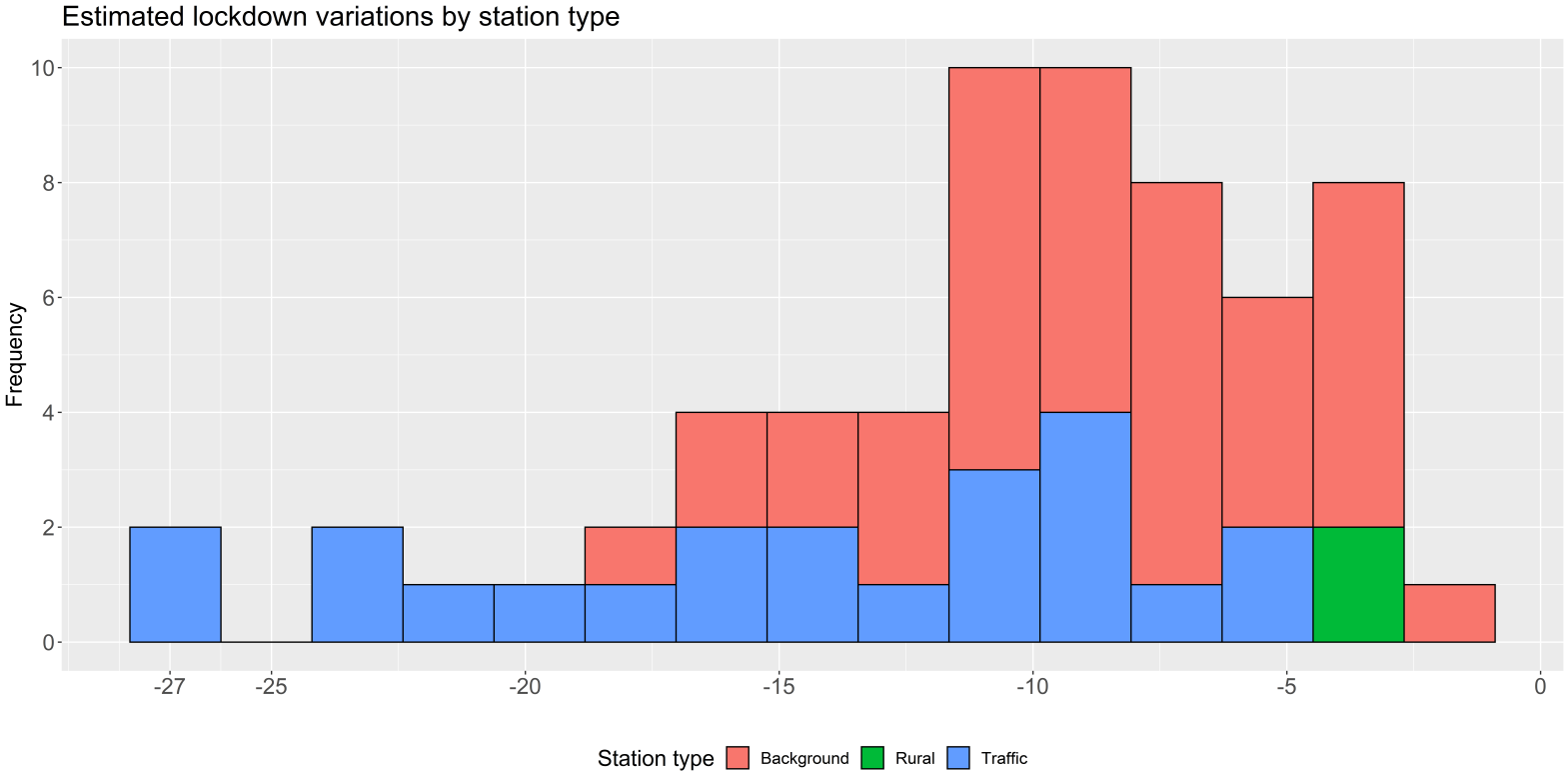


FIGURE 4.11: Frequency distribution of the model-based lockdown variations, in $\mu g/m^3$, by station type

Figure 4.11 shows the empirical distribution of estimated lockdown effects for the 63 units by the type of station. Estimates in traffic sites show a high variability range, ranging from approximately -5$\mu g/m^3$ to -27$\mu g/m^3$, while in background sites there were variations between -1$\mu g/m^3$ and -17$\mu g/m^3$. The figure also highlights that all the stations registered a negative variation, supporting the hypothesis of a generalised improvement of air quality in Lombardy.

4.8.2 Geographical distribution of the lockdown effect

To give a synthetic representation of the territorial heterogeneity of the effect of lockdown measures on oxide concentrations, we aggregated the estimated coefficients by province. Observed and estimated lockdown impacts by province are reported in Figure 4.12. The observed average variations are computed by averaging the observed differences of NO_2 previously, and during the 2020 lockdown, whereas to compute the estimated provincial variations, we considered as null the lockdown impact for all the stations in which the LASSO algorithm did not include the dummy variable among the final dataset. See Figure 4.10 for the geolocation of these stations. The figure's left panel shows the observed average reductions by province,

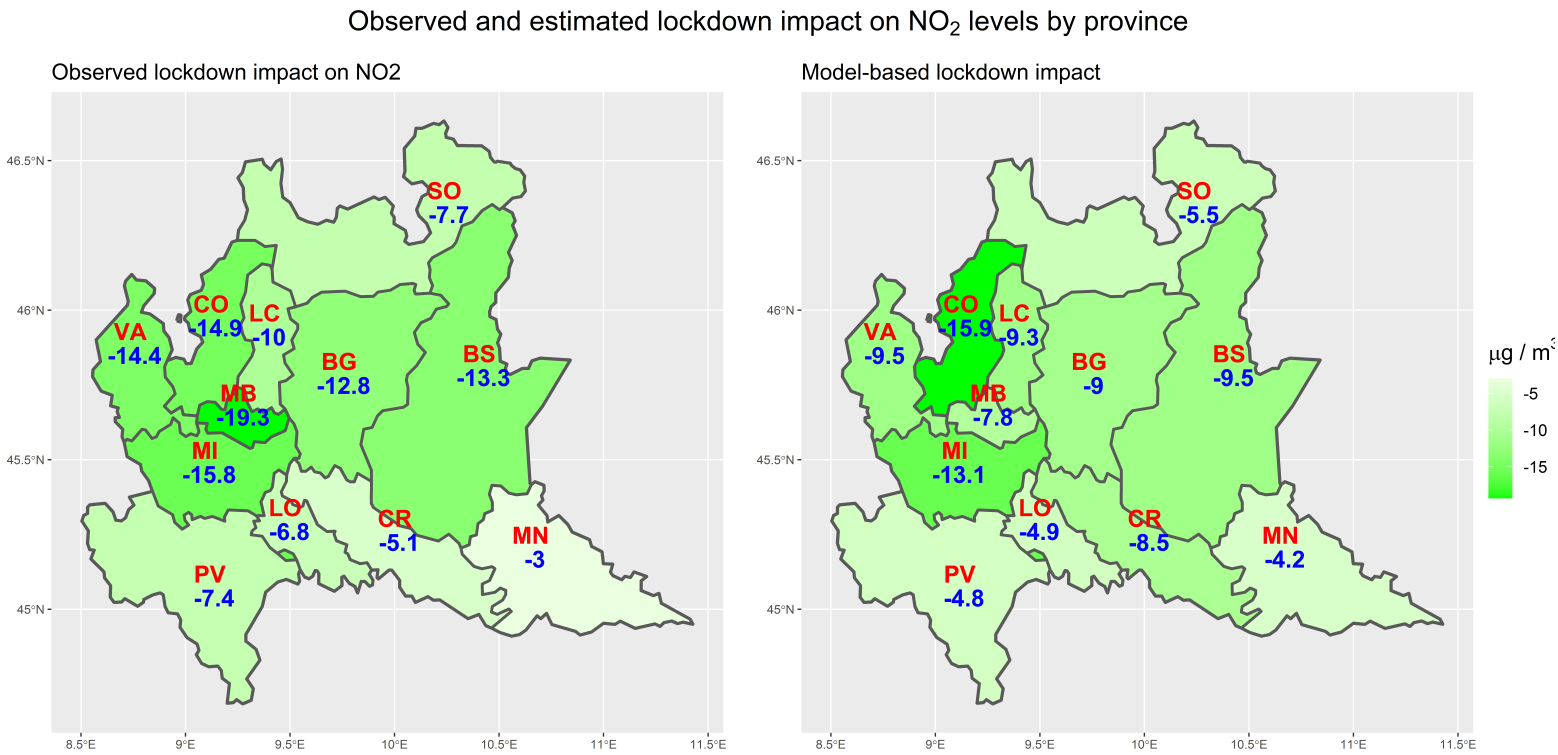


FIGURE 4.12: Variation of NO_2 levels by province. Left panel: Observed average variation computed as the raw difference for the period 9th March-18th May in 2020 and 2015-2019. Right panel: model-based variation for the period 9th March-18th May 2020.

Provinces: Varese (VA), Como (CO), Lecco (LC), Sondrio (SO), Bergamo (BG), Brescia (BS), Milano (MI), Monza e Brianza (MB), Pavia (PV), Lodi (LO), Cremona (CR) and Mantova (MN)

which are compared with the average provincial reductions estimated by our model, reported in the right panel. The two plots are in line with what has been said above regarding the differences by station types and the gap between rural and urban areas. In almost all the provinces, the estimated impacts are lower than those observed, meaning that, according to our model, the estimated lockdown effect is mitigated by seasonal trends and the local meteorology.
The most remarkable reductions were observed and estimated in the Milan metropolitan area, which is notoriously strongly urbanised, followed by the province of Como. In the latter case, the estimated NO_2 reduction was -15.8$\mu g/m^3$ (-55.9%). A relevant factor for air quality improvement in the Como area could have been the interruption of cross-border traffic to and

from Switzerland. In 2018, about 18 thousands vehicles per month crossed the border at Chiasso custom (Ticino Statistics Office, 2020), and about 70 thousand Italian cross-border commuters were recorded (Switzerland Federal Statistics Office, 2019). According to estimates by the Swiss statistical office, in April 2020, there was an average daily decrease in the number of vehicles entering and leaving the Chiasso border of 79% compared to April 2019, while in May 2020 the decrease was 66% compared to the previous year. The observed reductions are mainly due to the breakdown in the number of transits of commercial or industrial vehicles (-81% in April and -66% in May).
The estimated impact and the observed variation in Milan are quite similar. The observed reduction was -15.8$\mu g/m^3$ and the estimated variation amounted to -13.1$\mu g/m^3$ (-36.1%). This result is consistent with the estimates of the provincial agency for the mobility of Milan (AMAT), which reported reductions in road traffic up to 77% (private vehicles) and 66% (commercial vehicles) in the Milan metropolitan area during the lockdown period (Agresti et al., 2020).
Interesting results were obtained for the Monza (MB), Bergamo (BG) and Brescia (BS) provinces. The three provinces have large, densely populated and congested urban centres, especially the provincial capital cities, but they present some peculiarities that can be argued to explain the empirical discrepancies. Regarding Monza, there exists a substantial gap between the observed average variation and the estimated reduction. Indeed, the province passes from an observed reduction of -19.3$\mu g/m^3$ to an estimate of -7.8$\mu g/m^3$ (-18.1%). On the Monza territory are located only three monitoring stations, one urban traffic station and two background stations. Both the background stations are located within the Monza city borders, while the urban ground unit is located in Meda. The LASSO model selection algorithm provided a non-null estimate for the background units, as it did not select the dummy at the urban site. The model-based estimate has been fixed at zero despite the observed variation at the traffic station -17.2$\mu g/m^3$ (-46.2%). In Section 4.8, we discussed the case study of Meda, claiming that it can be taken as an excellent example of air quality improvement due to changes in emission sources. Local meteorology is also a contributing factor. The ML post-LASSO estimator provided negative and statistically significant coefficients for the temperature during spring, winter and autumn, and a significant negative trend of -1.09 $\mu g/m^3$ per week. Rainfall and wind blowing from North-East contribute significantly to reducing concentrations during the winter. Similar results, in particular regarding the decreasing trend and the cleaning effect of rainfall and wind, hold for both the background stations in Monza city. At the two background sites, the estimated lockdown impact is negative but not statistically significant. In conclusion, regarding the specific case of the province of Monza, it is reasonable to think that the observed reduction can be explained by meteorological factors included in the model, which can have affected the air quality more than the lockdown restrictions.
Concerning the cases of Bergamo and Brescia, some geographical and morphological reasons can be addressed. The estimated impacts are smaller than the observed reductions, although with a smaller gap than in the case of Monza. In Bergamo we estimated a variation of -9$\mu g/m^3$ (-32.1%), while in Brescia -9.5$\mu g/m^3$ (-32.7%). The gaps can be attributed to the strongly heterogeneous physical conformations of both areas. Both provinces are characterised by a northern mountain area, therefore sparsely inhabited and with high environmental quality, by the flat area in the south with high industrialised districts and medium-large sized lakes. According to the stations' typology, both the observed and estimated reductions may depend on the geomorphological structure of the territory and its reliefs. Considering the cities of Brescia and Bergamo, the urban traffic sites estimated variations of -27.2$\mu g/m^3$ and -22.5$\mu g/m^3$, respectively, while near less urbanised areas of the provinces the estimated reductions are always lower (in absolute value) than 14$\mu g/m^3$ and often not statistically significant. To have a better insight into the geographical effect, we consider three stations in the province of Brescia characterised by different territories. The urban station of Brescia city, the urban background station

of Lonato del Garda overlooking Lake Garda and the suburban background station of Odolo at 350 meters altitude. At Brescia site, the decisive weather variables are the wind blowing from South-East, which reduces the concentrations significantly during all the seasons (-12.2$\mu g/m^3$ in winter and -11.1$\mu g/m^3$ in summer), and the summer rainfall (-17.5$\mu g/m^3$). The trend is negative (-1.2$\mu g/m^3$ per week) and statistically significant. In Odolo, the estimated lockdown reduction is 7.8$\mu g/m^3$, but it is not statistically significant. The trend is absent, and except for the winter rainfall and the summer temperature, all the weather covariates are not significant. In Lonato, the estimated variation is -4.8$\mu g/m^3$. The negative trend effect is weak but significant, while the main reducing drivers are the winds coming from the Garda Lake, i.e. from North and East (around -8$\mu g/m^3$ during summer and -5.5$\mu g/m^3$ in spring).
The results associated with the southern provinces - that is, Pavia (PV), Lodi (LO), and Cremona (CR) - were in line with expectations: given the local agricultural and rural context, the observed and estimated reductions of NO_2 levels were very modest. In Pavia, the estimated reduction of NO_2 is about 4.8$\mu g/m^3$ (-20.6%), in Lodi it is 4.9$\mu g/m^3$ (-19.7%), while in Cremona it is 4.2$\mu g/m^3$ (-24.4%). The only exception is represented by Mantova (MN), where the estimated NO_2 reduction reached 47.5%.
As shown in Figure 4.4, the pre-lockdown NO_2 concentrations in the southern provinces during spring were considerably below the regional average and those of the other provinces. Thus, starting from relatively low levels, the observed reductions are small. In Cremona and Mantova, the estimated lockdown impact appeared even stronger than the observed raw differences. This last fact can be explained by the use of meteorological variables to adjust the estimates. The inclusion of meteorological covariates in the models allows obtaining unbiased estimates of the impact of restrictions on NO_2 concentrations while controlling for possible climatic variations.
It is plausible that the meteorology hid part of the reduction caused by the restrictions during the lockdown period. Indeed, thanks to the available data, it is possible to observe that during the lockdown period (March-May 2020) in both provinces the observed average temperature is about 0.30 to 0.40 Celsius higher than the average in the same period between 2015 and 2019. Moreover, the rainfall occurring in 2020 is on average lower than in the period 2015-2019, as well as the wind speed coming from the North-East and North-West, i.e. from the Alps. These facts are compatible with a reduction in the atmospheric cleaning capacity in the area. Therefore, the higher temperatures and the failure to recycle the air may have increased the quantities of oxides in the atmosphere, compensating for the considerable reduction generated by the restrictions. In Section 4.8, we introduced the case of the rural ground station in Schivenoglia as an example of null lockdown impact. The associated plot showed that the NO_2 concentrations in the area were already deficient before the lockdown. From the estimated model at that station, we deduce that the meteorology provides a limited effect in reducing concentrations. Indeed, most of the coefficients are shrunk toward low values and are often not statistically significant. In particular, the wind blowing from all the directions are associated with coefficients around -1$\mu g/m^3$ to -5$\mu g/m^3$. As stated in Section 4.7.2 commenting the estimated coefficients associated to the weather covariates, the wind is often a fundamental factor in explaining the reductions. The only significant effect is due to the winter rainfall, which reduces the concentrations of -17.3$\mu g/m^3$.
The previous results on oxide reduction in Lombardy were consistent with the study conducted by Agresti et al., 2020, in which the impacts of vehicular traffic on weekly NO_2 concentrations in Lombardy and the metropolitan area of Milan were investigated by implementing physical-chemical models based on emission inventories. Their findings suggest that the lockdown restrictions generated reductions in NO_2 concentrations in March 2020 of 31% in Milan, 35% in Como, 27% in Bergamo, and 23% in Brescia compared to the data observed during previous years.

4.9 Conclusions and future developments

Our findings suggest that during the Italian COVID-19 lockdown, the air quality in Lombardy improved noteworthy due to the restrictions imposed by the Government. In general, all provinces experienced remarkable improvements, especially densely populated, and congested and trade areas.
The adopted time series model provided evidence of statistically significant reductions of NO_2 levels at 47 over 74 ground sites. Overall, the analysis of residuals and diagnostic checks confirmed the goodness of the models used. The estimated models report very high fitting values (often above 80%) and low prediction errors. Moreover, the residuals are weakly or not autocorrelated at many stations, and where some serial correlation persists, it is very poor. The estimates can therefore be considered unbiased and meaningful.
Compared to the observed average reduction, the estimated lockdown impact was mitigated by the decreasing temporal trend and local meteorology. The observed raw differences by provinces are almost everywhere larger than the estimated average impact. However, the reductions appear not to be ubiquitous all over the regional territory. The aggregation of the estimated impacts by province showed that all provinces experienced a relevant reduction. However, the effect was stronger in large urban areas located in the central industrialised belt (-36.1% in Milan and -32.7% in Brescia), while the southern territories, characterised by flat rural landscape, showed a significant but weaker reduction of NO_2 concentrations (-20.6% in Pavia and -19.7% in Lodi). Moreover, considering the type of area covered by the stations, the highest reductions in nitrogen dioxide levels were associated with traffic control units (-38.2%), while reductions in rural areas were smaller (-21.7%). These results are consistent with other studies showing that the lockdown restrictions, which affected mobility and production activities, drastically lowered traffic and therefore stopped the primary sources of oxide emissions.
In this paper, we considered the spatial dimension indirectly, as it has been used to characterise the results achieved in terms of estimated variations and not directly at the estimation stage. However, we recognise that spatial and spatio-temporal modelling would improve the quality of the estimates considerably. Crucial would be the spatial prediction of variations in sites where there is no direct monitoring and the mapping of the occurred variations over the whole region. This would improve the accuracy of the reduction estimates by province, which currently depends on the number of stations located within the provincial boundaries.

TABLE 4.2: Summary statistics for the coefficients estimates of the weather covariates

	Negative sign		**Positive sign**		**Null**				
	Signif.	Not signif.	Signif.	Not signif.	**coefficients**	**Min**	**Mean**	**Max**	**S**
	%	%	%	%	%	$\mu g/m^3$	$\mu g/m^3$	$\mu g/m^3$	
					Winter				
Temperature	71.6	9.5	1.4	2.7	14.9	-1.8	-0.9	0.6	
Humidity	14.9	25.7	13.5	33.8	12.2	-0.5	-0	0.2	
Wind speed NE	27	12.2	1.4	18.9	40.5	-102.6	-6.7	10.4	
Wind speed SE	43.2	12.2	1.4	8.1	35.1	-39.9	-10.6	19.7	
Wind speed SW	10.8	12.2	5.4	29.7	41.9	-33	-1.2	20.4	
Wind speed NW	41.9	9.5	2.7	9.5	36.5	-19.9	-5	25.9	
Rainfall	47.3	32.4	0	4.1	16.2	-43.7	-16.9	8.9	
					Spring				
Temperature	70.3	16.2	0	0	13.5	-1.1	-0.5	-0.1	
Humidity	24.3	14.9	0	0	60.8	-0.5	-0.1	-0	
Wind speed NE	35.1	36.5	0	0	28.4	-41.2	-6.7	-0.8	
Wind speed SE	45.9	33.8	0	0	20.3	-20.2	-7.1	-0.8	
Wind speed SW	32.4	28.4	0	1.4	37.8	-41.7	-8	1.1	
Wind speed NW	47.3	31.1	0	0	21.6	-12.7	-5.7	-1	
Rainfall	9.5	67.6	0	8.1	14.9	-16.5	-4.7	12.9	
					Summer				
Temperature	20.3	27	0	1.4	51.4	-1	-0.4	0	
Humidity	63.5	27	0	0	9.5	-0.6	-0.2	-0	
Wind speed NE	21.6	59.5	0	0	18.9	-24.4	-5.9	-0.9	
Wind speed SE	31.1	50	1.4	0	17.6	-25.2	-5.8	17.2	
Wind speed SW	24.3	40.5	0	0	35.1	-20.4	-6.4	-0.8	
Wind speed NW	21.6	51.4	0	0	27	-16.3	-5.9	-0.7	
Rainfall	1.4	44.6	1.4	10.8	41.9	-17.5	-1.9	8.6	
					Autumn				
Temperature	75.7	6.8	0	0	17.6	-1.2	-0.6	-0.1	
Humidity	21.6	0	1.4	5.4	71.6	-0.5	-0.2	0.1	
Wind speed NE	24.3	32.4	1.4	1.4	40.5	-51.7	-7.5	27.7	
Wind speed SE	36.5	32.4	2.7	2.7	25.7	-31.8	-7.4	20.2	
Wind speed SW	9.5	21.6	1.4	5.4	62.2	-33.9	-6	11.6	
Wind speed qNW	14.9	13.5	1.4	12.2	58.1	-17.8	-5.1	7.1	
Rainfall	18.9	62.2	0	5.4	13.5	-13	-5.5	13.8	

Note: we considered as statistically significant the ML post-LASSO coefficients showing p-values below 5%. Percentages are calculated with respect to the 74 involved stations. The column *Null coefficients* reports the percentage of stations with null estimates, that is those stations in which the LASSO did not selected the covariate and set up to zero by the ML post-LASSO estimator.

TABLE 4.3: Lockdown variations of NO_2 concentration classified by station type.

Stations type	**Number of stations**	**Pre-lockdown NO_2**		**In-Lockdown** NO_2	**Observed differences**	**Model-based differences**	
		Average	Std. Dev.	Average	Average	Estimate	Std. Error
Stations with non-null effect							
Background	39	25.3	3.4	15.4	-9.8	-8.7	4.2
Rural	2	16.1	3.5	11.8	-4.3	-3.5	2.8
Traffic	22	38.5	4.9	20.9	-17.5	-14.7	5.1
	63	29.6	3.9	17.2	-12.3	-10.6	4.5
Stations with null effect							
Background	3	15.6	2.0	10.4	-5.2	-	-
Rural	5	19.6	2.9	16.5	-3.1	-	-
Traffic	3	36.4	7.0	24.7	-11.7	-	-
	11	23.1	3.8	17.1	-6.0	-	-
Overall	74	28.6	3.9	17.2	-11.4		

Note: all the reported values are measured in $\mu g/m^3$. The "pre-lockdown" average and the standard error are computed using the period 9th March- 18th May of years 2015-2019, while the "lockdown average" uses the period 9th March- 18th May, 2020. The observed difference is the difference between these averages. The model-based difference and its standard error are obtained by the statistical model discussed above.

Chapter 5

Rank-based nonparametric tests for event studies for cross-sectional correlated air quality time series

Based on:

Pelagatti, M., & Maranzano, P.
'Nonparametric tests for event studies under cross-sectional dependence'

5.1 A quick overview of the contribution of event studies in the econometric and statistical literature

5.1.1 The origins: the role of event studies in empirical finance

Event studies research framework is a sub-field of econometrics which involves statistical methods used to analyse the determinants of time series reaction to shocks caused by distinct event types. Event studies are used for the practical purpose of testing if the abnormal effects generated by the event of interests are significantly different from zero, and therefore not the result of pure randomness. This assessment will be performed by statistical hypothesis testing. Event study results are oftentimes used as dependent variables in regression analyses.
The first stage of event studies literature dates back to the 60s and 70s. The seminal studies by Ball and Brown, 1968 and Fama et al., 1969 introduced the methodology that is essentially still in use today, which is based on the concepts of *abnormal returns* and *market regression models*.
Starting from the 80s of the last century, the statistical literature on these methods has found fertile ground in empirical financial economics and empirical corporate finance. Readers interested in survey articles on event studies may refer to Campbell et al. (1996), MacKinlay (1997b), and Cowan (1992). Nonetheless they are not recent, they contain the necessary information. An engaging article that provides economic insights on event study is that of Prabhala (1997). Finally, the web page `https://www.eventstudytools.com/significance-tests` is an up-to-date repository of many event study test statistics. Within the empirical corporate finance framework, event studies statistics are mainly used to examine the behaviour of a firm's stock prices around corporate events. In a corporate context, the usefulness of event studies arises from the fact that the magnitude of abnormal performance at the time of an event provides a measure of the unanticipated impact of this event on the wealth of the firms' shareholders (Kothari and Warner, 2007). In fact, given the assumption of rationality in the marketplace, the

effect of an event is reflected immediately in asset prices. Therefore the event's economic impact can be measured using asset prices observed over a relatively short time period (Campbell et al., 1998).
In empirical finance applications, event studies yield as an outcome *abnormal returns*, which can be aggregated to draw overall inferences for the event of interest. The aggregation concerns two dimensions: the time dimension and the cross-sectional dimension. Abnormal returns can be cumulated over time to obtain *cumulated abnormal returns*, can be averaged over time and, in the presence of multiple time series, they can also be aggregated in cross-sectional sums or averages. When both aggregation over time and cross-sectional dimensions are computed, the abnormal returns assume the form of *cumulated or averaged cross-sectional abnormal returns*.
Event studies can focus both on short-horizon and long-horizons impacts. Example of the former is those studies focusing on announcement effects for a short-horizon around an event which provide evidence relevant for understanding corporate policy decisions. On the other hand, event studies focusing on long-horizons following an event can provide key evidence on market efficiency (Brown and Warner, 1980; Fama, 1991). In fact, persistent non-zero abnormal returns observed after some particular events are inconsistent with market efficiency.
Important innovations and modifications to the original seminal work have been proposed from the 80s up to nowadays. Between the 1980s and 1990s, when computational capacity was still limited, the literature on event studies handled with the issue of highly-frequency data, moving from monthly data to daily and hourly observations. Today these two are the analytical standard for any application in the field. As an example in Brown and Warner, 1980, the authors considered implementation issues for data sampled at a monthly interval, while the same authors in the 1985 paper deal with daily data. Until then, most of the models used were parametric and were subject to some heavy statistical assumptions, such as the linearity of the parameters or the fully-parametric formulation of the statistics involved.
Since the 2000s, part of the literature has focused on the violation of the classical statistical assumption underlying parametric event studies models. In particular, there has been a shift from a purely parametric modelling approach to one that includes nonparametric statistics, such as statistics based on ranks and signs.

5.1.2 Event studies in other related domains

Beyond empirical financial econometrics, event studies are useful in other areas. For example, in law and economics, event studies are used to examine the effect of corporate governance mechanism, public and companies regulations, and assess the effects of merging and acquisition (M&A) and takeovers as in Fiordelisi and Martelli (2011). See, for example, the not recent but extensive literature review proposed by Bhagat and Romano (2002a) and Bhagat and Romano (2002b) concerning the application of event studies methods in the corporate law research field.

5.1.3 Our scientific contribution

This chapter of the manuscript aims to provide a scientific contribution both methodologically and applicatively to event studies research.
On the methodological side, two novel nonparametric statistics based on ranks are presented and discussed. The proposed statistics enjoy the benefit of dealing with multivariate time series characterised by high cross-sectional correlation. The cross-sectional dependence topic is crucial in the modern literature on event studies, as many real cases present the characteristics from the correlation between series, and the use of standard methods would generate biased conclusions far from reality. For one of these tests (P1), we carry out a rigorous derivation of its

asymptotic distribution under the null hypothesis of no shift in the event-window returns. The P1 test is distribution-free, valid under cross-sectional dependence for any length of the event window and any number of stocks in the analysis. Unlike some of the existing nonparametric tests, the asymptotic approximation improves as the estimation window length increases regardless of the event window size. This implies that, under the typical event study set-up, the asymptotic approximation is extremely accurate as our simulation experiments prove. The asymptotic distribution of our second proposal, P2, is derived under the assumption that a central limit theorems (CLT) under cross-sectional dependence applies for abnormal returns (from which common factors have been removed). These two tests' performance is compared to those of the other 16 tests selected in the relevant literature in a large simulation study. We believe this empirical comparison is one of the most complete as we included all tests that take cross-sectional dependence into account and tests that can be easily modified to work under this assumption. Indeed, we modify the Corrado test with Tuckey's lambda transform Corrado (2011) to make it reliable also under dependence.
Secondly, it is intended to extend the methodology's area of application to previously unexplored but highly relevant domains. Specifically, we refer to the time series of pollutant concentrations, already presented and analysed in the previous chapters. There are no examples of such applications in environmetrics literature, even though the environmental series presents some characteristics very similar to the financial world, such as the non-normality of the data, the high temporal persistence and the cross-sectional dependence between the measures. In particular, we are interested in showing event studies methodologies' capability to detect and assess shocks in observed concentrations of pollutants following abnormal events. Such events may be either artificial, e.g. the introduction of a new policy, or natural or near-natural, e.g. the lockdown restriction due to COVID-19 in 2020. We will present an empirical example of event study using the lockdown restrictions imposed in the Lombardy region by the national authorities to face the COVID-19 disease spread. The aim is to test whether traffic and mobility limitations have affected average levels of nitrogen oxide concentrations. Expectations on research hypotheses are for a significant reduction in average levels supported by various indicators.

5.2 Outline of an event study

As stated by Campbell et al. (1998), although there is no unique structure, the literature generally suggests seeing the event study as a procedure having four main steps.

- Event definition and observation subsetting
- Regression analysis and abnormal residuals definition
- Estimation procedure
- Statistical testing

5.2.1 Event definition and observation subsettings

The initial step concerns the definition of the event of interest and identifies the period over which the time series involved will be examined and tested. This period is commonly named *event window*. For example, in a financial framework, if one is looking at the information content of a macroeconomic news announcement with daily data, the event will be the announcement date, and the event window might be the week following that day. This is done to capture the price effects of announcements which occur after the stock market closes on the announcement

day. Similarly, in the case of unexpected socio-economic events, such as the lockdown imposed in response to the COVID-19 crisis (see Chapter 4), or the introduction of public policies such as Area B (see Chapter 3), it is possible to identify the event with the date of occurrence and the event window as the period thereafter.
The period before the event may also be of interest and included in the analysis. For example, in the previous macro news announcement case, the market may acquire information about the novelty before the actual announcement, allowing the researcher to examine pre-event returns. The same holds for the lockdown restrictions, as rumours of possible public countermeasures were already circulating a few days before the decree was approved.
We now introduce some simple temporal notations useful to understand the concepts described in the following sections.

- Let $\tau = 0$ be the event date;
- Let T_0 be the time index for the first available observation;
- Let T_1 be the time index for the last observation included in the estimation window;
- Let T_2 be the time index for the last observation in the event window;
- Let T_3 be the time index for the last available observation;
- The set $\Omega_0 = T_0 + 1, T_0 + 2, ..., T_1$, of cardinality n_0, denotes the set of time indexes for the *estimation window*, that is, the sample on which the market model is estimated.;
- The set $\Omega_1 = T_1 + 1, T_1 + 2, ..., T_2$, of cardinality n_1, denotes the *event window*, that is, the time indexes at which we want to test the presence of abnormal movements on returns.
- $\Omega = \Omega_0 \cup \Omega_1$ has cardinality $n = n_0 + n_1$.

We will refer to *single-event window* when the event window is composed by just one observation, whereas we will talk of *multi-event window* when the event window includes multiple consequent observations.
Generally, in financial applications $n_0 \in [150, 250]$ and $n_1 \in [1, 10]$ daily returns (see for example Brown and Warner, 1985; Lummer and McConnel, 1989; Kolari and Pynnönen, 2010).

Parallelism between event studies timing and statistical learning principles

The estimation window and the event window are used in event studies according to a training-test principle commonly used in the statistical learning framework. Drawing a similarity, the event window can be viewed as the validation test, while the estimation window is seen as the training set. From this perspective, the abnormal residuals are comparable to the prediction errors produced by the model estimated with the training data. Indeed, model parameters are estimated using the observations contained in the estimation window. Then, using the parameter estimates, the estimated normal residuals for the event window are computed. Finally, the abnormal residuals in the event window are calculated as the difference between the estimated normal residuals and the observed residuals.

Event timeline and estimation scheme

The timing sequence and the estimation pipeline are illustrated in Figure 5.1. The timeline (upper plot) shows three main sub-periods: the estimation window, the event window (which includes both values prior and posterior to the event date, and the post-event window. In our

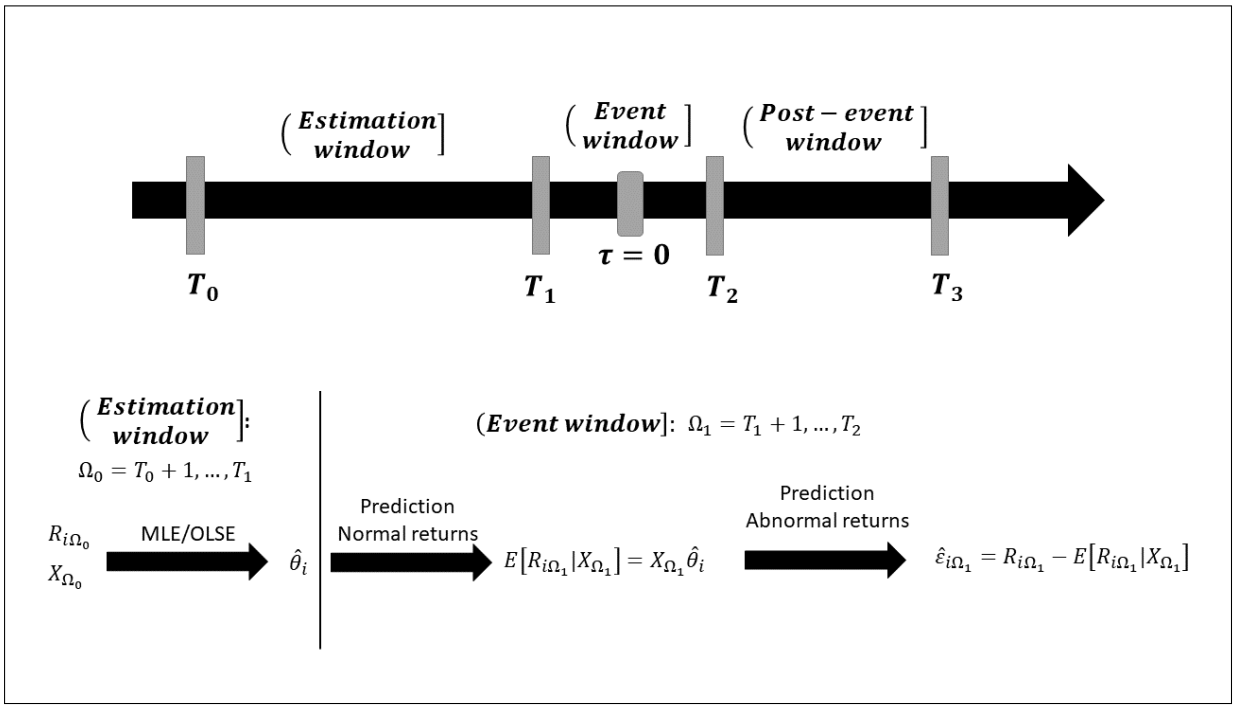


FIGURE 5.1: Event studies time sequence.

future purposes, the post-event window will be ignored since it does not provide any information about the series's behaviour. Hence, our timeline will be split just into the estimation window and post-event window.
In the lower plot, we represented a simplified version of the estimation scheme, which involves both the estimation and the event windows. The graph's purpose is to illustrate clearly which are the roles covered by the two windows and how they can be read in a statistical learning key, as said in the previous paragraph. The notation and the analytical definition of normal and abnormal residuals, as well as the estimation techniques details, will be explained in the following sections.

5.2.2 Regression analysis and abnormal returns in financial econometrics contexts

Despite the statistical methodologies in event studies have reached high levels of analytical quality, and at the same time also high complexity degrees, approaching more and more faithfully the jagged reality they intend to analyse, the fundamentals have remained unchanged. Financial event studies are based on some simple statistical concepts, such as *normal returns*, *abnormal returns* and their variants. The concept of residue derives directly from the regression techniques used in the computational process of event studies, in which the residuals are computed as the difference between the observed dependent variable and the values estimated through the empirical model.
Let i be the index for cross-sectional units, with $i = 1, ..., m$ and let t be the time index $t = T_0, T_0 + 1, ..., T_2$. In the context of financial economics, the prices or returns of stocks and securities quoted on the financial markets are considered as cross-sectional units i. The *abnormal returns* are defined as the actual ex-post return of the stock i over the event window minus the firm's normal return over the event window (Ω_1). The *normal returns* are defined as the return expected if the event did not occur.

Let P_{it} be the price of stock i at the end of day t and let X_t be the conditioning information for the normal performance model. Classical examples of conditioning informations are market indices and market portfolios. For each stock i the actual return at time t is

$$R_{it} = \frac{P_{it} - P_{it-1}}{P_{it-1}} \qquad \forall t \in \Omega. \tag{5.1}$$

Using the information given by the return for a specific stock i at time t and by a set of conditioning informations at time t the ormal returns at time t for stock i can be defined as the expected value of the returns conditioned to the informations:

$$NR_{it} = E[R_{it}|X_t] \qquad \forall t \in \Omega. \tag{5.2}$$

Consequently, the abormal returns at time t for stock i (AR_{it} or ε_{it}) are defined as the differences between the observed return at time t and the normal returns at time t for stock i:

$$AR_{it} = \varepsilon_{it} = R_{it} - E[R_{it}|X_t] \qquad \forall t \in \Omega, \tag{5.3}$$

where $E[R_{it}|X_t]$ are the normal returns for time period t.
The basic idea behind event studies statistics in the empirical finance field is to remove the general stock market price movements and separate confounding events.
There are two common choices for modeling the normal return: the *constant-mean-retum model* where the conditioning information X_t is a constant, and the *market model* where X_t is the market index return. As the name implies, the constant-mean-return model assumes that the mean return of a given stock is constant through time. The market model assumes a statistical relationship, both linear and non-linear, between the market return and the security return. We will focus just on the case of linear regression models between the response and the covariates for our purposes.

Distributive assumption for abnormal returns

It is conventional to assume that asset returns are jointly multivariate normal with mean 0 and variance-covariance matrix Σ and independently and identically distributed (I.ID.) over time. Sometimes, stock returns are referred to be weak white noise Gaussian random variables with:

$$E[R_{i,t}] = \mu_i \qquad \forall t$$

$$VAR[R_{i,t}] = \sigma_i^2 \qquad \forall t$$

$$COV[R_{i,t_1}, R_{i,t_2}] = 0 \qquad \forall t_1 \neq t_2.$$

This distributional assumption is sufficient for the constant-mean-return model and the market model to be correctly specified and permits the development of exact finite-sample distributional results for the estimators and statistics. Inferences using the normal return models are robust to deviations from the assumption (Campbell et al., 1998).
As often reported in the statistical air quality literature, airborne pollutants time series are affected by strong serial correlation over time, strong right-skewness and, in some cases, heteroscedasticity. Econometric theory suggests that these facts lead to biased and inconsistent estimates. However, as demonstrated by other applications in this manuscript and other papers, when environmental series are modelled by simple linear approaches and by persistence models (e.g. ARIMA or ARIMAX models), the residuals are white noise processes with Gaussian distribution.

Constant-mean-return model

The constant-mean-return model is simply defined as a linear regression model where the returns are a linear function of the intercept plus a random noise with Gaussian and I.ID. distribution. Then, for each stock i we have:

$$NR_{it} = E[R_{it}|R_t] : R_{it} = \alpha_i + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i}), \tag{5.4}$$

where R_{it} is the return for stock i at time t and ε_{it} is a white noise random componenent.

Adjusting returns for general market conditions

The *market model* is defined as a linear regression model where the returns are a linear function of the intercept and the market index portfolio plus a random noise with Gaussian and I.ID. distribution. The linearity assumption implies that the model inherits all the properties typical of linear regression, in particular the gaussian distribution of the coefficients. Then, for each stock i we have:

$$NR_{it} = E[R_{it}|M_t] : R_{it} = \alpha_i + \beta_i M_t + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i}), \tag{5.5}$$

where R_{it} is the return for stock i at time t, M_t is the market index at time t and ε_{it} is a white noise random componenent.

Multiple market regression models

The simple market model can be easily extended in order to include further covariates, leading to a multivariate market regression model. A well-known multivariate model of returns adjustment is the one proposed by Eugene and French, 1992; Fama and French, 1993; Fama and French, 1996 and called multi-factor model:

$$NR_{it} = E[R_{it}|M_t] : R_{it} = \alpha_i + \beta_M M_t + \beta_{SB} SMB_t + \beta_{HL} HML_t + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i}), \tag{5.6}$$

where R_{it} is the return for stock i at time t, M_t is the market index at time t, SMB_t is the small-minus-big market capitalization factor return and HML_t is the high-minus-low book-to-market equity factor return and ε_{it} is a white noise random componenent.
Factor models potentially benefit from reducing the variance of the abnormal return by explaining more of the variation in the normal return. Typically the factors are portfolios of traded securities. The market model is an example of a single-factor model, but in a multifactor model one could include further industry indexes in addition to the market index.

5.2.3 Generalisation to the concept of abnormal residuals

Dealing with the specific case of the financial literature of the event studies, we have seen how, in reality, the modelling is essentially based on a few simple statistical and econometric concepts. In particular, we have seen how abnormal returns derive from the linear regression of a dependent variable on one or more covariates. Therefore it is possible to generalise this situation to a wider context in which the key role is held by the *abnormal residuals*.

The linear regression model for the general case

The literature on event studies in financial context proposed many linear regression approaches to model the relationship between the response variable and the conditioning information.

Let i be the index for cross-sectional units, with $i = 1, ..., m$ and let t be the time index $t = T_0, T_0 + 1, ..., T_2$. Let's suppose to have conditioning information set X_{it} consisting of multiple covariates and a single response variable Y_{it}. Let's also assume that for each cross-sectional unit i, the information set may vary in terms of covariates.
Let there exists a statistical relationship between the dependent variable Y_{it} and the conditioning information X_{it}. For each unit i, the expected values of variable Y at time t given the set X_{it} are represented by the conditional expected value of Y on X:

$$\hat{Y}_{it} = E[Y_{it}|X_{it}]. \tag{5.7}$$

Assuming a linear relation between Y and X, the conditional expectation can be easily expressed by a linear regression function with multiple covariates and parameters:

$$Y_{it} = E[Y_{it}|X_{it}] = \alpha_i + \beta_1 X_{1it} + \beta_2 X_{2it} + ... + \beta_P X_{Pit} + \varepsilon_{it}, \tag{5.8}$$

or equivalently, in a more compact notation,

$$Y_i = X_i'\theta_i + \varepsilon_i, \tag{5.9}$$

where Y_i is the (nx1) vector of response variable for unit i, X_i is the (nxP) matrix of covariates for unit i, $\theta_i = (\alpha_i, \beta_{1i}, ...\beta_{Pi})'$ is a $(p+1)$-dimensional parameter vector and ε_i is a (Ωx1) vector of regression residuals.

Abnormal residuals

With the term *abnormal residuals*, hereafter AR_{it} or ε_{it}, we define the residuals of the regression of the dependent variable Y_{it} on a set of conditioning information X_{it} calculated using the event window Ω_0 for unit i. AR_{it} are then computed as the difference between the actual response values Y_{it} and the fitted values obtained through the linear regression model in 5.9. In this context, what we have called normal returns in the financial context are nothing more than the expected or fitted values, which are now interpreted as the value of Y that would be expected if the event did not take place:

$$\hat{Y}_{it} = X_{it}'\hat{\theta}_i, \tag{5.10}$$

where $\hat{\theta}_i$ is a consistent estimator of the parameter vector θ computed using the estimation window Ω_0 The abnormal residuals for unit i at time t are then computed as follows:

$$AR_{it} = \hat{\varepsilon}_{it} = Y_{it} - \hat{Y}_{it} = Y_{it} - X_{it}'\hat{\theta}_i \qquad \forall t \in \Omega. \tag{5.11}$$

Considering the abnormal residuals within the event-window, we denote them as:

$$AR_{i\Omega_1} = \hat{\varepsilon}_{i\Omega_1} = Y_{i\Omega_1} - X_{i\Omega_1}'\hat{\theta}_i \qquad \forall t \in \Omega_1. \tag{5.12}$$

The abnormal residuals in the event window Ω_1 can be interpreted as the abnormal values of Y not explained by the model and potentially generated by the event of interest.

Cumulated abnormal residuals

By the term *cumulated abnormal residuals* ($\text{CAR}_{i,\tau}$) we mean the cumulative sum of abnormal residuals (AR_{it}) in a given time window denoted as τ, with $T_1 \leq \tau \leq T_2$. When the cumulative window coincides with the event-window, we refer to cumulated abnormal residuals in the event-window, hereafter $\text{CAR}_{i\Omega_1}$. Cumulated abnormal residuals accomodate the presence of multiple-sampling intervals within the event window, i.e. the event-window is composed by

more than one observation. Recalling that $\varepsilon_{it} = AR_{it}$ are the abnormal residuals for cross-sectional unit i, the cumulated abnormal residuals over the event window Ω_1 are defined as:

$$CAR_{i\Omega_1} = \sum_{t=T_1+1}^{T_2} \hat{\varepsilon}_{it} = \sum_{t=T_1+1}^{T_2} AR_{it}. \tag{5.13}$$

Standardized abnormal returns

In a real-world application and a simulation scheme, abnormal and cumulated abnormal residuals are often standardised or scaled with respect to their variability. The standardisation is obtained by dividing the abnormal or cumulated abnormal residuals by their standard deviation. Given the abnormal residuals sequence, that is $\text{AR}_{it} \quad \forall t \in \Omega$, the standardized abnormal residuals (SAR_{it}) of unit i at time t are defined as follows:

$$SAR_{it} = \frac{AR_{it}}{\hat{\sigma}_{it}}, \tag{5.14}$$

where $\hat{\sigma}_{it}$ is a consistent estimator of the standard deviation of ε_{it} for the stock i estimated using the estimation set, that is Ω_0. If we call X_0 and X_1 the design matrices for the subsamples Ω_0 and Ω_1, and estimate the coefficients β by ordinary least squares, then it is possible to prove that the variance of the i-th regression residual has form $\sigma^2(1 + c_i)$, where, for $i \in \Omega_0$, c_i is the i-th element of the diagonal of $X_0(X_0^\top X_0)^{-1}X_0^\top$, while for $i \in \Omega_1$, c_i is given by the corresponding element on the diagonal of $X_1(X_0^\top X_0)^{-1}X_1^\top$. Under the usual conditions, using $\hat{\sigma}_{it}^2$ or $\hat{\sigma}^2(1 + c_i)$ for standardising the residuals is asymptotically equivalent, and from our simulations on finite samples, for the sample dimensions used in event studies, there is no practical difference between the two choices.
From Kolari and Pynnönen (2010) we can state that the basic intuition of this operation is that standardisation weights individual observations by the inverse of the standard deviation, which implies that more volatile, i.e. noisier, observations get less weight in the averaging than the less volatile or more reliable observations. Moreover, while scaled abnormal returns are more difficult to interpret than raw returns, they have been proven to exhibit better statistical properties. Thus, scaled returns should be used only for statistical testing purposes as signal detection tools of the event effect, while raw returns carry the economic information for interpretation purposes when a signal is detected.
As in the case of single returns, it is also possible to define the formula for the standardized abnormal returns ($\text{SCAR}_{i\Omega_1}$) for stock i in the event window Ω_1 as follows:

$$SCAR_{i\Omega_1} = \frac{CAR_{i\Omega_1}}{\hat{\sigma}_{i\Omega_1}} = \frac{\sum_{t\in\Omega_1} AR_{it}}{\hat{\sigma}_{i\Omega_1}}, \tag{5.15}$$

where $\hat{\sigma}_{i\Omega_1}$ is a consistent estimator of the standard deviation of $\text{CAR}_{i\Omega_1}$ estimated using the set Ω_0.
Since SARs are simply correct versions of abnormal returns, it is possible to define their cumulative version over an interval of time. As before, when the aggregation time interval coincides with the event window Ω_1 we will talk about cumulated standardized abnormal returns ($\text{CSAR}_{i\Omega_1}$) within the event-window for stock i, which are defined as:

$$CSAR_{i\Omega_1} = \sum_{t\in\Omega_1} SAR_{it}. \tag{5.16}$$

5.2.4 Estimation procedure and statistical properties

Once the most suitable model to represent the relationship between Y and X has been identified, the model's parameters must be estimated using the subset of the original data belonging to the estimation window. Given the linear structure of the equations listed above, the models can be expressed simply using linear regression functions with Gaussian errors. Thus, the parameters can be easily estimated using maximum likelihood estimators (MLE) or ordinary least squares estimators (OLSE) in the estimation window. Under general conditions, both OLSEs and MLEs are consistent estimator of the unknown parameters.

Statistical properties of abnormal residuals and cumulated abnormal residuals

Without loss of generality, we assume to have chosen a linear regression specification with a single covariate and the intercept and to have obtained consistent estimates of unknown parameters. Let $\hat{\theta}_i = [\hat{\alpha}_i, \hat{\beta}_i]'$ be the OLSEs of the unknown parameters α and β. Let be the event window the set of indices from $T_1 + 1$ to T_2, i.e. $\Omega_1 = T_1 + 1, ..., T_2$. For each unit i let Y_{i,Ω_1} be the (n_1x1) vector of the response variable and $X_{i\Omega_1}$ be the (n_1x1) vector of actual covariate in the event-window Ω_1. Then, the (n_1x1) sample vector of estimated abnormal residuals within the event-window is calculated as follows:

$$\hat{\varepsilon}_{i\Omega_1} = Y_{i\Omega_1} - \hat{\alpha}_i - \hat{\beta}_i X_{i\Omega_1} = Y_{i\Omega_1} - X_{i\Omega_1}\hat{\theta}_i, \tag{5.17}$$

where $\hat{\theta}_i = [\hat{\alpha}_i, \hat{\beta}_i]'$ is the parameter vector estimate. Conditioning to the observed values of the covariate in the event window it is possible to demonstrate that the abnormal residuals in the event-window Ω_1 have conditional expected value

$$\mu_{\varepsilon_{\Omega_1}} = E[\hat{\varepsilon}_{\Omega_1}|X_{i\Omega_1}] = E[Y_{i\Omega_1} - \hat{\alpha}_i - \hat{\beta}_i X_{i\Omega_1}|X_{i\Omega_1}] = 0, \tag{5.18}$$

and conditional variance equals to

$$\begin{aligned} V_i = VAR(\hat{\varepsilon}_{\Omega_1}|X_{i\Omega_1}) &= E[\hat{\varepsilon}_{\Omega_1}\hat{\varepsilon}'_{\Omega_1}|X_{i\Omega_1}] = \sigma^2_{\varepsilon_i} + X_{\Omega_1}(X'_{\Omega_0}X_{\Omega_0})^{-1}X'_{\Omega_1}\sigma^2_{X_i} \\ &= \sigma^2_{\varepsilon_i} + \frac{1}{n_1}[1 + \frac{(X_{i\Omega_1} - \hat{\alpha}_i - \hat{\beta}_i X_{i\Omega_1})^2}{\hat{\sigma^2}_{X_i}}], \end{aligned} \tag{5.19}$$

where X_{Ω_0} and X_{Ω_1} are the design matrices for the sub-samples Ω_0 and Ω_1 and $\sigma^2_{\varepsilon_i}$ and $\sigma^2_{X_i}$ are the disturbance variance and the sampling error variance, respectively. See Section 4.4.2 of Campbell et al., 1998 and Section 5B in MacKinlay, 1997a for further details about the proofs of equations 5.18 and 5.19.

From Equation 5.18 we notice that the abnormal residuals vector, having the conditional expectation of zero, is unbiased of the unknown parameter vector. The conditional covariance matrix of the abnormal residuals vector from 5.19 is composed of two parts. The first term, i.e. $\sigma^2_{\varepsilon_i}$, in the sum is the variance due to the future disturbances and the second term is the additional variance due to the sampling error in $\hat{\alpha}_i$ and $\hat{\beta}_i$, i.e. $\hat{\theta}_i$. This sampling error, which is common for all the event window observations, leads to a serial correlation of the abnormal residuals even though the true disturbances are independent over time. As the length of the estimation window n_0 becomes large, the second term approaches zero as the parameters' sampling error vanishes. Thus, the variance of the abnormal residuals will tend to $\sigma^2_{\varepsilon_i}$ and the abnormal residuals observations will become independent asymptotically. In practice, the estimation window can usually be chosen to be large enough to make it reasonable to assume a null contribution of the second term.

Under the null hypothesis of absence of impact of the event on the behaviour of the results, both in terms of variance and average, the abnormal residuals' distributional properties can be used to draw inferences over any period within the event window. Let us assume a Gaussian distribution for the regression residuals. Under the null hypothesis, the abnormal residuals have Gaussian distribution with mean and variance expressed by 5.18 and 5.19. Hence, the abnormal residuals $\hat{\varepsilon}_{i,\Omega_1}$ have distribution:

$$\hat{\varepsilon}_{i,\Omega_1} \sim N(0, V_i). \tag{5.20}$$

Furthermore, let us assume serial independence among the abnormal returns in the event window.
From the two assumptions of serial independence and Gaussian distribution of the abnormal residuals, we can state that the sequence $\hat{\varepsilon}_{i\Omega_1} \quad \forall t \in \Omega_1$ are independent and identically distributed (I.ID.) random variables. Thus, it is straightforward to derive some important distributional properties also for the cumulated abnormal returns. Recall that the cumulated abnormal residuals are defined in 5.13 as:

$$CAR_{i\Omega_1} = \sum_{t=T_1+1}^{T_2} \varepsilon_{it} = \sum_{t=T_1+1}^{T_2} AR_{it}.$$

From probability theory we know that the variance of a linear combination of I.ID. components is simply the linear combination of the common variance and that the expected value of the linear combination is the linear combination of the expectations (Stock and Watson, 2015). Hence, we can show that the cumulated abnormal residuals have expected value

$$E[\sum_{t\in\Omega_1} \hat{\varepsilon}_t | X_{i\Omega_1}] = \sum_{t\in\Omega_1} E[\hat{\varepsilon}_t | X_{i\Omega_1}] = 0, \tag{5.21}$$

and variance

$$V_{i,CAR} = VAR(\sum_{t\in\Omega_1} \hat{\varepsilon}_t | X_{i\Omega_1}) = \sum_{t\in\Omega_1} VAR(\hat{\varepsilon}_t | X_{i\Omega_1}) = n_1 V_i. \tag{5.22}$$

The probability theory also provide us the tools for assessing the distribution of cumulated abnormal residuals. Since we showed that abnormal residuals follow a gaussian distribution, the sum of independent gaussian random variables will follows a Gaussian distribution as well (Stock and Watson, 2015). Hence, the cumulated abnormal residuals within the event-window for unit i have distribution:

$$\hat{CAR}_{i,\Omega_1} \sim N(0, V_{i,CAR}). \tag{5.23}$$

5.2.5 Statistical testing

Recall from Section 5.1 that the main purpose of event studies methodologies is to statistically test whether the event window's abnormal effect is null or if it is different from zero. Therefore, the null hypothesis H_0 is to be interpreted as the absence of abnormal variations in residuals after the event of interest. In this situation, the variations are, therefore, to be attributed to randomness in the data.
According to general principles of inferential statistics, the null hypothesis (H_0) states that there are no abnormal residuals within the event window, whereas the alternative hypothesis (H_1) suggests the existence of abnormal residuals within the event window.

Testing abnormal residuals

When dealing with single-event windows, i.e. event-windows with just one observation, the only possibility is to test if that single estimated residual is abnormal or not. The same testing procedure is also feasible when dealing with multi-event windows. In this case, the statistical test is iterated individually for all the observations included in the event sample, producing a set of independent test statistics. Each statistics checks if the estimated residual for the time-stamp is abnormal or not. In a linear regression context, under the assumption of Gaussianity of the regression residuals, it is straightforward to show that the standardised test statistics for the average value of the distribution are distributed according to at-Student law with degrees of freedom equal to the number of estimated parameters (Stock and Watson, 2015).
Recalling that $\mu_{\varepsilon_{\Omega_1}}$ is the conditional expected value of the abnormal residuals within the event window, the hypothesis testing system for single-days can be expressed as follows:

$$H_0 : \mu_{\varepsilon_{\Omega_1}} = 0$$

$$H_1 : \mu_{\varepsilon_{\Omega_1}} \neq 0.$$

Considering single-event windows, the expectation coincides with the unique estimated abnormal residual. Hence, the previous system can be expressed equivalently substituting the conditional expectation with the estimated abnormal residual for stock i at time Ω_1:

$$H_0 : \mu_{\varepsilon_{\Omega_1}} = AR_{i\Omega_1} = 0$$

$$H_1 : \mu_{\varepsilon_{\Omega_1}} = AR_{i\Omega_1} \neq 0.$$

Rejecting the null hypothesis, the data suggest that the abnormal residual for a specific time-point was affected by some shocks. In contrast, a failure in rejecting the hypothesis suggests that the observed residual remains aligned with the previous period and those eventual variations were generated by randomness in the data.

Testing cumulated abnormal residuals

When working with multi-event windows, one can test that any one of the abnormal residuals in the window is zero or different from zero, or you can test that their aggregation, that is the cumulative version, have null mean in the event window. Similarly to what has been done for the single residuals, to test the significance of the shocks on the residuals, it is possible to set the following statistical hypothesis system for their cumulative sum:

$$H_0 : CAR_{i\Omega_1} = 0$$

$$H_1 : CAR_{i\Omega_1} \neq 0.$$

In the case of a test on cumulated residuals, the test statistics is a unique value for the whole event period. The test statistic is then compared to some critical values extracted by its distribution, allowing us to decide whatever the hypothesis can be rejected or not. Rejecting the null hypothesis, the data suggest that the abnormal residuals cumulated over time were affected by some perturbation along the event-window. At the opposite, a fail in rejecting the hypothesis suggests that the residuals were stable during the period, and their variations were totally random.

How to test the hypothesis

From a statistical perspective, the two hypothesis systems discussed above belong to the statistical significance tests class, as they provide information about the statistical significance or no significant variations in the average level of estimated returns. Generally, significance tests can be grouped in parametric and nonparametric tests. Parametric tests assume that individual time series are normally distributed, whereas nonparametric tests do not rely on any such assumptions.
We will discuss and compare the main test statistics available from event studies literature in the next sections. We will begin introducing the basic parametric statistics, both for the single-event and the multi-event cases, provided during the 80s and the 90s of the last century. We will then discuss available statistics for cumulated abnormal returns based on cross-sectional independence between the time series of interest. This assumption can be very limiting in contexts where cross-sectional dependence is a relevant factor that must be taken into account for economic or natural reasons. Therefore, in addition to discussing the cross-sectional statistics already existing in the literature, we will present two new nonparametric statistics based on the ranks of observations and account for a correction against strong cross-sectional dependence between the series.

5.3 A definition of abnormal airborne pollutant concentrations and their statistical properties

The event studies approach applied to the case of airborne pollutant concentrations allows us to estimate the concentration of the pollutant in excess of a hypothetical situation of normality and analyse if their variations in a time window are due to randomness in the data, i.e. no event effect, or if they are consequences of a specific event, i.e. presence of an event-induced effect.
Given the premises and definitions stated in Sections 5.1 and 5.2 about financial returns, we now propose a definition of normal, abnormal, standardised and cumulated abnormal airborne pollutant concentrations. We will refer to these concepts as *normal concentrations* (NC_{it}), *abnormal concentrations* (AC_{it}), *standardized concentrations* (SC_{it}) and *cumulated abnormal concentrations* (CAC_{it}) respectively.
The analogy is derived simply by assimilating the concentrations of a specific pollutant observed in a given time t at a given monitoring station i to the returns observed for a stock i at time t.
Let C_{it} be the observed airborne pollutant concentrations observed at time t at the monitoring station i. Moreover, let X_{it} be a set of conditioning information collected at time t at the same station i. Please note that the set X_{it} can include station-specific information and information common to all the sensor, such as calendar dates or other variables measured at aggregate geographical level. For each station and time stamp, conditioning the observed concentrations to the available informations, the normal concentrations at time t for station i can be defined as the conditional expected value:

$$NC_{it} = E[C_{it}|X_{it}] \qquad \forall t \in \Omega. \tag{5.24}$$

Thus, the abormal concentrations at time t for station i (AC_{it} or ε_{it}) are defined as the differences between the observed concentrations at time t and the normal concentrations at time t for station i:

$$AC_{it} = \varepsilon_{it} = C_{it} - E[C_{it}|X_{it}] \qquad \forall t \in \Omega, \tag{5.25}$$

where $E[C_{it}|X_{it}]$ are the normal concentrations for time period t.
Given the abnormal concentrations sequence, that is $AC_{it} \quad \forall t \in \Omega$, the standardized abnormal concentrations (SAC_{it}) of station i at time t are defined as follows:

$$SAC_{it} = \frac{AC_{it}}{\hat{\sigma}_{it}}, \tag{5.26}$$

where $\hat{\sigma}_{it}$ is a consistent estimator of the standard deviation of ε_{it} estimated using the set Ω_0.
Finally, the cumulated abnormal concentrations for station i over the event window Ω_1 are defined as:

$$CAC_{i\Omega_1} = \sum_{t=T_1+1}^{T_2} \varepsilon_{it} = \sum_{t=T_1+1}^{T_2} AC_{it}. \tag{5.27}$$

By applying a linear modelling approach, we can further specify the abnormal concentrations, and then both the standardized and cumulated abnormal concentrations, through the following set of equations. Let's model the observed concentrations C_{it} as a linear function of the set of conditioning information X_{it}:

$$C_{it} = \alpha_i + X_{it}\theta_i' + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon_i}). \tag{5.28}$$

The abnormal concentrations are the regression residuals obtained by implementing OLS or ML to estimate the parameter vector $[\hat{\alpha}, \hat{\theta}]'$:

$$\hat{\varepsilon}_{it} = C_{it} - \hat{\alpha}_i - X_{it}\hat{\theta}_i', \tag{5.29}$$

where $\hat{\alpha}$ and $\hat{\theta}'$ are the parameter estimates using the observations in the estimation window Ω_0. Considering the abnormal concentrations within the event-window Ω_1, the estimated abnormal concentrations for station i in the event-window are

$$\hat{\varepsilon}_{i\Omega_1} = \hat{\varepsilon}_{it} \qquad \forall t \in \Omega_1. \tag{5.30}$$

We point out that it is necessary to respect some statistical conditions to achieve meaningful and well-defined results. First of all, it is necessary to choose a suitable set of conditioning information for air quality assessment, that means to use weather covariates, calendar regressors or ad-hoc regressors for seasonal adjustment; secondly, it is essential to define a proper functional specification for the linear regression model. This last task has been solved in the empirical application by implementing models belonging to the ARIMAX class or seasonal models or regression models with ARMA residuals. The application presented in Chapters 3 and 4 of this manuscript could give the reader a detailed overview and motivation for their use. Therefore, considering these mandatory issues fulfilled, the regression residuals should enjoy the statistical properties established for abnormal returns: Gaussian distribution, serial independence and homoscedasticity. we will work on the cross-sectional dependence instead. The use of ML or OLS estimators for parameters estimation Abnormal concentrations, standardized abnormal concentrations and cumulated abnormal concentrations inherits all the statistical properties described in Section 5.2.4. Hence, they can be considered

5.4 Critical issues arising from cross-correlation and non-gaussianity in event studies

The use of rank-based nonparametric tests for event studies has a long history in the financial literature (McConnell and Muscarella, 1985; Corrado, 1989; Lummer and McConnel, 1989;

Zivney and Thompson, 1989; Corrado and Zivney, 1992; Cowan, 1992). The need for nonparametric (distribution-free) tests is mainly due to the non-Gaussian nature of financial returns, which tend to be highly leptokurtic and often negatively skewed. Indeed, Gaussian parametric tests such as the Patell t-test (Patell, 1976; Dodd and Warner, 1983) and the BMP t-test (Boehmer et al., 1991; Sanders and Robins, 1991) are often found over-sized for real-world stock returns (Corrado and Truong, 2008; Bartholdy et al., 2007).
However, many existing nonparametric tests have been designed to work under the assumption of cross-sectional independence of abnormal returns (Corrado, 2011; Ataullah et al., 2011, for instance), which does not hold when the study analyses a panel of contemporaneous time series with common event date. As it will be discussed later in this manuscript, the (average) correlation among stock returns can be strong enough to oversize any statistical test built under the assumption of cross-sectional independence even after the dependence on common factors has been eliminated.
Cross-sectional dependence and non-normality are serious issues when conducting event studies. As we will show in this section, both these typical features of both financial returns and air quality measurements work in the same direction to make practitioners find event effects even when no real effect is present in the data.
The classical Patell (1976) statistic used to carry out event studies on a sample of m firms on a single event date is given by

$$T_p = \bar{A}\sqrt{\frac{m(n_0-4)}{n_0-2}}, \tag{5.31}$$

where $\bar{A}$ is the average of the standardised abnormal returns and n_0 is the sample size of the estimation period. Under the assumption of normality and uncorrelatedness of the abnormal returns and for a moderate or large n_0 (> 30) the Patell statistic (or its large sample equivalent $\bar{A}\sqrt{m}$) is approximately distributed as a standard normal under the null of no event effect.
If the assumption of uncorrelatedness does not hold, as it is often the case when the event date is the same for all the series in the sample, then using the standard normal can be an extremely bad approximation to the real distribution of the test statistic. Indeed, as it can be easily derived, the variance of $\bar{A}$ is given by

$$V = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}\rho_{ij} = 1 + 2(m-1)\bar{\rho},$$

where ρ_{ij} represents the correlation between the returns i and j and

$$\bar{\rho} = \frac{1}{m(m-1)}\sum_{i=2}^{m}\sum_{j=1}^{i}\rho_{ij}$$

is the average correlation among the returns. Thus, T_p is approximately distributed as $N(0, V)$ under the null of no event effect and, if $\bar{\rho}$ is positive, the Patell test is oversized (i.e., it rejects too often) with the actual size of the test converging to 1 (i.e., the null hypothesis is always rejected) as the number of firms increases. In the rare case in which $\bar{\rho}$ is negative, the Patell test results undersized (i.e., it rejects too little). Figure 5.2 depicts the actual size of the Patell test given a nominal size of 5% for four different values of $\bar{\rho}$ as a function of the number of firms in the sample: even minimal deviations from zero of the average correlation can cause serious distortions in the size of the test in typical portfolio sizes.
The average correlation of the S&P100 stock returns in our sample is around 0.22, while that of NASDAQ is some 0.14. Of course, computing the abnormal returns as residuals of a regression

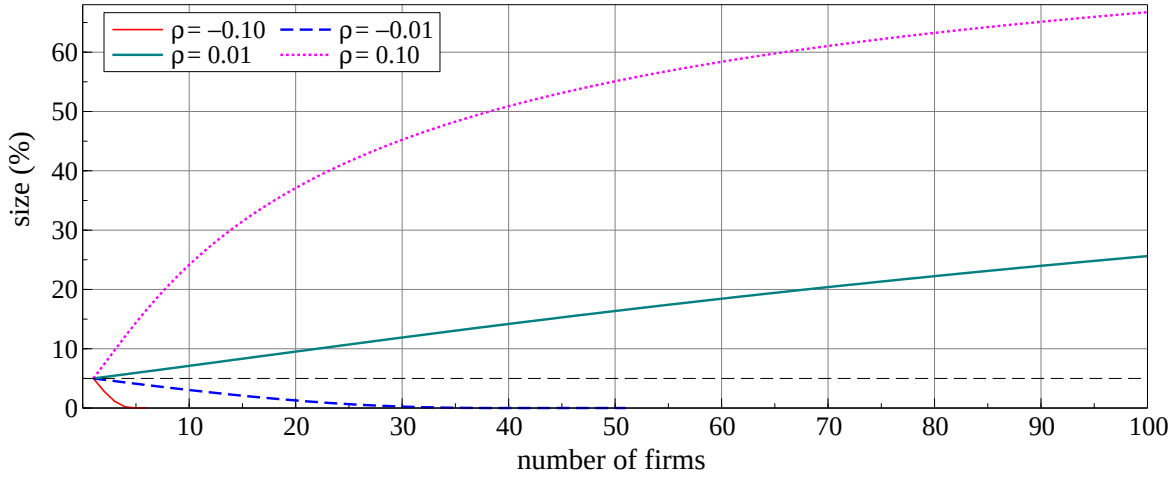


FIGURE 5.2: Actual size of the Patell test as a function of the number of firms for different values of the average correlation at a nominal 5% size.

of the stock returns on a relevant index reduces this average correlation significantly. However, if the returns considered in the study belong to a specific sector (e.g. banking, automotive industry, etc.), even after regressing them on a stock index there may be enough average correlation to seriously distort the size of the Patell test or any test built under the assumption of uncorrelatedness.

Another serious issue for many parametric event study tests is the assumption of normality of returns. By observing Table 5.1, which reports some descriptive statistics about the skewness and kurtosis of the stock returns belonging to the S&P100 and NASDAQ indexes, it is clear how stock returns tend to be negatively skewed and extremely leptokurtic. Indeed, by using the Doornik and Hansen (2008) version of the Jarque-Bera test, the null of normality is rejected 100% of the times for S&P100 stocks and 98% of the times for NASDAQ stocks.

TABLE 5.1: Descriptive statistics for the skewness and kurtosis of the stock returns belonging to the S&P100 and NASDAQ indexes.

		Percentiles				
	Mean	Min	25%	Median	75%	Max
Skew(SP100)	-0.09	-2.30	-0.31	-0.03	0.14	1.60
Skew(NASDAQ)	-0.11	-5.50	-0.23	-0.03	0.21	1.13
Kurt(SP100)	13.50	6.22	8.92	11.49	14.82	52.54
Kurt(NASDAQ)	13.83	3.12	8.65	10.84	14.59	123.63

It may be argued that, since Patell-like tests take averages of several random quantities, some central limit theorem (CLT) could intervene to make the statistic normal even when the returns are not normal. This is certainly a good point if the conditions for a CLT hold, but the extreme thickness of the tails of the return distributions may slow down the convergence to a normal. For example, if we average the daily returns of the 100 firms belonging to the S&P100, the resulting time series has still a kurtosis of 14.3, which is well above the kurtosis of the normal distribution (i.e., 3). Thus, the actual size of a test statistic based on average abnormal returns, possibly corrected for cross-correlation, will be larger than the nominal size (i.e., it will reject

too often under the null). Furthermore, without the assumption of independence of the returns across firms, it is not easy to find a workable condition under which the CLT holds.
From the above reasoning, it appears clear that a practitioner's negligence in taking into proper account a positive average correlation of the returns and/or the presence of fat-tails in their distribution leads to find event effects too often.

5.5 Review of the main parametric and nonparametric tests

The literature on event study test statistics is wide and covers both parametric and nonparametric approaches. Parametric tests assume joint Gaussianity of the firms' abnormal returns, whereas nonparametric tests make no distributive assumptions. Nonparametric tests are often used as a complement to parametric statistics to check that parametric results are not affected by data anomalies.
The most popular approach to nonparametric tests for event studies is based on rank and sign statistics. However, parametric tests can be turned into nonparametric by using bootstrap techniques (Hein and Westfall, 2004). In this review section and in the simulation experiment, we limit our attention to nonparametric tests based on ranks, which are generally simple to implement, robust and statistically powerful, when confronted with real financial returns.
Assume that the relationship between Y_{it} and X_{it} can be well approximated by Equation 5.9 and that the abnormal residuals can be computed as in Equation 5.11. This, the abnormal residuals and the cumulated abnormal residuals within the event window are computed as in Equations 5.12 and 5.13, respectively.

5.5.1 Parametric tests

Parametric test statistics are built to test the null hypothesis that all abnormal returns in the event window have zero mean. The basic t-test for both single-day and clustered events is based on the ratio between the single firm's abnormal return and its standard error, comparing it with the quantiles of a Student's-t distribution. The literature proposes numerous test statistics accounting for multiple stocks in a multi-period event window.
Some simple multivariate modification of the basic t-test define the statistics as the ratio between the cross-sectionally averaged cumulated abnormal returns and their cross-sectional standard deviation (Brown and Warner, 1985). The *cross-sectional t-test* (t_{CS}) estimates the variance using only the cumulated abnormal returns in the event window and is given by:

$$t_{CS} = \sqrt{m}\frac{\overline{CAR}}{S_{CAAR}}, \tag{5.32}$$

where

$$\overline{CAR} = \frac{\sum_{i=1}^{m} CAR_i}{m}$$

and

$$S^2_{CAAR} = \frac{\sum_{i=1}^{m}\left(CAR_i - \overline{CAR}\right)^2}{m}$$

are respectively, the cross-sectional average and variance of the cumulative abnormal returns.

Alternatively it is possible to define another statistic, which bases the estimate of the cross-sectional variance using the full sample, i.e. both in the temporal and cross-sectional dimensions. This statistic is denoted by t_{CD} and is computed as:

$$t_{CD} = \frac{\overline{CAR}}{\sqrt{n_1}\, S_{AAR}} \tag{5.33}$$

where

$$AAR_t = \frac{\sum_{i=1}^{m} AR_{it}}{m},$$
$$\overline{AAR} = \frac{\sum_{t\in\Omega} AAR_t}{n},$$

and

$$S^2_{AAR} = \frac{\sum_{t\in\Omega} \left(AAR_t - \overline{AAR}\right)^2}{n-2},$$

are respectively the cross-sectionally averaged abnormal returns at time t, their mean and standard deviation over the estimation window.
A more sophisticated version of the parametric t-test was introduced by Patell, 1976 with the aim to take into account the forecast-error generated by the market model regression. Patell's statistic considers a standardized version of abnormal return (SAR_{it}) generated by dividing each daily abnormal return by the forecast-error corrected standard deviation $(S^2_{AR_{it}})$ with market model adjustment, i.e.

$$S^2_{AR_{it}} = S^2_{AR_i}\left(1 + \frac{1}{n_0} + \frac{(x_t - \bar{x})^2}{\sum_{t\in\Omega_0}(x_t - \bar{x})^2}\right).$$

where x_t is the return of the market index at time t and $\bar{x}$ is the average market index return over Ω_0.
Denoting as $CSAR_i = \sum_{t\in\Omega_1} SAR_{it}$ the cumulated standardized abnormal returns for stock i, under the assumption of cross-sectionally independent returns, Patell's Z statistic for multi-period event window is

$$z_{PAT} = \frac{\sqrt{m}}{S_{CSAR}} \sum_{i=1}^{m} \frac{CSAR_i}{m},$$

where $S^2_{CSAR} = n_1 \frac{n_0-2}{n_0-4}$ is the cross-sectional variance – equal for all the series – obtained by the market model regression. Under the null, the z_{PAT} statistic follows a standard normal distribution. Note that assuming independence and homogeneity of one-period return variances, $S^2_{CAR_i} = n_1 S^2_{AR_{it}}$, the z_{PAT} statistic can be rewritten as (Kolari and Pynnönen, 2010)

$$z_{PAT} = \sqrt{\frac{m(n_0-4)}{n_0-2}} \sum_{i=1}^{m} \frac{SCAR_i}{m}.$$

Another parametric test based on standardized abnormal returns has been proposed by Boehmer et al., 1991, hereafter BMP, with the aim to extend Patell's statistic controlling for event-induced volatility. While Patell's test assumes implicitly that all the returns have constant variance, the BMP statistic adjusts for event-induced volatility estimating the average event-day-volatility with the cross-sectional standard deviation (Kolari and Pynnönen, 2010). Denoting as $SCAR_i = \frac{CAR_i}{S_{CAR_i}}$ the standardized cumulated abnormal returns in the event window, under the assumption of cross-sectionally independent returns, the BMP Z statistic to test for the mean cumulated

event effect is defined as follows:

$$z_{BMP} = \frac{\sqrt{m}}{S_{SCAR}} \sum_{i=1}^{m} \frac{SCAR_i}{m} \tag{5.34}$$

where $S^2_{SCAR} = \frac{1}{m-1} \sum_{i=1}^{m} (SCAR_i - \bar{SCAR})^2$ and $SCAR_i = \frac{CAR_i}{S_{CAR_i}}$. The forecast-error corrected variance $(S^2_{CAR_i})$ is computed applying Mikkelson and Partch's 1988 market model correction, which adjusts each firm's statistic for its serial correlation in the returns, i.e.

$$S^2_{CAR_i} = S^2_{AR_i} \left(n_1 + \frac{n_1^2}{n_0} + \frac{\sum_{t \in \Omega_1} (x_t - \bar{x}_m)^2}{\sum_{t \in \Omega_0} (x_t - \bar{x}_m)^2} \right).$$

As in the case of cross-sectional t-test, neither Patell's statistic nor BMP statistic account explicitly for cross-correlation among returns. By including the component given by the covariance in the calculation of the theoretical variance of the estimators, Kolari and Pynnönen (2010) proposed a simple correction of BMP and Patell statistics capable of tening into account the cross-correlation among multiple series of returns. As shown in Section 5.4 the presence of even a minimal correlation between the securities can strongly affect the size of the tests. The cross-correlation adjusted Patell's statistic $(z_{PAT_{adj}})$ and the cross-correlation adjusted BMP statistic $(z_{BMP_{adj}})$ are defined as follows:

$$z_{PAT_{adj}} = z_{PAT} \sqrt{\frac{1 - \bar{r}}{1 + (m-1)\bar{r}}}, \tag{5.35}$$

$$z_{BMP_{adj}} = z_{BMP} \sqrt{\frac{1 - \bar{r}}{1 + (m-1)\bar{r}}}, \tag{5.36}$$

where $\bar{r}$ is the average of the sample cross-correlations of the estimation-period abnormal returns $AR_{it} \in \Omega_0$. Often the literature refers to $\sqrt{\frac{1-\bar{r}}{1+(m-1)\bar{r}}}$ as correlation correction factor.

5.5.2 Nonparametric tests

Event study tests following the nonparametric approach consider two different classes of statistics: the first based on ranks and the second based on signs.
Given a set of observations $Y = \{y_1, y_2, \ldots, y_n\}$, we define the *rank* of $y_j \in Y$ as it position index once the observations in Y have been ordered or, formally, as

$$rank(y_j) = \sum_{k=1}^{n} \mathbb{1}_{(-\infty, y_j]}(y_k),$$

with $\mathbb{1}_A(\cdot)$ indicator function for set A.
Let $R_{it} = \text{rank}(SAR_{it})$ be the ranks of the standardized abnormal returns over the complete observation set Ω and let $K_{it} = R_{it}/(n+1) \in (0,1)$ be the normalized ranks sequence (for the formal definitions see Equation (5.42)). The cumulated ranks in the event window for stock i are defined as $U_{i\Omega_1} = \sum_{t \in \Omega_1} K_{it}$, while the cross-sectionally averaged normalized cumulated rank is defined as $\bar{U}_{\Omega_1} = \frac{1}{m} \sum_{i=1}^{m} U_{i\Omega_1}$.
Under the assumption of no event effect and cross-sectionally dependent returns, the quantity $\bar{U}_{\Omega_1}$ has expected value

$$E(\bar{U}_{\Omega_1}) = \frac{n_1}{2}$$

and variance

$$Var(\bar{U}_{\Omega_1}) = \frac{n_1 n_0}{12(n+1)m}\left(1+(m-1)\bar{\rho}\right)$$

where $\bar{\rho}$ is the average correlation among the ranks of different stocks. These facts allow to consider a standardized statistic which uses the theoretical expected value and the theoretical variance. We will refer to this statistic as $CumRank_Z$ and it is computed as follows:

$$CumRank_Z = \frac{\bar{U}_{\Omega_1} - \frac{n_1}{2}}{\sqrt{\frac{n_1 n_0}{12(n+1)m}(1+(m-1)\bar{\rho})}}.$$

Instead of using the theoretical properties of the standardized cumulated ranks, Corrado (Corrado, 1989; Corrado and Zivney, 1992) suggested to estimate the variance of $\bar{U}_{\Omega_1}$ using the cross-section in sample Ω_1, by substituting the theoretical variance with its cross-sectional estimate (cf. Equation (5.37)). The original Corrado's rank statistic considers the single-day event case and it can be easily extended to the multi-period case as suggested by Campbell and Wasley, 1992 by computing the cumulative ranks in the event window. The Campbell-Wasley statistic is denoted in the text as *CumRank* and it is computed as follows:

$$CumRank = \frac{\bar{U}_{\Omega_1} - \frac{n_1}{2}}{\sqrt{n_1 \frac{n-n_1}{n-1} s^2_{\bar{K}}}}.$$

where

$$s^2_{\bar{K}} = \frac{\sum_{t\in\Omega}\left(\bar{K}_t - \frac{1}{2}\right)^2}{n} \tag{5.37}$$

is the sample cross-sectional variance of the cumulated ranks.
When dealing with small samples, in terms of the number of temporal observations, the *CumRank* statistic can be adjusted by including the following correction:

$$CumRank_T = CumRank\sqrt{\frac{n-2}{n-1-CumRank^2}}.$$

As stated by Luoma, 2011 and Hagnäs and Pynnonen, 2014, the main advantage of the *CumRank* and $CumRank_T$ over the $CumRank_Z$ is that the former statistics use an estimate of the variance that implicitly takes into account the cross-correlation among stocks, leading to a more robust statistic against cross-sectional correlation, thus, with better size properties. At the opposite, their disadvantage concerns a loss in term of power.
Moreover, considering short event-windows, hence with small n_1, it is possible to gain a better approximation of the test statistic to Gaussianity by transforming the normalized ranks using Tuckey's lambda transform (Corrado, 2011). Thus, we adjust the idea of Corrado (2011) to the case of cross-sectionally dependent abnormal returns over a multi-day window as follows

$$CT = \frac{4.91}{\sqrt{n_1 m(1+(m-1)\bar{\rho}_V)}}\sum_{i=1}^{m}\sum_{t\in\Omega_1} V_{it}, \tag{5.38}$$

where

$$V_{it} = K_{it}^{0.14} - (1-K_{it})^{0.14}$$

and $\bar{\rho}_V$ is the average correlation of V_{it} across stocks, which can be estimated by its sample counterpart over Ω_0. Under the null, the distribution of CT should be very close to standard Gaussian for every value of n_1 (Joiner and Rosenblatt, 1971). We will refer to this statistic with

the name of *correlation adjusted Corrado-Tuckey*.
A complementary approach to the cumulative rank statistics is represented by the class of *generalized rank statistics* (Kolari and Pynnönen, 2011). For each stock, this second block of statistics collapses the observed abnormal returns in the event window into a single cumulated observation, i.e. the standardized abnormal returns in the event window are substituted by their sum. The rank of this composite observation is computed with respect to the whole sample, which is now a sequence of $n_0 + 1$ values. Recalling that $SCAR_i = CAR_i / S_{CAR_i}$, the generalized standardized abnormal returns for stock i at time $t \in \{1, 2, \ldots, n_0, n_0 + 1\}$ are computed as

$$GSAR_{it} = \begin{cases} SAR_{it} & \text{if } t \in \Omega_0, \\ SCAR_i = \frac{CAR_i}{S_{CAR_i}} & \text{for } t = n_0 + 1. \end{cases} \tag{5.39}$$

Applying the rank operator to the $GSAR_{it}$ sequence, we obtain the demeaned normalised ranks of the generalised standardised abnormal returns as

$$G_{it} = \frac{\text{rank}(GSAR_{it})}{n_0 + 2} - \frac{1}{2}.$$

Consequently, the cross-sectionally averaged statistic is simply obtained applying the mean over the stocks: $\bar{G}_t = \sum_{i=1}^{m} G_{it}/m$. Under the assumption of no event effect, the cross-sectionally averaged generalized rank in the event window, denoted by $\bar{G}_{\Omega_1}$, has expected value $E(\bar{G}_{\Omega_1}) = 0$ and variance

$$Var(\bar{G}_{\Omega_1}) = \frac{n_0}{12(n_0 + 2)m}(1 + (m - 1)\bar{\rho}_g),$$

where $\bar{\rho}_g$ is the average cross-correlation of the ranks applied to the $GSAR_{it}$. As for the cumulated ranks, it is possible to define a standardized statistic based on the theoretical variance and expected value denoted by $GRank_Z$:

$$GRank_Z = \frac{\bar{G}_{\Omega_1}}{\sqrt{\frac{n_0}{12(n_0+2)m}(1 + (m - 1)\bar{\rho}_g)}}.$$

Substituting the theoretical variance with an unbiased estimate it can be defined an alternative test statistic named $GRank$:

$$GRank = \frac{\bar{G}_{\Omega_1}}{\sqrt{S^2_{\bar{G}}}},$$

where $S^2_{\bar{G}} = \sum_{t \in \Omega} \bar{G}_t^2 / (n_0 + 1)$ is the sample cross-sectional variance of the averaged generalized ranks. Moreover, in order to adjust for small-sample temporal sizes, the $GRank$ statistic can be modified as follows:

$$GRank_T = GRank \sqrt{\frac{n_0 - 1}{n_0 - GRank^2}}.$$

Generalised ranks enjoy similar properties to the cumulated rank statistics. In fact, even small cross-sectional correlations among the stocks can generate significant misspecification of the estimators and lead to a loss in terms of tests sizes (Kolari and Pynnönen, 2010). According to this, the $GRank_T$ and $GRank$ statistics are expected to be more robust to cross-correlation than the $GRank_Z$ statistic (Kolari and Pynnönen, 2011).

5.5.3 Asymptotic distribution of the statistics

Considering large temporal windows (large n), for all the previous statistics, both based on parametric and nonparametric approaches, it is possible to (more or less) prove their asymptotical normality. In particular, analytical details about the cumulated ranks statistics' asymptotic behaviours and the generalised rank statistics are provided respectively by Hagnäs and Pynnonen, 2014, and Kolari and Pynnönen, 2011.

5.6 New nonparametric tests

5.6.1 The P1 test

We propose a novel test statistic for the null hypothesis that the linear model in equation 5.9 holds for $t \in \Omega_1$ with the same distributional assumptions on the error term ε_{it}. Under the alternative hypothesis, the regression errors undergo a positive or negative shift in the location parameter, such as the median or the mean, if it exists.
Now, recalling the above notation, let $\hat{\theta}$ be a consistent estimator of the coefficients vector θ computed only on the observations in the estimation window Ω_0. The regression residuals, which in event studies context are called abnormal residuals, are given by

$$AR_{it} = \hat{\varepsilon}_{it} = y_{it} - x_{it}^{\top}\hat{\theta}, \qquad \text{for } t \in \Omega, \tag{5.40}$$

and, since the scale parameter (standard deviation if it exists) of each $\hat{\varepsilon}_{it}$ may vary over i, let us define the standardized abnormal returns as in

$$SAR_{it} = \frac{\hat{\varepsilon}_{it}}{\hat{\sigma}_i}, \tag{5.41}$$

where $\hat{\sigma}_i$ is a consistent estimator for the scale parameter σ_i of the regression error ε_{it}.
Take the cross-sectional sums of the standardized residuals, $SAR_t = \sum_{i=1}^{m} SAR_{it}$, and their ranks for the complete sample Ω,

$$R_{nt} = \sum_{k \in \Omega} \mathbb{1}_{(-\infty, SAR_t]}(SAR_k), \tag{5.42}$$

with $\mathbb{1}_A(\cdot)$ indicator function for the set A.
The first test statistic we propose is

$$P = \frac{1}{\sqrt{n_1}} \sum_{t \in \Omega_1} \Phi^{-1}\left(\frac{R_{nt}}{n+1}\right), \tag{5.43}$$

where $\Phi^{-1}(\cdot)$ is the standard normal quantile function. Under the null hypothesis supplemented with mild regularity conditions (see below) P is approximately distributed as a standard normal, whatever the distribution of ε_{it} and x_{it}.

Theorem 1. *Assume that (5.9) holds with i.i.d. vectors of error terms* $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})^{\top}$ *and*

1. *the distribution function of* $\varepsilon_t = \sum_{i=1}^{m} \varepsilon_{it}/\sigma_i$, $F(\cdot)$, *is absolutely continuous with absolutely continuous density* $f(\cdot)$ *and support* Θ;
2. $I_0(f) = \int_{\Theta} [f'(x)/f(x)]^2 f(x)\,\mathrm{d}x < \infty$ *(finite information for location);*
3. $\hat{\sigma}_i = \sigma_i + O_p(n^{-1/2})$, *for* $i = 1, \ldots, m$;

4. $\max_{1\le t\le n_0} x_{jit}^2 / \left(\sum_{t\in\Omega_0} x_{jit}^2\right) \to 0$ *as* n_0 *diverges, for* $j = 1,\dots,k$, i=1,...,m;

5. $\sqrt{\sum_{t\in\Omega_0} x_{jit}^2}\,(\hat\beta_{ji} - \beta_{ji}) = O_p(1)$ *as* n_0 *diverges, for* $j = 1,\dots,k$, i=1,...,m;

where x_{jit}, β_{ji} *and* $\hat\beta_{ji}$ *are the j-th elements of the respective vectors* x_{it}, β_i *and* $\hat\beta_i$.
Then, for the statistic defined in equation (5.43) it holds

$$P \xrightarrow{d} N(0,1), \qquad as\ n_0 \to \infty.$$

See Appendix A for a detailed proof of the Theorem's results.
The conditions in the theorem are rather mild. Condition 1 requires that the marginal distribution of the sum of the regression residuals has a continuous density, which is constant over time; the joint stationarity of the returns and their densities' continuity are sufficient for this condition. Condition 2 requires the existence of the Fisher Information for location (as in maximum likelihood estimation theory). Condition 3 asks for a (root-n or faster) consistent estimator for the regression error scale parameter. Notice that, actually consistency is not needed, but just convergence in probability to some constant. For Condition 4, no single observation of the regressors should (asymptotically) dominate the sum (of squares). Condition 5 allows for any estimator of the regression coefficients with the same rate of convergence as the OLS estimator (e.g., SURE, LAD, rank, Theil-Sen and other robust estimators can be used instead of OLS).
To improve the approximation of the test statistic's finite-sample distribution, we can apply two small corrections to P. Notice that

$$T = \sum_{t\in\Omega_1} \Phi^{-1}\left(\frac{R_{nt}}{n+1}\right) = \sum_{t=1}^{n} c_t a_{R_{ni}}$$

is a linear rank statistic with

$$c_t = \begin{cases} 0 & \text{for } t \in \Omega_0 \\ 1 & \text{for } t \in \Omega_1 \end{cases}, \qquad a_{R_{nt}} = \Phi^{-1}\left(\frac{R_{nt}}{n+1}\right),$$

and, under the null hypothesis, T has mean and variance (cf. Vaart, 1998, Lemma 13.1)

$$ET = n\bar c\bar a = 0, \qquad VarT = \frac{1}{n-1}\sum_{t=1}^{n}(c_t-\bar c)^2 \sum_{t=1}^{n}(a_t-\bar a)^2 = \frac{n_0 n_1}{n} V_n$$

with

$$V_n = \frac{1}{n-1}\sum_{t=1}^{n}\left[\Phi^{-1}\left(\frac{t}{n+1}\right)\right]^2.$$

Of course, as $n \to \infty$ the variance V_n converges to

$$\int_0^1 \Phi^{-1}(u)^2\,\mathrm{d}u = \int_{-\infty}^{\infty} x^2\phi(x)\,\mathrm{d}x = 1,$$

where $\phi(.)$ denotes the standard normal density function, but, for finite n, V_n tends to be smaller than 1.
If instead of the normalized ranks $R_{nt}/(n+1)$, we use the asymptotically equivalent quantities $(R_{nt}-0.5)/n$, then the variance V_n is very close to 1 also for moderate samples (see Figure 5.3).
Considering the specific case, it is not possible to apply standard linear rank statistic asymptotics to P. Indeed, there are two reasons for this: i) we apply the ranks to estimated regression

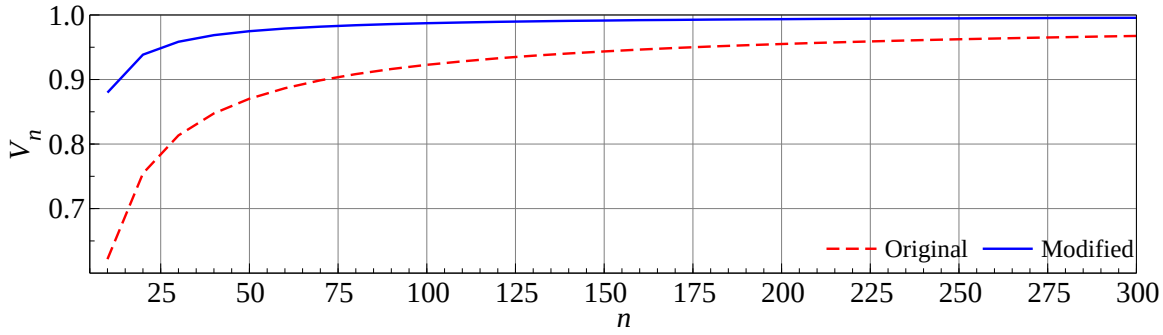


FIGURE 5.3: Variance V_n as a function of the sample size for the original and modified normalised ranks.

residuals and not to raw observations, ii) our asymptotic results should work for $n_0 \to \infty$ and n_1 fixed, while for linear rank statistics both sample sizes have to diverge at the same rate.
This reasoning suggests the modification of the test statistic (5.43) with the asymptotically equivalent (as $n_0 \to \infty$)

$$P_1 = \sqrt{\frac{n}{n_0 n_1}} \sum_{t \in \Omega_1} \Phi^{-1}\left(\frac{R_{nt} - 0.5}{n}\right), \tag{5.44}$$

obtained by dividing T by its (approximate) standard deviation $\sqrt{n_0 n_1 / n}$. Simulations (not reported) confirm that the test based on P_1 has a more accurate size in small samples than P.
A nice feature of this test is that it can also be used to find event-induced volatility if in equation (5.41) standardized residuals (abnormal returns) are substituted by their absolute values:

$$SAR_{it} = \frac{|AR_{it}|}{\hat{\sigma}_i}.$$

Indeed, Theorem 1 holds whatever the distribution of $SAR_t = \sum_i SAR_{it}$ (i.e., the statistics P and P_1 are distribution-free). If we work with the absolute abnormal returns, a high value of the statistic indicates either a positive or negative mean-shift or an increment in the returns' volatility. If our test applied to the abnormal returns cannot reject the null, but the right-tailed version of our test applied to the absolute abnormal returns rejects its null, then there is evidence of an increase in the volatility of the returns in the event window.
Figure 5.4 depicts the distributions of ε_{it} and $|\varepsilon_{it}|$ for two different values of the variance: it makes clear how a shift in the volatility of ε_{it} turns into a shift in the mean of $|\varepsilon_{it}|$.

5.6.2 The P2 test

In our first proposal, the m return time series are combined into just one sequence by linear combination, on which the ranks are then computed. This trick overcomes the problem of cross-dependent returns, but we expect this approach to make the test statistic less sensitive to deviation from the null hypothesis, especially when only a few stocks witness a shift in the mean of the abnormal returns.
For our second test statistic, instead, we compute the normalised ranks of the abnormal returns of each stock and, then, after turning the obtained random quantities normal using the Gaussian quantile function, we sum them over the cross-section. Such a test statistic is expected to have better power properties in situations in which only some stocks are affected by event

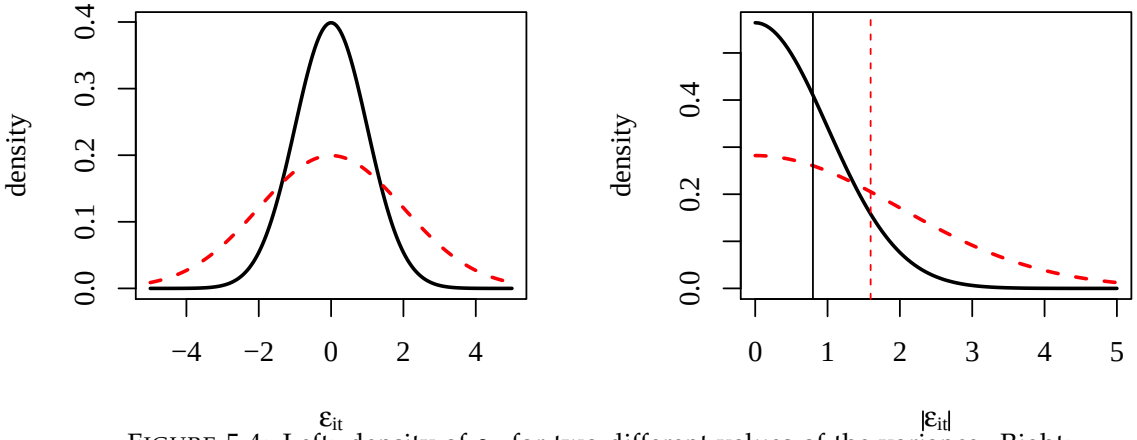


FIGURE 5.4: Left: density of ε_{it} for two different values of the variance. Right: resulting densities of $|\varepsilon_{it}|$ and relative means (vertical lines).

induced shifts, but the statistic's Gaussianity can be proven only if a proper central limit theorem (CLT) for dependent random variables applies. Indeed, central limit theorems (CLTs) for dependent variables exist, but they were developed for stochastic processes, thus, with ordered indexing of the random variables with respect to the sum. Some form of mixing (i.e., asymptotic independence) and the existence of higher-order moments are the usual assumptions under which these CLTs hold. We will derive our test statistic's asymptotic distribution assuming such a CLT applies and comment on its choice in the remarks below. We empirically verify if the Gaussian approximation is reliable by bootstrapping stock returns belonging to the S&P100 and NASDAQ indexes in Section 5.7.
The second test statistic we propose is

$$P_2 = \frac{1}{\sqrt{n_1}} \sum_{t \in \Omega_1} \sum_{i=1}^{m} \frac{\Phi^{-1}(U_{it})}{\hat{\rho}^{1/2}}, \tag{5.45}$$

where

$$\hat{\rho} = \sum_{i=1}^{m} \sum_{j=1}^{m} \hat{\rho}_{ij},$$

with $\hat{\rho}_{ij}$ sample correlation between $\Phi^{-1}(U_{it})$ and $\Phi^{-1}(U_{jt})$ computed using the sub-sample Ω_0.
For $m = 1$ the test statistic P_2 is identical to P of equation (5.43) and converges in distribution to the standard normal as n_0 diverges, regardless of the size of n_1. For $m > 1$, under general cross-sectional dependence, we need some CLT to justify the normal approximation: indeed, we are summing dependent (approximately) normal variates that are not necessarily jointly normal.

Theorem 2. *Assume that (5.9) holds with i.i.d. vectors of error terms* $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})^\top$ *and*

1. *the distribution function of* ε_{it}, $F^{(i)}(\cdot)$, $i = 1, \ldots, m$, *is absolutely continuous with absolutely continuous density* $f^{(i)}(\cdot)$ *and support* Θ_i;
2. $I_0(f^{(i)}) = \int_{\Theta_i} \left[f^{(i)\prime}(x) / f^{(i)}(x) \right]^2 f^{(i)}(x)\, \mathrm{d}x < \infty$ *(finite information for location),* $i = 1, \ldots, m$;
3. $\max_{1 \le t \le n_0} x_{jit}^2 / \left(\sum_{t \in \Omega_0} x_{jit}^2 \right) \to 0$ *as* n_0 *diverges, for* $j = 1, \ldots, k,\ i = 1, \ldots, m$;

4. $\sqrt{\sum_{t\in\Omega_0} x_{jit}^2}\,(\hat{\beta}_{ji} - \beta_{ji}) = O_p(1)$ *as* n_0 *diverges, for* $j = 1,\dots,k,\ i = 1,\dots,m$;

5. *the Gaussian random variables* $Z_{it} := \Phi^{-1}(F^{(i)}(\varepsilon_{it}))$ *obey a central limit theorem with respect to the index i:*

$$T_{mt} := m^{-1/2} \sum_{i=1}^{m} Z_{it} \xrightarrow{p} N(0,\rho), \quad \text{as } m \to \infty,$$

where

$$\rho = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m}\sum_{j=1}^{m} \rho_{ij} < \infty,$$

with $\rho_{ij} = E(Z_{it}Z_{jt})$.

As above, x_{jit}, β_{ji} *and* $\hat{\beta}_{ji}$ *are the j-th elements of the respective vectors* x_{it}, β_i *and* $\hat{\beta}_i$.
Then, for the statistic defined in equation (5.45) it holds

$$P_2 \xrightarrow{d} N(0,1), \quad \text{as } n_0 \to \infty,\ m \to \infty,\ \frac{m}{n_0} \to 0.$$

See Appendix B for a detailed proof of the Theorem's results.
In assumption 5., the covariance sum limit can be rewritten as $\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}\rho_{ij} = 1 + (m-1)\bar{\rho}_m$, with $\bar{\rho}_m$ mean correlation among the Z_{it}, and it clearly implies that $\bar{\rho}_m = O(m^{-1})$, that is, mean correlation must decrease at speed m or faster. The correlations in Table 5.2 seem to support this assumption.
We decided to assume the existence of a CLT instead of selecting one of the many CLT for dependent observations that are available in the literature since very often their conditions are very technical and hard to verify on cross-sectional data (cf. Dvoretsky, 1972; McLeish, 1974; Kuersteiner and Prucha, 2013, for example). Indeed, these theorems are generally proven for stochastic processes and, thus, for random variables with a natural ordering. In our framework, one of the most reasonable conditions for a CLT is m-dependence (Hoeffding and Robbins, 1948; Diananda, 1955; Orey, 1958; Berk, 1973). In our context, we can interpret m-dependence as the dependence within groups of stocks belonging to the same industrial sector, after removing the common factors implied by the (possibly multi-factor) market model as in Equation (5.40). Moreover, the fact that Z_{it} is Gaussian and, thus, possesses finite moments of any order, should help to weaken the assumptions on dependence and increase the speed of convergence to the normality of the test statistic.
Notice that our two tests can be easily modified to include the adjustment proposed by Boehmer et al. (1991) (see also Kolari and Pynnönen, 2010) that makes the tests robust to event-induced volatility and that of Kolari and Pynnönen (2011), which increases the power of the tests when only a small portion of the returns in the event window has a shifted distribution. However, we limit our simulation experiment to our original formulations: in fact, rank-based tests are well known to be not too sensitive to event induced volatility.

5.7 Simulation experiment

The finite sample behaviour of our two tests and all the tests cited in the literature review of Section 5.5 is investigated by bootstrapping the daily returns of stocks selected from the S&P100 and NASDAQ indexes. The daily returns were recorded over the period of February 2003–February 2013 (S&P100) and April 2003–April 2013 (NASDAQ).
The estimation window is fixed at $n_0 = 250$ observations and three event windows of $n_1 = 1, 5, 10$ days are considered. We apply the tests to different numbers of stocks: $m = \{5, 10, 20, 40, 80\}$.

The stocks are chosen according to a maximum correlation criterion, that is, for each selected value of m, we select the subsamples of stocks whose returns have the largest average correlation. Therefore, the mean correlation in each set of stocks decreases with m, as illustrated in Table 5.2. For each experiment, returns are bootstrapped 10,000 times.

TABLE 5.2: Mean correlation of abnormal returns for the different samples of stocks in S&P100 and NASDAQ.

Number of stocks (m)	5	10	20	40	80
Mean correlation SP100	0.272	0.184	0.138	0.044	0.002
Mean correlation NASDAQ	0.405	0.238	0.116	0.055	0.025

The tests are applied to the (OLS) regression residuals of the stock returns on their respective stock indexes (i.e., the abnormal returns concerning the market model).
Table 5.3 reports the worst (largest) empirical size of each test with respect to the 15 ($= 3 \times 5$) considered cases. The empirical sizes are computed for nominal sizes of both 5% and 1%.
Figures 5.5 and 5.6 depict the rejection rates for all cases of the simulation experiment at nominal 5% size. The first column reports the test sizes estimated using a single-day event window, i.e. $n_1 = 1$, the second reports the sizes for $n_1 = 5$ days, whereas the third shows the estimated sizes for the case $n_1 = 10$ days.
Looking at the corresponding bar plots, the statistics CorradoTuckey, Crude_dep_T, CumRank family, P1 and P2 provide empirical sizes conforming to the nominal size, regardless of the cross-sectional size of the sample.
Our P1 and P2 tests have sizes that are very close to the nominal ones. In particular, P1 has excellent empirical sizes for both nominal sizes (best size in the 1% case). The CumRank family of statistics and the CorradoTuckey test have excellent size properties as well. Among the parametric tests, only the Crude_dep_T has an acceptable size for the 5% case.
We decided to carry out the power analysis only for those tests whose worst empirical size is below 0.065 in both markets for the nominal 5% case. Of course, 0.065 is an arbitrary number, and it may sound funny because 6.5% is not as round as 6%. However, we fixed 0.065 for three reasons: (i) the same set of tests is selected for S&P100 and NASDAQ returns, (ii) the size estimates based on the worst-case are positively biased, (iii) at least one parametric test is in the selected battery.
Figures 5.7–5.8 depict the size-adjusted power curves of the seven test statistics with acceptable empirical sizes (CorradoTuckey, Crud_dep_T, CumRank, CumRank_T, CumRank_Z_adj, P1, P2) under the considered set-ups. For better readability, we omitted the case $m = 80$, which is really similar to the case $m = 40$. The frequency of rejection is computed using empirical critical values that leave 2.5% of probability mass on both tails of the distribution under the null. The drift in the mean in the event window is expressed in the number of standard deviations, as all the stock returns receive a drift that is proportional to their respective standard deviations.
We can clearly notice two clusters of tests from the power curves: P1 and Crude_dep_T on one side, with lower power, and the other five tests, with better power properties. Based on our simulations results, we can state that all CumRank statistics, the CorradoTuckey and our P2 test can be used interchangeably as they have good size and high power. Within the "top" cluster, the difference among powers is always less than a percentage point. However, the CumRank tests show virtually identical power, which on average tends to be slightly higher than that of P2 and CorradoTuckey. The latter two tests have almost identical power in all cases, with P2 being slightly better.
The only parametric test that passed the size-based selection, that is, Crud_dep_T, is the one that performs worst in terms of power. As we expected, also P1 has lower power than the

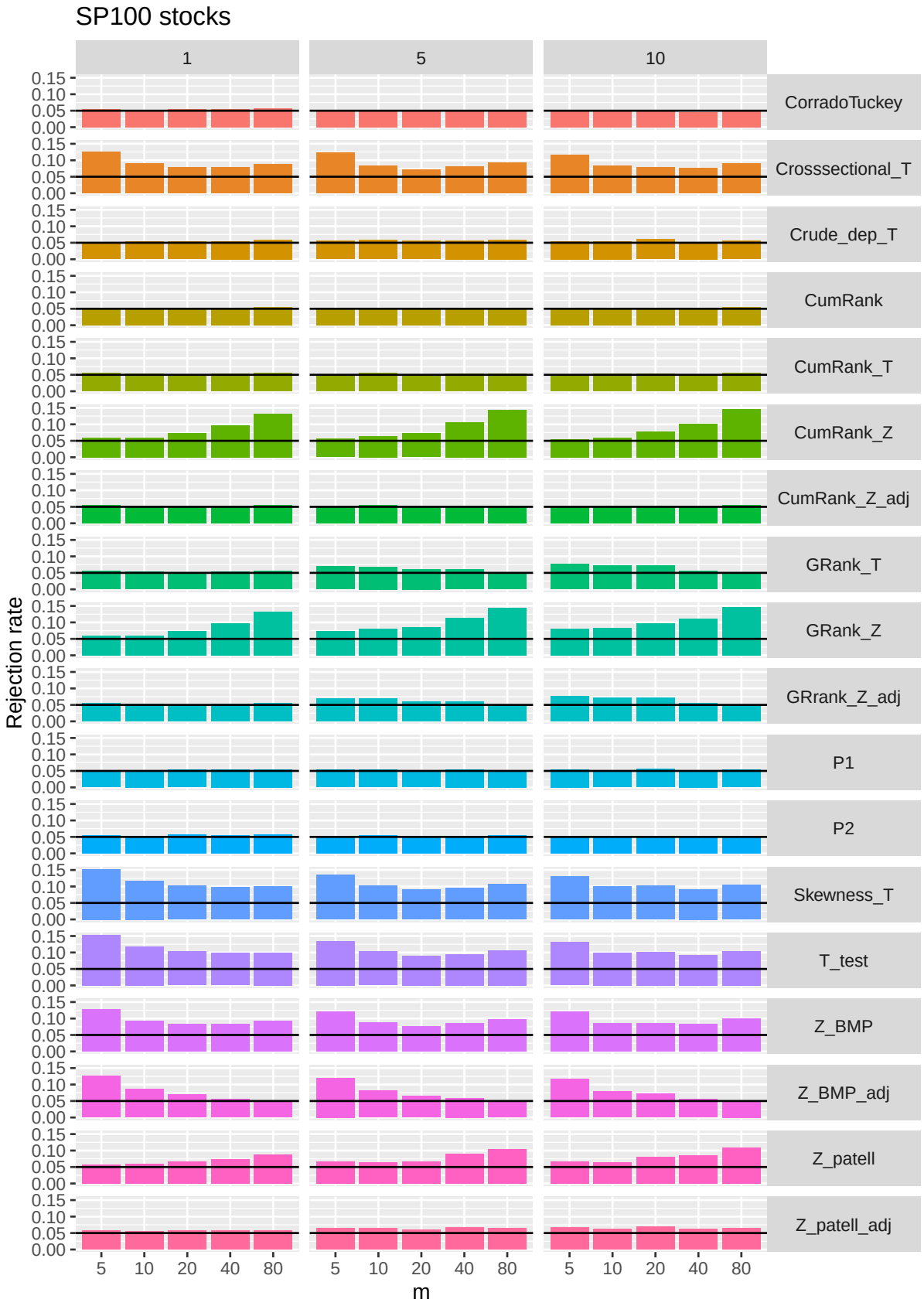


FIGURE 5.5: Power curves of the tests computed on abnormal returns taken from the m most correlated SP100 constituents for an event windows of n_1 days. The columns report the test sizes for $n_1 = 1$, $n_1 = 5$ and $n_1 = 10$, respectively.

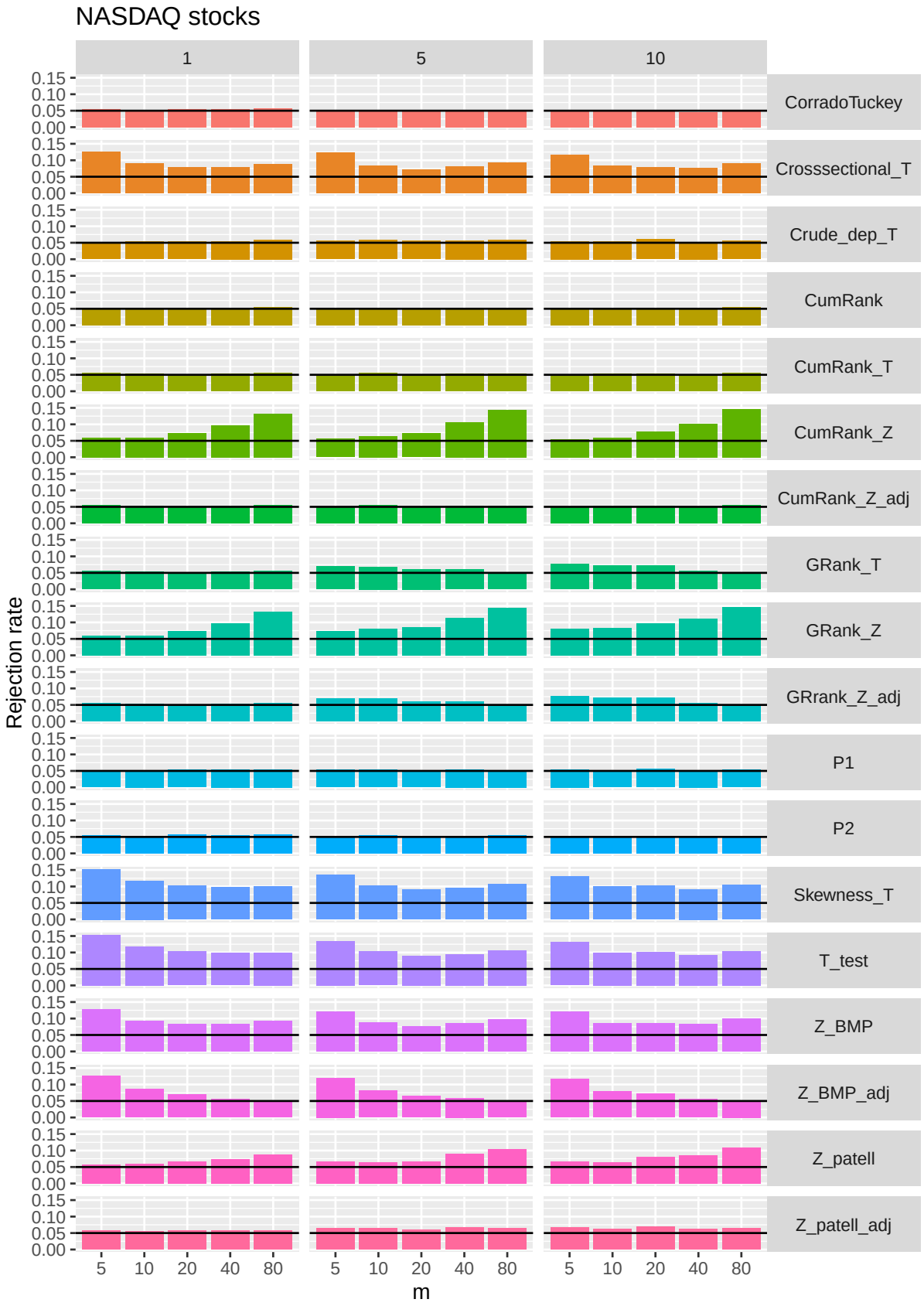


FIGURE 5.6: Size of the test statistics computed on abnormal returns taken from the m most correlated NASDAQ constituents for an event windows of n_1 days. The columns report the test sizes for $n_1 = 1$, $n_1 = 5$ and $n_1 = 10$, respectively.

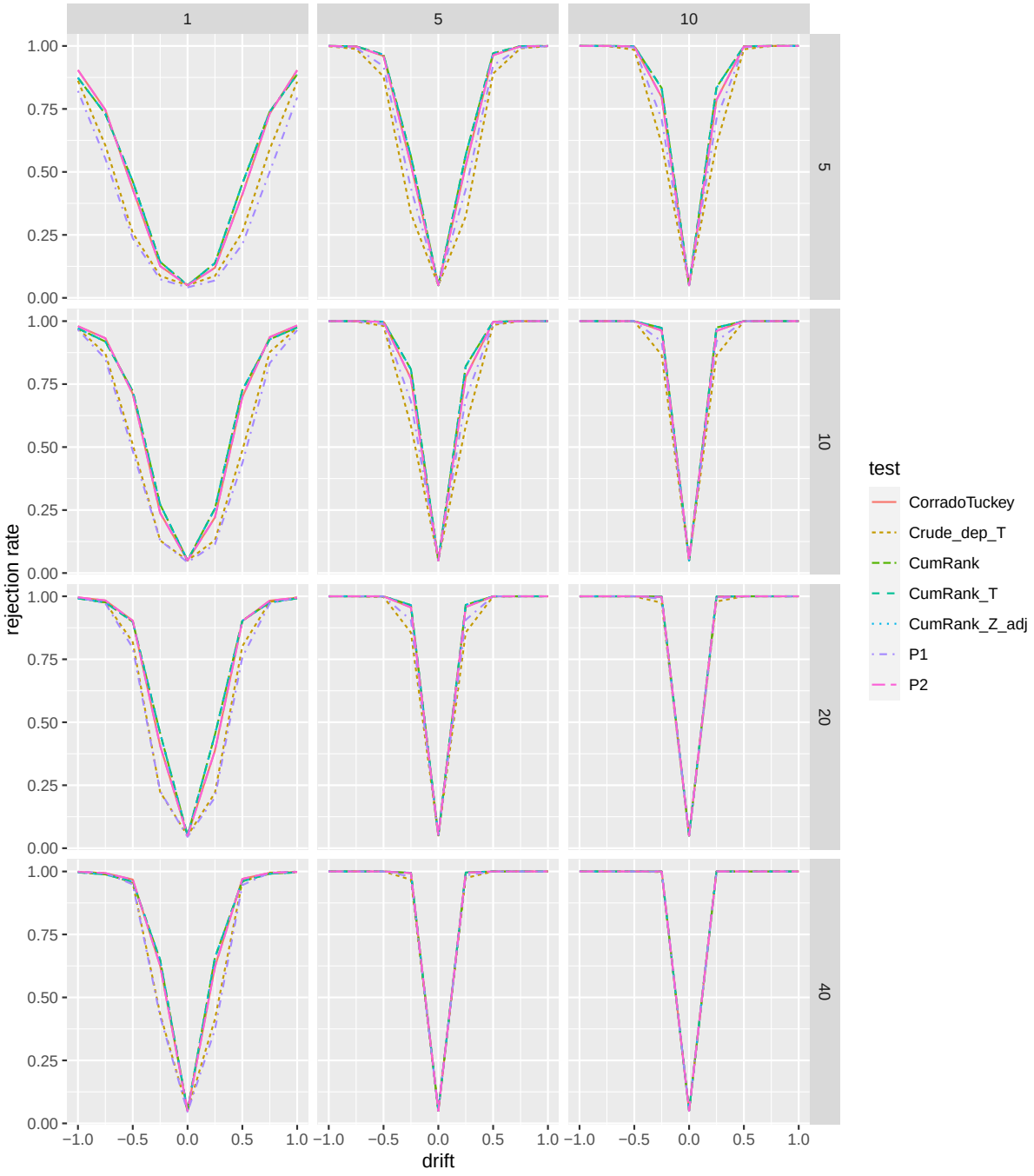


FIGURE 5.7: Power curves of the tests computed on abnormal returns taken from the m most correlated SP100 constituents for an event windows of n_1 days.

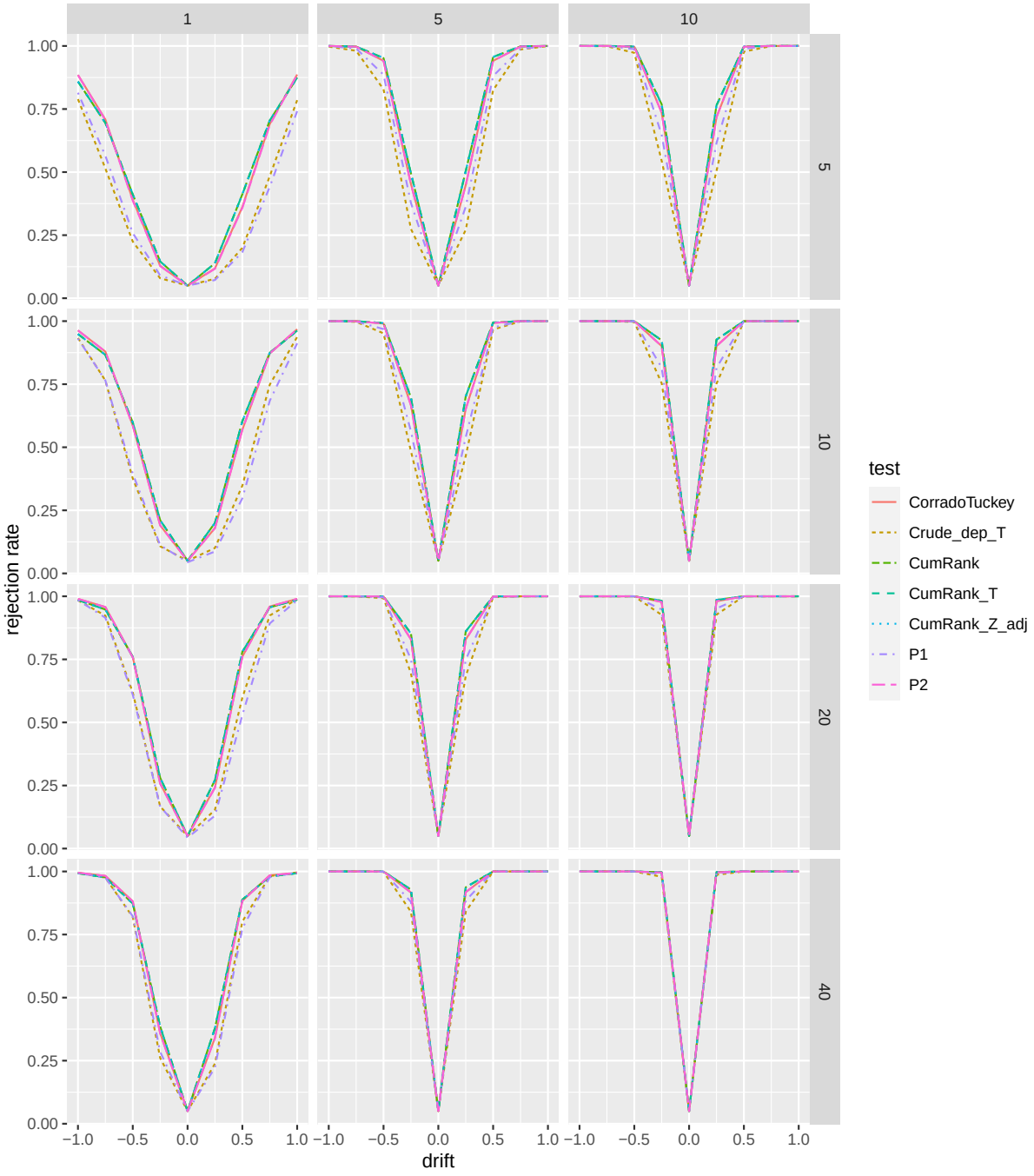


FIGURE 5.8: Power curves of the tests computed on abnormal returns taken from the m most correlated NASDAQ constituents for an event windows of n_1 days.

TABLE 5.3: Worst (largest) empirical size of each test for a nominal size of 5% and 1%. Rows are ordered with respect to the SP100/5% size.

	Nominal 5% size		Nominal 1% size	
Test	SP100	NASDAQ	SP100	NASDAQ
CumRank	0.054	0.057	0.022	0.018
CumRank_T	0.055	0.058	0.023	0.018
CumRank_Z_adj	0.056	0.058	0.023	0.018
P1	0.056	0.057	0.012	0.012
CorradoTuckey	0.056	0.057	0.025	0.020
P2	0.057	0.058	0.025	0.020
Crude_dep_T	0.061	0.058	0.030	0.024
Z_patell_adj	0.070	0.066	0.027	0.026
GRank_Z_adj	0.076	0.069	0.023	0.018
GRank_T	0.076	0.069	0.023	0.018
Z_patell	0.110	0.280	0.045	0.160
Z_BMP_adj	0.126	0.125	0.064	0.062
Crosssectional_T	0.126	0.259	0.063	0.140
Z_BMP	0.129	0.281	0.064	0.157
CumRank_Z	0.146	0.353	0.062	0.221
GRank_Z	0.147	0.347	0.061	0.211
Skewness_T	0.153	0.268	0.082	0.152
T_test	0.153	0.268	0.082	0.152

other nonparametric tests considered in the power comparison. However, because of the solid approximation theory behind P1, one can always be sure that its actual size is reliable at any nominal size, and practitioners should include it in their battery of tests as a benchmark.

5.8 Application: the impact of lockdown restriction on air quality in Lombardy in 2020

5.8.1 Sample and event definition

To test the effectiveness of the two novel statistics presented above, we conducted an empirical analysis applied to air quality data. In addition to evaluating the performance of the statistics, which have already been rigorously addressed in the simulative study, the application aims to test the behaviour of the event studies statistics in contexts outside the financial and economic spheres, which are well known in the literature.
We considered as the event of interest the lockdown due to the COVID-19 disease in Lombardy during 2020 and its effect on the pollutant concentrations. The null hypothesis that we are testing is that the restrictions did not generate any effect on the cumulative oxide concentrations during the lockdown period, as against the alternative that the restrictions produced significant variations in the concentrations by changing the average level.
Recalling that the nationwide restrictions to mobility and economic activities were active from the 8th March to 18th May 2020, we conducted the analysis using daily observations from 1st January 2014 to 20th May 2020. We considered as estimation-window the period between the 1st January 2014 and the 1st March 2020. The event-window is instead composed of daily observations from the 2nd March 2020 and the 20th May 2020. The estimation-window has

cardinality $n_0 = 2252$ days, while the event-window has cardinality $n_1 = 80$ days. Although the event window is very long, it can be assumed that there were no other overlapping events capable of hiding the impact of the restrictive measures during the period. Moreover, by controlling through regression for possible exogenous effects, several confounding factors typical of environmental series, such as changes in meteorology or calendar events, are isolated.
We considered the observed concentrations of total nitrogen oxide (NO_X) and nitrogen dioxide (NO_2) collected in $m = 83$ ground stations belonging to the ARPA Lombardia network. Figure 5.9 illustrates the map of station locations across the regional territory.

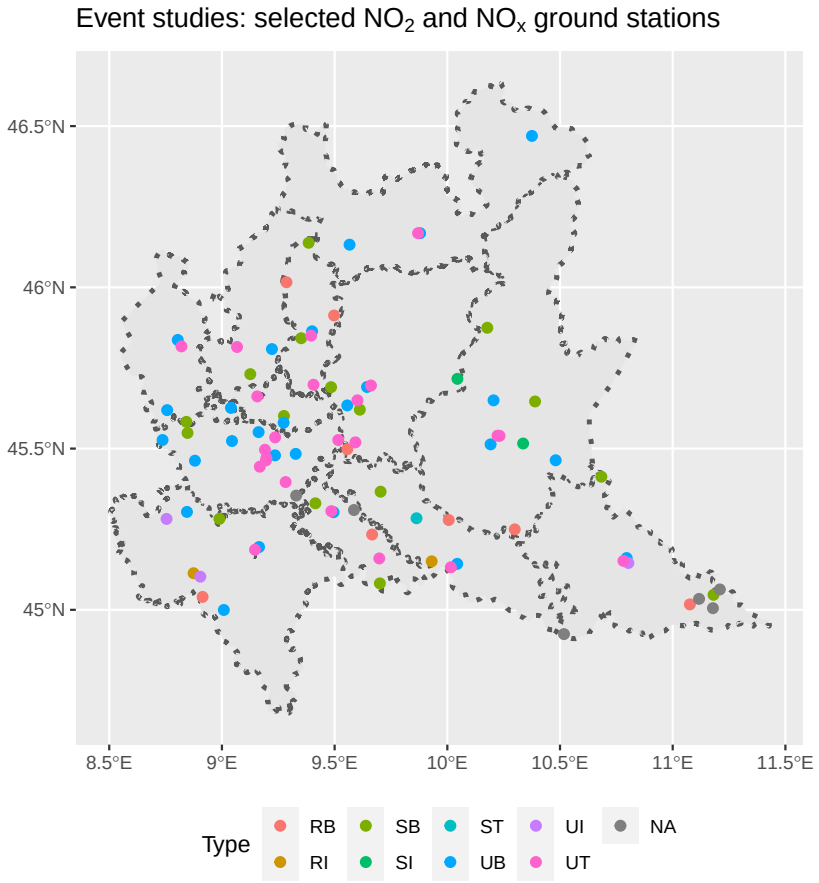


FIGURE 5.9: Selected air quality ground stations by type of monitored area.

The stations were selected using the following criteria: stations having both NO_X and NO_2 sensors, installed before 2014 and active up to the end of November 2020.
We applied to the airborne pollutant series all the statistics for cumulated abnormal residuals (CAR) presented above, both those corrected for cross-sectional dependence and those not corrected. Expectations for the results are twofold. Given the enormous scale of the phenomenon analysed, statistically significant negative variations are expected during the lockdown period. This implies that the sign of the statistics is negative. About the dependence, it is expected that the cross-sectionally adjusted statistics will be more moderate in absolute terms than the unadjusted statistics. As we will see later, the empirical correlation of pollutants is considerable so that the unadjusted statistics would show overestimated negative shocks compared to the real changes occurred.

5.8.2 Statistical characteristics of the sample

Tables 5.4 and 5.5 report some descriptive statistics for the linear correlation index (ρ), skewness, kurtosis and variability coefficient (cv) of NO_2 and NO_x, respectively, and calculated using the sample. The average cross-sectional correlation for NO_2 is 0.579, whereas for NO_x

TABLE 5.4: Descriptive statistics for NO_2 concentrations observed at the 83 stations included in the sample and regarding the period January 2014 - November 2020.

	Min	25% perc	mean	Median	75% perc	Max
ρ	-0.327	0.464	0.579	0.632	0.737	0.953
skewness	0.073	0.665	0.849	0.858	1.005	2.093
kurtosis	1.325	2.964	3.525	3.314	3.893	7.367
cv	0.286	0.456	0.558	0.520	0.573	4.121

TABLE 5.5: Descriptive statistics for NO_x concentrations observed at the 83 stations included in the sample and regarding the period January 2014 - November 2020.

	Min	25% perc	mean	Median	75% perc	Max
ρ	-0.171	0.601	0.687	0.721	0.806	0.969
skewness	0.714	1.648	1.825	1.826	2.142	2.840
kurtosis	2.619	5.595	6.663	6.651	7.847	9.695
cv	0.507	0.756	0.851	0.838	0.902	2.120

is 0.687. The simulations were conducted with a financial sample showing an average cross-sectional dependence of about 14%. The air quality sample showed an average positive correlation of more than 60%. This last value is an indicator of a very high linear correlation which must be taken into account in the proposed methodologies. The simulative scheme of the previous section and, in a more crude way, Figure 5.2 about the size of the Patell test in the presence of weak correlations between the series have amply demonstrated that ignoring the cross-sectional dependence in the construction of the test statistics for event studies would automatically lead such statistics towards values much higher than they actually are, with the direct consequence of always rejecting the null hypothesis of no effect of the event. Therefore, the introduction of correct statistics for cross-sectional dependence in event studies applied to air quality context is essential.
Regarding both the skewness and the kurtosis, the results are not surprising. Both NO_2 and NO_X always report positive values for the asymmetry index and kurtosis values above 3, a value that in the literature is used to highlight leptokurtic distributions. In the previous sections, we have often mentioned that the concentration data have strong positive asymmetries and extreme values. Indeed, the histograms of the distributions show these characteristics through long tails on the right side of the distribution.
The variability, here represented by the variability coefficient, that is $cv = \frac{\sigma}{\mu}$, is compact in most cases, showing poor dispersion of observations.

5.8.3 Regression analysis

Recalling the definitions described in Section 5.3, abnormal concentrations (AC) have been computed through a regression analysis framework. Control variables needed for regression

analysis were chosen among those registered by ground control units belonging to the ARPA weather network, that is temperature, cumulative rainfall, relative humidity and wind speed. Several air quality stations do not have meteorological measurements, and it is necessary to associate weather measurements as consistent and realistic as possible. For each air quality station, we associated the meteorological covariates measured by the nearest weather station among those available, where the distance in kilometres is computed using the Euclidean distance function. We also considered weekend effects and calendar events effects as a relevant factor to explain the evolution of observed concentrations. As for the case study on Area B discussed in Chapter 3 we included as covariates the dummy variables for weekends, that is one dummy fro Saturday and one dummy for Sunday, for the main Italian national holidays and the interaction among the weekend and holidays, i.e. *SaturdayHoliday* and *SundayHoliday*.
To account for the strong serial dependence and persistence usual of airborne concentrations, we modelled the air quality through a linear regression model with ARIMA disturbances, namely regARIMA models. This wide class of models represents the regression disturbances as ARMA-like filters composed of an autoregressive and a moving average component. This makes it possible to generalise classic linear regression, which takes on white noise residues, to more flexible and realistic versions. Assuming an ARMA specification for the residuals of linear regression models allows dealing with the persistent serial auto-correlation not handled by the covariates and typical of statistical models for air quality.
Considering one of NO_2 and NO_X, for each location $i = 1, ..., m = 74$ let Y_{it} be the observed airborne pollutant concentrations at time t and let X_{it} be the design matrix including weather and calendar covariates. The set of covariates is fixed for all the ground stations $i = 1, ..., 83$. The linear regression model with ARMA(p,q) structures of the residuals can be expressed by the following set of equations:

$$Y_{it} = \theta_i X_{it} + \eta_{it}, \tag{5.46}$$

$$\eta_{it} = \phi_1 \eta_{it-1} + ... + \phi_p \eta_{it-p} + \varepsilon_{it} + \psi_1 \varepsilon_{it-1} + ... + \psi_q \varepsilon_{it-q} \qquad \varepsilon_{it} \sim WN(0, \sigma_\varepsilon^2), \tag{5.47}$$

where ε is a I.I.D. sequence of white-noise innovations with null expected value and variance σ_ε^2, θ_i' is the parameter vector of the covariates, ϕ_1,...,ϕ_p are the autoregressive coefficients for the disturbances and ψ,...,ψ_q are the moving average coefficients for the disturbances. Empirical abnormal concentrations are computed as the linear regression model's innovations, that is, the regression residuals filtered using the ARMA specification. Hence, the estimated AC_{it} at time t at location i corresponds to the innovation ε_{it} at time t for location i.
Assuming a Gaussian distribution for the disturbances ε_{it}, the model parameters can be easily estimated using the corresponding Maximum Likelihood Estimators (MLEs).
Having fixed the set of covariates to be included in the model, for each location i, the ARMA specification is automatically detected using the Corrected Akaike Information Criterion (AICc). Hence, the ARMA filter's optimal order, i.e. p^* and q^*, is the one associated to the model with less AICc over all the possible models within fixed order constraints. Dealing with daily data to simplify the search for the optimal model by algorithm, we have restricted the ARMA specifications to those between ARMA(0,0) and ARMA(7,7). In this way, we allow the concentrations not explained by the regressors to be affected by their past values up to seven days and random shocks occurring within one week of the date. Moreover, the AICc-based selection algorithm allows modelling each measuring site using a customised model.

5.8.4 Results

Table 5.6 reports the results for all the statistics considered on both NO_2 and NO_X average concentrations. The null hypothesis is that the cumulated abnormal concentrations of oxide pollutants in Lombardy during the event-window ($CAC_{i\Omega_1}$) have null mean value, i.e. $CAC_{i\Omega_1} = 0$, whereas the the alternative hypothesis is that the cumulated abnormal concentrations in the event-window ($CAC_{i\Omega_1}$) have negative mean value, i.e. $CAC_{i\Omega_1} < 0$. Hence, the hypothesis test is unilateral on the left tail. The reported significance levels are computed using the quantiles of a standard Normal distribution at level α.

TABLE 5.6: Test statistics for H_0: $CAR_{\Omega_1} = 0$ VS H_1: $CAR_{\Omega_1} < 0$.

Cross-sectional adjusted	Statistic	NO_2		NO_x	
		Value	Signif.	Value	Signif.
Adjusted	P1	-2.238	**	-2.596	***
Adjusted	P2	-2.320	**	-2.654	***
Not-adjusted	cross_T_Test	-11.982	***	-12.511	***
Not-adjusted	crude_dep_T_Test	-2.240	**	-2.382	***
Not-adjusted	T_skew	-14.534	***	-20.713	***
Not-adjusted	Z_patell	-13.599	***	-16.962	***
Adjusted	Z_patell_adj	-2.565	***	-3.313	***
Not-adjusted	Z_BMP	-5.708	***	-8.145	***
Adjusted	Z_BMP_adj	-0.881	.	-1.217	.
Adjusted	T_grank	-2.133	**	-2.306	**
Not-adjusted	Z_grank	-11.124	***	-12.376	***
Adjusted	Z_grank_adj	-2.132	**	-2.306	**
Adjusted	CumRank	-2.240	**	-2.749	***
Adjusted	CumRank_mod	-2.279	**	-2.796	***
Adjusted	CumRank_T	-2.281	**	-2.800	***
Not-adjusted	CumRank_Z	-11.939	***	-15.015	***
Adjusted	CumRank_Z_adj	-2.288	**	-2.797	***
Adjusted	CorradoTuckey	-2.315	**	-2.653	***

'***', '**', '*' stand for 'significant at levels 0.01, 0.05, 0.10'. The symbol '.' stands for 'p-value greater than 0.10', hence not significant.

The results for both NO_2 and NO_x confirm all the anticipated expectations and convey several great messages.
Regarding the signs, all the estimated statistics report a negative value, meaning the presence of negative variations (reductions) in average concentration levels during the lockdown period. The absolute value of the statistics for NO_x appears a bit stronger than those calculated for NO_2. The latter consideration is in line with other research investigating the effect of lockdowns in various settings and where the greatest reductions were associated with NO_x compared to NO2. See in this regard Rossi et al. (2020) for the case of Padova (Italy), Nakada and Urban, 2020 about Sao Paulo (Brazil) and Higham et al. (2020) or Ropkins and Tate (2020) for the case of the United Kingdom. Besides, since nitrogen dioxide is a direct sub-component of total nitrogen oxide, it is credible that the reductions of NO_x levels are stronger than those on NO_2 because the other components may also have been reduced due to the restrictions.
However, another important message is that the choice of statistics adjusted for cross-sectional dependence seems to have had a very positive and evident effect. As already stated above, the results confirm that it is a necessary choice in situations like the one presented. In absolute values, all the unadjusted statistics are greater than the cross-sectionally adjusted statistics. This confirms the over-rejection issue generated by the cross-sectional dependence in event studies. Even though the value for corrected statistics appears to be more moderate, all the statistics except one reject the null hypothesis. The adjusted BMP test by Kolari and Pynnönen (2010) is

the only test in contrast to all the others: it indicates that the variation is negative, but is not statistically significant at every level of significance (p-value greater than 10%). An important issue about this indicator and valid for the Patell Patell (1976) test is that its distributions under the null are valid only under the hypothesis of the dependent variable's joint normality provided that a cross-sectional central limit theorem holds. All these considerations may lead us to conclude with reasonableness and credibility that the restrictions imposed on mobility and production activities have had a significant effect on oxide concentrations in Lombardy leading to a reduction in the average level in the medium run.

5.8.5 Model fitting and diagnostics

5.8.6 ARMA structure

Figure 5.10 summarises the pairwise combinations of p and q selected using the AICc criterion for NO_2 and NO_x on the 83 estimated regARIMA models.

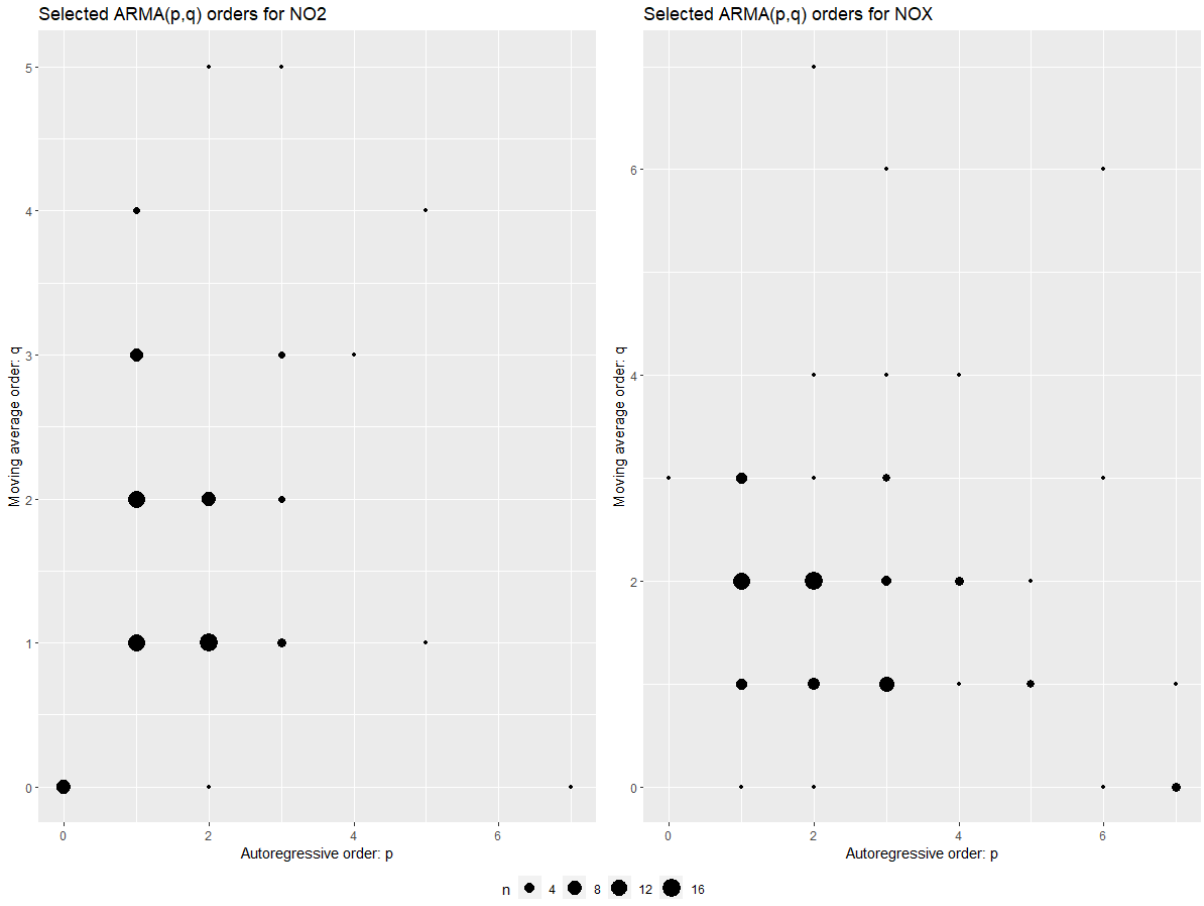


FIGURE 5.10: Selected ARMA orders for NO_2 (left panel) and NO_x (right panel) using the AICc criterion. Point size depends on the count of the pair p and q.

Using daily data, both in the case of NO_2 and NO_x, the criterion suggests in most cases simple models, included within the order (3,3). This is consistent with most applications of these models to environmental and economic issues. Indeed, most economical time series are commonly treated as simple ARMA models with degree lower or equal to 2 or 3. Just some cases present higher autoregressive terms, whereas almost none of the models used high moving-average orders.

Interestingly, the ARMA(1,0) model is rarely selected for both NO2 and NOx. This indicates that the regression component can only partially model the stochastic process underlying the concentrations and that the moving average component is persistent in the residuals. In fact, the number of MA components in the residuals is almost always greater than zero. This fact might contradict the choice made to estimate the changes in oxides due to lockdown by COVID-19 (Chapter 4), where the concentrations were represented in ARX(1) form by fixing the moving average components to zero. The data, however, support this choice. As shown in Section 4.6.2, the serial correlation of the estimated model innovations show that the residuals are white-noise processes. This contradiction could be explained by several factors: 1) regARIMA models assume the residuals to be stochastic ARMA processes, while ARMAX models represent the time series as the linear combination of deterministic components and ARMA processes and assume the residuals to be white-noise; 2) in Chapter 4 we use weekly data, whose persistence is naturally reduced by the seven-day aggregation, while in the current chapter the data used are daily. By definition, reducing frequency also reduces the degree of dependence of the observations on past shocks. As already extensively discussed in chapter 4, the choice of fixing a priori the number of lags of the ARX model is dictated by the will to privilege the interpretative simplicity of the results. In any case, given the results on the regARIMA models, we set the future goal of choosing through analytical and data-driven tools the optimal model order, just as we did in this chapter.

5.8.7 Abnormal concentrations

Tables 5.7 and 5.8 report some descriptive statistics for diagnostics statistics computed on the regression residuals, which can be interpreted as the abnormal concentrations at time t for location i. We considered the following set of diagnostics: linear correlation index (ρ), shape indices, that is skewness, kurtosis and variability (cv), Ljung-Box test for serial autocorrelation at various lags, i.e. $lbq1$ is the LB test at one-day lag, $lbq3$ is the LB test at three-day lag, $lbq7$ is the LB test at seven-day lag, $lbq14$ is the LB test at fourteen-day lag and $lbq21$ is the LB test at 21-day lag, the Bera-Jarque test for Normality (bj) and the full sample mean (μ).

TABLE 5.7: Descriptive statistics for abnormal concentrations (AC) of NO_2 estimated using the full sample and regarding the period January 2014 - November 2020.

	Min	25% perc	mean	Median	75% perc	Max
ρ	-0.029	0.257	0.342	0.338	0.420	0.835
skewness	-0.612	-0.046	0.080	0.059	0.191	1.769
kurtosis	3.490	4.449	5.267	5.031	5.731	8.352
cv	55.185	135.680	625.164	213.894	460.225	11656.911
bj	0.000	0.000	0.000	0.000	0.000	0.000
$lbq1$	0.000	0.555	0.680	0.555	0.555	0.680
$lbq3$	0.000	0.536	0.691	0.536	0.536	0.691
$lbq7$	0.000	0.000	0.135	0.000	0.000	0.135
$lbq14$	0.000	0.000	0.038	0.000	0.000	0.038
$lbq21$	0.000	0.000	0.030	0.000	0.000	0.030
μ	-0.132	-0.047	-0.034	-0.030	-0.014	0.022

The results give similar suggestions for both pollutants. First of all, the correlation between the series has dropped considerably, up to halve: indeed, the average and median values are about 0.33, compared with 0.60 in the pre-regression phase. However, some series appears to be again

TABLE 5.8: Descriptive statistics for abnormal concentrations (AC) of NO_x estimated using the full sample and regarding the period January 2014 - November 2020.

	Min	25% perc	mean	Median	75% perc	Max
ρ	-0.113	0.248	0.353	0.338	0.452	0.892
skew	-9.179	0.366	0.543	0.633	0.890	2.158
kurt	6.091	7.334	8.111	8.170	8.848	9.926
cv	49.088	130.551	909.795	202.617	301.276	29896.320
bj	0.000	0.000	0.000	0.000	0.000	0.000
*lbq*1	0.258	0.714	0.808	0.714	0.714	0.808
*lbq*3	0.068	0.721	0.806	0.721	0.721	0.806
*lbq*7	0.000	0.002	0.183	0.002	0.002	0.183
*lbq*14	0.000	0.000	0.041	0.000	0.000	0.041
*lbq*21	0.000	0.000	0.031	0.000	0.000	0.031
μ	-0.582	-0.171	-0.134	-0.093	-0.061	0.005

strongly correlated. The abnormal concentrations are symmetrical, but they do fit to a Gaussian distribution: the skewness is in fact poorly variable and centred on zero, but the Bera-Jarque test always rejects the hypothesis of normality. The correlation in the short term, i.e. within two weeks of the date, is statistically not significant, indicating that the regression models have managed to capture temporal dependence in the short term. Long-term autocorrelation, i.e. beyond two weeks, is still present. The average abnormal concentrations for NO_2 are zero centred and have little variability, while those for NO_x tend to negative values. Therefore, in the case of nitrogen dioxide, it could be said that models have returned estimates of unbiased abnormal concentrations, while they appear biased in the case of total oxides.
Not surprisingly, the regression diagnostics on daily data may seem somewhat weak. Typically, high-frequency air quality concentrations, i.e. daily and even more so hourly, are extremely variable and complex to model. Purely temporal models are not always able to model these characteristics in a good manner. In any case, given the simplicity of the models used and the good characteristics of symmetry, non-distortion and poor serial autocorrelation in the short term, the results can be considered satisfactory.

5.9 Conclusion and final remarks on event studies

This chapter discussed a class of statistical methodologies known as *event studies*. Event studies are positioned in a branch of econometrics that makes it possible to assess whether a particular event of interest has generated structural changes, e.g. changes in level or variability, in time series of interest. Event studies for a variable of interest Y are based on the concept of abnormal residuals, i.e. the residuals of the regression of the variable of interest on an informative set of covariates X. The regression aims to identify cyclical or seasonal behaviours of the response variable and purify it from the presence of possible effects of exogenous or confusing factors able to explain its statistical behaviour. The expected values of abnormal residuals are then statistically tested employing appropriate statistical tests. In particular, we discussed event studies methods applied to multivariate time series characterised by high cross-sectional dependence. It is well known in the literature that the presence of even a small cross-correlation can considerably bias the classical test statistics, leading to an over-rejection of the tests and thus to highlight anomalous events that do not actually exist.

We proposed two novel nonparametric test statistics for the case of cumulated abnormal residuals (CAR), namely $P1$ and $P2$, based on the ranks of abnormal residuals and accounting explicitly for cross-sectional dependence. We demonstrated the main statistical properties of the proposed statistics, that is the closed-formulation for their expected values and variances and their asymptotic distributions. In both cases, thanks to the statistical properties of rank functions and having defined a proper central limit theorem for cross-sectional units, the asymptotic distributions are Gaussian with finite expected value and variance.
The two test statistics are compared in terms of size and power of the corresponding test with other indicators introduced in the literature, some of which take into account cross-sectional dependence, while others do not. The comparison is performed through a simulation study and an empirical application to real-world data. The Monte Carlo simulation study aims at evaluating the size and the power of the hypothesis test under the assumption of cross-sectionally dependent units. The observations used for the Monte Carlo simulations come from a sample of daily stock prices quoted on the SP&100 and the NASDAQ stock exchanges and referring to the period from February 2003 to April 2013. The real-time series have been bootstrapped 10 thousand times in order to obtain always different but comparable samples. The simulative results show that all the cross-sectionally adjusted statistics provide sizes very close to the nominal value, whereas not-adjusted statistics do not. In particular, the statistic $P1$ has excellent empirical sizes for both 5% and 1% nominal sizes. The power functions for the statistics clearly highlight that the statistics $P2$ benefits of both good size and power, while the $P1$ statistic suffers from slightly lower power, but acceptable for large samples. For large cross-sectional samples, all the considered statistics are equivalent in terms of size and power.
Finally, all the statistics are applied to the empirical case of airborne pollutant concentrations registered in Lombardy during 2020. We considered as the event of interest the restrictions imposed on mobility and activities during the COVID-9 pandemic. The main interest is to state if the lockdown generated significant reductions in average concentrations of oxide (NO_2 and NO_x) or not. In the specific case of air quality data collected through ground monitoring network, such as the ARPA Lombardy monitoring system, it is plausible to assume strong correlations among the monitoring sites. Indeed, the original sample shows an average linear correlation around 0.60 for both NO_2 and NO_x. From our knowledge of the literature, no other applications of event studies methodologies to environmental and air quality issues are known. The results highlighted two main messages. The first one is that, dealing with highly cross-sectionally correlated series, avoid corrections will generate biases in statistic estimates and over-rejection of the statistical hypotheses. This reflects immediately in the erroneous information about the presence of significant anomalous events that actually did not occur. Compared to the unadjusted statistics, the cross-sectionally adjusted statistics are more moderate, but always statistically significant. The second message concerns the effect of lockdown. Since all the statistics show negative empirical value and are all strongly statistically significant, we can conclude that the restrictions due to COVID-19 generated consistent reductions in average oxide concentrations in Lombardy.
Some issues about event studies methodologies arise. Although they have been applied in a minimal fashion in fields other than financial econometrics, these methodologies reveal themselves to be important and useful analytical tools for impact analysis of policies and events in natural and/or economic-social contexts. In particular, they can be used as preliminary tools to assess whether events are significant or not. At the same time, a clear disadvantage is that statistics give a detailed indication only of the sign of the shock, but say nothing about its magnitude. As a last remark, it is to be noticed how ignoring the cross-sectional dependence typical of various realities, such as financial markets and air quality, in analytical contexts can have imposing and harmful distortive effects.

Chapter 6

Trend and long-run relations in noisy time series: why pre-filtering is inevitable

Based on:

Gianfreda, A., Maranzano, P., Pelagatti, M., & Visconti Parisio, L.,
'Trend and long-run relations in noisy time series: why pre-filtering is inevitable'
Book of Short Papers SIS 2020, pages 1405-1410, Pearson, ISBN 9788891910776

6.1 Introduction and motivation

Many applied papers study the dynamics of multiple series intending to find some long-run relationship governing the observed variables' behaviour. In this stream of research, cointegration among time series and stationarity, called also mean reversion, are handy tools to design successful analysis strategies in contexts characterized by significant uncertainty.

6.1.1 Noisy time series in economic contexts: the case of electricity markets

Real-world contexts in which stationarity and long-term relationships drive many analytical strategies are electricity and financial markets. Both markets are prominent examples of trading structures that produce series characterized by the kind of features we want to deal with. These two contexts are typically characterized by large fluctuations in both the prices of securities and the quantities traded, as well as by a high degree of uncertainty that affects the behaviour of agents. Therefore, it is essential to develop analysis strategies that are supported by data and remain as stable as possible by relying on the long-run relationships between the series of interest. From an empirical point of view, a technique that goes in this direction is the so-called pair-trading strategy, in which the cointegration analysis is used to identify pairs of stocks characterizes by high correlation and equilibrium price ratio. From the economic perspective, the idea is to highlight to the agents, i.e. traders, signals of change in trends or spreads and suggest greater or lesser exposure to one of the two securities.
Electricity prices are determined in the long-run mainly by demand level, fuel prices and the mix of generation technologies. However, short term prices are determined as equilibria of 24 hourly auction mechanisms where generators and consumers submit price or quantity offers. The short-run equilibrium is influenced by many factors like line congestions, firms' strategies

and, possibly, the exercise of market power at particular hours or seasons, plant maintenance (both programmed and unexpected), and start-up costs of power units. Moreover, in recent years it is possible to observe import/export effects on prices, due to a larger degree of integration of electricity markets, and a higher degree of price variability, due to the introduction of a large share of intermittent generation from Renewable Energy Sources (RES). For these reasons, hourly and daily observed electricity prices are determined by their market fundamentals and buried into high-variance leptokurtic noise. Several applied papers have dealt with this problem filtering price time series using different approaches. Bosco et al., 2010 and Gianfreda et al., 2016 used weekly means or medians, Bosco et al., 2010 and Pelagatti and Sen, 2013 developed robust integration and cointegration tests. Other works extracted the long-run component using UCM and Kalman filtering/smoothing (Gianfreda et al., 2016; Gianfreda et al., 2018).

6.1.2 Noisy time series in environmental contexts: the case of air quality

As in the case of prices in the electricity market, there are other empirical cases in which the processes present significant short-term deviations from the equilibrium, causing high variability and uncertainty in the data. In the case of environmental studies, two examples of such series are represented by air quality and weather conditions time series. In both cases, the factors governing the long-run and short-run processes are complex phenomena, some natural and others anthropogenic and artificial. Airborne pollutant concentrations result from complex physicochemical phenomena in the atmosphere in which interactions between natural and anthropogenic elements, such as local meteorology and shocks in emission sources, can perturb air quality considerably, generating high volatility and noise in the observed data.
Many factors can influence air pollutant concentrations. These include the quantity of air pollutants released by sources, the distance from the sources, and meteorological conditions such as air temperature, the stability of the air, wind speed and direction. The main drivers of infra-daily and annual seasonality of air quality are the temperature, wind, and rainfall air pressure. Some air pollutants can be carried by the wind and affect the air quality in locations hundreds to thousands of kilometres away from the sources. The higher the wind speed, the more airborne pollutants are dispersed, and the lower their concentration. In this context, the presence of abnormal or unexpected meteorological events due to climate change generate additional confounding noise that must be taken into account. Low atmospheric pressure brings wet and windy weather conditions. A passing storm front can wash pollutants out of the atmosphere or transport them to a new area, clearing the skies. In this case, the pollutants are not disappeared, rather they have been moved to a different location. At the opposite, under a high-pressure system, the air is often stagnant: airborne pollutants such as vehicle and factory emissions concentrate over the source area when the air stops moving. Air temperature affects the movement of air, and thus the movement of air pollution. In fact, since the sun's solar energy is absorbed by the terrestrial surface, the air close the ground is warmer than air that is further up in the troposphere. Due to the convection mechanism, the warmer air at the surface rises, and the cooler and heavier air in the upper troposphere goes down. This moves pollutants from the ground to higher altitudes. When the weather is cold, pollutants generated from vehicles or industries are more pronounced and visible, indicating high pollution levels. However, while industrial emissions remain constant throughout the year, particulate matter and carbon monoxide pollutants increase during the cold winter months. On the other side, other elements, such as ground-level ozone, grows considerably with hot weather due to physical reactions that require sunlight. In the summers and especially during extreme heatwaves, ozone often reaches dangerous levels in cities or nearby rural areas. Heatwaves often lead to poor air quality. The extreme heat and stagnant air can generate drought, triggering forest fires

and increasing ozone and particulate concentrations. Also, human activity heavily affect the air quality: traffic, industrial power plants, energy consumption and burning and heating are some of the possible pollution sources. Obviously, the introduction of restrictions and actions to reduce concentrations have the opposite effect. All of these sources contribute considerably to generating fluctuations and volatility in concentrations, with the direct consequence of increasing the noise in the data.

6.1.3 A definition of noisy time series and main contribution of the study

In the statistical-econometric literature, the relationship between the natural volatility of the series and the noise generated by the underlying processes, exemplified by the instances discussed above, is often represented in the form of a ratio and is called the signal-to-noise ratio. Define σ_v^2 as the variance of the underlying generating process and let σ_u^2 be the variance of the noise burying the data. Thus, the *noise-to-signal ratio* is defined as:

$$\lambda = \frac{\sigma_u^2}{\sigma_v^2}. \tag{6.1}$$

Note that the noise-to-signal ratio grows when the variance of the noise increases, while it reduces when the signal, i.e. the variability of the data generating process, increases. High values of the ratio are associated with extremely noisy series, represented graphically to hide the patterns characterizing the underlying processes.
This Chapter stresses that tests for stationarity of time series and cointegration may fail when data show some unpleasant characteristics like high variability, leptokurtosis, and extreme values. In particular, we show that standard ADF results and Johansen's tests are biased towards more stationarity and more generally, least-square based methods cannot be applied given the characteristics of time series. In these cases, cointegration analysis may produce unreliable results and generate false trading or policy signals. Therefore, it is essential to treat data appropriately when the scope of the analysis is to retrieve long-run components governing the behaviour of the series. We show that data filtering improves standard tests' performance and should become a good practise when dealing with very noisy datasets. We also provide results for different approaches to filtering and smoothing, comparing their effect on the statistical performances, i.e. evaluating the powers and the sizes, of stationarity and cointegration tests.
We prove the effectiveness of different filtering strategies using simulated series, and then we apply our procedure to air quality data.
Many empirical applications grounding on real-world data show how easy it can be to handle all or certain statistical features in the matter, that is the presence of high noise in the data or strong cross-sectional correlations among the units. For this reason, we stress that the filtering technique is useful in many other cases of datasets showing similar characteristics.

6.2 Filtering and stationarity issues in environmental time series

6.2.1 Some statistical characteristics of air quality time series

As a severe concern, air pollution has received considerable attention in recent years, and various air pollution assessment methodologies and approaches have been advanced. Many studies, concerning different locations worldwide have shown the presence of common characteristics of air quality time series for various pollutants, such as particulate matters, ozone or oxides. As stated by Horowitz and Barakat, 1979, air pollutant concentrations are not independently and identically distributed, and are generated by non-stationary autocorrelated

stochastic processes. In general, air quality series are characterized by seasonality, strong persistence, long-memory (Chelani, 2013; Varotsos et al., 2005) or non-stationarity (Ng and Yan, 2004), right skewness with uni-modal distribution (Windsor and Toumi, 2001), scale invariance and fractal behaviour (Lee, 2002). Moreover, environmental and hydrogeological data can also be affected by outliers and extreme values, which act as confounders in the models and require specific modelling techniques, such as the *generalized extreme value* (GEV) probability theory (Mudelsee, 2020) or *peaks-over-threshold* (POT) models (Gyarmati-Szabó et al., 2017). In particular, the characteristics of positive asymmetry requires particular attention since most of the classical statistical and econometric tools ground on the assumption of Gaussianity of the data. As already discussed in Chapters 3 and 5, even in the presence of large samples, simple statistics as the mean or median difference between air quality samples can easily fail within ordinary statistical contexts. For this reason, it is good practice to consider methodologies such as non-parametric bootstrap (Mudelsee, 2014) or non-parametric statistics to model the observations.

6.2.2 why we should not ignore such features in the models

All the previous characteristics influence the choice of the right model to use and the analysis's objectives. In the case of short-run forecasting of air quality data, literature based on ARMA filters and smoothers is well known and developed (Köppelová and Jindrová, 2019; Naseem et al., 2018; Zhang et al., 2017; Zhang et al., 2018), while the use of new approaches based on machine learning is growing (García Nieto et al., 2018; Zhu et al., 2018). Whereas the latter can deal with non-linearities and non-stationary data (Freeman et al., 2018), the first class of models requires a detailed stochastic processes analysis, including testing for stationarity and the presence of extreme values.
The selection of miss-specified models (Proietti, 2005; Crato and Taylor, 1996; Nelson, 1992; Nelson and Foster, 1995), or to ignore the presence of outliers in estimating the model (Ledolter, 1989), can lead to low forecasting performances due to overall increases in the forecasting error variance and the prediction interval width and generate bias in parameter estimation. As shown by Crato and Taylor, 1996, long-memory ARFIMA processes can be easily confounded with non-stationary ARIMA processes since both have similar and hardly distinguishable peculiarities. Both processes have slowly decaying autocorrelations, similarities in the frequency domain and present long non-periodic waves. Moreover, the ADF unit-root test has low power against fractional alternatives (Rudebush, 1991; Hassler and Wolters, 1994) with the direct consequence of highlighting false non-stationarity.
All these facts necessarily lead to considering the use of appropriate and robust analytical tools with respect to the various characteristics of air quality data. Therefore, focusing on the stationarity and cointegration issues, correct identification of the DGP through pre-filtering procedures, such as those proposed in this Chapter, could make the analysis strategies more robust and suggest more accurate considerations.

6.3 Why the ADF and related tests fail when integrated time series are observed with strong noise

Let us consider a simple random walk plus white noise model:

$$\begin{aligned} y_t &= x_t + \varepsilon_t, \qquad \varepsilon_t \sim WN(\sigma_\varepsilon^2) \\ x_t &= x_{t-1} + \eta_t, \quad \eta_t \sim WN(\sigma_\eta^2), \end{aligned} \tag{6.2}$$

where the notation $\varepsilon_t \sim WN(\sigma_\varepsilon^2)$ is to be read as "ε_t is a white noise sequence with variance σ_ε^2". Moreover, let

$$\lambda = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \geq 0$$

be the *signal-to-noise ratio*.
It is straightforward to prove that the process y_t has the reduced ARIMA(0, 1, 1) form

$$\Delta y_t = \eta_t + \varepsilon_t - \varepsilon_{t-1} = \zeta_t - \theta\zeta_{t-1}, \qquad \zeta_t \sim WN(\sigma^2) \tag{6.3}$$

with

$$\theta = 1 + \frac{\lambda - \sqrt{\lambda^2 + 4\lambda}}{2}, \qquad \sigma^2 = \frac{\sigma_\varepsilon^2}{\theta}.$$

When the signal-to-noise ratio is zero the MA coefficients is $\theta = 1$, the unit root operator cancels out with the MA operator and y_t turns out to be just a white noise sequence:

$$(1 - L)y_t = (1 - L)\zeta_t \quad \Leftrightarrow \quad y_t = \zeta_t.$$

When λ is close to zero, the MA coefficient approaches one from below. In this case, the exact cancellation does not occur; however, in small samples, the process y_t is almost indistinguishable from white noise. Such a MA process still has the purely AR representation

$$\Delta y_t = \zeta_t + \theta\Delta y_{t-1} + \theta^2\Delta y_{t-2} + \theta^3\Delta y_{t-3} + \ldots,$$

however, this representation cannot be well approximated by an AR(p) process with small p because θ^j approaches zero very slowly. Now, most unit root tests deriving from the Dickey-Fuller test such as ADF (Said and Dickey, 1984), ADF-GLS (Elliott et al., 1996), Johansen (Johansen, 1991) are based on autoregressive approximations and, if the y_t is generated as above with λ close to zero, then they are severely over-sized (see Galbraith and Zinde-Walsh, 1999, for example).
Table 6.1 gives an idea of the amount of size-distortion as a function of the signal-to-noise ratio and the moving average coefficient. The reported values are the results from a Monte Carlo experiment with the following characteristics: 10,000 replications of time series of length 200, the auxiliary AR(p) model order selected according to the Akaike Information Criterion (AIC) with p ranging from 0 to 20, the ADF test applied with a drift coefficient.[1]

TABLE 6.1: Actual ADF test size for different values of the signal-to-noise ratio for a nominal size of 5%.

λ	θ	size
0.001	0.97	1.00
0.010	0.90	0.89
0.100	0.73	0.34
1.000	0.38	0.10
10.000	0.08	0.05

[1]Simulations have been carried out using R 3.5.1 (R Core Team, 2018) with the ADF test implemented in the package *urca* (Pfaff, 2008).

6.4 Filtering

In various works on electricity prices, we found that two types of time series filters can improve the performance of unit root and cointegration tests:

1. reducing the frequency of the time series by taking averages, for instance by working on weakly means of daily or hourly electricity prices;
2. extracting the level component using the smoother in an unobserved component model (UCM) containing trend, noise and, possibly, seasonal components.

In this section, we analyze the effects of these filters on the reduced form of the model on time series generated by equation (6.2). The next section will illustrate the effect of these filters on the ADF and Johansen tests' size and power by Monte Carlo experiments.

6.4.1 Frequency reduction by averaging

Theorem 3. *For $t = 1, 2, \ldots$, let y_t be defined as in (6.2) and let*

$$\bar{y}_t = \frac{1}{m}\sum_{i=0}^{m-1} y_{t-i}, \qquad \bar{\eta}_t = \frac{1}{m}\sum_{i=0}^{m-1} \eta_{t-i}, \qquad \bar{\varepsilon}_t = \frac{1}{m}\sum_{i=0}^{m-1} \varepsilon_{t-i}, \tag{6.4}$$

where $\bar{y}_t$ is sampled over the set of time points $t \in \mathcal{T} := \{m, 2m, 3m, \ldots\}$. Then, over the time+ set $\mathcal{T}$

$$\bar{y}_t - \bar{y}_{t-m} = \bar{\eta}_t + \bar{\varepsilon}_t - \bar{\varepsilon}_{t-m},$$

with $\bar{\varepsilon}_t$ white noise sequence with variance σ_ε^2/m and $\bar{\eta}_t$ MA(1) process with

$$\begin{aligned} VAR\bar{\eta}_t &= \sigma_\eta^2\left[\frac{(m-1)(2m-1)}{3m} + 1\right], \\ COV\bar{\eta}_t, \bar{\eta}_{t-m} &= \sigma_\eta^2\frac{(m-1)(m+1)}{6m}, \\ COR\bar{\eta}_t, \bar{\eta}_{t-m} &= \frac{(m-1)(m+1)}{2(m-1)(2m-1)+3m}. \end{aligned}$$

Corollary 1. *Under the hypotheses of Theorem 3, the process $\bar{y}_t$ is ARIMA(0, 1, 1) for $t \in \mathcal{T}$ with first-order autocorrelation given by*

$$\rho = COR[(\bar{y}_t - \bar{y}_{t-m}), (\bar{y}_{t-m} - \bar{y}_{t-2m})] = \frac{\lambda\frac{(m-1)(m+1)}{6m} - \frac{1}{m}}{\lambda\left(\frac{(m-1)(2m-1)}{3m} + 1\right) + \frac{2}{m}} \tag{6.5}$$

and moving average coefficient given by

$$\theta = \frac{-1 + \sqrt{1 - 4\rho^2}}{2\rho}.$$

See Appendix C for a detailed proof of the previous Theorem's results.
The left panel of Figure 6.1 depicts the value of the MA coefficient θ as a function of the window size m, for various values of the noise-to-signal ratio λ^{-1}. The right panel of the same figure depicts the locus of points (λ^{-1}, m) for which $\theta = 0$. In principle, for any signal-to-noise ratio λ, there is a value m that annihilates the MA component. Of course, since m can only take integer values, θ will be in general close to zero and not exactly equal to zero. If the noise is

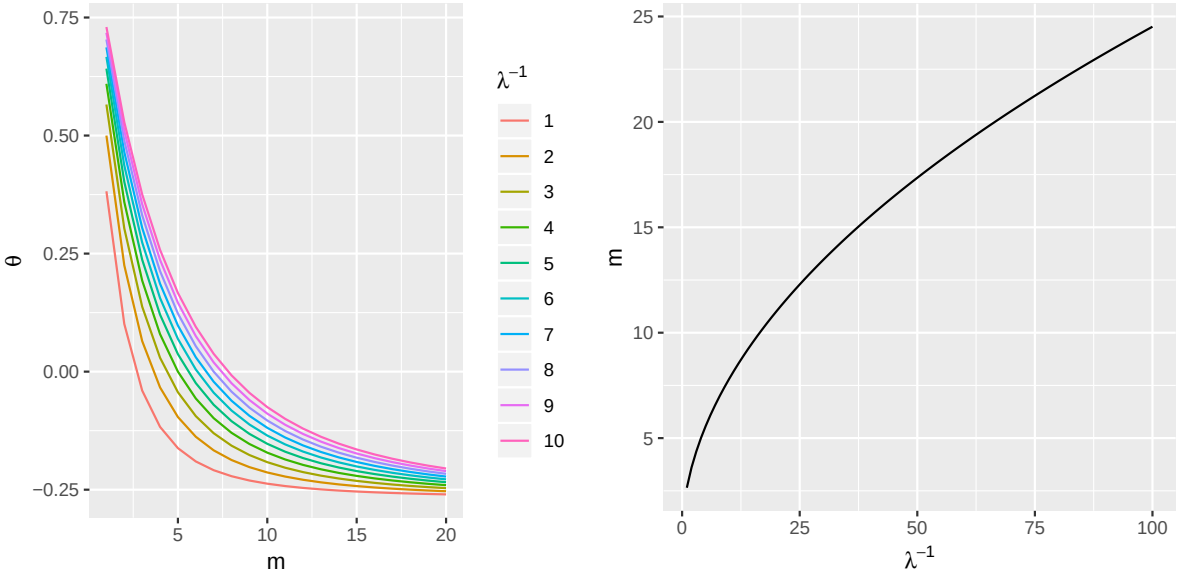


FIGURE 6.1: Left: MA coefficient of mean-filtered time series as a function of the window for different values of the noise-to-signal ratio. Right: value of the window size m that achieve an MA coefficient equal to zero as a function of the noise-to-signal ratio.

excessive, the original time series must be relatively long as the length of $\bar{y}_t$ is given by $\lfloor n/m \rfloor$, where n is the length of the time series y_t.

6.4.2 Signal extraction in a unobserved component model

An alternative way to reduce the noise in the process y_t defined in Equation 6.2 is by estimating the random walk component μ_t by projecting it on the linear span of $y_0, y_1, \ldots, y_s$ where s is either equal to t or to $n-1$. This operation can be easily carried out by stating the model in state-space form and running the Kalman filter (for $s = t$) and smoother (for $s = n-1$) on the level component μ_t (the voulmes by Harvey, 1989; Durbin and Koopman, 2001; Pelagatti, 2015b, are detailed references on the subject).
The Kalman filter and smoother for μ_t are a linear filters whose weights are different for every t: the former is only backward looking,

$$\mu_{t|t} = \sum_{i=0}^{t} w_{ti} y_{t-i}, \qquad t = 0, 1, \ldots, n-1, \tag{6.6}$$

while the latter is two-sided,

$$\mu_{t|n-1} = \sum_{i=t}^{t-n+1} w_{ti} y_{t-i}, \qquad t = 0, 1, \ldots, n-1. \tag{6.7}$$

Nonetheless, when t is not too close to 0, the Kalman filter for the RW plus noise model can be well approximated by its steady state version, say $\tilde{\mu}_t$, given by the recursive filter[2]

$$\tilde{\mu}_t = \gamma y_t + (1-\gamma)\tilde{\mu}_{t-1}, \qquad t = 0, 1, \ldots, n-1, \tag{6.8}$$

[2] A simple derivation of this result can be found in Example 5.12 of Pelagatti (2015b) and this particular form of the coefficient γ in Harvey (2006, Eq.13).

where

$$\gamma = \frac{\sqrt{\lambda^2 + 4\lambda} - \lambda}{2}. \tag{6.9}$$

Similarly, when t is not too close to 0 or $n-1$, the smoother can be well approximated by its steady-state version, say $\hat{\mu}_t$, which is given by the backward recursion on $\tilde{\mu}_t$

$$\hat{\mu}_t = \gamma\tilde{\mu}_t + (1-\gamma)\hat{\mu}_{t+1}, \qquad t = n-1, n-2, \ldots, 0. \tag{6.10}$$

Theorem 4. *Assume that the variances σ_ε^2 and σ_η^2 of the process defined in equation (6.2) are known and $\lambda = \sigma_\eta^2/\sigma_\varepsilon^2$ is the signal-to-noise ratio. Then, $\tilde{\mu}_t$ is a random walk and $\hat{\mu}_t$ is an ARIMA$(1,1,0)$ process with autoregressive coefficient $\phi = 1-\gamma$.*

See Appendix D for a detailed proof of the previous Theorem's results.
In applications, the unknown variances are replaced with their Gaussian (quasi) maximum likelihood estimates $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\eta^2$ and the population signal-to-noise ratio with its estimate $\hat{\lambda} = \hat{\sigma}_\eta^2/\hat{\sigma}_\varepsilon^2$.

6.5 Simulation experiments

To verify the empirical effects of the frequency-reduction and UCM filters on the ADF and Johansen tests, we performed a Monte Carlo simulation set. Each experiment evaluates the performances of ADF and Johansen tests both under the null and alternative hypothesis to evaluate the size and power of the tests.

6.5.1 ADF test

The first set of experiments aims at assessing the performance of the ADF test under the random walk plus (leptokurtic) noise model. We simulate time series from a random walk buried in leptokurtic noise and each simulation experiment is characterized by a different combination of noise-to-signal ratio and kurtosis value. The data generating process (DGP) for the observation y_t is

$$\begin{aligned} y_t &= x_t + \sqrt{c}\,\varepsilon_t, & \varepsilon_t &\sim i.i.d.(0,1), \\ x_t &= x_{t-1} + \eta_t & \eta_t &\sim \text{NID}(0,1), \end{aligned}$$

where ε_t is the leptokurtic noise generated by a standardized Student's t with ν degrees of freedom (DF) and $c = \lambda^{-1}$ is the fixed parameter identifying the noise-to-signal ratio. The number of DF governs the thickness of the tails of the noise component: the lower the DF, the larger the kurtosis.
For each experiment, we simulate 10000 paths of length 1095, corresponding to 3 years of daily observations, for all of the possible pairs of noise-to-signal ratio c in $\{0, 1, 2, ..., 10\}$ and degrees of freedom ν in $\{3, 6, 9, 12\}$. On each of these time series we apply the mean-filter (mean), the Kalman filter (ucm_flt) and the smoother (ucm_smo).
The ADF statistic (with drift and number of lags selected by corrected AIC) was computed on every simulated time series, and the empirical rejection rates for a nominal size of 5% are represented in Figure 6.2. In the presence of high level of noisy, the classic ADF test should reject the non-stationarity hypothesis much more often than is tolerated by the theoretical values (i.e. large increase in test's size) as well it should suggest accepting the stationarity hypothesis excessively (i.e. loss of test's power); Looking at the four graphs, we can conclude that

1. the ADF applied to the filtered time series has size close to the nominal one;

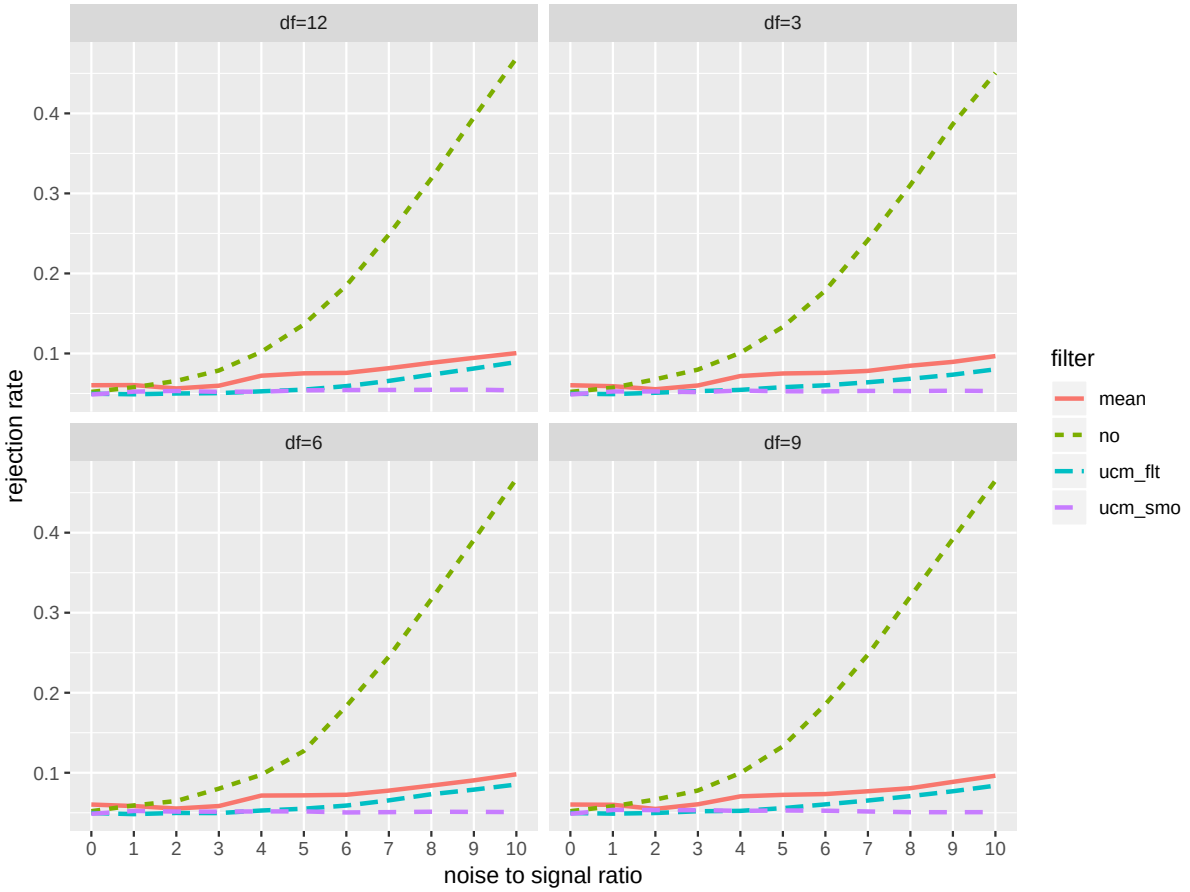


FIGURE 6.2: Actual size of the ADF test for a nominal size of 5%.

2. the size of the ADF test applied to the smoother (ucm_smo) time series is equal to the nominal size for all the considered noise-to-signal ratios;

3. the thickness of the tails of the noise distribution has virtually no effect on the size of the ADF.

The latter remark may sound surprising; however, the ADF test's asymptotic distribution is derived under the hypothesis of finite second moments. In our simulations, the variance of the noise ε_t is always finite, and the sample size large ($n = 1095$).
Analyzing the size-adjusted powers of the simulated ADF, here reported in Figure 6.3, we notice that the noise-to-signal ratio clearly influences the performance of the filters applied to the data and the corresponding tests, whereas the thickness of the tails (kurtosis) seems to operate very poor influence. This is very similar to what we learned earlier about the size.
Simulations of the ADF statistics show that, for low or non-existent noise-to-signal ratio, i.e. in the presence of weak noise, using the test on raw data is equivalent in terms of power to the tests applied after filtering and smoothing the series. The estimated power is in fact almost identical to the nominal power (in the 95% simulations). Unfortunately, by increasing the signal-to-noise ratio even by a small fraction, tests on filtered and smoothed data lose power rapidly, reaching 50% for values of the ratio close to 10. On the contrary, the raw data test overestimates the power as it approaches values closer and closer to 1. The size-adjusted power of the averaged data shows very weak performances. It always underestimates the power of the test and never approaches the nominal threshold of 95%. Both the noise-to-signal ration and the thickness do not influence it.

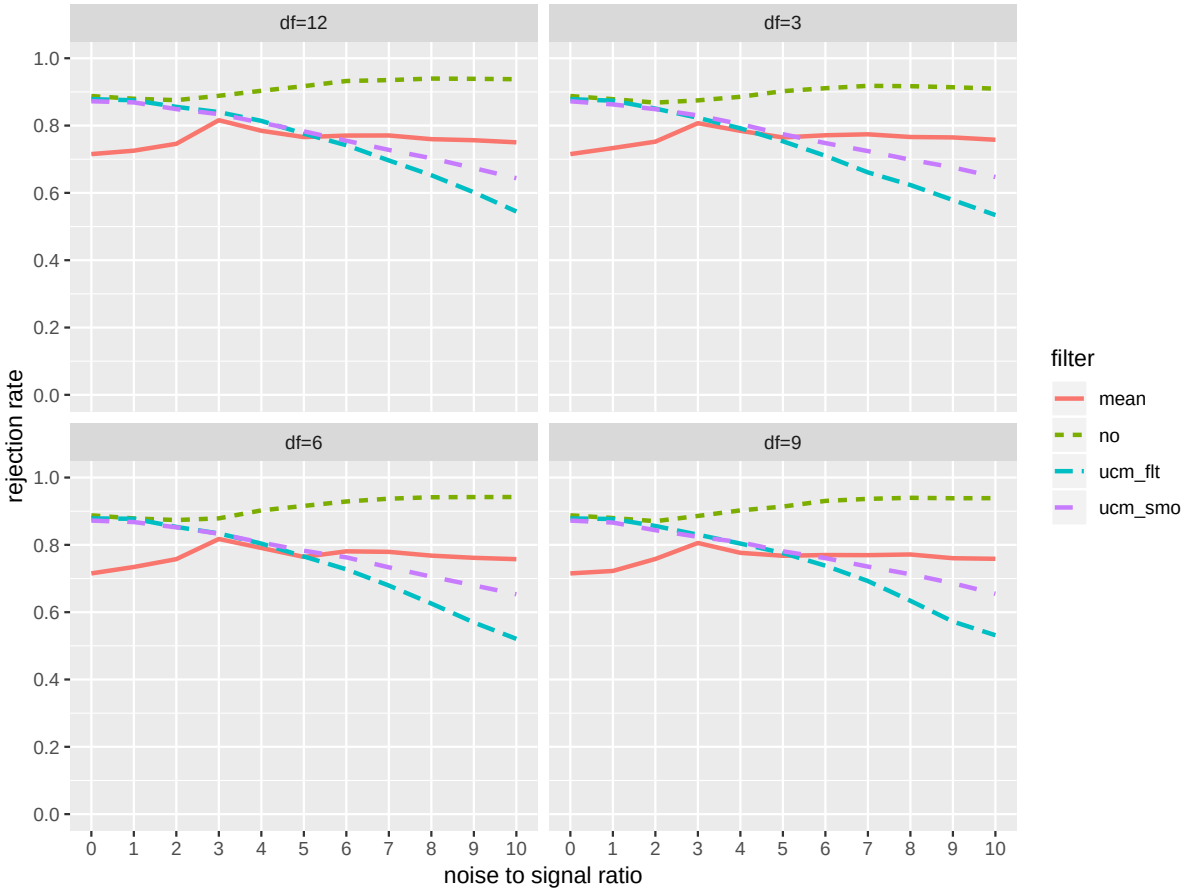


FIGURE 6.3: Size-adjusted power of the ADF test for the size of 5% and data generating process being AR(1) plus Student's t noise with autoregressive coefficient $\phi = 0.98$.

Summing up, from the analysis of the sizes and of the powers of the ADF test, we can remark the following considerations:

- the ADF applied to unfiltered time series maintains good performances in terms of power in the presence of weak noise in the data, but tends to overestimate the absence of stationarity in the series (over-rejection of the null hypothesis) even for modest values of noise;
- both filtered and smoothed time series perform very well in terms of size even for large and relevant noises. In particular, the smoother seems to keep excellent performances for all the simulated values of the noise-to-signal ratio. However, in the case of stationary series, the smoothed and filtered ADF test tends to detect non-stationarity in the series more often than the theoretical value (power loss) even in the presence of a low signal-to-noise ratio;
- averaging filter allows to keep under control the size of the test, similarly to the Kalman filter and the smoother, but heavily underperforms in terms of power of the ADF;
- the thickness of the tails of the noise distribution has virtually no effect on both the size and size-adjusted ADF power. In fact, all the simulated statistics show similar patterns for increasing thickness.

6.5.2 Johansen test

Similarly to the univariate case, we developed a simulation scheme for multivariate time series that include both integrated and cointegrated processes to evaluate the statistical properties of Johansen's cointegration test in the presence of leptokurtic noise.
Data are simulated according to a vector error correction model (VECM) with r cointegrating relations and $k = 4$ underlying times series augmented by a leptokurtic noise term. The noise is randomly generated by a standardized Student's t random variable with ν degrees of freedom and directly affects the VECM through the noise-to-signal ratio. We performed the simulation analysis on the number of cointegrating relations detected by the test, considering the case of $r = 1$ and $r = 2$.
For the simulation experiments involving the Johansen test we used the following two DGP: the first with $r = 1$ cointegrating relations,

$$r = 1: \quad \Delta x_t = \begin{bmatrix} 0.0 \\ 0.1 \\ -0.1 \\ 0.1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix} x_{t-1} + \varepsilon_t, \tag{6.11}$$

and the second with $r = 2$ cointegrating relations,

$$r = 2: \quad \Delta x_t = \begin{bmatrix} 0.0 & 0.0 \\ 0.1 & 0.0 \\ 0.2 & 0.0 \\ 0.0 & 0.2 \end{bmatrix} \begin{bmatrix} 1.0 & -0.5 & -0.5 & 0.0 \\ 0.0 & 1.0 & 0.0 & -1.0 \end{bmatrix} x_{t-1} + \varepsilon_t. \tag{6.12}$$

and the noisy series

$$y_{it} = x_{it} + \gamma_i z_{it} \qquad \gamma_i^2 = c \frac{VAR(\Delta y_{it})}{VAR(x_{it})} \tag{6.13}$$

where z_{it} is a leptokurtic noise and c is a fixed parameter representing the noise-to-signal ratio. Transforming the previous models from VECMs form to VAR(1) form, it's easy to assess that the characteristics roots of the processes are (1, 1, 1, 0.7) for equation 6.11 and (1, 1, 0.85, 0.80) for equation 6.12.
As in the previous experiment, we simulated 10,000 time series paths of length 1095 for all the paired-combinations of noise-to-signal ratio $c = 0, 1, 2, ..., 10$ and degrees of freedom $\nu = 3, 6, 9, 12$. The three linear filters are then applied to the series, rising three further series sharing a common underlying process. The number of cointegrating vectors is finally tested on each simulated quartet using the Johansen's trace test.
The trace test is a sequential test which sets under the null hypothesis a number of cointegrating vectors r lower than its maximum k, i.e. $H_0 : r = r^* < k$, against the alternative hypothesis $H_1 : r = k$. The value $r^* = 1, 2, ..., k-1$ is updated sequentially from 1 to $k-1$ producing a sequence of tests whose first non-rejection of the null hypothesis can be considered an estimate of r. The test rejects H_0 when the test statistic exceeds the tabulated critical values.
For each pair of degrees of freedom and noise-to-signal ratio and each sequential value r^*, we computed the test's size (rejection rate) as the proportion of tests rejecting the null hypothesis over the total number of simulations. For values of r^* lower than the real one, the expected rejection rate should be the closest possible to 1; while it should tend towards zero approaching or overcoming the actual value. Since the existence of high noise overstates the stationarity of the series, the Johansen's cointegration test will overestimate the cointegration rank, concluding that the number of cointegration relations is higher than the effective number (i.e. large

increase in test's size and reduction of the test's power). It is also possible to estimate the selection rate as the unitary complement with respect to the rejection rate and use it to compare the test performance. Considerations on how the selection rate should approach the sequence are opposite to the rejection rate: high selection proportions should be associated to values of r^* close to the true one and low when r^* is enough lower from it.

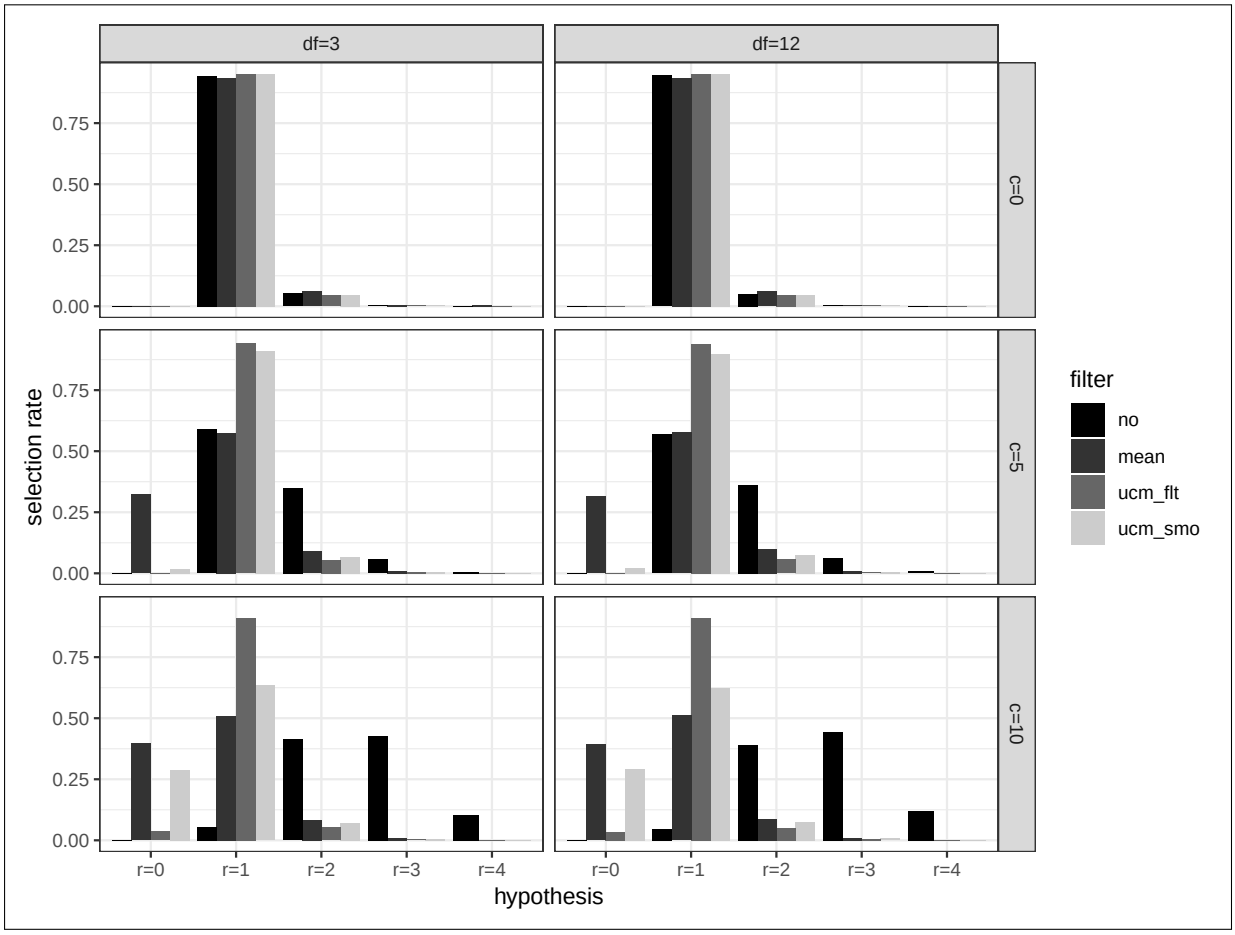


FIGURE 6.4: **Selection rate when** $r = 1$.

Figures 6.4 and 6.5 show the empirical selection rates when the real number of cointegrating vector is $r = 1$ and $r = 2$. Until the noise-to-signal ratio remains close to zero and independently by the degrees of freedom, both filtered and raw data perform similarly and no real advantages in pre-filtering are visible. However, as soon as the signal-to-noise ratio gets unbalanced, the selection rates change considerably. Kalman filter and smoother are able to guarantee selection proportions in all situations, in particular, the filter remains stable with values above 90% for $r = 1$ and over 80% when $r = 2$. The figures also show that for highly unbalanced noise-to-signal ratios, just the filter can maintain acceptable performances, while the other filters, and obviously raw data, are not very effective.
In Appendix E, we provide the representation of the simulated rejection rates at increasing values of the noise-to-signal ratio. Figures E.1 to E.8 show the rejection rates for each of the alternative hypotheses of the test when the true number of cointegration relationships are r=1 and r=2. When the hypothesis is equal to the real value r, i.e. testing $r = 1$ when $r = 1$ and testing $r = 2$ when $r = 2$, the tests applied to unfiltered series (*no*) tend to over-reject the hypothesis even for weak noise values, while the three filters are almost equivalent. However, the Kalman filter (*ucm_flt*) and the smoother (*ucm_smo*) provide rejection rates very close to

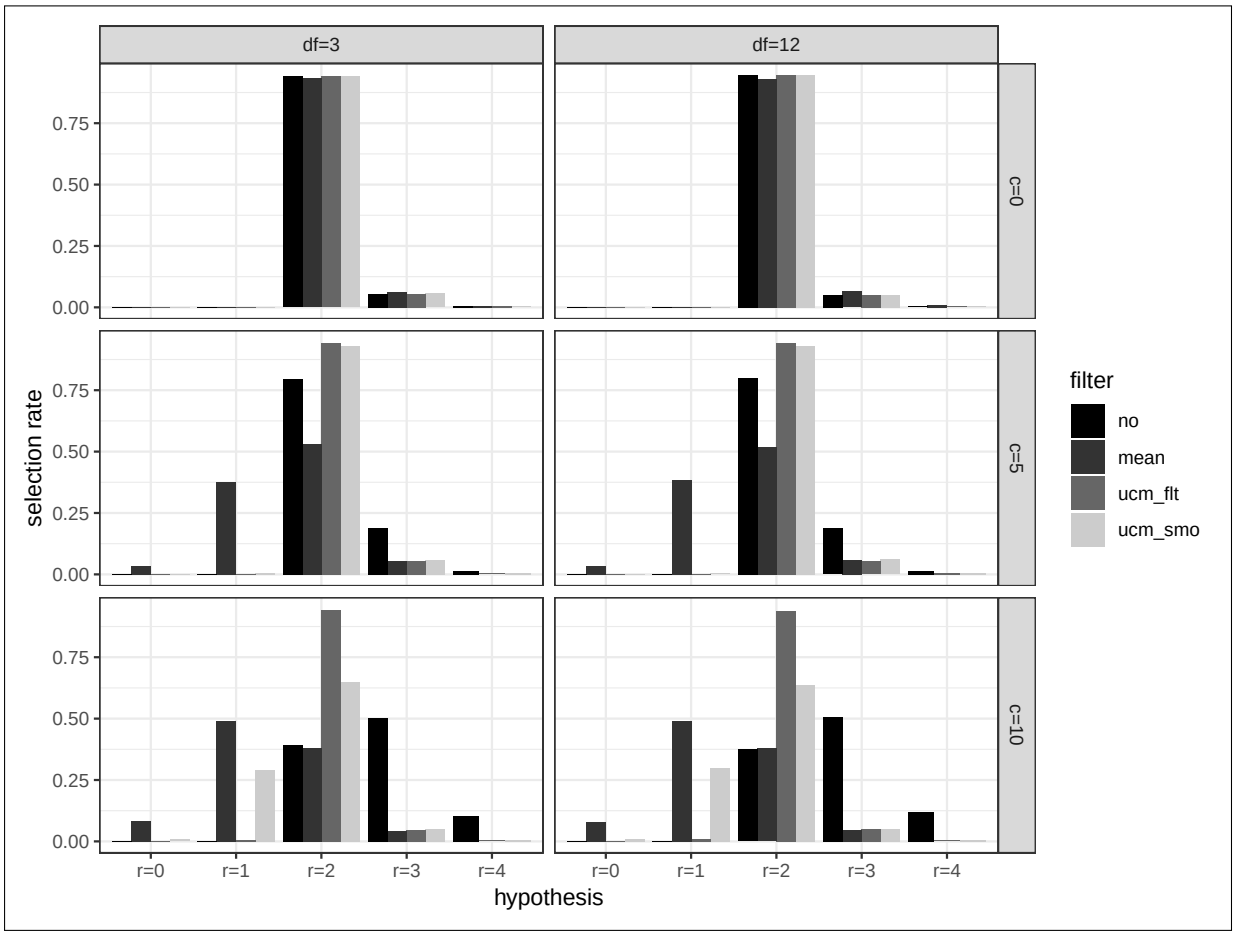


FIGURE 6.5: **Selection rate when** $r = 2$.

the nominal value (5%), thus preserving the real confidence level. Regarding the power of the test, when testing hypothesis greater than the true value of r, the filters are associated with rejection rates equal to the nominal value, independently by the noise-to-signal strength. At the opposite, considering the raw data, even for modest noise-to-signal ratios, the rejection rate increases sharply to high values. When the hypothesis is lower than the true value, only the Kalman filter succeeds in keeping the rejection rate close to nominal, as the moving average (*mean*) accepts the hypothesis even small values of the noise and the smoother begins to under-reject for medium to high values. Therefore, we can state that the filters are able to preserve the test's power even in the case of noisy data.

6.6 Application: stationarity and cointegration analysis of air quality time series in Milan, Italy

6.6.1 Sample description: NO_2 and NO_x concentrations in Milan

To test the performances of filtering techniques on empirical contexts, we decided to apply the ADF test for stationarity and the Johansen's test for cointegration to the case of airborne oxide concentrations in Milan, Lombardy. The application concerns both nitrogen dioxide (NO_2) and total nitrogen oxides (NO_x) levels. Within the Milan territory, five ground stations are installed for air quality monitoring managed by the regional agency for environmental protection (ARPA Lombardia). Figure 6.6 shows the geographical location of each station.

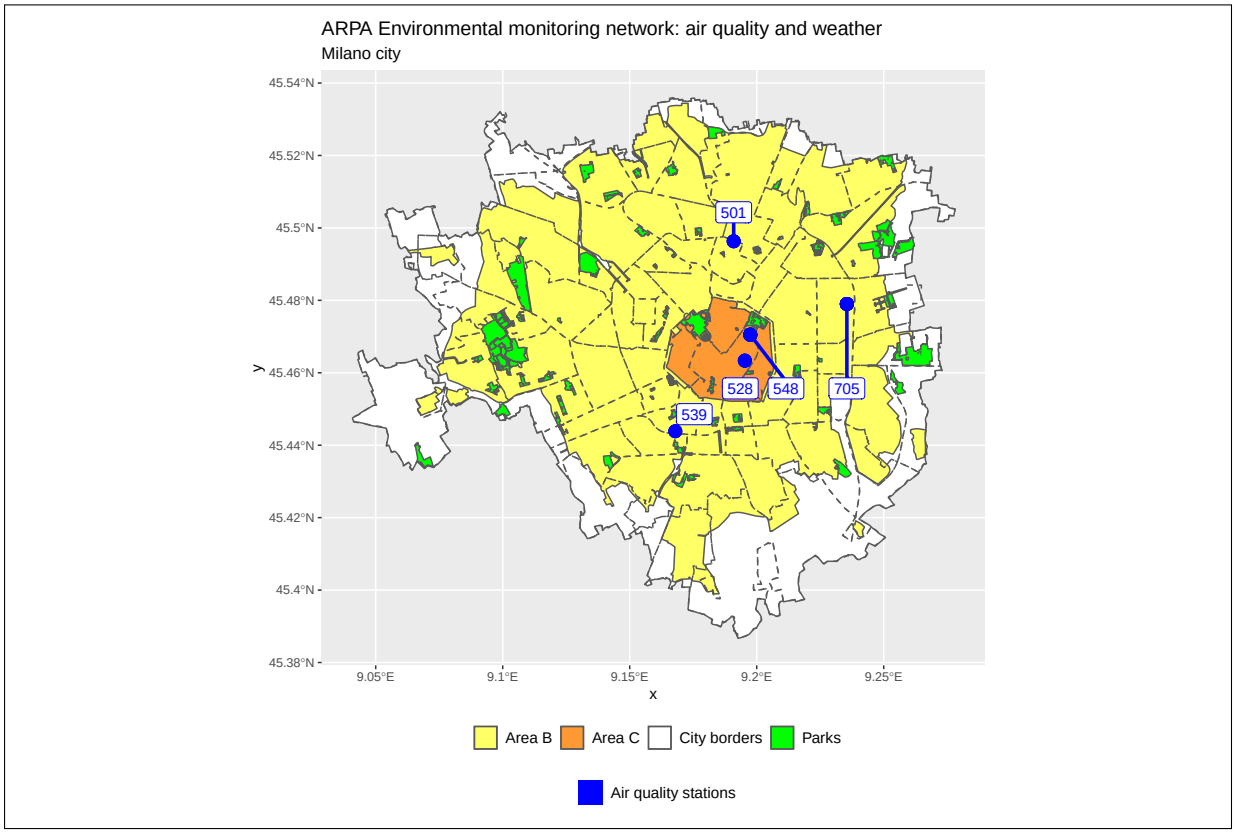


FIGURE 6.6: ARPA Lombardia air quality ground monitoring stations in Milan

Three of them, that is Liguria (539), Marche (501) and Pascal-Città Studi (705) are located near the borders of the city, whereas Senato (548) and Verziere (528) stations cover the historical city centre. According to the international criteria, all the stations are classified as urban units. Moreover, Verziere, Liguria, Marche and Senato are traffic stations, while Pascal is a background monitoring unit.
For all the stations, sensors able to detect oxide concentrations were installed before 2011 and are currently active; hence, they provide exceptionally long information on these airborne pollutants' long-run evolution. Filtering and tests are applied considering the longest possible available window: we used average daily observations of NO_2 and NO_x concentrations from the 1st January 2011 to the 31st December 2019. We decided to stop the data collection before 2020 to avoid confounding effects generated by the recent COVID-19 pandemic, which affected Milan's air quality and all Lombardy region heavily. Recall the results presented in Chapter 4 for details on this event.

6.6.2 Estimation and testing mechanisms

Given the proximity of the monitoring stations, they are expected to be affected by statistical common trends. Therefore, cointegration analysis should highlight one or more long-term relationships. Regarding the stationarity, air quality time series are often referred to be stationary. As shown by the previous analyses in this manuscript, Lombardy's oxides are affected by a downward trend and a strong annual and weekly seasonality, which is very smooth and apparently deterministic. For this reason, before applying the filters and tests, we will proceed to seasonal adjustment and detrending of the univariate time series using deterministic linear components. We will pre-treat the series by running a linear regression consisting of a linear

trend and a set of trigonometric covariates at various frequencies to identify the seasonal and long-run component. The deterministic harmonic regression is run using a double seasonality component: one for annual and one for weekly cyclicity. The weekly seasonality is identified using three sinusoids with period seven days, whereas 40 harmonics captures the annual seasonality with period 365 days. As we do not want to use information from other observed covariates, such as temperature, the number of annual sine waves must necessarily be high for a correct seasonal adjustment. However, we recall that the maximum number of harmonics with a 365-day periodicity is 182.
After running the regression, the residuals, which can be interpreted as the deseasonalized and detrended concentrations, are passed to the Kalman filter, to the smoother and to the moving-average filter. Finally, the ADF and the Johansen's tests are applied both to unfiltered and to filtered series.
We analyzed the case of stationarity and cointegration of the five-time series using two estimation mechanisms. In the first case, we use a rolling window approach in which the original series has been detrended and deseasonalized, and then filtered, using data referring to a time span of 5 years. In this way, we fed the filters and the tests with around 1825 daily observations at each iteration (recall that 2012 and 2016 were leaping years). The algorithm proceeds as follows: it estimates the linear regression; it then filters the data; it applies the ADF and cointegration tests to both the filtered and unfiltered data; finally, it moves forward the window of one day keeping fixed the estimation window. In the second case, we use a recurrent window approach in which the estimation window is updated at each iteration adding a new observation and repeating all the previous operations with an increasingly larger window.
To give an idea of the effect that filters have on the original series, Figure 2 shows the plot of the unfiltered and filtered time series of NO_2 concentrations for the Pascal-Città Studi station from 2016 to 2019.
All the filters maintain the original structure of the detrended and deseasonalized concentrations, that is the upper and lower peaks and the range, but at the same time are able to remove large amounts of noise that could affect test results. The filtering capacity of the three filters is increasingly noticeable: in ascending order, the Kalman filter, then the smoother and finally the moving average reduce more and more noise. Besides, in contrast to the Kalman approach, the moving average significantly reduces the number of observations.

6.6.3 Results

This section presents the main results obtained from the application of stationarity and cointegration tests on filtered and unfiltered NO_2 and NO_x series. The unfiltered series are indicated by the acronym *RAW*, while the filtered ones are *FLT*, *SMO* and *AVG* in the respective cases of Kalman Filter, Smoother and seven-day moving average. For both tests, the conclusions drawn using filtered and raw data are almost the same. Concerning stationarity, in general, all filters tend to strongly reject the non-stationary hypothesis of concentrations, both with the rolling window and with the recurrent. This holds for the whole length of the time series, meaning that airborne pollutants' stationarity does not be affected by the period of interest. The number of cointegration relationships identified, on the other hand, seems to depend on the period considered, but the conclusions reached by the filters are consistent with those suggested by the raw data.

Stationarity

In Figures 6.8 and 6.9 are reported the estimated ADF statistics (τ-statistic) for NO_2 using the rolling window sample and the reccurent window, respectively. Figures 6.10 and 6.11 show the

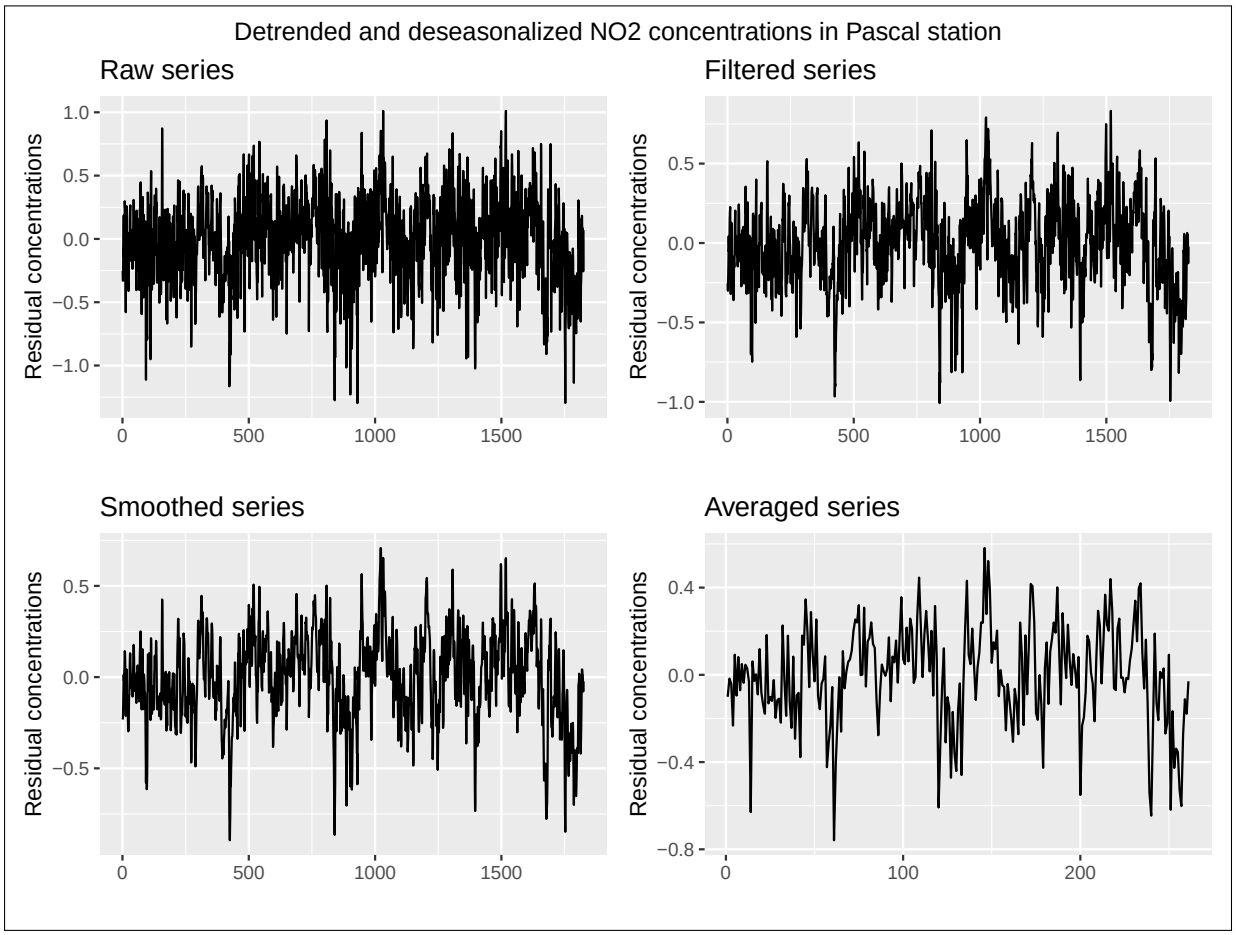


FIGURE 6.7: Raw detrended and deseasonalized concentrations (upper-left panel); detrended and deseasonalized concentrations passed into the Kalman's filter (upper-right panel); detrended and deseasonalized concentrations passed into the Kalman's smoother (lower-left panel); detrended and deseasonalized concentrations passed into the average weekly filter (lower-right panel).

same statistics but calculated on NOx concentrations. With the exception of a few specific moments when the test statistics are slightly above the critical 10% (see e.g. Pascal's rolling window in mid-2017 or Marche's recurrent window in early 2017), both the filtered and the original series lean towards the hypothesis of stationary concentrations. The Kalman Filter and smoothed series move similarly to the raw series. The smoother and filter often have higher values of the statistics than the raw ones, showing how the rejection decision is less strong. A detailed analysis of the average filter statistics shows that the values absorb the short-term seasonality (7-10 days) of the original series. This could be a sign of weakness of the moving average in the presence of strongly seasonal series such as those under analysis, despite the fact that they have already been adjusted for this component. This weakness, instead, is not evident with the Kalman filter and the smoother. In any case, the estimated values of the ADF statistics are well below the critical values. Finally, note that as the sample size of the recurrent window increases, the test statistics tend to be farther and farther away from the critical values, highlighting the stationarity property. This could be because, in long time series, possible non-stationarity is hidden by short-term variations, becoming more relevant as the sample size increases.

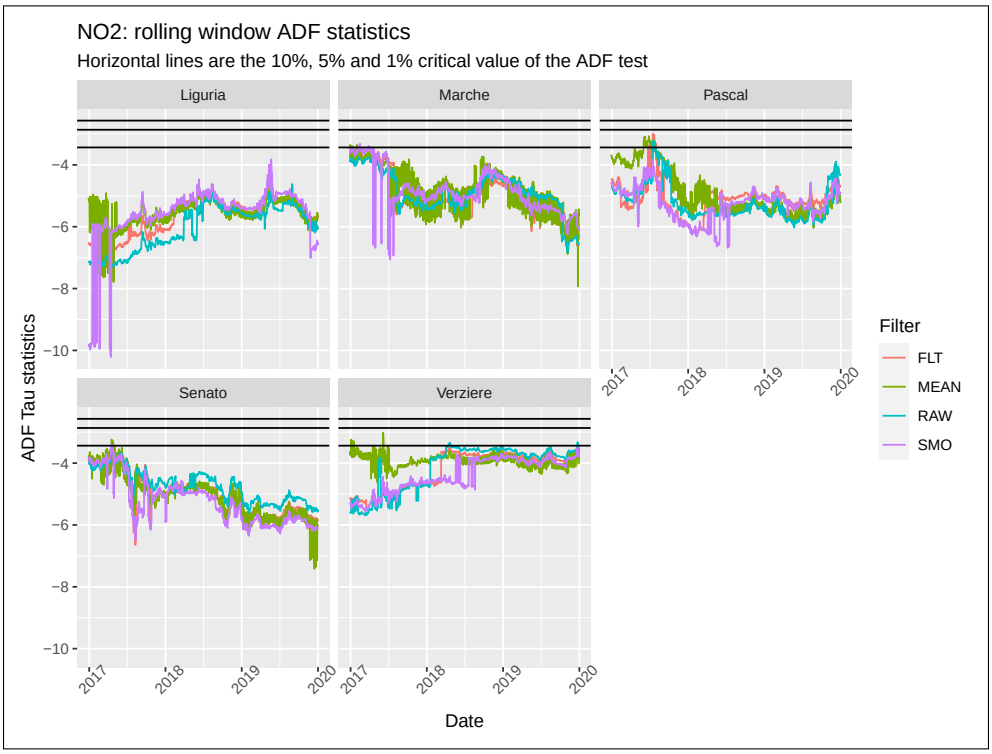


FIGURE 6.8: ADF Tau statistics for NO_2 by station. Rolling window of length 5 years.

Cointegration

In Figures 6.12 and 6.13 are reported the estimated trace statistics (τ-statistic) for NO_2 using the rolling window sample and the reccurent window, respectively. Figures 6.14 and 6.15 show the same statistics but calculated on NOx concentrations. Similarly to what has been discussed regarding stationarity, cointegration tests applied to filtered and unfiltered series also seem to be quite in line with each other. In contrast to the former case, however, the estimated trace statistics do not suffer from large variations as the sample varies. Recall that the Johansen trace test is sequential: we test the null hypothesis that the number of relations is exactly a value, against the alternative that it is less than that value, and then repeat the test up to the maximum number of possible relations, in our case 4. The number of cointegrating relationships for the oxide concentrations in Milan seems to depend on the sample which is being processed, Indeed, it can be seen that up to mid-2017 the trace statistics suggest between 1 and 3 cointegrating relationships for NO_2 and between 2 and 3 for NO_x. The filtered and unfiltered data are in agreement with each other. However, from the second half of 2017 onwards, all series reject the sequential null hypotheses by identifying a number of relationships greater than 4. This suggests that there are as many as five cointegrating relationships, so all stations in the city would be stationary, i.e. I(0). These results seem valid both when pre-filtering is applied to the series and when using the raw data. This is in line with the previous results obtained about the stationarity of the series: starting from the second half of 2017, the ADF statistics calculated on both filtered and unfiltered data suggest that all the time series of oxide concentrations in Milan are strongly stationary.

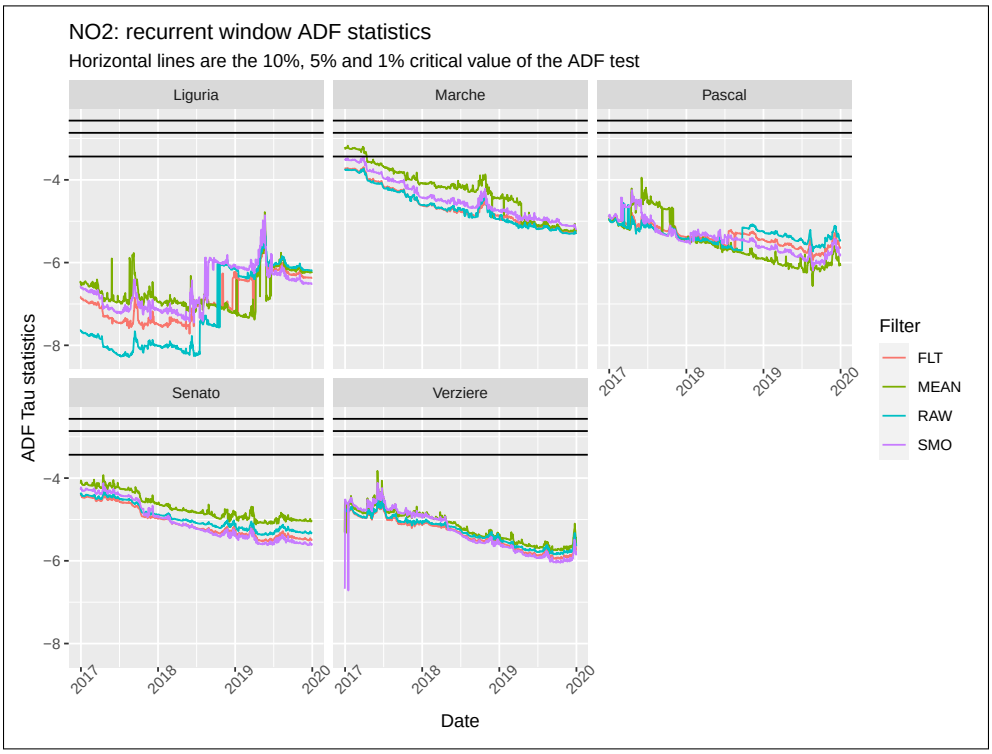


FIGURE 6.9: ADF Tau statistics for NO_2 by station. Recurrent window starting at 5 years.

6.7 Conclusions and final remarks on pre-filtering in stationarity and cointegration issues

In this Chapter, we investigated the effect of pre-filtering techniques on noisy time series. In particular, we have analyzed the issues of stationarity and cointegration of multivariate time series affected by high confounding noise, a phenomenon capable of leading to substantial biases in selecting appropriate models and data generating processes. Specifically, we assumed that the series were non-stationary and buried by a high noise, leading the classic stationarity tests (ADF) to reject the stationarity hypothesis more times than necessary and therefore to erroneous conclusions on the true nature of the data. We have shown theoretically and through simulations the beneficial effect that pre-filtering can bring to the Dickey-Fuller stationarity and Johansen cointegration tests in contexts such as those just described. Three filters are proposed: the Kalman filter, the Kalman smoother and aggregation employing k-period averaging. In the presence of high noise, all three filters can preserve the size of the two tests, but lose their effectiveness when looking at the size-adjusted power. In any case, they prove to be powerful tools to support the classical approaches and reinforce or deny the raw data's suggestions.
Finally, we decided to show an empirical application of pre-filtering by analyzing the case of oxide concentrations (NO2 and NOx) over Milan. The choice of concentrations is driven by the fact that pollutants, deriving from very complex atmospheric Physico-chemical processes involving many natural and anthropogenic factors, are characterized by strong variability and therefore by confounding noise overlapping the real data signal. The application showed that pre-filtering could be a useful tool for at least two reasons: the tests performed on the pre-filtered data support the hypothesis identified by the raw data, namely that the univariate time series of concentrations are stationary; concerning cointegration, on the other hand, it appears that the number of relationships identified varies according to the chosen temporal

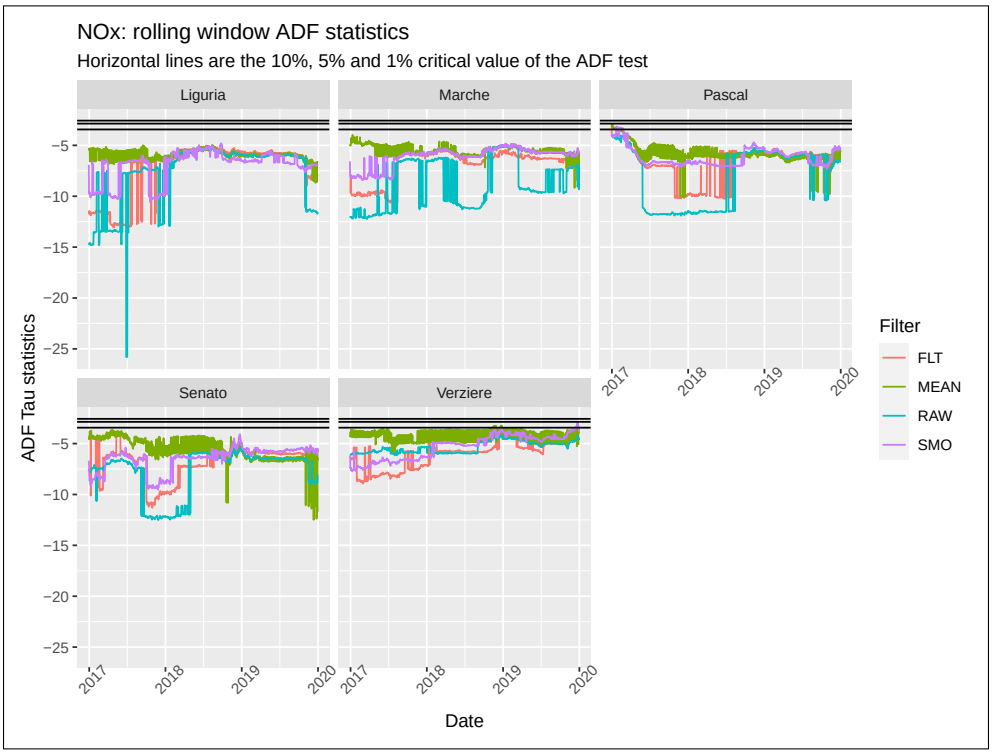


FIGURE 6.10: ADF Tau statistics for NO_x by station. Rolling window of length 5 years.

sample. For data before 2017, both pre-filtered and raw data reveal only between 2 and 4 cointegration relationships, whereas after 2017, as the series are strongly stationary, the number of relationships is equal to the number of stations analyzed.

The present research offers some important development opportunities. First of all, it would be interesting to develop other theoretical and simulative studies using structural models for time series (BSM), i.e. models which consider both the long-term (trend) and seasonal components. In this case, the signal would be composed of the trend component signal, discussed in the present work, and the stochastic seasonality signal. Moreover, as discussed in the sections 6.1.1 and 6.1.2, there are many empirical contexts in which data show high confusing noise that buries the real signal and that if not adequately considered can lead to immediate wrong conclusions about the appropriate models to describe the phenomenon. Applying pre-filtering methodologies in such situations could strengthen the scientific validity of applied research and, in the most extreme case, lead to the revision of some hypotheses that are now considered consolidated in the literature.

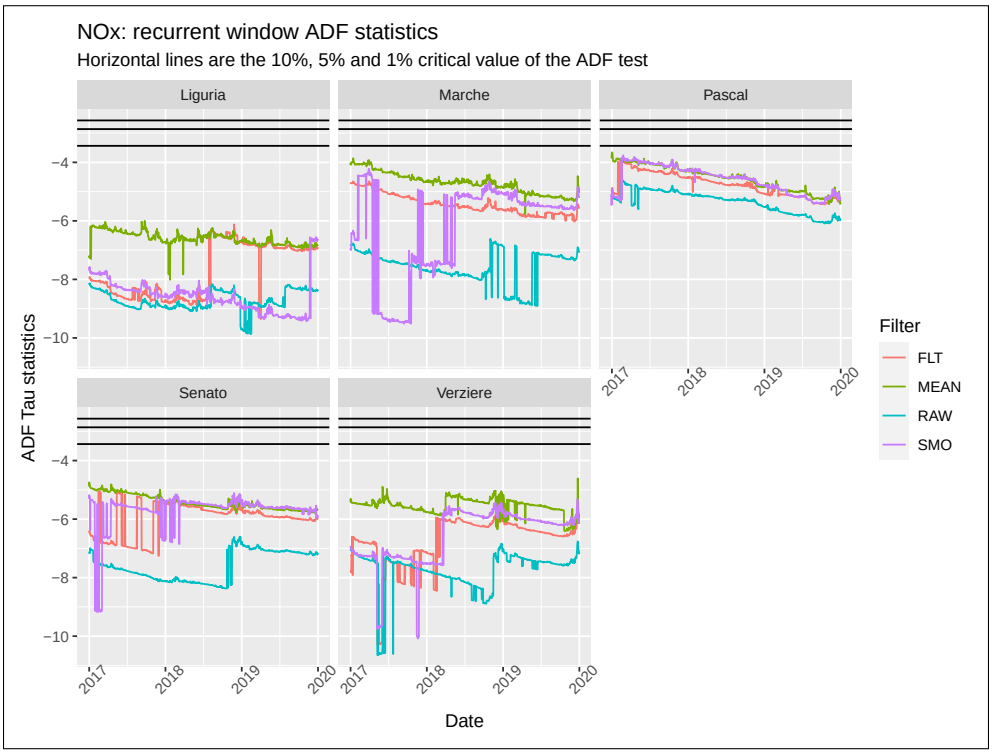


FIGURE 6.11: ADF Tau statistics for NO_x by station. Recurrent window starting at 5 years.

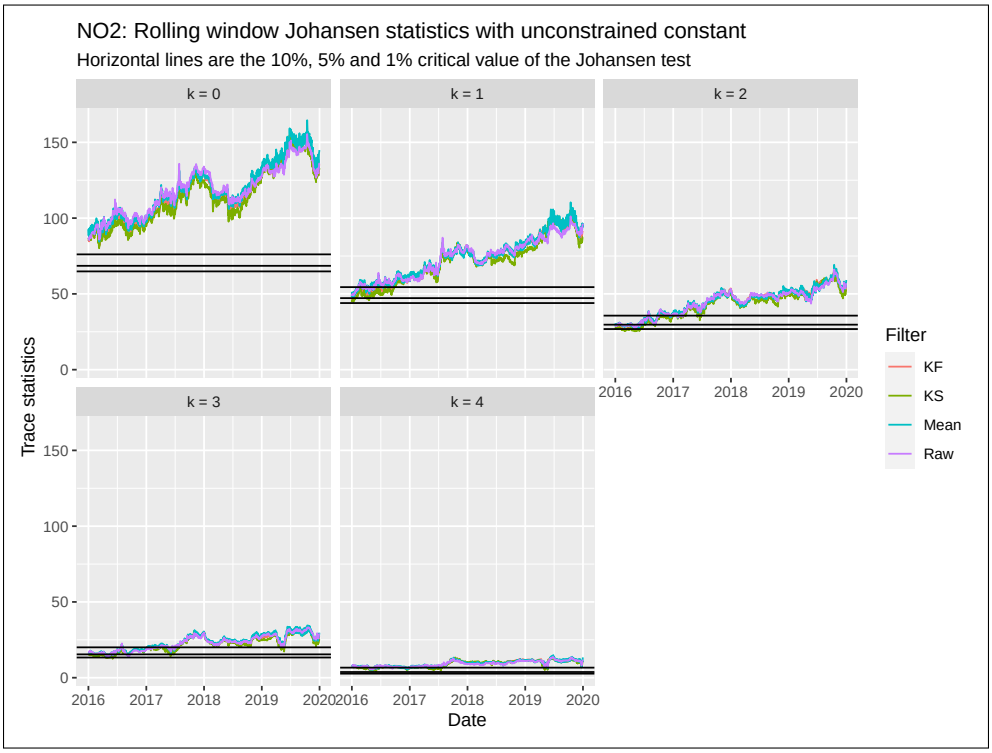


FIGURE 6.12: Johansen's trace statistics for NO_2 by the number of cointegration relationships. Rolling window of length five years.

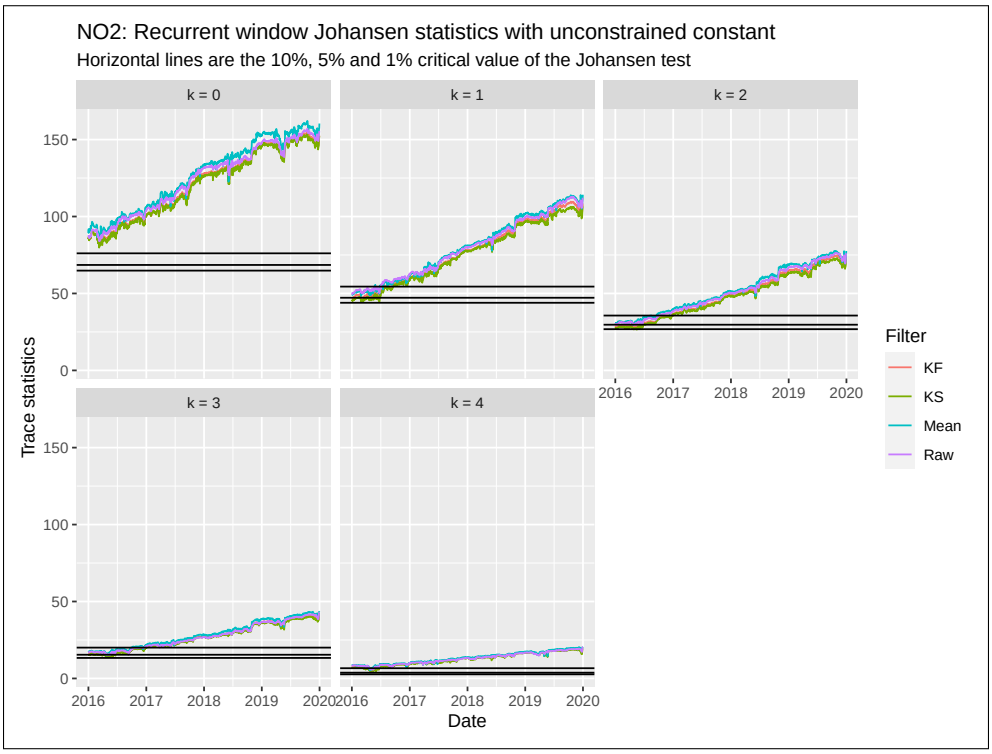


FIGURE 6.13: Johansen's trace statistics for NO_2 by the number of cointegration relationships. Recurrent window starting at five years.

FIGURE 6.14: Johansen's trace statistics for NO_x by the number of cointegration relationships. Rolling window of length five years.

FIGURE 6.15: Johansen's trace statistics for NO_x by the number of cointegration relationships. Recurrent window starting at five years.

Chapter 7

Conclusions and final remarks

This research thesis has discussed and analyzed various statistical methodologies aimed at analyzing air quality and available data on air pollution from both a theoretical and applied point of view. In particular, we focused on univariate and multivariate time series methodologies and models to analyze the concentrations of some of the primary air pollutants, such as fine particulate matter (PM_{10} and $PM_{2.5}$) and nitrogen oxides (NO_2 and NO_x).
In addition to some relevant theoretical results, several empirical applications based on real data were presented. These applications focused on the case study of the concentrations of airborne pollutants observed in the Italian region of Lombardy, one of the most polluted and economically advanced regions in Europe. Lombardy is considered one of Italy's economic leaders, but it suffers from decades of poor air quality and high environmental degradation. Many scientific studies show evidence of connections between high pollutant concentrations, damage to citizens' health, and environmental damage. In a socio-environmental context such as Lombardy, scientific research on environmental and economic issues has an important social function. One of the objectives of this thesis is to contribute to improving this severe situation by offering precise and adequate analytical tools to analyze the phenomenon of air pollution and enabling policymakers to make decisions based on reliable, scientifically supported facts.
In Chapter 2 we have contextualized Lombardy within the European and national forum of environmental protection institutions. We proposed a full overview of the European and Italian air quality monitoring system, namely EIONET, describing the air quality conditions in Europe and Lombardy. In particular, the chapter analyzed the case of Lombardy explaining its main geographical and physical characteristics, the composition of the economic and social network and the consequent effects on air pollution. The description also included descriptive and graphical statistics on the main air pollutants detected in the region and the main meteorological and weather factors, such as temperature, wind or rainfall. In the chapter, we have also described the source of the data used in the empirical applications and some simulation studies to support the methodology developed. Air quality and weather data have been collected from the Regional Agency for the Protection of the Environment (ARPA Lombardia), which currently manages and maintains the monitoring network and analyses all the region's environmental data. Finally, we discussed some side issues related to advanced techniques to analyze weather and air quality data properly, laying the foundations for possible extensions in that direction.
In Chapter 3, we discussed air quality time series analysis using a purely-time series approach based on Unobservable Component Models (UCM) in State Space form and bootstrap resampling algorithms. State-space models are well known and recognized in the statistical and econometric literature, which consider them adequate tools for the contemporary study of time series characteristics and the evaluation of the impact of policies and shocks on the variables of interest. We also showed empirically that bootstrap techniques in time series context make it possible to correctly identify any variations in the average levels of the concentrations, taking

into account the characteristics of positive asymmetry (right skewness) typical of pollutant concentrations. Among the others, the analysis's main contribution was to introduce a multi-stage model selection algorithm based on both fitting and forecasting performances criteria. The algorithm allows selecting the most relevant observable and non-observable features to explain the pollution concentrations before estimating the policy's unbiased effect. We presented the case study concerning the introduction of a new restricted traffic zone in Milan, called Area B, and its possible effect on concentrations of pollutants. The concentrations of nitrogen oxides (NO_x) and nitrogen dioxide (NO_2), which are mainly primary pollutants with high reactivity to changes in emission sources, have been considered as proxies of pollution emissions. We concluded that, although environmental protection policies are a fundamental step for sustainability improvement, in some cases, they may not be sufficient, or their implementation may be misleading. Indeed, compared to the common trend of the considered area, Milan air quality is improving slowly, and, in this sense, the first phase of Area B seems to harm air quality. Given the results obtained, we have highlighted some possible developments to improve on what has been done so far. We suggested to include in the models further data concerning the traffic of vehicle movements crossing the access points. Moreover, we suggested considering multivariate approaches, i.e. analyzing jointly multiple pollutants, and Spatio-temporal models to highlight spatial patterns effects. Finally, the extension to hourly data could analyze further intra-daily effects and explain the spatial dynamics related to traffic.
In Chapter 4, we have dealt with a very recent case study of great social and economic relevance in the environmental field. We have focused on the effect of restrictions on productive activities and mobility imposed by the Italian government, called in jargon lockdown, on concentrations in Lombardy during the COVID-19 pandemic in 2020. This is a preliminary analysis of the air quality in Lombardy during this event and is intended to give a first insight into what happened. Airborne pollutant concentrations time series were investigated through autoregressive with exogenous covariate models (ARX) models. Model tuning and variable selection were performed by a LASSO algorithm that accounts for seasonal and non-seasonal components. The lockdown event was defined through a dummy variable which expresses the change in average levels of pollution due to the restrictive measures after being depurated by the effect of local meteorology. This study is part of the recent literature investigating the relationship between virus containment policies and environmental and social effects in Europe and actively contributes to highlighting the positive effects generated by this critical situation. The results confirmed a generalized NO_2 reduction due to the lockdown. Compared to the observed average reduction, the lockdown effect was mitigated by the decreasing temporal trend and meteorology. The aggregation of the estimated impacts by province shows that all provinces experienced a relevant reduction, but the effect was stronger in large urban areas in the central belt. In contrast, the southern territories, characterized by flat rural landscape, showed a significant but weaker reduction of NO_2. Moreover, considering the type of area covered by the stations, the highest reductions in nitrogen dioxide levels were associated with the traffic control units.
In Chapter 5 we discussed a class of statistical methodologies known as *event studies*. Event studies allow assessing whether a particular event of interest has generated structural changes, e.g. changes in level or variability, in time series of interest. In particular, we discussed event studies methods applied to multivariate time series characterized by high cross-sectional dependence. The contribution of the study was twofold. On the one hand, we tried to extend the above methodology to non-typical areas, such as environmental statistics; indeed, as far as we know, there is no literature investigating the airborne pollutant time series through event studies methods. On the other hand, we proposed two novel nonparametric test statistics for the case of cumulated abnormal residuals (CAR), namely $P1$ and $P2$, based on the ranks of abnormal residuals and accounting explicitly for cross-sectional dependence. We demonstrated

the main statistical properties of the proposed statistics, that is the closed-formulation for their expected values and variances and their asymptotic distributions. The two test statistics have been compared in terms of size and power with other indicators introduced in the literature. The comparison was performed through a simulation study and an empirical application to real-world data. The simulative results showed that all the cross-sectionally adjusted statistics provide sizes very close to the nominal value, whereas not-adjusted statistics do not. The novel statistics provided excellent empirical findings, showing how they are valid tools for the case study of interest. Finally, all the statistics are applied to the empirical case of airborne pollutant concentrations registered in Lombardy during 2020. We considered as an event of interest the restrictions imposed on mobility and activities during the COVID-9 pandemic. The main interest was to state if the lockdown generated significant reductions in average concentrations of oxide (NO_2 and NO_x) or not. The results highlighted that dealing with highly cross-sectionally correlated series, to avoid corrections would generate biases in statistic estimates and over-rejection of the statistical hypotheses. Moreover, since all the statistics show negative empirical value and are all strongly statistically significant, we can conclude that the restrictions due to COVID-19 generated statistically significant reductions of the average oxide concentrations in Lombardy. These results are fully consistent with the findings reported in Chapter 4. When analysing in aggregate the coefficients associated with lockdown and estimated using ARX models for each of the available stations, it is clear that most of them are statistically significant and negative. All stations have a negative coefficient, but some of them are not significant. The consistency of these results shows that event studies methodologies can be integrated into a policy or event impact assessment context and provide preliminary information to support the estimation of effects. In conclusion, although event studies methods have been applied in a minimal fashion in fields other than financial econometrics, these methodologies reveal themselves to be essential and useful analytical tools for impact analysis of policies and events in natural and/or economic-social contexts. In particular, they can be used as preliminary tools to assess whether events are statistically significant or not and the sign of the occurred variation. Unfortunately, these methods do not provide an empirical estimate of the effect generated by the event of interest, but only the sign. This precludes their all-encompassing use as a substitute for other estimation techniques, but makes them a powerful complementary and integrative tool.

In Chapter 6, we investigated the effect of pre-filtering techniques on noisy time series. In particular, we have analyzed the issues of stationarity and cointegration of multivariate time series affected by high confounding noise. We have shown theoretically and through simulations the beneficial effect that pre-filtering can bring to the Dickey-Fuller stationarity and Johansen cointegration tests in contexts such as those just described. Three filters have been proposed: the Kalman filter, the Kalman smoother and aggregation employing k-period averaging. Our findings suggest that, in the presence of high confunding noise, all three filters can preserve the size of the two tests, but lose their effectiveness when looking at the size-adjusted power. In any case, they prove to be powerful tools to support the classical approaches and to reinforce or deny the suggestions provided by the raw data. We have also decided to show an empirical application of pre-filtering by analyzing the case of oxide concentrations (NO_2 and NO_x) over the city of Milan. The decision to work with airborne concentrations is driven by the fact that pollutants, deriving from very complex atmospheric physico-chemical processes involving many natural and anthropogenic factors, are naturally characterized by strong variability and therefore by confounding noise overlapping the real data signal. The application showed that pre-filtering could be a useful tool for at least two reasons: the tests performed on the pre-filtered data support the hypothesis identified by the raw data, i.e. the univariate time series of concentrations are stationary. Concerning cointegration, on the other hand, it appears that the number of relationships identified varies according to the chosen temporal sample.

The study offers some important development opportunities. It would be interesting to develop other theoretical and simulative studies using structural models for time series (BSM), i.e. models which consider both the long-term (trend) and seasonal components. Moreover, there are many empirical contexts in which data show high confusing noise that buries the real signal and that if not adequately considered can lead to immediate wrong conclusions about the appropriate models to describe the phenomenon. Applying pre-filtering methodologies in such situations could strengthen the scientific validity of applied research and, in the most extreme case, lead to the revision of some hypotheses that are now considered consolidated in the literature.
As mentioned in the introductory chapter, the most crucial objective we set for this thesis was to develop an integrated approach (pipeline) in which various methodologies are nested and accumulated to ensure consistent and robust results. The methodologies presented here cover the main phases of a statistical evaluation of public policies or event analysis: identification of the characteristics of the phenomenon (stationarity and cointegration - Chapter 6), preliminary assessment of the effects generated by the event both at the individual and aggregate level (event studies - Chapter 5) and impact estimation. In this regard, it should be noted that the proposed models, i.e. regression models with ARIMA and ARMAX errors and latent component models, differ profoundly in the way they represent the historical series of interest but are partially united by the method of estimation (maximum likelihood). Depending on the research's objectives, the combination of these methods guarantees a complete and unbiased statistical evaluation of the phenomena. For example, in Chapters 3 and 6, we discussed the case study of oxide concentrations in Milan, Italy. Chapter 3 focused on the phase of model selection and estimation of the effects generated by introducing a limited traffic zone within the municipal boundaries. In chapter 6, we studied the stationarity and the long-term relationship of the same time series. In this case, the study of stationarity using linear filters confirms that concentrations in the city are stationary and do not follow stochastic trends. Although marginal for policy analysis, this information strengthens the results obtained in the third chapter, which in turn are based on the hypothesis of stationarity of the data. In contrast, considering the case study related to the effects of lockdown restrictions for COVID-19 on Lombardy concentrations, the fourth and fifth chapters' results appear complementary and consistent with each other. The restrictions led to large reductions in average concentrations across much of the region, with a few identified and justified exceptions. As is well known, event studies indicate the sign of the variation generated by the event and its statistical significance but do not quantify the variations. These considerations allow us to conclude positively about the ideas that motivated and guided this doctoral thesis. Individually, every statistical technique has a precise purpose and well-defined functioning. The result does not necessarily lie in its independent application but in the researcher's ability to integrate it and communicate with other methods to achieve a more complex but complete goal.

Appendix A

Appendix A: Proof of Theorem 1

First, notice, that n_1 is a finite fixed integer number, so that $n_0 \to \infty \Leftrightarrow n \to \infty$. Let $F_n(x) = n^{-1} \sum_{t \in \Omega} \mathbb{1}_{(-\infty,x]}(\varepsilon_t)$ be the empirical distribution function of ε_t. By standard empirical process theory, $F_n(x) = F(x) + O_p(n^{-1/2})$ uniformly in x.
By Taylor expansion of $\hat{\sigma}_i^{-1}$ in a neighbourhood of σ_i,

$$\begin{aligned}
\frac{AR_{it}}{\hat{\sigma}_i} &= \frac{\varepsilon_{it}}{\hat{\sigma}_i} - \frac{\boldsymbol{x}_{it}^\top(\hat{\beta}_i - \beta)}{\hat{\sigma}_i} \\
&= \frac{\varepsilon_{it}}{\sigma_i} + \frac{\varepsilon_{it}}{\sigma_i^2}(\hat{\sigma}_i - \sigma_i) - \frac{\boldsymbol{x}_{it}^\top(\hat{\beta}_i - \beta)}{\hat{\sigma}_i} + o_p(n^{-1/2}) \\
&= \frac{\varepsilon_{it}}{\sigma_i} + \delta_{ni},
\end{aligned}$$

where δ_{ni} is equal to the last three addends of the second line of the above equations and, by assumptions 3., 4., 5., $\delta_{ni} = O_p(n^{-1/2})$.
Then,

$$SAR_t = \sum_{i=i}^{m} \frac{AR_{it}}{\hat{\sigma}_i} = \sum_{i=i}^{m} \left(\frac{\varepsilon_{it}}{\sigma_i} + \delta_{ni} \right) = \varepsilon_t + \delta_n,$$

where $\delta_n = \sum_{i=1}^{m} \delta_{ni} = O_p(n^{-1/2})$. By Taylor's theorem, for some $h_{\varepsilon,\delta} \in [0,1]$

$$F(SAR_t) - F(\varepsilon_t) - f(\varepsilon_t)\delta_n = f(\varepsilon_t + h_{\varepsilon,\delta}\delta_n) - f(\varepsilon_t).$$

For any ε and δ, by applying the Cauchy-Schwarz inequality we can write

$$\begin{aligned}
|f(\varepsilon + h\delta) - f(\varepsilon)| &= \left| \int_0^{h\delta} f'(\varepsilon + u)\,\mathrm{d}u \right| \leq \int_0^{\delta} \left| f'(\varepsilon + u) \right| \mathrm{d}u \\
&= \int_0^{\delta} \frac{|f'(\varepsilon + u)|}{\sqrt{f(\varepsilon + u)}} \sqrt{f(\varepsilon + u)}\,\mathrm{d}u \\
&\leq \left[\int_0^{\delta} \left(\frac{f'(\varepsilon + u)}{f(\varepsilon + u)} \right)^2 f(\varepsilon + u)\,\mathrm{d}u \cdot \int_0^{\delta} f(\varepsilon + u)\,\mathrm{d}u \right]^{\frac{1}{2}} \\
&\leq \left[\int_0^{\delta} \left(\frac{f'(\varepsilon + u)}{f(\varepsilon + u)} \right)^2 f(\varepsilon + u)\,\mathrm{d}u \cdot [F(\varepsilon + \delta) - F(\varepsilon)] \right]^{\frac{1}{2}} \\
&\leq \sqrt{I_0(f)} \cdot \sqrt{F(\varepsilon + \delta) - F(\varepsilon)},
\end{aligned}$$

which by assumptions 1. and 2. is bounded. Thus, we have

$$F(SAR_t) = F(\varepsilon_t) + O_p(n^{-1/2}).$$

Putting the pieces together we have

$$U_{nt} = \frac{R_{nt}}{n+1} = \frac{n+1}{n} F_n(SAR_t) = F(SAR_t) + O_p(n^{-1/2}) = F(\varepsilon_t) + O_p(n^{-1/2}),$$

which converges in probability to a uniform random variable, say U_t. Thus, by the probability integral transform, as $n_0 \to \infty$

$$\Phi^{-1}(U_{nt}) \xrightarrow{p} \Phi^{-1}(U_t) \sim N(0,1).$$

Moreover, as n_0 diverges ranks become independent and, thus, for fixed n_1

$$\sum_{t \in \Omega_1} \Phi^{-1}(U_{nt}) \xrightarrow{p} \sum_{t \in \Omega_1} \Phi^{-1}(U_t) \quad \text{as } n_0 \to \infty$$

which is the sum of n_1 independent standard normal random variables.

Appendix B

Appendix B: Proof of Theorem 2

First, notice, that n_1 is a finite fixed integer number, so that $n_0 \to \infty \Leftrightarrow n \to \infty$. Let $F_n^{(i)}(x) = n^{-1} \sum_{t \in \Omega_i} \mathbb{1}_{(-\infty,x]}(\varepsilon_{it})$ be the empirical distribution function of ε_{it}. By standard empirical process theory, $F_n^{(i)}(x) = F^{(i)}(x) + O_p(n^{-1/2})$ uniformly in x.
Under the assumptions of the theorem,

$$AR_{it} = \varepsilon_{it} - \boldsymbol{x}_{it}^{\top}(\hat{\beta}_i - \beta) = \varepsilon_{it} + \delta_{ni},$$

where $\delta_{ni} = O_p(n^{-1/2})$.
Analogously to the previous proof, by Taylor's theorem, for some $h_{\varepsilon,\delta} \in [0,1]$

$$F^{(i)}(AR_{it}) - F^{(i)}(\varepsilon_{it}) - f(\varepsilon_{it})\delta_{in} = f^{(i)}(\varepsilon_{it} + h_{\varepsilon,\delta}\delta_{in}) - f^{(i)}(\varepsilon_i)$$

and

$$\left| f^{(i)}(\varepsilon_{it} + h\delta) - f(\varepsilon_{it}) \right| \leq \sqrt{I_0(f^{(i)})} \cdot \sqrt{F^{(i)}(\varepsilon_{it} + \delta) - F^{(i)}(\varepsilon_{it})},$$

which by assumptions 1. and 2. is bounded. Thus, we have

$$F^{(i)}(AR_{it}) = F^{(i)}(\varepsilon_{it}) + O_p(n^{-1/2}),$$

where $F^{(i)}(\varepsilon_{it})$ are uniform random variable that are i.i.d. with respect to the time index t.
Consequently, we have

$$U_{nit} = \frac{R_{nit}}{n+1} = \frac{n+1}{n} F_n^{(i)}(\varepsilon_{it}) = F^{(i)}(\varepsilon_{it}) + O_p(n^{-1/2})$$

and, by applying Taylor's theorem to the normal quantile function $\Phi^{-1}()$,

$$Z_{nit} := \Phi^{-1}(U_{nit}) = \Phi^{-1}\left(F^{(i)}(\varepsilon_{it})\right) + O_p(n^{-1/2}).$$

Thus, since $F^{(i)}(\varepsilon_{it})$ is a uniform random variable,

$$Z_{nit} \xrightarrow{p} Z_{it} := \Phi^{-1}\left(F^{(i)}(\varepsilon_{it})\right) \sim N(0,1).$$

Taking the sum with respect to the index i, we get

$$\hat{T}_{mt} := m^{-1/2} \sum_{i=1}^{m} Z_{nit} = m^{-1/2} \sum_{i=1}^{m} Z_{it} + O_p\left((m/n)^{1/2}\right) = T_{mt} + O_p\left((m/n)^{1/2}\right),$$

and, by assumption 4., if $m, n_0 \to \infty$ e $m/n_0 \to 0$

$$\hat{T}_{mt} \xrightarrow{p} T_{mt} \sim N(0, \rho).$$

Finally, since all moments of $\hat{T}_{mt}$ and T_{mt} exist and $\{Z_{mit}\}_{i=1,\dots,m}$ is i.i.d. with respect to the index t, we can consistently estimate ρ using the law of large number: as n_0 diverges,

$$\hat{\rho}_{ij} = \frac{1}{n_0}\sum_{t=1}^{n_0} Z_{nit}Z_{njt} = \frac{1}{n_0}\sum_{t=1}^{n_0} Z_{it}Z_{jt} + O_p(m/n) \xrightarrow{p} \rho_{ij} := E(Z_{it}Z_{jt}),$$

and the classical CLT implies that $\hat{\rho}_{ij} - \rho_{ij} = O_p(n_0^{-1/2})$. Now, by taking the average correlation:

$$\begin{aligned}
\hat{\bar{\rho}} &:= \frac{1}{m(m-1)}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\hat{\rho}_{ij} \\
&= \frac{1}{m(m-1)}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}(\rho_{ij} + (\hat{\rho}_{ij} - \rho_{ij})) \\
&= \bar{\rho} + O_p(m^{-1}n_0^{-1/2}),
\end{aligned}$$

because the bias of $\hat{\rho}_{ij}$ as estimator of ρ_{ij} is of order n_0^{-1}. By assumption, as m diverges $m^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}\rho_{ij} \to \rho < \infty$, and, thus, for $n_0, m \to \infty$, $m/n \to 0$, we have

$$\begin{aligned}
P_t &= \frac{\sum_{i=1}^{m} Z_{nit}}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{\rho}_{ij}}} \\
&= \frac{m^{-1/2}\sum_{i=1}^{m} Z_{nit}}{\sqrt{m^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{\rho}_{ij}}} \\
&= \frac{m^{-1/2}\sum_{i=1}^{m} Z_{nit}}{\sqrt{1 + (m-1)\hat{\bar{\rho}}}} \\
&= \frac{m^{-1/2}\sum_{i=1}^{m} Z_{nit}}{\sqrt{1 + (m-1)\bar{\rho} + O_p(m^{-1}n_0^{-1/2})}}. \xrightarrow{d} N(0,1).
\end{aligned}$$

Since P_t is asymptotically i.i.d., while n_1 is finite and fixed, the result of the theorem is easily obtained.

Appendix C

Appendix C: Proof of Theorem 3 and Corollary 1

Let y_t be a random walk plus noise with $t \in \{1, 2, 3, \dots, \}$ with given starting value x_0:

$$\begin{aligned} y_t &= x_t + \varepsilon_t, & \varepsilon_t &\sim WN(\sigma_\varepsilon^2) \\ x_t &= x_{t-1} + \eta_t, & \eta_t &\sim WNa_\eta^2). \end{aligned}$$

We can rewrite the random walk as

$$x_t = x_{t-m} + \sum_{j=0}^{m-1} \eta_{t-j},$$

and taking the average over m subsequent observations, we obtain

$$\frac{1}{m}\sum_{i=0}^{m-1} x_{t-i} = \frac{1}{m}\sum_{i=0}^{m-1} x_{t-m} + \frac{1}{m}\sum_{i=0}^{m-1}\sum_{j=0}^{m-1} \eta_{t-j-i},$$

which we can restate as

$$\bar{x}_t = \bar{x}_{t-m} + \bar{\eta}_t,$$

where $\bar{x}_t = \frac{1}{m}\sum_{i=0}^{n-1} x_{t-i}$ and

$$\begin{aligned} \bar{\eta}_t = \frac{1}{m}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} \eta_{t-j-i} = &\frac{1}{m}\big[\eta_t + 2\eta_{t-1} + 3\eta_{t-2} + \ldots + m\eta_{t-m+1} + \\ &(m-1)\eta_{t-m} + (m-2)\eta_{t-m-1} + \ldots + 2\eta_{t-2m+1} + \eta_{t-2m+2}\big]. \end{aligned}$$

Now, let us consider the process for $t \in \mathcal{T} := \{m, 2m, 3m, \ldots\}$. Using the above formulae, it is simple to derive the autocovariance function (ACF) of $\bar{\eta}_t$ over the set $\mathcal{T}$:

$$E\bar{\eta}_t\bar{\eta}_{t-km} = \begin{cases} \left[\frac{(m-1)(2m-1)}{3m} + 1\right]\sigma_\eta^2 & \text{for } k = 0, \\ \frac{(m-1)(m+1)}{6m}\sigma_\eta^2 & \text{for } |k| = 1, \\ 0 & \text{for } |k| = 2, 3, \ldots \end{cases},$$

which is the ACF of a MA(1) process. Consequently, the process

$$\bar{y}_t - \bar{y}_{t-m} = \bar{\eta}_t + \bar{\varepsilon}_t - \bar{\varepsilon}_{t-1},$$

with $\bar{\varepsilon}_t := \frac{1}{m}\sum_{i=0}^{m-1}\varepsilon_{t-i}$, is also MA(1) with ACF

$$\begin{cases} \left[\frac{(m-1)(2m-1)}{3m}+1\right]\sigma_\eta^2+\frac{2}{m}\sigma_\varepsilon^2 & \text{for } k=0, \\ \frac{(m-1)(m+1)}{6m}\sigma_\eta^2+\frac{1}{m}\sigma_\varepsilon^2 & \text{for } |k|=1, \\ 0 & \text{for } |k|=2,3,\dots \end{cases}.$$

Therefore, its order-1 autocorrelation over $\mathcal{T}$ is

$$\rho := \frac{\frac{(m-1)(m+1)}{6m}\sigma_\eta^2+\frac{1}{m}\sigma_\varepsilon^2}{\left[\frac{(m-1)(2m-1)}{3m}+1\right]\sigma_\eta^2+\frac{2}{m}\sigma_\varepsilon^2} = \frac{\frac{(m-1)(m+1)}{6m}\lambda-\frac{1}{m}}{\left[\frac{(m-1)(2m-1)}{3m}+1\right]\lambda+\frac{2}{m}},$$

where $\lambda := \sigma_\eta^2/\sigma_\varepsilon^2$ is the signal-to-noise ratio.
We can conclude that, over the set $\mathcal{T}$, $\bar{y}_t$ is ARIMA(0, 1, 1) and the MA coefficient can be easily obtained from ρ using the formula in Corollary 1.

Appendix D

Appendix D: Proof of Theorem 4

By assumption y_t is defined by the ARIMA(0, 1, 1) process in equation (6.3). Thus,

$$\begin{aligned}\tilde{\mu}_t &= \gamma y_t + (1-\gamma)\tilde{\mu}_{t-1} \\ \big(1-(1-\gamma)L\big)\tilde{\mu}_t &= \gamma y_t \\ \big(1-(1-\gamma)L\big)(1-L)\tilde{\mu}_t &= \gamma(1-\theta L)\zeta_t.\end{aligned}$$

It is straightforward to check that $\theta = 1-\gamma$, which implies that the operator $\big(1-(1-\gamma)L\big)$ annihilates $(1-\theta L)$ and

$$(1-L)\tilde{\mu}_t = \gamma\zeta_t,$$

defines a random walk.
In order to obtain the approximate smoother $\hat{\mu}_t$ one has to pass a backward autoregressive filter on $\tilde{\mu}_t$. The second-order properties of the resulting process are the same if the filter is run backward or forward and, thus, the process $\hat{\mu}_t$ has the same autocorrelation function as

$$\big(1-(1-\gamma)L\big)(1-L)\tilde{\mu}_t = \gamma\zeta_t,$$

which defines an $ARIMA(1,1,0)$ process with autoregressive coefficient $\phi = 1-\gamma$.

Appendix E

Appendix E: simulated rejection rates

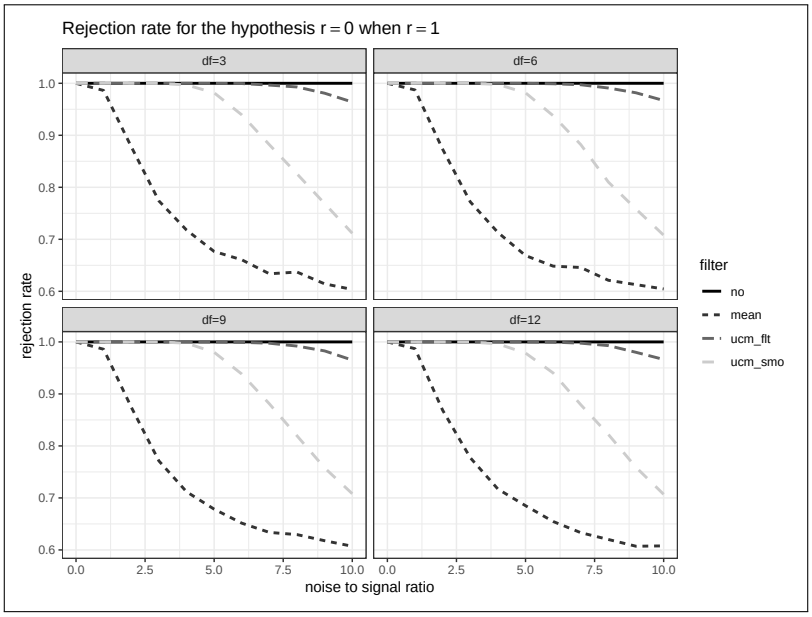


FIGURE E.1: Rejection Rates coint1_n_rejections_r0.

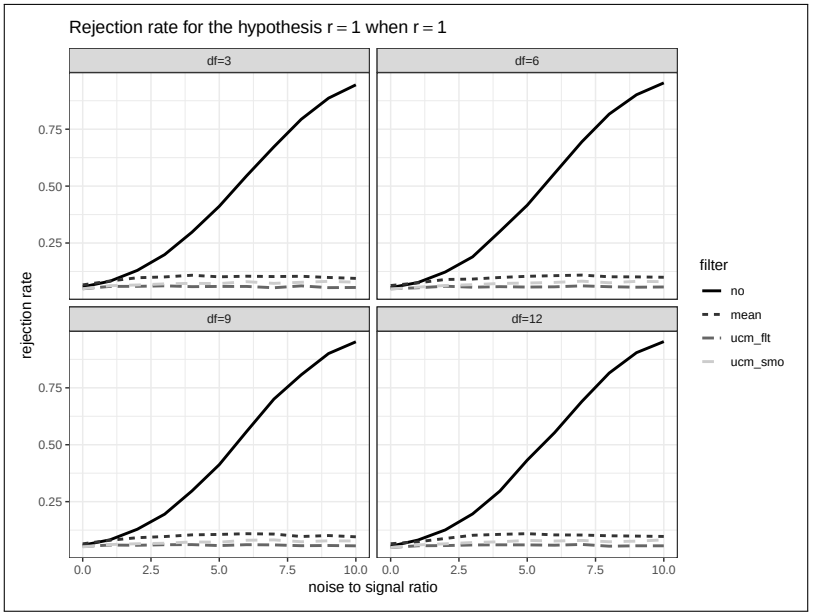


FIGURE E.2: Rejection Rates coint1_n_rejections_r1.

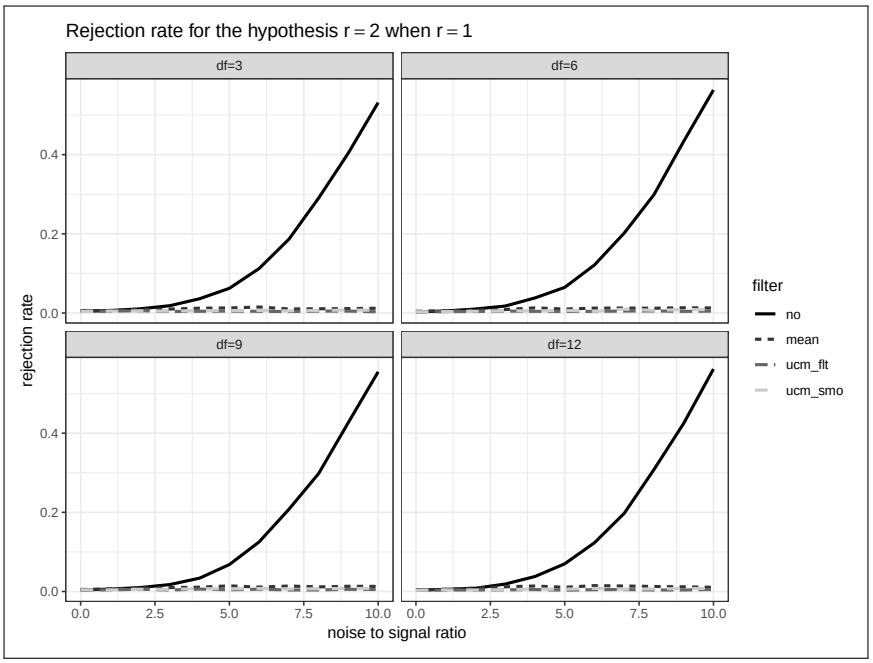


FIGURE E.3: Rejection Rates coint1_n_rejections_r2.

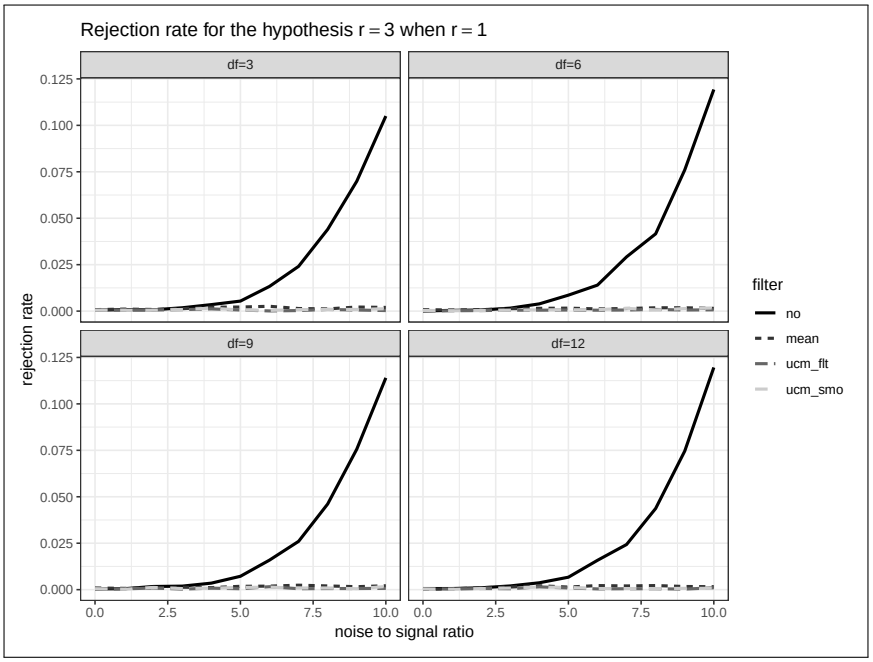


FIGURE E.4: Rejection Rates coint1_n_rejections_r3.

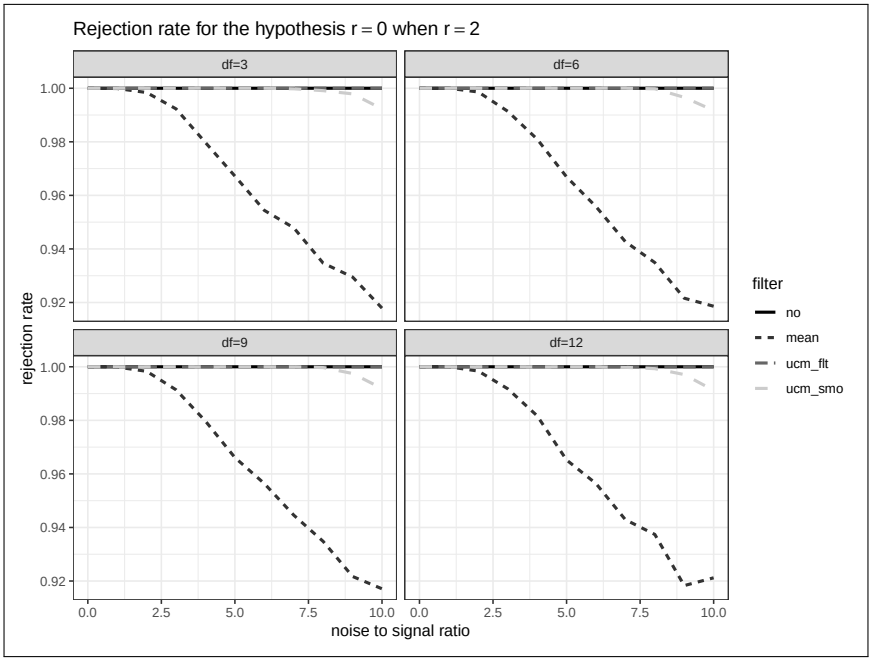


FIGURE E.5: Rejection Rates coint2_n_rejections_r0.

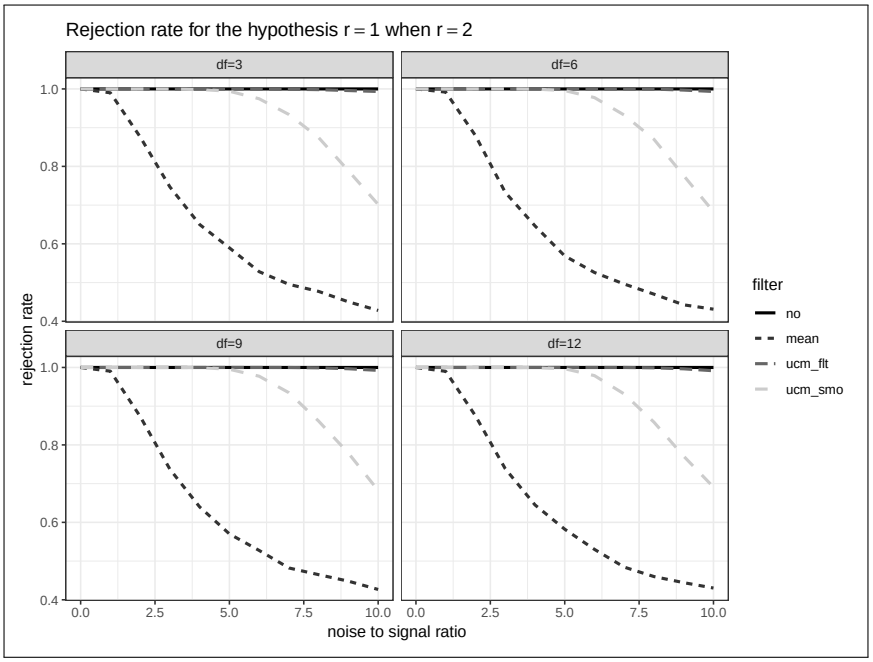


FIGURE E.6: Rejection Rates coint2_n_rejections_r1.

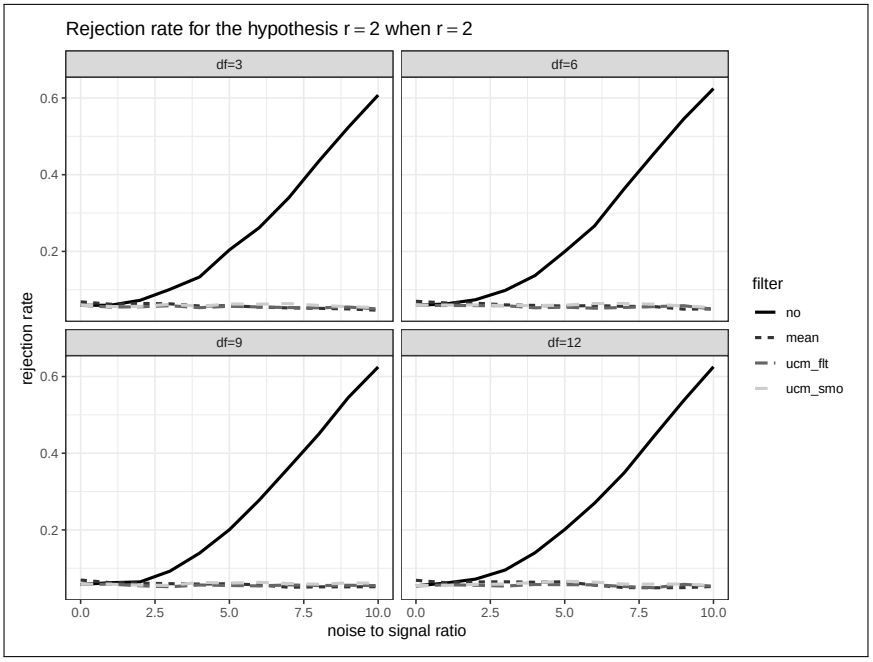


FIGURE E.7: Rejection Rates coint2_n_rejections_r2.

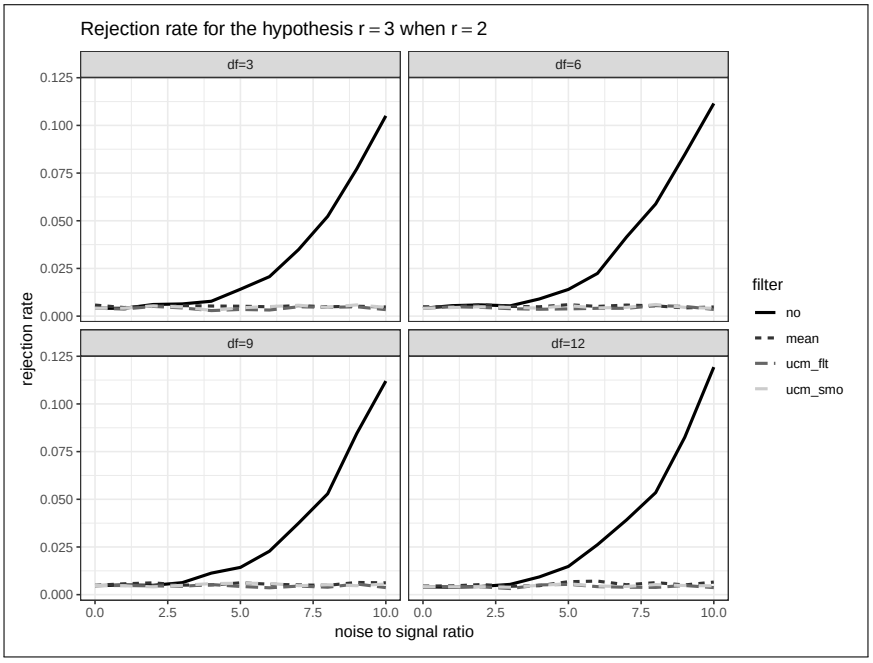


FIGURE E.8: Rejection Rates coint2_n_rejections_r3.

Acknowledgements

Bibliography

[1] A. Adhikari and J. Yin. "Short-Term Effects of Ambient Ozone, PM2.5, and Meteorological Factors on COVID-19 Confirmed Cases and Deaths in Queens, New York https://www.mdpi.com/1660-4601/17/11/4047". In: *International Journal of Environmental Research and Public Health* 17.11 (2020), p. 4047.

[2] U. E. P. Agency. "Technical Assistance Document for the Reporting of Daily Air-Quality—The Air-Quality Index (AQI). 2018, Accessed 09-15-2020." In: (2018).

[3] V. Agresti, A. Balzarini, R. Bonanno, E. Collino, F. Colzi, Lacavalla, M. Pirovano, G. Riva, A. Toppetti, F. Riva, and A. Piccoli. "Gli effetti del lockdown sulla qualità dell'aria a Milano e in Lombardia". In: (2020).

[4] A. Ataullah, X. Song, and M. Tippett. "A modified Corrado test for assessing abnormal security returns". In: *The European Journal of Finance* 17.7 (2011), pp. 589–601.

[5] M. Baccini, A. Biggeri, P. Grillo, D. Consonni, and P. A. Bertazzi. "Health impact assessment of fine particle pollution at the regional level". In: *American journal of epidemiology* 174.12 (2011), pp. 1396–1405.

[6] M. Baccini, L. Grisotto, D. Catelan, D. Consonni, P. A. Bertazzi, and A. Biggeri. "Commuting-adjusted short-term health impact assessment of airborne fine particles with uncertainty quantification via Monte Carlo simulation". In: *Environmental health perspectives* 123.1 (2015), pp. 27–33.

[7] L. Bai, J. Wang, X. Ma, and H. Lu. "Air pollution forecasts: An overview". In: *International journal of environmental research and public health* 15.4 (2018), p. 780.

[8] J. M. Baldasano. "COVID-19 lockdown effects on air quality by NO2 in the cities of Barcelona and Madrid (Spain)". In: *Science of The Total Environment* 741 (2020), p. 140353.

[9] R. Ball and P. Brown. "An empirical evaluation of accounting income numbers". In: *Journal of accounting research* (1968), pp. 159–178.

[10] J. Bartholdy, D. Olson, and P. Peare. "Conducting Event Studies on a Small Stock Exchange". In: *The European Journal of Finance* 13.3 (2007), pp. 227–252.

[11] T. Belyakhina, A. Zaslavsky, K. Mitra, S. Saguna, and P. P. Jayaraman. "DisCPAQ: Distributed context acquisition and reasoning for personalized indoor air quality monitoring in IoT-based systems". In: *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*. Springer, 2017, pp. 75–86.

[12] C. Bergmeir, R. J. Hyndman, and B. Koo. "A note on the validity of cross-validation for evaluating time series prediction". In: *Monash University Department of Econometrics and Business Statistics Working Paper* 10 (2015), p. 15.

[13] C. Bergmeir, R. J. Hyndman, and B. Koo. "A note on the validity of cross-validation for evaluating autoregressive time series prediction". In: *Computational Statistics & Data Analysis* 120 (2018), pp. 70–83.

[14] K. N. Berk. "A Central Limit Theorem for m-Dependent Random Variables with Unbounded m". In: *The Annals of Probability* 1.2 (1973), pp. 352–354.

[15] J. D. Berman and K. Ebisu. "Changes in U.S. air pollution during the COVID-19 pandemic". In: *Science of The Total Environment* 739 (2020), p. 139864.

[16] S. Bhagat and R. Romano. "Event studies and the law: Part I: Technique and corporate litigation". In: *American Law and Economics Review* 4.1 (2002), pp. 141–168.

[17] S. Bhagat and R. Romano. "Event studies and the law: Part II: Empirical studies of corporate law". In: *American Law and Economics Review* 4.2 (2002), pp. 380–423.

[18] S. Bittanti. *Model identification and data analysis*. Wiley Online Library, 2019.

[19] M. Blangiardo, M. Cameletti, M. Pirani, G. Corsetti, M. Battaglini, and G. Baio. "Estimating weekly excess mortality at sub-national level in Italy during the COVID-19 pandemic". In: *PLOS ONE* 15.10 (2020), e0240286.

[20] E. Boehmer, J. Musumeci, and A. B. Poulsen. "Event-study methodology under conditions of event-induced variance". In: *Journal of Financial Economics* 30.2 (1991), pp. 253–272.

[21] E. Bontempi. "First data analysis about possible COVID-19 virus airborne diffusion due to air particulate matter (PM): The case of Lombardy (Italy)". In: *Environmental Research* 186 (2020), p. 109639.

[22] B. Bosco, L. Parisio, M. Pelagatti, and F. Baldi. "Long-run relations in European electricity prices". In: *Journal of Applied Econometrics* 25.5 (2010), pp. 805–832.

[23] S. Brown and J. Warner. "Using daily stock returns: the case of event studies". In: *Journal of Financial Economics* 14 (1985), pp. 3–31.

[24] S. J. Brown and J. B. Warner. "Measuring security price performance". In: *Journal of financial economics* 8.3 (1980), pp. 205–258.

[25] G. H. Brundtland. "Report of the World Commission on Environment and Development: Our Common Future. United Nations General Assembly". In: (1987).

[26] C. Calculli, A. Fassò, F. Finazzi, A. Pollice, and A. Turnone. "Maximum likelihood estimation of the multivariate hidden dynamic geostatistical model with application to air quality in Apulia, Italy". In: *Environmetrics* 26.6 (2015), pp. 406–417.

[27] M. Cameletti. "The Effect of Corona Virus Lockdown on Air Pollution: Evidence from the City of Brescia in Lombardia Region (Italy)". In: *Atmospheric Environment* 239 (2020), p. 117794.

[28] M. Cameletti, V. Gómez-Rubio, and M. Blangiardo. "Bayesian modelling for spatially misaligned health and air pollution data through the INLA-SPDE approach". In: *Spatial Statistics* 31 (2019), p. 100353.

[29] C. J. Campbell and C. E. Wasley. "Measuring security price performance using daily NASDAQ returns". In: *Journal of Financial Economics* 33.1 (1992), pp. 73–92.

[30] J. Y. Campbell, A. W. Lo, and A. C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1996.

[31] J. Y. Campbell, A. W. Lo, A. C. MacKinlay, and R. F. Whitelaw. "The econometrics of financial markets". In: *Macroeconomic Dynamics* 2.4 (1998), pp. 559–562.

[32] C. Carnevale, G. Finzi, E. Pisoni, M. Volta, G. Guariso, R. Gianfreda, G. Maffeis, P. Thunis, L. White, and G. Triacchini. "An integrated assessment tool to define effective air quality policies at regional scale". In: *Environmental Modelling & Software* 38 (2012), pp. 306–315.

[33] M. Carugno, D. Consonni, P. A. Bertazzi, A. Biggeri, and M. Baccini. "Temporal trends of PM10 and its impact on mortality in Lombardy, Italy". In: *Environmental Pollution* 227 (2017), pp. 280–286.

[34] M. Carugno, D. Consonni, G. Randi, D. Catelan, L. Grisotto, P. A. Bertazzi, A. Biggeri, and M. Baccini. "Air pollution exposure, cause-specific deaths and hospitalizations in a highly polluted Italian region". In: *Environmental research* 147 (2016), pp. 415–424.

[35] S. Caserini, A. Fraccaroli, A. M. Monguzzi, M. Moretti, A. Giudici, E. Angelino, G. Fossati, and G Gurrieri. "A detailed Emission Inventory for air quality planning at local scale: the Lombardy (Italy) experience". In: *13th Emission Inventory Conference of the US Environmental Protection Agency, Clearwater, FL*. 2004.

[36] A. B. Chelani. "Statistical characteristics of ambient PM2.5 concentration at a traffic site in Delhi: source identification using persistence analysis and nonparametric wind regression". In: *Aerosol and Air Quality Research* 13.6 (2013), pp. 1768–1778.

[37] L. Chen, B. Guo, J. Huang, J. He, H. Wang, S. Zhang, and S. X. Chen. "Assessing air-quality in Beijing-Tianjin-Hebei region: The method and mixed tales of PM2. 5 and O3". In: *Atmospheric environment* 193 (2018), pp. 290–301.

[38] M. Coccia. "Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID". In: *Science of The Total Environment* 729 (2020), p. 138474.

[39] E. S. Coker, L. Cavalli, E. Fabrizi, G. Guastella, E. Lippo, M. L. Parisi, N. Pontarollo, M. Rizzati, A. Varacca, and S. Vergalli. "The effects of air pollution on COVID-19 related mortality in northern Italy". In: *Environmental and Resource Economics* 76.4 (2020), pp. 611–634.

[40] M. C. Collivignarelli, A. Abbà, G. Bertanza, R. Pedrazzani, P. Ricciardi, and M. Carnevale Miino. "Lockdown for CoViD-2019 in Milan: What are the effects on air quality?" In: *Science of The Total Environment* 732 (2020), p. 139280.

[41] P. Connerton, J. Vicente de Assunção, R. Maura de Miranda, A. Dorothée Slovic, P. José Pérez-Martínez, and H. Ribeiro. "Air Quality during COVID-19 in Four Megacities: Lessons and Challenges for Public Health". In: *International Journal of Environmental Research and Public Health* 17.14 (2020), p. 5067.

[42] C. Copat, A. Cristaldi, M. Fiore, A. Grasso, P. Zuccarello, S. S. Signorelli, G. O. Conti, and M. Ferrante. "The role of air pollution (PM and NO2) in COVID-19 spread and lethality: A systematic review". In: *Environmental Research* 191 (2020), p. 110129.

[43] C. J. Corrado. "A nonparametric test for abnormal security-price performance in event studies". In: *Journal of Financial Economics* 23 (1989), pp. 385–395.

[44] C. J. Corrado and C. Truong. "Conducting event studies with Asia-Pacific security market data". In: *Pacific-Basin Finance Journal* 16 (2008), pp. 493–521.

[45] C. J. Corrado and T. L. Zivney. "The specification and power of the sign test in event study hypothesis tests using daily stock returns". In: *Journal of Financial and Quantitative Analysis* 27.3 (1992), pp. 465–478.

[46] C. J. Corrado. "Event studies: a methodology review". In: *Accounting & Finance* 51.1 (2011), pp. 207–234.

[47] A. R. Cowan. "Nonparametric event study tests". In: *Review of Quantitative Finance and Accounting* 2 (1992), pp. 343–358.

[48] N. Crato and H. M. Taylor. "Stationary persistent time series misspecified as nonstationary arima". In: *Statistical Papers* 37.3 (1996), p. 215.

[49] C. Cuvelier, P. Thunis, R. Vautard, M. Amann, B. Bessagnet, M. Bedogni, R. Berkowicz, J. Brandt, F. Brocheton, P. Builtjes, C. Carnavale, A. Coppalle, B. Denby, J. Douros, A. Graf, O. Hellmuth, A. Hodzic, C. Honoré, J. Jonson, A. Kerschbaumer, F. de Leeuw, E. Minguzzi, N. Moussiopoulos, C. Pertot, V. H. Peuch, G. Pirovano, L. Rouil, F. Sauter, M. Schaap, R. Stern, L. Tarrason, E. Vignati, M. Volta, L. White, P. Wind, and A. Zuber. "CityDelta: A model intercomparison study to explore the impact of emission reductions in European cities in 2010". In: *Atmospheric Environment* 41.1 (2007), pp. 189–207.

[50] A. C. Davison and D. V. Hinkley. *Bootstrap Methods and their Application*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 1997.

[51] M. Di Marzio, A. Panzera, and C. C. Taylor. "Local polynomial regression for circular predictors". In: *Statistics & Probability Letters* 79.19 (2009), pp. 2066–2075.

[52] P. H. Diananda. "The central limit theorem for m-dependent variables". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 51.1 (1955), pp. 92–95.

[53] P. Dodd and J. B. Warner. "On corporate governance: a study of proxy contests". In: *Journal of Financial Economics 11(1), 401.* 11.1 (1983), pp. 401–438.

[54] J. L. Domingo, M. Marquès, and J. Rovira. "Influence of airborne transmission of SARS-CoV-2 on COVID-19 pandemic. A review". In: *Environmental Research* 188 (2020), p. 109861.

[55] A. Donnelly, B. Misstear, and B. Broderick. "Application of nonparametric regression methods to study the relationship between NO2 concentrations and local wind direction and speed at background sites". In: *Science of the Total Environment* 409.6 (2011), pp. 1134–1144.

[56] J. A. Doornik and H. Hansen. "An Omnibus Test for Univariate and Multivariate Normality". In: *Oxford Bulletin of Economics and Statistics* 70 (2008), pp. 927–939.

[57] J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, 2001.

[58] J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, 2012.

[59] A. Dvoretsky. "Central limit theorems for dependent random variables". In: *Proceedings of the Sixth Berkeley Symposimum on Mathematical Statistics and Probability*. University of California Press, 1972, pp. 513–535.

[60] A. M. Ecosystem. "Guide to the millennium assessment reports". In: *www. millenniu massessment. org/en/index. aspx (accessed June 26, 2008)* (2013).

[61] G. Elliott, T. J. Rothenberg, and J. H. Stock. "Efficient Tests for an Autoregressive Unit Root". In: *Econometrica* 64.4 (1996), pp. 813–836.

[62] P. D. Elrod and D. D. Tippett. "The "death valley" of change". In: *Journal of organizational change management* (2002).

[63] U. Environment. *Global Environment Outlook – GEO-6: Healthy Planet, Healthy People*. Cambridge: Cambridge University Press, 2019.

[64] E. Erdem and J. Shi. "ARMA based approaches for forecasting the tuple of wind speed and direction". In: *Applied Energy* 88.4 (2011), pp. 1405–1414.

[65] F. Eugene and K. French. "The cross-section of expected stock returns". In: *Journal of Finance* 47.2 (1992), pp. 427–465.

[66] E. European Commission. *Special Eurobarometer 2017, Attitudes of European citizens towards the environment, 468*. Report. 2017.

[67] E. European Environmental Agency. *Europe's urban air quality — re-assessing implementation challenges in cities, EEA Report No 24/2018*. Report 24/2018. 2018.

[68] E. European Environmental Agency. *Air quality in Europe - 2019 report*. Report. 2019.

[69] E. European Environmental Agency. *EMEP/EEA air pollutant emission inventory guidebook 2019, EEA Report No 13/2019*. Report 13/2019. 2019.

[70] E. P. C. of the European Union. *Decision No 1600/2002/EC of the European Parliament and of the Council of 22 July 2002 laying down the Sixth Community Environment Action Programme*. Government Document. 2002.

[71] E. P. C. of the European Union. "Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 relating to arsenic, cadmium, mercury, nickel and polycyclic aromatic hydrocarbons in ambient air". In: (2004).

[72] E. P. C. of the European Union. "Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe". In: (2008).

[73] E. P. C. of the European Union. *Regulation (EC) No 401/2009*. Legal Rule or Regulation. 2009.

[74] E. P. C. of the European Union. *Decision No 1386/2013/EU of the European Parliament and of the Council of 20 November 2013 on a General Union Environment Action Programme to 2020 'Living well, within the limits of our planet' Text with EEA relevance*. Government Document. 2013.

[75] G. R. Factors. "Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013". In: *Lancet (London, England)* 386.10010 (2015), p. 2287.

[76] E. F. Fama. "Efficient capital markets: II". In: *The journal of finance* 46.5 (1991), pp. 1575–1617.

[77] E. F. Fama, L. Fisher, M. C. Jensen, and R. Roll. "The adjustment of stock prices to new information". In: *International economic review* 10.1 (1969), pp. 1–21.

[78] E. F. Fama and K. R. French. "Common risk factors in the returns on stocks and bonds". In: *Journal of* (1993).

[79] E. F. Fama and K. R. French. "Multifactor explanations of asset pricing anomalies". In: *The journal of finance* 51.1 (1996), pp. 55–84.

[80] A. Fassò. "Statistical assessment of air quality interventions". In: *Stochastic environmental research and risk assessment* 27.7 (2013), pp. 1651–1660.

[81] A. Fassò and E. Porcu. "Latent variables and space-time models for environmental problems". In: *Stochastic Environmental Research and Risk Assessment* 29.2 (2015), pp. 323–324.

[82] D. Fattorini and F. Regoli. "Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy". In: *Environmental Pollution* 264 (2020), p. 114732.

[83] M. Fazzini, C. Baresi, C. Bisci, C. Bna, A. Cecili, A. Giuliacci, S. Illuminati, F. Pregliasco, and E. Miccadei. "Preliminary analysis of relationships between covid19 and climate, morphology, and urbanization in the lombardy region (Northern Italy)". In: *International Journal of Environmental Research and Public Health* 17.19 (2020), p. 6955.

[84] F. Finazzi and A. Fassò. "The impact of the Covid-19 pandemic on Italian mobility". In: *Significance (Oxford, England)* 17.3 (2020), p. 17.

[85] F. Finazzi and L. Paci. "Kernel-based estimation of individual location densities from smartphone data". In: *Statistical Modelling* (2019), p. 1471082X17870331.

[86] F. Fiordelisi and D. Martelli. "Corporate Culture and M&A Results in Banking". In: *Available at SSRN 1784016* (2011).

[87] B. S. Freeman, G. Taylor, B. Gharabaghi, and J. Thé. "Forecasting air quality time series using deep learning". In: *Journal of the Air & Waste Management Association* 68.8 (2018), pp. 866–886.

[88] J. W. Galbraith and V. Zinde-Walsh. "On the distributions of Augmented Dickey–Fuller statistics in processes with moving average components". In: *Journal of Econometrics* 93.1 (1999), pp. 25–47.

[89] J. Gao, Z. Yuan, X. Liu, X. Xia, X. Huang, and Z. Dong. "Improving air pollution control policy in China—A perspective based on cost–benefit analysis". In: *Science of The Total Environment* 543 (2016), pp. 307–314.

[90] P. J. García Nieto, F. Sánchez Lasheras, E. García-Gonzalo, and F. J. de Cos Juez. "PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study". In: *Science of The Total Environment* 621 (2018), pp. 753–761.

[91] A. Gianfreda, L. Parisio, and M. Pelagatti. "The impact of RES in the Italian Day-Ahead and Balancing Markets". In: *The Energy Journal* 37.SI2 (2016).

[92] A. Gianfreda, L. Parisio, and M. Pelagatti. "A Review of Balancing Costs in Italy before and after RES introduction". In: *Renewable and Sustainable Energy Reviews* 91 (2018), pp. 549–563.

[93] I. Government. *Legge 132/2016, Sistema nazionale a rete per la protezione dell'ambiente*. Government Document. 2016.

[94] S. K. Grange and D. C. Carslaw. "Using meteorological normalisation to detect interventions in air quality time series". In: *Science of The Total Environment* 653 (2019), pp. 578–588.

[95] S. K. Grange, D. C. Carslaw, A. Lewis, E. Boleti, and C. Heuglin. "Random forest meteorological normalisation models for Swiss PM10 trend analysis". In: *Atmospheric Chemistry and Physics Discussions* (2018).

[96] J. Gyarmati-Szabó, L. V. Bogachev, and H. Chen. "Nonstationary POT modelling of air pollution concentrations: Statistical analysis of the traffic and meteorological impact". In: *Environmetrics* 28.5 (2017), e2449.

[97] T. Hagnäs and S. Pynnonen. "Testing for Cumulative Abnormal Returns in Event Studies with the Rank Test". In: *Available at SSRN 2479228* (2014).

[98] A. C. Harvey. *Forecasting Structural Time Series and the Kalman Filter*. Cambridge University Press, 1989.

[99] A. Harvey. "Chapter 7 Forecasting with Unobserved Components Time Series Models". In: ed. by G. Elliott, C. W. J. Granger, and A. Timmermann. Vol. 1. Handbook of Economic Forecasting. Elsevier, 2006, pp. 327–412.

[100] A. Harvey and S. J. Koopman. "Structural time series models". In: *Wiley StatsRef: Statistics Reference Online* (2014).

[101] A. C. Harvey and P. H. J. Todd. "Forecasting economic time series with structural and Box-Jenkins models: A case study". In: *Journal of Business & Economic Statistics* 1.4 (1983), pp. 299–307.

[102] U. Hassler and J. Wolters. "On the power of unit root tests against fractional alternatives". In: *Economics Letters* 45.1 (1994), pp. 1–5.

[103] S. E. Hein and P. Westfall. "Improving Tests of Abnormal Returns by Bootstrapping the Multivariate Regression Model with Event Parameters". In: *Journal of Financial Econometrics* 2.3 (2004), pp. 451–471.

[104] J. Helske. "KFAS: Exponential family state space models in R". In: *arXiv preprint arXiv:1612.01907* (2016).

[105] J. Higham, C. A. Ramírez, M. Green, and A. Morse. "UK COVID-19 lockdown: 100 days of air pollution reduction?" In: *Air Quality, Atmosphere & Health* (2020), pp. 1–8.

[106] O. Hänninen, A. B. Knol, M. Jantunen, T.-A. Lim, A. Conrad, M. Rappolder, P. Carrer, A.-C. Fanetti, R. Kim, and J. Buekers. "Environmental burden of disease in Europe: assessing nine risk factors in six countries". In: *Environmental health perspectives* 122.5 (2014), pp. 439–446.

[107] W. Hoeffding and H. Robbins. "The central limit theorem for dependent random variables". In: *Duke Math. J.* 15.3 (1948), pp. 773–780.

[108] J. Horowitz and S. Barakat. "Statistical analysis of the maximum concentration of an air pollutant: Effects of autocorrelation and non-stationarity". In: *Atmospheric Environment (1967)* 13.6 (1979), pp. 811–818.

[109] A. N. Hristov. "Contribution of ammonia emitted from livestock to atmospheric fine particulate matter (PM2. 5) in the United States". In: *Journal of dairy science* 94.6 (2011), pp. 3130–3136.

[110] A. INEMAR ARPA Lombardia Settore Aria. "INEMAR Emission Inventory 2017: emission in Lombardy Region - final data - public revision". In: (2020).

[111] G. Invernizzi, A. Ruprecht, R. Mazza, C. De Marco, G. Močnik, C. Sioutas, and D. Westerdahl. "Measurement of black carbon concentration as an indicator of air quality benefits of traffic restriction policies within the ecopass zone in Milan, Italy". In: *Atmospheric Environment* 45.21 (2011), pp. 3522–3527.

[112] S. R. Jammalamadaka and A. Sengupta. *Topics in circular statistics*. Vol. 5. world scientific, 2001.

[113] S. R. Jammalamadaka and U. J. Lund. "The effect of wind direction on ozone levels: a case study". In: *Environmental and Ecological Statistics* 13.3 (2006), pp. 287–298.

[114] C. M. Jarque and A. K. Bera. "A test for normality of observations and regression residuals". In: *International Statistical Review/Revue Internationale de Statistique* (1987), pp. 163–172.

[115] Y. Jiang, X.-J. Wu, and Y.-J. Guan. "Effect of ambient air pollutants and meteorological variables on COVID-19 incidence". In: *Infection Control & Hospital Epidemiology* 41.9 (2020), pp. 1011–1015.

[116] S. Johansen. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models". In: *Econometrica* 59.6 (1991), pp. 1551–1580.

[117] B. L. Joiner and J. R. Rosenblatt. "Some properties of the range in samples from Tukey's symmetric lambda distributions". In: *Journal of the American Statistical Association* 66.334 (1971), pp. 394–399.

[118] A. M. Jones, R. M. Harrison, B. Barratt, and G. Fuller. "A large reduction in airborne particle number concentrations at the time of the introduction of "sulphur free" diesel and the London Low Emission Zone". In: *Atmospheric Environment* 50 (2012), pp. 129–138.

[119] J. Kang and K.-I. Hwang. "A comprehensive real-time indoor air-quality level indicator". In: *Sustainability* 8.9 (2016), p. 881.

[120] A. Kerimray, N. Baimatova, O. P. Ibragimova, B. Bukenov, B. Kenessov, P. Plotitsyn, and F. Karaca. "Assessing air quality changes in large cities during COVID-19 lockdowns: The impacts of traffic-free urban conditions in Almaty, Kazakhstan". In: *Science of The Total Environment* 730 (2020), p. 139179.

[121] H. Kobori, J. L. Dickinson, I. Washitani, R. Sakurai, T. Amano, N. Komatsu, W. Kitamura, S. Takagawa, K. Koyama, and T. Ogawara. "Citizen science: a new approach to advance ecology, education, and conservation". In: *Ecological research* 31.1 (2016), pp. 1–19.

[122] J. W. Kolari and S. Pynnönen. "Event study testing with cross-sectional correlation of abnormal returns". In: *The Review of financial studies* 23.11 (2010), pp. 3996–4025.

[123] J. W. Kolari and S. Pynnönen. "Event Study Testing with Cross-sectional Correlation of Abnormal Returns". In: *Review of Financial Studies* 23.11 (2010), pp. 3996–4025.

[124] J. W. Kolari and S. Pynnönen. "Nonparametric rank tests for event studies". In: *Journal of Empirical Finance* 18.5 (2011), pp. 953–971.

[125] S. P. Kothari and J. B. Warner. "Econometrics of event studies". In: *Handbook of empirical corporate finance*. Elsevier, 2007, pp. 3–36.

[126] J. Köppelová and A. Jindrová. "Application of exponential smoothing models and arima models in time series analysis from telco area". In: *AGRIS on-line Papers in Economics and Informatics* 11.665-2019-4145 (2019), pp. 73–84.

[127] G. M. Kuersteiner and I. R. Prucha. "Limit theory for panel data models with cross sectional dependence and sequential exogeneity". In: *Journal of Econometrics* 174.2 (2013), pp. 107 –126.

[128] S. N. Lahiri. *Resampling Methods for Dependent Data*. Springer Series in Statistics. Springer-Verlag New York, 2003, pp. XIV, 374.

[129] J. Ledolter. "The effect of additive outliers on the forecasts from ARIMA models". In: *International Journal of Forecasting* 5.2 (1989), pp. 231–240.

[130] C.-K. Lee. "Multifractal characteristics in air pollutant concentration time series". In: *Water, Air, and Soil Pollution* 135.1-4 (2002), pp. 389–409.

[131] H. Li, X.-L. Xu, D.-W. Dai, Z.-Y. Huang, Z. Ma, and Y.-J. Guan. "Air pollution and temperature are associated with increased COVID-19 incidence: A time series study". In: *International Journal of Infectious Diseases* 97 (2020), pp. 278–282.

[132] Z. Liu, L. Zhao, C. Wang, Y. Yang, J. Xue, X. Bo, D. Li, and D. Liu. "An Actuarial Pricing Method for Air Quality Index Options". In: *International Journal of Environmental Research and Public Health* 16.24 (2019), p. 4882.

[133] G. M. Ljung and G. E. Box. "On a measure of lack of fit in time series models". In: *Biometrika* 65.2 (1978), pp. 297–303.

[134] S. Lolli, Y.-C. Chen, S.-H. Wang, and G. Vivone. "Impact of meteorology and air pollution on Covid-19 pandemic transmission in Lombardy region, Northern Italy". In: (2012).

[135] S. Lummer and J. McConnel. "Further evidence on the bank lending process and the capital-market response to bank loan agreements". In: *Journal of Financial Economics* 25.1 (1989), pp. 99–122.

[136] T. Luoma. "Nonparametric event study tests for testing cumulative abnormal returns". In: *Acta Wasaensia* 254 (2011).

[137] A. C. MacKinlay. "Event Studies in Economics and Finance". In: *Journal of Economic Literature* 35 (1997), pp. 13–39.

[138] A. C. MacKinlay. "Event studies in economics and finance". In: *Journal of Economic Literature* 35.1 (1997), pp. 13–39.

[139] P. Maranzano, A. Fassò, M. Pelagatti, and M. Mudelsee. "Statistical Modeling of the Early-Stage Impact of a New Traffic Policy in Milan, Italy". In: *International Journal of Environmental Research and Public Health* 17.3 (2020), p. 1088.

[140] K. V. Mardia and P. E. Jupp. *Directional statistics*. Vol. 494. John Wiley & Sons, 2009.

[141] L. Martelletti and P. Martelletti. "Air pollution and the novel Covid-19 disease: a putative disease risk factor". In: *SN Comprehensive Clinical Medicine* (2020), pp. 1–5.

[142] A. Ma'rufatin and H. Kusnoputranto. "Investigating the relationship between variation of greenhouse gases concentrations and humidex". In: *E3S Web of Conferences*. Vol. 74. EDP Sciences, 2018, p. 11004.

[143] J. M. Masterton and F. A. Richardson. *Humidex: a method of quantifying human discomfort due to excessive heat and humidity*. Environment Canada, Atmospheric Environment, 1979.

[144] J. McConnell and C. Muscarella. "Corporate capital expenditure decisions and the market value of the firm". In: *Journal of Financial Economics* 14.3 (1985), pp. 399–422.

[145] D. L. McLeish. "Dependent Central Limit Theorems and Invariance Principles". In: *Annals of Probability* 2.4 (1974), pp. 620–628.

[146] H Menezes Piairo, P Garcia-Soidan, and I Sousa. "Spatial-temporal modellization of the NO2 concentration data through geostatistical tools". In: *Stat Methods Appl. doi* 10 (2015).

[147] W. H. Mikkelson and M. M. Partch. "Withdrawn Security Offerings". In: *Journal of Financial and Quantitative Analysis* 23.2 (1988), pp. 119–133.

[148] M. Comune di Milano. "Delibera 1366/2018". In: (2018), p. 276.

[149] M. Mudelsee and M. Alkio. "Quantifying effects in two-sample environmental experiments using bootstrap confidence intervals". In: *Environmental Modelling & Software* 22.1 (2007), pp. 84–96.

[150] M. Mudelsee. *Climate Time Series Analysis - Classical Statistical and Bootstrap Methods*. 2nd ed. Atmospheric and Oceanographic Sciences Library. Springer International Publishing, 2014, pp. XXXII, 454.

[151] M. Mudelsee. *Statistical Analysis of Climate Extremes*. Cambridge University Press, 2020.

[152] L. Y. K. Nakada and R. C. Urban. "COVID-19 pandemic: Impacts on the air quality during the partial lockdown in São Paulo state, Brazil". In: *Science of The Total Environment* 730 (2020), p. 139087.

[153] F. Naseem, A. Rashid, T. Izhar, M. I. Khawar, S. Bano, A. Ashraf, and M. N. Adnan. "An integrated approach to air pollution modeling from climate change perspective using ARIMA forecasting". In: *Journal of Applied Agriculture and Biotechnology* 2.2 (2018), pp. 37–44.

[154] D. B. Nelson. "Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model". In: *Journal of Econometrics* 52.1 (1992), pp. 61–90.

[155] D. B. Nelson and D. P. Foster. "Filtering and forecasting with misspecified ARCH models II: Making the right forecast with the wrong model". In: *Journal of Econometrics* 67.2 (1995), pp. 303–335.

[156] C. N. Ng and T. L. Yan. "Recursive estimation of model parameters with sharp discontinuity in non-stationary air quality data". In: *Environmental Modelling & Software* 19.1 (2004), pp. 19–25.

[157] Y. Ogen. "Assessing nitrogen dioxide (NO2) levels as a contributing factor to coronavirus (COVID-19) fatality". In: *Science of The Total Environment* 726 (2020), p. 138605.

[158] S. Orey. "A central limit theorem for m-dependent random variables". In: *Duke Math. J.* 25.4 (1958), pp. 543–546.

[159] J. M. Patell. "Corporate forecasts of earnings per share and stock price behavior: empirical tests". In: *Journal of Accounting Research* 14 (1976), pp. 246–276.

[160] E. Pebesma. "sf: Simple Features for R". In: *R package version 0.6-1* (2018).

[161] M. M. Pelagatti. *Time series modelling with unobserved components*. Chapman and Hall/CRC, 2015.

[162] M. M. Pelagatti. *Time Series Modelling with Unobserved Components*. Chapman and Hall / CRC, 2015.

[163] M. M. Pelagatti and P. K. Sen. "Rank tests for short memory stationarity". In: *Journal of Econometrics* 172.1 (2013), pp. 90–105.

[164] M. Percoco. "The effect of road pricing on traffic composition: Evidence from a natural experiment in Milan, Italy". In: *Transport Policy* 31 (2014), pp. 55–60.

[165] D. Perlman and G. J. Takacs. "The 10 Stages of Change: To cope with change". In: *Nursing management* 21.4 (1990), pp. 33–38.

[166] D. Pernigotti, E. Georgieva, P. Thunis, and B. Bessagnet. "Impact of meteorology on air quality modeling over the Po valley in northern Italy". In: *Atmospheric environment* 51 (2012), pp. 303–310.

[167] H. Petetin, D. Bowdalo, A. Soret, M. Guevara, O. Jorba, K. Serradell, and C. Pérez García-Pando. "Meteorology-normalized impact of the COVID-19 lockdown upon NO 2 pollution in Spain". In: *Atmospheric Chemistry and Physics* 20.18 (2020), pp. 11119–11141.

[168] B. Pfaff. *Analysis of Integrated and Cointegrated Time Series with R*. Second. ISBN 0-387-27960-1. New York: Springer, 2008.

[169] N. R. Prabhala. "Conditional Methods in Event Studies and an Equilibrium Justification for Standard Event-Study Procedures". In: *The Review of Financial Studies* 10.1 (1997), pp. 1–38.

[170] T. Proietti. "Forecasting and signal extraction with misspecified models". In: *Journal of Forecasting* 24.8 (2005), pp. 539–556.

[171] R. Qadir, G Abbaszade, J. Schnelle-Kreis, J. Chow, and R Zimmermann. "Concentrations and source contributions of particulate organic matter before and after implementation of a low emission zone in Munich, Germany". In: *Environmental pollution* 175 (2013), pp. 158–167.

[172] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2018.

[173] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020.

[174] K. Raffaelli, M. Deserti, M. Stortini, R. Amorati, M. Vasconi, and G. Giovannini. "Improving Air Quality in the Po Valley, Italy: Some Results by the LIFE-IP-PREPAIR Project". In: *Atmosphere* 11.4 (2020), p. 429.

[175] R. Regional Statistical Yearbook. *Lombardia Regional Statistical Yearbook 2017/2018*. Report. 2017.

[176] R. Regional Statistical Yearbook. *Regional Statistical Yearbook of Lombardia in Europe 2017/2018*. Report. 2017.

[177] K. Ropkins and J. E. Tate. "Early Observations on the impact of the COVID-19 Lockdown on Air Quality Trends across the UK". In: *Science of The Total Environment* 754 (2020), p. 142374.

[178] R. Rossi, R. Ceccato, and M. Gastaldi. "Effect of Road Traffic on Air Pollution. Experimental Evidence from COVID-19 Lockdown". In: *Sustainability* 12.21 (2020), p. 8984.

[179] G. Rudebush. "On the power of Dickey-Fuller tests against fractional alternatives". In: *Economics letters* 35 (1991), pp. 155–160.

[180] S. Saad, A. Shakaff, A. Saad, A. Yusof, A. Andrew, A Zakaria, and A. Adom. "Development of indoor environmental index: Air quality index and thermal comfort index". In: *AIP Conference Proceedings*. Vol. 1808. AIP Publishing LLC, 2017, p. 020043.

[181] S. E. Said and D. A. Dickey. "Testing for unit roots in autoregressive-moving average models of unknown order". In: *Biometrika* 71.3 (1984), pp. 599–607.

[182] R. Salcedo, M. A. Ferraz, C. Alves, and F. Martins. "Time-series analysis of air pollution data". In: *Atmospheric Environment* 33.15 (1999), pp. 2361–2372.

[183] R. W. J. Sanders and R. P. Robins. "Discriminating between wealth and information effects in event studies in accounting and finance research". In: *Review of Quantitative Finance and Accounting* 1.3 (1991), pp. 307–330.

[184] S. Sharma, M. Zhang, Anshika, J. Gao, H. Zhang, and S. H. Kota. "Effect of restricted emissions during COVID-19 on air quality in India". In: *Science of The Total Environment* 728 (2020), p. 138878.

[185] R. P. Singh and A. Chauhan. "Impact of lockdown on air quality in India during COVID-19 pandemic". In: *Air Quality, Atmosphere & Health* 13.8 (2020), pp. 921–928.

[186] J. Stock and M. Watson. *Introduction to Econometrics, Update, Global Edition*. Pearson Education Limited, 2015.

[187] U. Switzerland Federal Statistics Office. *Cross-border commuters statistics in 2018*. Report. 2019.

[188] S. M. Taghavi-Shahri, A. Fassò, B. Mahaki, and H. Amini. "Concurrent spatiotemporal daily land use regression modeling and missing data imputation of fine particulate matter using distributed space-time Expectation Maximization". In: *Atmospheric Environment* (2019), p. 117202.

[189] P. Thunis, A. Miranda, J. M. Baldasano, N. Blond, J. Douros, A. Graff, S. Janssen, K. Juda-Rezler, N. Karvosenoja, G. Maffeis, A. Martilli, M. Rasoloharimahefa, E. Real, P. Viaene, M. Volta, and L. White. "Overview of current regional and local scale air quality modelling practices: Assessment and planning tools in the EU". In: *Environmental Science & Policy* 65 (2016), pp. 13–21.

[190] U. Ticino Statistics Office. *Traffic in selected stations of Ticino historical data*. Report. 2020.

[191] A. Tobías, C. Carnerero, C. Reche, J. Massagué, M. Via, M. C. Minguillón, A. Alastuey, and X. Querol. "Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic". In: *Science of The Total Environment* 726 (2020), p. 138540.

[192] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge series in statistical and probablistic mathematics. The Edinburgh Building, Cambridge CB2 2RU, UK: Cambridge University Press, 1998.

[193] C. Varotsos, J. Ondov, and M. Efstathiou. "Scaling properties of air pollution in Athens, Greece and Baltimore, Maryland". In: *Atmospheric Environment* 39.22 (2005), pp. 4041–4047.

[194] V. Vasquez-Apestegui, E. Parras-Garrido, V. Tapia, V. M. Paz-Aparicio, J. P. Rojas, O. R. Sánchez-Ccoyllo, and G. F. Gonzales. "Association Between Air Pollution in Lima and the High Incidence of COVID-19: Findings from a Post Hoc Analysis". In: *Research square* (2020).

[195] R. Vautard, P. H. J. Builtjes, P. Thunis, C. Cuvelier, M. Bedogni, B. Bessagnet, C. Honoré, N. Moussiopoulos, G. Pirovano, and M. Schaap. "Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry transport models over four European cities within the CityDelta project". In: *Atmospheric Environment* 41.1 (2007), pp. 173–188.

[196] C. Vlachokostas, C. Achillas, Moussiopoulos, E. Hourdakis, G. Tsilingiridis, L. Ntziachristos, G. Banias, N. Stavrakakis, and C. Sidiropoulos. "Decision support system for the evaluation of urban air pollution control options: Application for particulate pollution in Thessaloniki, Greece". In: *Science of the Total Environment* 407.23 (2009), pp. 5937–5948.

[197] R Von Mises. "Ueber die" ganzzaligkeit" der atomgewicht und verwandte fragen". In: *Physikal* 19 (1918), pp. 490–500.

[198] P. T. Vu, T. V. Larson, and A. A. Szpiro. "Probabilistic Predictive Principal Component Analysis for Spatially-Misaligned and High-Dimensional Air Pollution Data with Missing Observations". In: *arXiv preprint arXiv:1905.00393* (2019).

[199] W. H. O. WHO. "Air quality guidelines for Europe". In: (2000).

[200] W. H. O. WHO. *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. World Health Organization, 2006.

[201] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[202] H. L. Windsor and R. Toumi. "Scaling and persistence of UK pollution". In: *Atmospheric Environment* 35.27 (2001), pp. 4545–4556.

[203] X. Wu, R. C. Nethery, B. M. Sabath, D. Braun, and F. Dominici. "Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study". In: *medRxiv* (2020), p. 2020.04.05.20054502.

[204] X. Xu, D. Dong, Y. Wang, and S. Wang. "The Impacts of Different Air Pollutants on Domestic and Inbound Tourism in China". In: *International Journal of Environmental Research and Public Health* 16.24 (2019), p. 5127.

[205] W. Yao, C. Zhang, X. Wang, J. Sheng, Y. Zhu, and S. Zhang. "The research of new daily diffuse solar radiation models modified by air quality index (AQI) in the region with heavy fog and haze". In: *Energy Conversion and Management* 139 (2017), pp. 140–150.

[206] K. Yu, Y. Cheung, T Cheung, and R. C. Henry. "Identifying the impact of large urban airports on local air quality by nonparametric regression". In: *Atmospheric Environment* 38.27 (2004), pp. 4501–4507.

[207] S. Zangari, D. T. Hill, A. T. Charette, and J. E. Mirowsky. "Air quality changes in New York City during the COVID-19 pandemic". In: *Science of The Total Environment* 742 (2020), p. 140496.

[208] C. Zhang, C. Shen, Q. Yang, S. Wei, G. Lv, and C. Sun. "An investigation on the attenuation effect of air pollution on regional solar radiation". In: *Renewable Energy* (2020).

[209] H. Zhang, S. Zhang, P. Wang, Y. Qin, and H. Wang. "Forecasting of particulate matter time series using wavelet analysis and wavelet-ARMA/ARIMA model in Taiyuan, China". In: *Journal of the Air & Waste Management Association* 67.7 (2017), pp. 776–788.

[210] L. Zhang, J. Lin, R. Qiu, X. Hu, H. Zhang, Q. Chen, H. Tan, D. Lin, and J. Wang. "Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model". In: *Ecological Indicators* 95 (2018), pp. 702–710.

[211] J. Zhu, P. Wu, H. Chen, L. Zhou, and Z. Tao. "A hybrid forecasting approach to air quality time series based on endpoint condition and combined forecasting model". In: *International journal of environmental research and public health* 15.9 (2018), p. 1941.

[212] Y. Zhu, J. Xie, F. Huang, and L. Cao. "Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China". In: *Science of The Total Environment* 727 (2020), p. 138704.

[213] Y. Zhu, J. Xie, F. Huang, and L. Cao. "The mediating effect of air quality on the association between human mobility and COVID-19 infection in China". In: *Environmental research* 189 (2020), p. 109911.

[214] T. L. Zivney and D. J. Thompson. "The specification and power of the sign test in measuring security price performance: comments and analysis". In: *The Financial Review* 24 (1989), pp. 581–588.

[215] M. A. Zoran, R. S. Savastru, D. M. Savastru, and M. N. Tautan. "Assessing the relationship between ground levels of ozone (O3) and nitrogen dioxide (NO2) with coronavirus (COVID-19) in Milan, Italy". In: *Science of The Total Environment* 740 (2020), p. 140005.

[216] M. A. Zoran, R. S. Savastru, D. M. Savastru, and M. N. Tautan. "Assessing the relationship between surface levels of PM2.5 and PM10 particulate matter impact on COVID-19 in Milan, Italy". In: *Science of The Total Environment* 738 (2020), p. 139825.