WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Can satellite data on air pollution predict industrial production?

Jean-Charles Bricongne[1], Baptiste Meunier[2] & Thomas Pical[3]

November 2021, WP #847

ABSTRACT

The Covid-19 crisis has highlighted innovative high-frequency dataset allowing to measure in real-time the economic impact. In this vein, we explore how satellite data measuring the concentration of nitrogen dioxide (NO_2, a pollutant emitted mainly by industrial activity) in the troposphere can help predict industrial production. We first show how such data must be adjusted for meteorological patterns which can alter data quality and pollutant emissions. We use machine learning techniques to better account for non-linearities and interactions between variables. We then find evidence that nowcasting performances for monthly industrial production are significantly improved when relying on daily NO_2 data compared to benchmark models based on PMIs and auto-regressive (AR) terms. We also find evidence of heterogeneities suggesting that the contribution of daily pollution data is particularly important during "crisis" episodes and that the elasticity of NO_2 pollution to industrial production for a country depends on the share of manufacturing in the value added. Available daily, free-to-use, granular and covering all countries including those with limited statistics, this paper illustrates the potential of satellite-based data for air pollution in enhancing the real-time monitoring of economic activity.

Keywords: Data Science, Big Data, Satellite Data, Nowcasting, Machine Learning, Industrial Production
JEL classification: C51, C81, E23, E37

[1] Banque de France, Paris I University, LEO, LIEPP, jean-charles.bricongne@banque-france.fr
[2] Banque de France, European Central Bank, AMSE, baptiste.meunier@ecb.europa.eu
[3] Equancy, AMSE, tpical@equancy.com

We are very grateful to Juri Marcucci, Sébastien Laurent, Ewen Gallic, Nicolas Woloszko, Hervé Mignot, as well as participants to the FRB – Banca d'Italia joint conference on non-traditional data & statistical learning (Nov. 2020), the OECD-BdF conference on new approaches to macroeconomic forecasting (June 2021), the Banque de France seminar (June 2021), and the BCEAO-BdF seminar on forecasting during crisis episodes (Sept. 2021) for useful comments. We are indebted to Marcel Voia for sharing code on panel-MIDAS. We also thank for their excellent assistance, Michel Arnault, Juan Porras, Eduardo Leite-Kropiwiec, Alvaro Sanchez and Ana-Gabriela Prada regarding the satellite data, as well as Matthieu Dertilis, Fabien Lebreton, and Aude Le Métayer regarding macroeconomic data. The views expressed in this paper are those of the authors, and not those of the Banque de France, the European Central Bank, the Laboratoire d'Économie d'Orléans, the LIEPP, Paris I University or the Aix-Marseille School of Economics.

NON-TECHNICAL SUMMARY

The sudden shock of the Covid-19 crisis – with some countries shutting down almost entirely in a matter of days – has put new emphasis on high-frequency data. Weekly, daily, or even hourly data have been extensively used to assess in real-time the impact of the pandemics. In particular, satellite measurements of air pollution have been put forward on several occasions to show the alleged dramatic effect of the shutdown of factories on air pollution.

Against this background, we assess whether satellite data for tropospheric pollution can help predicting industrial production. We focus on nitrogen dioxide (NO_2), a pollutant mainly emitted by industrial activity. Compared to official or alternative indicators, such data present advantages of timeliness, global coverage – including over developing countries with limited official statistics, granularity, and free use. While an avenue for future research could be to compare predictive performances of NO_2 satellite data with other high-frequency indicators, the global and uniform coverage of satellite data appears a key advantage.

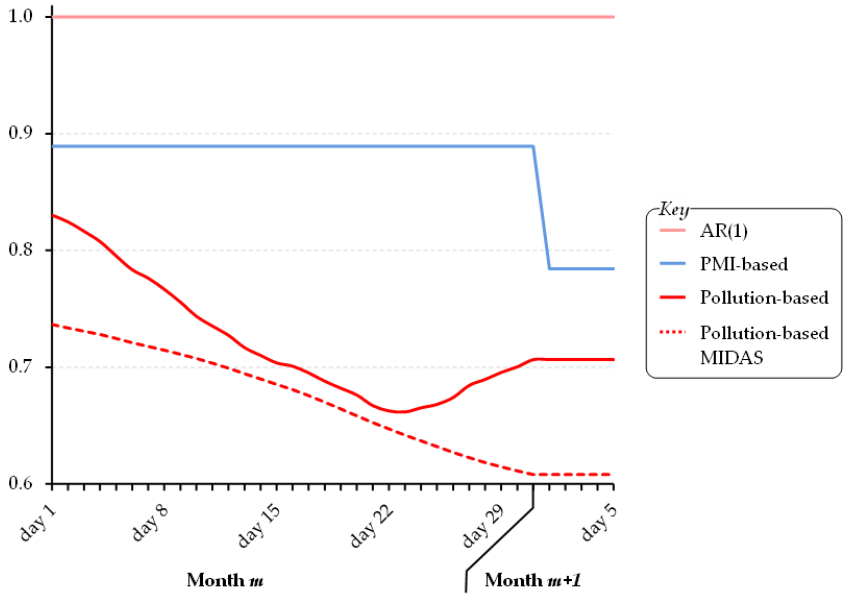
Raw satellite data are however far from ready-to-use. Our first step is fetching the data and making it easier to process: we select and aggregate relevant data at ZIP code, allowing to go from a daily download of 4 Gb in multiple files into a single 20 Mb *csv* file. As the quality of satellite data can be altered over cloudy or snowy areas, we also clean the data. This results however in a large amount of missing points. Since missing data at local level might result in undesired composition effects when aggregating at national level, we interpolate them. We rely on a machine learning technique (the k-nearest neighbours' algorithm) that allows interpolation to account both for spatial and temporal correlations. Finally, NO_2 pollution heavily depends on meteorological factors (temperature, wind, humidity). Given that their effect is non-linear and features interactions between variables, we use a random forest algorithm. Data is then aggregated at national level to match the granularity of official statistics, most notably the industrial production that we intend to nowcast.

Our second step checks the relevance of NO_2 pollution from a forecasting standpoint. We rely on panel regression over 17 emerging and 16 advanced countries to make up for the limited available timespan (since Dec. 2018 only). We find evidence that a model based on daily NO_2 pollution data over-performs benchmark models based on survey data (Purchasing Managers' Index) or auto-regressive (AR) terms. Mimicking a real-time set-up from March 2020 to December 2020, we find that this over-performance holds for all days of the month: the model based on daily NO_2 pollution data over-performs benchmarks at every day – with evidence that gains in predictive accuracy are greater as the month advances and more daily data become then available (see **figure 1**). There are additional accuracy gains when relying on a MIxed DAta Sampling (MIDAS) approach to predict monthly industrial production using daily NO_2 pollution – using a panel-MIDAS recently introduced in the literature.

We finally find evidence for heterogeneities. First, accuracy gains are greater for countries that have been more affected by the Covid-19 crisis, suggesting that the contribution of high-frequency data is more important during "crisis" episodes than during "normal" times. Second, the elasticity of pollution to industrial production appears to be greater for countries with a larger share of manufacturing in the value added. In the end, we turn to business cycle detection. Building on a Markov-switching framework, we find that daily NO_2 pollution data allow for a swifter detection of turning points compared with relying on monthly official data, with a lead time of around 2.5 months for the former over the latter.

Figure 1. Models' performance (out-of-sample RMSE) relative to the AR(1) benchmark

(Sources: ESA, authors' calculation)



Les données satellite de pollution de l'air prédisent-elles la production industrielle ?

RÉSUMÉ

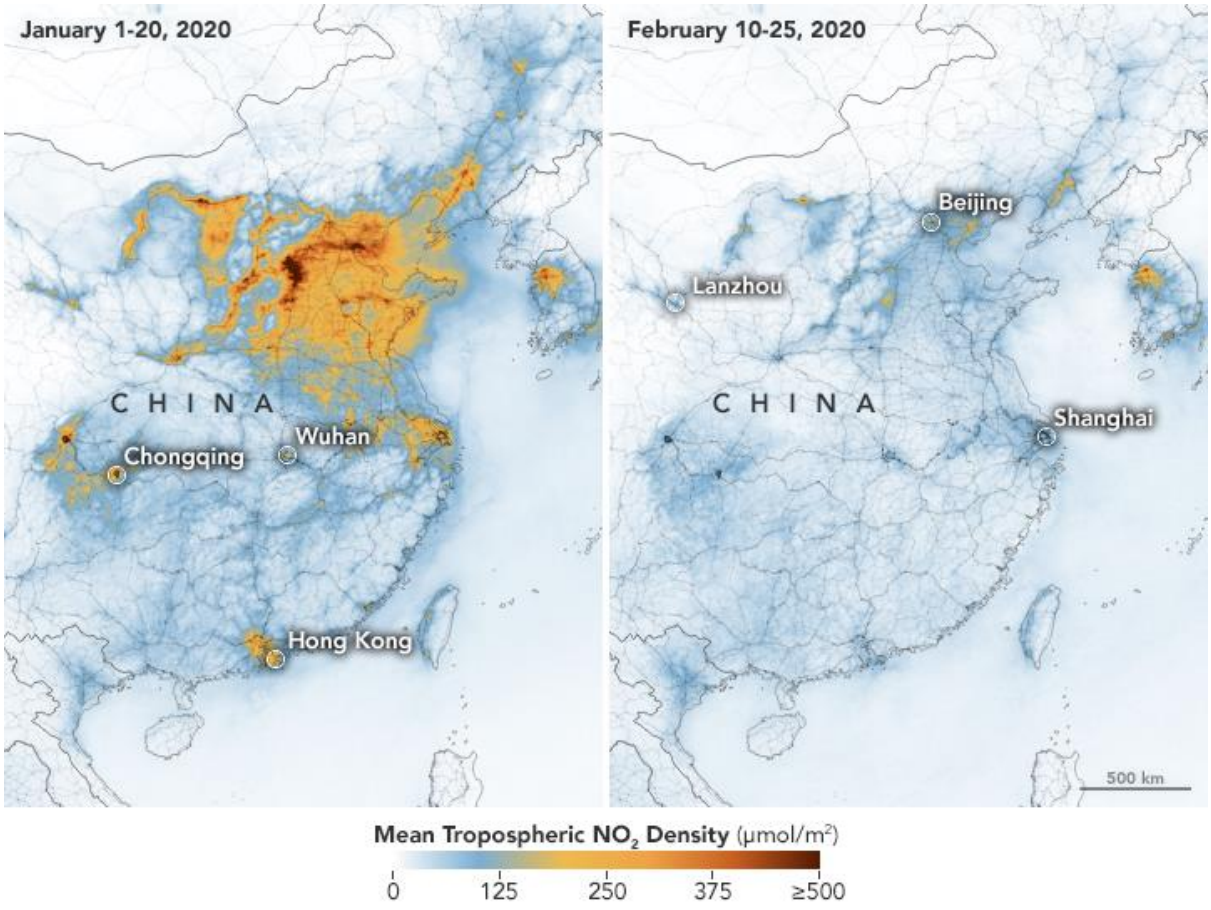
La crise de la Covid-19 a illustré le rôle des données haute-fréquence pour mesurer en temps réel l'activité économique. Dans cet esprit, nous étudions si les données satellite de pollution au dioxyde d'azote (NO_2, un polluant émis principalement par l'activité industrielle) dans la troposphère permettent de prédire la production industrielle. Nous montrons d'abord que ces données doivent être corrigées par les facteurs météorologiques qui affectent la qualité des données et le niveau de pollution. Nous utilisons des techniques de *machine learning* pour tenir compte des non-linéarités et interactions entre variables. Nous montrons ensuite que les prévisions en temps réel ont de meilleures performances en utilisant les données de pollution par rapport à des modèles basés sur les enquêtes PMI ou des termes autorégressifs. L'analyse fait aussi apparaitre des hétérogénéités avec une contribution plus significative des données de pollution pendant les épisodes de crise, et avec une élasticité de la pollution au NO_2 par rapport à l'activité plus importante dans les pays ou la part de l'industrie dans la valeur ajoutée est la plus forte. Disponible quotidiennement, a un niveau de précision élevée, et une couverture uniforme de tous pays – y compris ceux aux systèmes statistiques peu développés – cette étude montre le potentiel des données de satellite pour améliorer le suivi en temps réel de l'économie.

Mots-clés : science des données, big data, données satellite, prévision en temps réel, machine learning, production industrielle

Introduction

The sudden shock of the Covid-19 crisis – with some economies shutting down almost entirely in a matter of days – has put new emphasis on high-frequency data. Weekly, daily, or even hourly data have been extensively used to assess in real-time the impact of the Great Lockdown on economic activity in a way that could permanently disrupt how official statistics are constructed (Veronese *et al.*, 2020). Among all these high-frequency data, satellite measurements of tropospheric[1] pollution have been highlighted in a number of press articles displaying graphic comparisons of how the shutdown of factories had supposedly reduced air pollution in dramatic proportions (see for example China in **figure 1**).

Figure 1. Tropospheric pollution over China: January 2020 *vs.* February 2020
Source: Earth observatory



Against this background, this paper aims at assessing whether satellite data for tropospheric pollution can help predicting industrial production. We focus on nitrogen dioxide (NO_2), a pollutant emitted by industrial activity and transportation which presents the advantage of having a limited duration period in the troposphere – making it appropriate for real-time

[1] Earth atmosphere has a series of layers. Moving upward from ground level, these layers are the troposphere (0 to around 15 km), stratosphere, mesosphere, thermosphere and exosphere.

monitoring. Compared to other official or alternative statistics, these data present a number of additional advantages: timeliness, global coverage, granularity, and free use. Available daily, relying on these data can make up for the publication lags of official statistics – published monthly for industrial production and generally 20-30 days after month end. Also, satellite data has a global coverage, including in areas where statistics are scarce and for which such data can provide an alternative indicator. The uniformity of this coverage is also an advantage, alleviating potential biases at local level due to idiosyncratic errors or to the arbitrary locations of sensors. In addition, the high granularity – data are given for 7×3.5 km^2 areas – allows to capture events at regional level while official statistics are usually aggregated at national level. Not to finally mention that satellite data are free-to-use.

Raw satellite data are however far from ready-to-use and a first step is data retrieval, cleaning, and weather correction. We first fetch the data, check their consistency and make it more easily to process (from a daily 4 Gb in multiple files to a single 20 Mb *csv* file) by aggregating data at ZIP code level and dropping data out of land areas. We then clean data based on their quality index since satellite measurements can be altered over cloudy or snowy areas, but also by solar zenith or viewing angles. This results however in a large share of missing points, so we must interpolate missing data to avoid undesired composition effects when aggregating at regional or national level. We rely on the KNN (K-Nearest Neighbours) algorithm which allows to account for both spatial and temporal correlations while still having a limited computational cost. Finally, NO_2 pollution depends heavily on meteorological factors (temperature, wind, humidity). Since they affect air pollution in a non-linear fashion and feature interactions between variables, we use the random forest algorithm – in line with recent advances in the geo-statistical literature. In the end, we aggregate pollution data at a national level to match the granularity of official statistics – most notably of the industrial production that we intend to nowcast.

The second step checks the relevance of satellite data on air pollution from a forecasting standpoint. We rely on a panel regression over several advanced and emerging countries to make up for the limited timespan (satellite was launched in 2017 so our dataset starts in 2018). We find evidence that nowcasting models for industrial production based on daily pollution data over-perform benchmark models based on survey data (Purchasing Managers' Index) or auto-regressive (AR) terms. Mimicking a real-time set-up, we find that this holds for all days in a given month: nowcasting with the latest data available for NO_2 pollution (i.e. up to the preceding day) over-performs benchmark models for every day in the month – with evidence that gains in forecasting accuracy are greater as the month advances and more daily data become then available for the corresponding month. While these results are obtained using a simple averaging of daily data at monthly frequency, we show that there are additional accuracy gains when relying on a MIxed DAta Sampling (MIDAS) approach – where different weights are attributed to the different lags of high-frequency regressors. To do so, we build on recent advances in the literature by Khalaf *et al.* (2021) who are the first to combine MIDAS

and panel. We also find evidence of heterogeneities. First, accuracy gains are greater for countries more affected by the Covid-19 crisis, suggesting that the contribution of high-frequency data is more important during "crisis" episodes than during "normal" times. Second, the elasticity of pollution to industrial production is greater for countries with a larger share of manufacturing in the value added. In the end, we turn to the detection of business cycles. Building on a Markov-switching framework, we find evidence that the high-frequency satellite data allow for a swifter detection of turning points in the economy compared with a method relying on official monthly statistics. We indeed estimate that the former detects turning points around 2.5 months earlier than the latter.

This paper is at the crossroads of the literature on high-frequency data and on satellite data. It contributes to the former by exploring the potential of NO_2 pollution data, which had not been analysed yet and which, in contrast to most other high-frequency indicators, has the advantage of a global and uniform coverage. The paper also explores whether high-frequency data enhance nowcasting performances as well as the detection of turning points, contributing to the on-going debate over the value added of such data to forecasting. In addition, our paper contributes to the literature on satellite data by exploring their benefits for economic monitoring. Most of the literature to date has shown how economic crisis has an impact on air pollution, but very few have explored whether air pollution can be an early indicator for economic conditions. In that vein, closest to us is the literature using "night lights" to measure economic activity. We contribute to this strand of the literature not only by exploring a new dataset and setting a method for its usage in economics, but also by providing such satellite-based indicators at a daily frequency in contrast with the literature on "night lights" relying mostly on yearly averages. Also, while recent advances in the literature of "night lights" have shown that these data are uncorrelated with economic growth for advanced economies, evidence in our paper suggests that NO_2 pollution is still a valid indicator for industrial production in those economies – albeit with a lower elasticity than for more manufacturing-based emerging economies.

The rest of the paper is organized as follows: **section 1** reviews the related literature, **section 2** describes satellite data and explains our choice to rely on such data largely uncharted in the economics literature, **section 3** details the corrections applied to the raw data, while **section 4** compares nowcasting performances across different models.

Section 1: Literature review

Our paper first contributes to the burgeoning literature on alternative high-frequency data. In the wake of the Covid-19 crisis, a number of new datasets have emerged such as daily credit card spending (Carvalho *et al.*, 2020), daily housing online listings (Bricongne *et al.*, 2021) or

hourly electricity consumption (Chen *et al.*, 2020).[2] This paper first adds to this list, but in a greater contribution, it goes a step further in two directions. The first one is related to the global coverage of satellite-based pollution data, uniform across countries. This is not the case in most of the literature which generally focuses on a particular geography. Even for data with a wide-spectrum coverage – e.g. Google Trends or Google mobility data, a number of countries remain missing (notably China due to Google's ban) and the coverage is of heterogeneous quality depending on Google's market penetration (in particular for developing economies with a subpar Google's penetration).[3] In contrast, our data cover the entire world with a uniform quality. The second contribution is that besides delivering an innovative indicator, this paper explores to which extent such data enhance the real-time forecasting as well as the detection of turning points in economic activity.

In that vein, this paper contributes to the on-going debate on whether high-frequency data improve nowcasting performances. The general trade-off between timeliness and accuracy (Ahnert and Bier, 2001) makes it non-trivial since such data, while highly timely, can prove noisy. The INSEE (2020) has tried augmenting models with high-frequency data but did not generally report an improvement in out-of-sample performances. On the other hand, Ferrara *et al.* (2020) show that a nowcasting of the US economy based on high-frequency data provided more reliable forecasts than others based on standard monthly indicators. This paper contributes with a cross-country effort. In this endeavour, we are close to Woloszko (2020) who uses Google trends to predict GDP growth in OECD countries. More generally, this paper contributes to the literature on forecasting industrial production by exploring the potential of pollution data while the bulk of such literature relies on survey data and particularly on Purchasing Managers' Index (PMIs).[4]

Our paper also contributes to the literature using NO_2 pollution to observe economic events. A number of papers have shown that economic growth increases NO_2 pollution in the longer-run (e.g. Cui *et al.*, 2020) including at city level (Zhou *et al.*, 2018). While this literature points to a sense of causality in which economic activity drives pollution, this link can conversely be used in to detect large economic recessions that lead to a drop in NO_2 pollution. Some have documented the fall in pollution due to the Great Financial Crisis (GFC) in Europe (Boersma and Castellanos, 2012), the US (Russell *et al.*, 2012), or China (Du and Xie, 2017). Similar studies have been conducted during the Great Lockdown over China (Le *et al.*, 2020) and Europe (e.g. Tobias *et al.*, 2020) but also for less developed economies such as Kazakhstan (Baimatova *et al.*, 2020). Keola and Hayakawa (2021) also document the changes in NO_2 pollution following

[2] It is also worth mentioning Chetty *et al.* (2020) who have provided a vast amount of data on US economic activity – which were privately held until now. For France, Bricongne *et al.* (2020) have compiled a number of innovative indicators based on alternative data available for free and at higher frequency than official statistics.

[3] It might also be added that Google Trends data are by design prone to substantial revisions in subsequent data releases given the scaling of the series between 0 and 100 (its maximum). Pollution data are, on the contrary, not subject to revisions.

[4] This strand of the literature is developed in greater details in **section 4** as we explain our choice of benchmark models.

lockdown measures.[5] In a further effort, Franke *et al.* (2009) have shown how shipping lanes can be detected with satellite data while de Ruyter de Wilt *et al.* (2012) have documented how the fall in world trade during the GFC resulted in a 12-36% decline in NO_2 pollution over those shipping lanes. Albeit some have investigated the link between PMIs and NO_2 pollution (e.g. Diamond and Wood, 2020), we are not aware of papers which would have the causality and explored how NO_2 pollution could predict industrial production or other macroeconomic indicators. The bulk of this literature falls indeed outside the economics field; our paper then bridges a gap between this literature in geo-physical sciences and economics.

Our paper more broadly relates to the literature using satellite data for economic analysis. Among this rich literature (see Donaldson and Storeygard, 2016 for an extensive review), our paper is closely related to studies using night lights to measure economic activity. Following the seminal work of Henderson *et al.* (2012), several studies have used these data to provide alternative measures for GDP and GDP-per-capita in particular for countries whose statistical system is deficient (Keola *et al.*, 2015; Jean *et al.*, 2016; Pinkovskiy and Sala-i-Martin, 2016) or to track the impact of peculiar events such as India's demonetization (Chodorow-Reich *et al.*, 2020). The relationship also exists at more local level (Montgomery and Holloway, 2018). Recently, night lights have been used to examine the impact of Covid-19 crisis in India (Beyer *et al.*, 2021). Compared to this literature, our contribution is twofold. First, using NO_2 pollution data allows us to achieve higher frequency as such data are available daily *vs.* monthly at best for night lights. Second, and more important, Hu and Yao (2019) recently showed that the relationship between night lights and GDP depends on the stages of development and estimate a null elasticity for advanced economies – similar to the empirical findings of the World Bank (2017). In contrast, evidence in this paper suggests that the elasticity of industrial production to daily pollution is still significant for advanced countries, and that such data allow for an enhanced nowcasting of economic conditions including in those countries.

Section 2: Why satellite data?

Data come from satellite observations of tropospheric pollution made by the TROPOMI instrument, on-board the Sentinel-5P satellite launched by the European Spatial Agency (ESA). Technically, pollution data are obtained through a nadir-viewing (in which satellites can observe the atmosphere by looking straight down rather than diagonally), imaging spectrometer covering wavelength bands between the ultraviolet and the shortwave infrared. To attain an assessment of tropospheric pollution, the instrument uses passive remote sensing techniques by measuring the solar radiation reflected by and radiated from Earth. Sentinel-5P

[5] While the authors of this study document a fall in NO_2 pollution in emerging economies, they don't find similar evidence over the developed economies. It should be however noted that, in contrast with our paper, the authors don't apply weather-correction to the remote data – which might partly explain to some extent differences with our results.

covers the entire surface of Earth each day in Sun-synchronous orbit, making its daily passing at the same time (around 13:35 local Sun time, enabling to capture homogenously the middle of the day which appear intuitively relevant to follow industrial production in real time) above each location. Finally, it should also be noted that a number of papers have been validating the quality of TROPOMI pollution data against other well-tested measurement techniques (see e.g. Wang *et al.*, 2019).

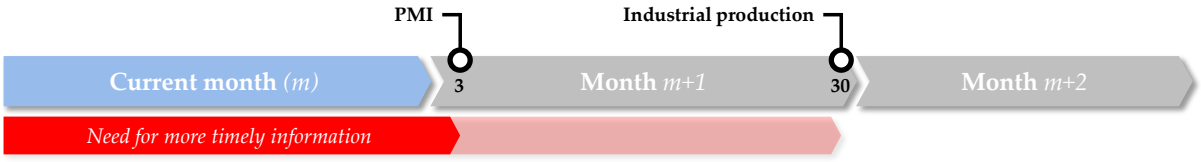
The TROPOMI instrument measures the concentration of various pollutants in the troposphere but we focus on nitrogen dioxide – abbreviated as NO_2 – which appears to be the most appropriate for our objective. First, its duration period in the troposphere is relatively small compared to other pollutants (Lamsal *et al.*, 2011), making it relevant for a real-time monitoring. Second, NO_2 is emitted by the combustion of fossil energies in industrial sites, transportation, and coal-fired power plants – all strongly correlated with industrial activity. Lastly and importantly, NO_2 is a precursor for other pollutants (Henneman *et al.*, 2017) and hence, particularly suitable as an "early" indicator.

Other data sources on pollution would however have been available, but our choice of turning to TROPOMI's is directed by the key advantages of global coverage, uniqueness of the sensor and enhanced precision. The former two mark a key difference with observations from ground sensors as satellite data reach a global coverage including over developing countries where ground sensors are scarce. The uniqueness of the sensor has also the double advantage of: *(i)* alleviating the risk of idiosyncratic errors due to the malfunction of some of the ground sensors; and *(ii)* delivering a uniform coverage over an area while data from pooled ground sensors would depend on the arbitrary choice in the location of these sensors. This is in line with conclusions of the geophysical literature demonstrating that for large-scale studies, it is more efficient to use remote sensing than ground observations despite the higher accuracy of the latter (Remer *et al.*, 2005). Finally, the enhanced precision is the reason why we have selected data from TROPOMI over other satellite measurements – in particular NASA's OMI which has been affected by various malfunctions over the past years (Wang *et al.*, 2020). TROPOMI data have the additional advantage of unprecedented spatial resolution (7×3.5 km^2); one of the reasons why studies have highlighted its higher quality over competing satellite data (Griffin *et al.*, 2019).

In our effort to monitor industrial production in real-time, the greater timeliness offered by such daily data is key as represented in **figure 2** (example for Japanese statistics). At any month *m*, the first estimation of industrial production is published only after 20 to 30 days of month *m+1*. Similarly, PMIs – which are the standard indicator for industrial activity – for a month *m* are not published before the first days of month *m+1*. Therefore, during month *m*, no indicator is available to assess in real-time the fluctuations of industrial production. While this might not be an issue in "normal" times where economic conditions remain broadly stable over time, this might be more of an issue when a sudden crisis occurs.

Figure 2. Timeline: standard macroeconomic indicators for industrial production

Sources: Thomson Reuters Datastream, IHS Markit, authors



Example for Japan – March 2020

Section 3: Data retrieval, cleaning, weather-normalization, and aggregation

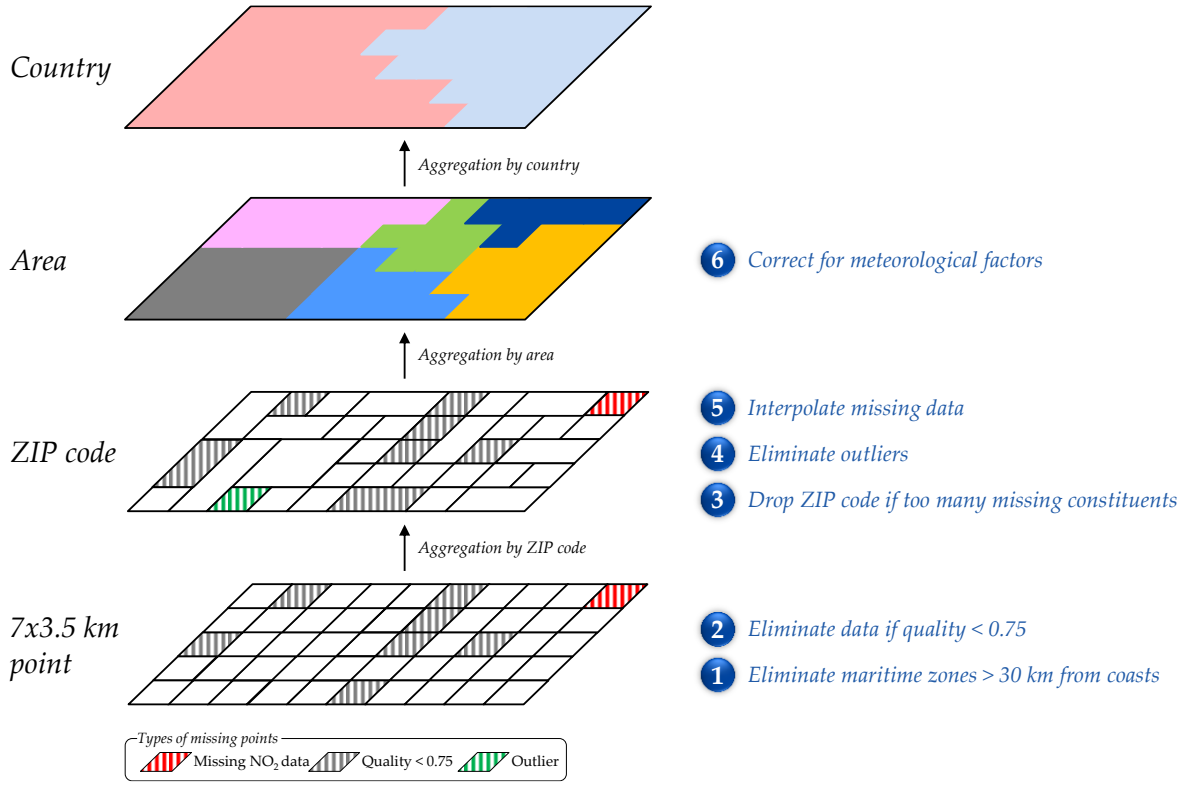
3.1. Data retrieval

The first step relates to fetching the data and making it exploitable considering the large volume of raw data (3-4 Gb per day). The retrieval process is automatized on a deviant server automatically fetching data from the CREODIAS platform. Among the various files that are retrieved, the variables of interest for our study include the NO_2 concentration for each data point of size 7×3.5 km² along with the coordinates for the centroid of the data point (longitude and latitude) and an index of data quality. The latter is key since the presence of snow, clouds, or changes in solar zenith and viewing angles can significantly alter TROPOMI's estimations (Wang *et al.*, 2020). As TROPOMI was launched only very recently, our dataset starts as of December 2018.

Once the data obtained from the ESA, multiple steps are applied to clean, interpolate, correct for weather and seasonal factors, and aggregate raw data at a level comparable to industrial production statistics (i.e. country level). Those steps are shown in **figure 3**.

Figure 3. Steps for data cleaning, interpolation, correction and aggregation

Source: authors



The first step relates to checking the overall consistency of the data and delivering a file that can be more easily exploited in further analysis. This is performed upon retrieving as the retrieval program checks for the consistency of entering files – notably whether NO_2 data are available – and flattens the data. Based on longitude and latitude, the algorithm detects land areas and automatically tags as "sea" the points that are more than 30 km away from coasts. It also deletes observations whose quality index stands below 0.75 – as per ESA recommendations. Finally, based on the latitude and longitude of the data point, observations are merged into a single municipality (i.e. a ZIP code) using a Python package providing the corresponding ZIP code for any set of coordinates. Data are aggregated by computing the unweighted average of all corresponding observations for a ZIP code. Doing so, the program adds a variable counting the number of data points included in the ZIP code average. All-in-all, from 3-4 Gb of raw data retrieved each day, we get a 10-20 Mb single *csv* file more easily exploitable thereafter. A guide to obtain the raw data, as well as a description of such data, is provided in **Technical appendix** – data and scripts can be retrieved from GitHub.[6]

[6] https://github.com/thomaspical/Sentinel5_NO2

3.2. Data cleaning

The second step, at ZIP code level, relates to data cleaning and interpolation. The purpose is to obtain complete NO_2 pollution series for all ZIP codes. First, an observation at a date t is removed if the number of non-missing data points for a ZIP code is too low – more precisely below a third of the number of constituents.[7] This alleviates potential concerns over composition effects at ZIP code level. We then remove outliers by winsorizing at 10% for each ZIP code. As a result, the share of missing data – missing from start, eliminated due to poor quality, or removed because the ZIP code had too few non-missing data points – can be relatively high, around 50%. To avoid composition effects that might arise upon aggregating at area level, we extrapolate missing points.

This topic is a whole literature strand in geo-physical sciences with methods ranging from linear extrapolation (Zhang *et al.*, 2017) to gradient boosting (Chi *et al.*, 2020), random forest (Bi *et al.*, 2019) or neural networks (Fouladgar and Främling, 2020). Following Peng *et al.* (2020), three streams can be identified in the literature depending on whether the proposed method relies on: *(i)* external data such as ground observations, *(ii)* spatial correlation in the data, or *(iii)* both spatial and temporal correlations. The latter two share the idea that missing data can be computed as a weighted average of the measurements at surrounding observations. Among them, spatial "kriging" is a very popular and accurate method (Laslett, 1994) which has been extended into the spatiotemporal dimension (e.g. Tadic *et al.*, 2017; Shao *et al.*, 2020) with "spatiotemporal kriging" considered state-of-the-art statistical methods. Another class of "gap-fill" algorithms, which interpolate missing data based on closest spatial or temporal neighbours, have also been developed (e.g. Weiss *et al.*, 2014).

Due to our high frequency and global coverage, interpolation faces daunting computational costs. "Kriging" methods are in particular impractical to large datasets due to the need to invert the spatial covariance matrix (Kianan *et al.*, 2021).[8] We empirically verify that the same goes with "gap-fill" algorithms using the R package developed by Gerber *et al.* (2018). As such, a number of papers which focus on larger scales and higher-frequency (daily or infra-daily) have used simpler techniques like linear interpolation or local average and have found those of sufficient accuracy (e.g. Noor *et al.*, 2006; Hirabayashi and Kroll, 2017). On the other hand however, numerous studies – such as Yang and Hua (2018) which compares a broad range of interpolation methods – have shown that spatiotemporal methods yield better performances.

To balance those considerations, we implement a K-Nearest Neighbour (KNN) algorithm that takes into account both the spatial and temporal dimensions. The use of KNN for geo-

[7] E.g. if a municipality fills 10 7×3.5 km^2 data points, all observations for which the number of non-missing data points is below 4 will be excluded.

[8] While a "lattice kriging" has been developed to overcome this issue over large datasets (Nychka *et al.*, 2015), it remains important when dealing with world data at daily frequency.

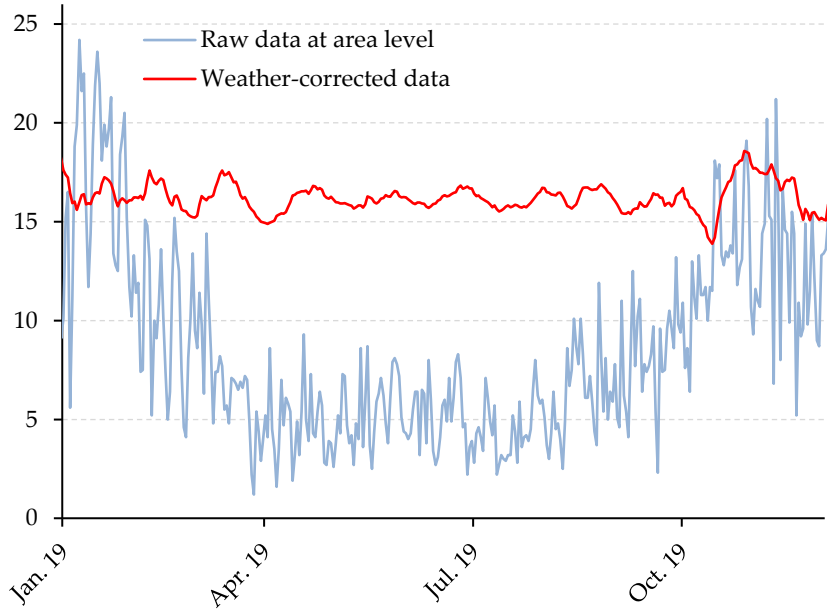
statistical interpolation has notably been proposed by Poloczek *et al.* (2014) which have shown its superior performances *vs.* comparable spatial methods. Under this algorithm, the value of a missing observation is replaced by a weighted average of its *K* closest neighbours.[9] To account for temporal dependency, we introduce time as one of the dimension in our algorithm – meaning that two points can be "neighbours" either spatially or temporally, or a combination of both. In our application, spatial and temporal distance are equally weighted. The KNN algorithm has the advantage of a relatively low computational cost while keeping a spatiotemporal approach. Once this algorithm has been applied and all ZIP code series have been interpolated, data are aggregated at area level using a weighted average of ZIP code series with the weight corresponding to the maximum number of data points belonging to this ZIP code.

3.3. Weather correction

The third step, at area level, relates to weather normalization. Air pollution is indeed very sensitive to weather which not only affects the chemical process of pollutant formation, but also influences human polluting behaviours – making weather normalization a long-lasting topic in the literature (Rao and Zurbenko, 1994). The influence of weather, wind and temperature in particular, will often be greater than the effect of policy interventions or economic events thereby complicating attempts to isolate such effect (Anh *et al.*, 1997). Alix-Garcia and Millimet (2021) present how weather can distort the inputs of remote data to the extent that it can result in misleading policy conclusions. Increased heating and usage of cars will for example result in higher pollution during winter – a pattern clearly visible on **figure 4** (raw data in blue for the Grenoble area, France – known for experiencing large temperature variations over the year).

[9] K (the number of neighbours considered) is the key hyper-parameter in KNN algorithm which needs to be set by the user. As is standard in the literature, the calibration is based on a 10-fold cross-validation with the optimal K chosen as the one which minimizes the out-of-sample RMSE on the evaluation sample. In the case of this paper, due to the multiplicity of locations, the computation is made over several representative areas accounting for a large number of ground types (seaside, mountains, plains, etc.), climates (mild, tropical, desertic, etc.), and economies (advanced, emerging, developing). Empirically, this approach suggests an optimal number of K=26 neighbours.

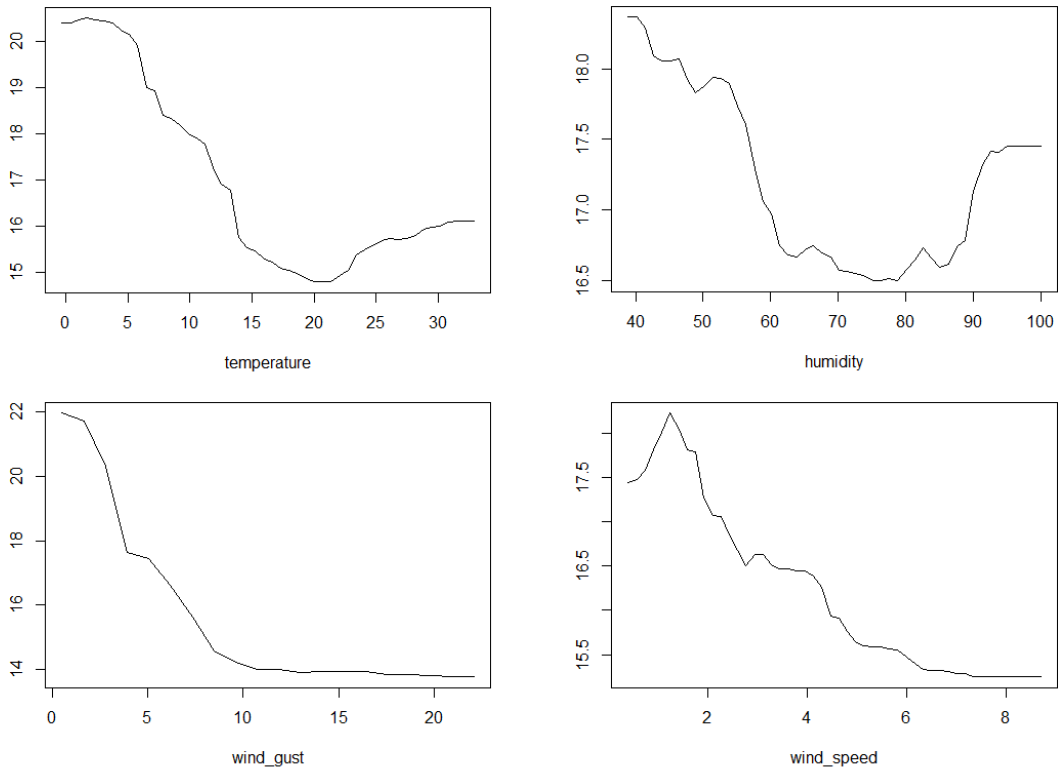
Figure 4. NO_2 pollution in 2019 (Grenoble, France)
Sources: ESA, authors' calculations



We correct for weather using a random forest algorithm which allows to take into account non-linearities. Weather-normalization is also a whole strand of the literature in which the application of random forests has been pioneered by Grange and Carslaw (2018) and has then received a large attention due to its versatility and ease, followed by various applications and refinements (e.g. Vu *et al.*, 2020). Other techniques exist, from classic linear regressions (Henneman *et al.*, 2015) to non-parametric methods (Libiseller *et al.*, 2005), gradient boosting (Petetin *et al.*, 2020), or neural networks (Gardner and Dorling, 2000). We turn to random forest over other techniques since: *(i)* it makes it possible to account for non-linearities and interactions between meteorological variables that are well documented in the literature (Barmpadimos *et al.*, 2011); *(ii)* compared to other machine learning algorithms, random forest are relatively not sensitive to the tuning of hyper-parameters (Biau and Scornet, 2016) which makes it highly versatile when applying it automatically to a broad range of areas,[10] and *(iii)* have a rather limited computational cost. **Figure 5** shows partial dependency plots that bring evidences of non-linearities in our case.

[10] We therefore calibrate the hyper-parameters of the random forest (number of trees, maximum number of features considered for splitting a node, maximum number of levels in each decision tree) using 5-fold cross-validation on a few number of heterogeneous areas, and then apply those hyper-parameters to all our sample.

Figure 5. Partial dependencies plots for meteorological variables
Source: ESA, WAQI, NOAA, authors' calculations



Our choice to apply weather normalization at area level – rather than at ZIP code or country level – results of a trade-off between computational costs, availability of weather data, and need to correct for weather at a granular level. Grange *et al.* (2018) among others have shown that weather normalization varies depending on the topology of the area under consideration (e.g. presence of mountains, forests, wind corridors, etc.) and on the location of polluting sources in that area. In addition, as meteorological variables can display important variations across regions, a robust weather normalization should be performed at local level (Liu *et al.*, 2020). This is however balanced by the availability of weather data that can be retrieved for a limited number of points over a country.[11] This is also balanced by the need to keep reasonable computational costs – in a set-up covering the globe at daily frequency.

At the end, weather normalization is made using for explanatory variables: humidity, temperature, wind speed, and wind gust. We also add day-of-week dummies for Saturday and Sunday – as well as on bank holidays – to account for the reduction of industrial activity during those days. Following Derwent *et al.* (1995) that showed that pollutant emissions or chemical processes may greatly vary by season, we also introduce monthly dummies. To a certain extent, this also allows for a correction of monthly seasonality in the data. We regress NO_2 pollution over these variables and calibrate the algorithm over pre-Covid-19

[11] For example, around 30 for France in the dataset from the World Air Quality Index (WAQI) used in our analysis for advanced economies, but also in the dataset from the National Oceanic and Atmospheric Administration (NOAA) that we use for countries and regions not covered by the WAQI.

observations (i.e. 2018-2019). Using the coefficients that are determined, we then apply this procedure for 2020 and 2021 observations. The weather-corrected series is the part of NO_2 pollution left unexplained by these variables, i.e. the residual of the regression. An example of the resulting weather-corrected series is given for the Grenoble area in **figure 4** (in red). Further results for various areas are shown in **Technical appendix**. Finally, data for all areas are aggregated at country level through a weighted average – with weights being the total number of data points belonging to the area.

Section 4: Nowcasting industrial production

We now turn to whether pollution data can enhance the real-time forecasting of industrial production. To this aim, we build a nowcasting model using daily pollution data and compare its performances to two benchmarks: *(i)* an AR(1) model; and *(ii)* a model based on Purchasing Managers' Index (PMIs[12]). PMIs are widely used in the literature to forecast industrial production (Bruno and Lupi, 2003; Tsuchiya, 2014) not only in advanced economies but also more recently in emerging markets such as Turkey (Akdag *et al.*, 2020) or India (Herwadkar and Ghosh, 2020). This index is widely used as an early indicator in forecasting models (Lahiri and Monokroussos, 2013) as well as in policy briefs. Studies have shown that nowcasting models incorporating PMIs have better performances than other comparable models (Bulligan *et al.*, 2010; d'Agostino and Schnatz, 2012).

4.1. Comparing nowcasting performances

We compare out-of-sample root mean squared errors (RMSE) across models and mimic a real-time exercise for nowcasting industrial production at month *m*. At each day *d* of this month *m*, we estimate the model up to month *m-1* and then forecast industrial production for month *m* using all information available at day *d*. For the pollution-based model, we then use daily pollution data available up to day *d-1*. In order to account for the frequency mismatch between daily pollution and monthly industrial production, we average pollution data over the last 31 days.[13] Industrial production and pollution are taken in log differences.[14] The PMI-based model also relies on the latest PMI data available at day *d*, i.e. the PMI for month *m*-1. We take

[12] This monthly index is constructed as a diffusion index on purchasing managers' answers to how activity in their firms has changed this month compared to the previous one. This index covers around 40 geographies – making it largely appropriate for our cross-country purpose.

[13] For the sake of simplicity, we allocate 31 days for every month. For those months with fewer days, the last points (the 31st in most cases; the 29th, the 30th, and the 31st for February) are linearly interpolated. While some literature rather advises to allocate 30 days to each month (e.g. Ollech, 2018), choosing 31 days allows for not losing information pertaining to the last day of the month.

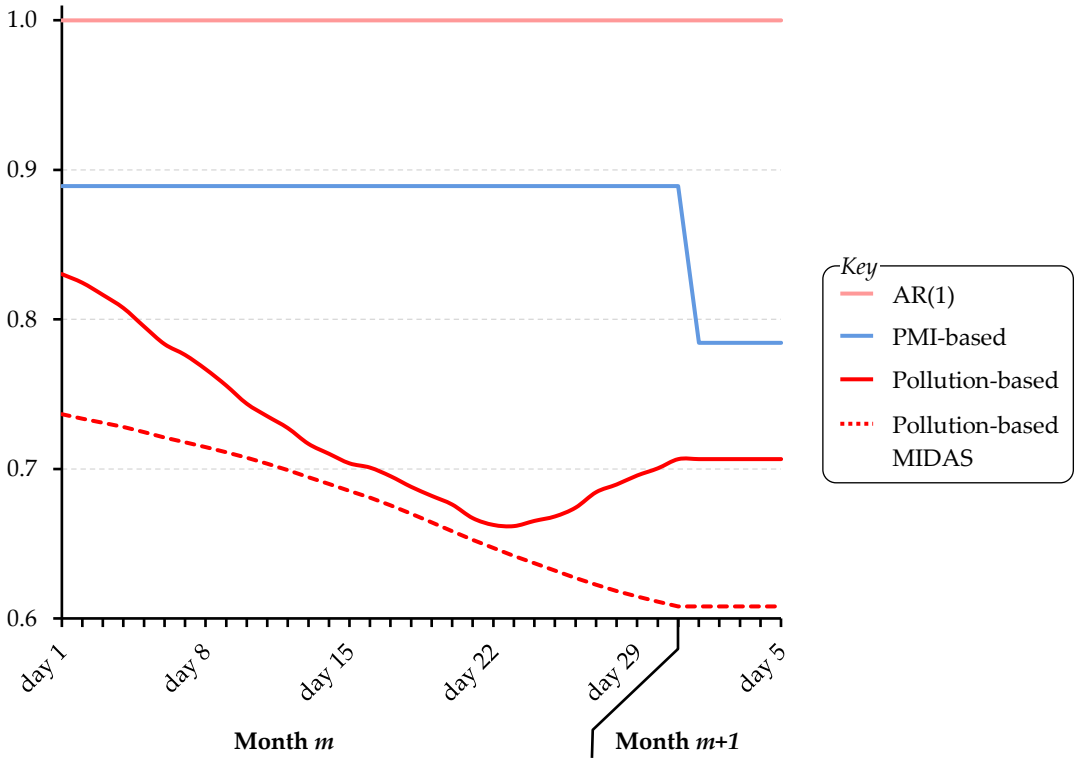
[14] For pollution data, this is the log difference of the average over the last 31 days relative to the average on the 31 days beforehand (i.e. *d-32* to *d-63*).

a panel approach on the 33 economies (16 advanced[15] and 17 emerging[16]) for which pollution data, PMIs, and harmonized statistics for industrial production are all available. To account for time-invariant patterns[17] such as the share of polluting industries, traffic pollution over cities, etc. we include country fixed effects (f_i). We cluster standard errors by country. Formally, our panel specification for nowcasting industrial production (*IPI*) in countries i at month m using pollution data (*Pol*) available at a day d is the following:

$$dlog(IPI_{i,m}) = \alpha_0 + \alpha_1 dlog\left(\sum_{k=1}^{31} Pol_{i,d-k}\right) + f_i + \varepsilon_{i,d}$$

We run these out-of-sample one-period ahead forecasts from March 2020 to December 2020. The results are shown in **figure 6** which compares out-of-sample RMSE of PMI-based and pollution-based models relatively to the AR model. The model based on daily NO_2 pollution over-performs the two benchmark models for all days of the month and displays up to a 30% improvement in out-of-sample RMSE compared with an AR model.

Figure 6. Out-of-sample RMSE relative to the AR(1) model
Source: authors' calculations



[15] Advanced economies: Australia, Austria, Canada, Denmark, France, Germany, Greece, Ireland, Israel, Italy, Japan, New Zealand, Spain, Switzerland, the United Kingdom, and the United States.
[16] Emerging economies: Brazil, China, Colombia, Czech Republic, India, Indonesia, Kazakhstan, Korea, Malaysia, Mexico, Philippines, Poland, Russia, Taiwan, Thailand, Turkey, and Vietnam.
[17] Given the relatively limited timespan – our data start in December 2018 – those factors can be considered time-invariant.

There is clear evidence that nowcasting performances are enhanced as more daily data become available for the corresponding month *m*. A peculiarity lies however in the fact that performances slightly deteriorate after the 23th day, a pattern that we can't empirically attribute to a particular outlier in terms of country or time period. Potential explanations include the fact that the first days of a month might matter more for month-on-month growth than the last days – in a similar way that the first quarters of the year are known to contribute more to the year-on-year growth rate than the last ones. Such a scheme enters in contradiction with our unweighted average of daily data over the last 31 days. Adopting instead a weighted average might therefore eliminate this peculiarity.

We therefore consider a MIxed DAta Sampling (MIDAS) approach which allocates different weights to the different lags of the high frequency regressors. Following the seminal work of Ghysels *et al.* (2004), numerous papers have shown that such set-up performs better than a flat aggregation when dealing with a frequency mismatch between a lower-frequency dependent variable and a higher-frequency explanatory variable. We build on the recent advances of Khalaf *et al.* (2021) who have been the first to propose a panel-MIDAS. To allocate the weights, we use a beta-weighting function, introduced by Ghysels *et al.* (2006), as it allows for a broad range of shapes – which seems particularly important as our hypothesis above suggests that weights would be rather different for a MIDAS in the first days of the month *vs.* a MIDAS in the last days of it. Formally, the beta-weighing function g attributes a weight to lag k depending on two parameters θ_1 and θ_2

$$g(k, \theta_1, \theta_2) = \left(\frac{1}{k+1}\right)^{\theta_1 - 1} \cdot \left(1 - \frac{1}{k+1}\right)^{\theta_2 - 1} \Bigg/ \sum_{p=0}^{K} \left(\frac{1}{p+1}\right)^{\theta_1 - 1} \cdot \left(1 - \frac{1}{p+1}\right)^{\theta_2 - 1}$$

Results are visible in **figure 6** where the dashed red line represents the out-of-sample RMSE obtained when using a MIDAS set-up for each day d of the month. Performances are even more greatly enhanced compared to our previous specification with accuracy gains that reach up to 40% relatively to the AR model. Importantly, there is no decline in performances after day 23.

4.2. Exploring heterogeneities

However, the literature suggests that while high-frequency data enhance forecasting performances during "crisis" episodes, such data are only of second order during normal episodes (Jardet and Meunier, 2020). Given our limited timespan, the real-time out-of-sample above starts only in March 2020. To test for this heterogeneity however, our sample does not allow to robustly break down the sample between "crisis" and "non-crisis" episodes[18], we

[18] We run this type of exercise taking the period from October 2019 until January 2020 (both included) as the "non-crisis" period. Despite the aforementioned limits on robustness, the out-of-sample RMSE for the model with daily

therefore run regressions by quartile depending on how severely a country was affected by the Covid-19 crisis. We break down our sample based on the larger month-on-month slump in industrial production growth over 2020, which ranges from -47.9% (India in April 2020) to -0.9% (Australia in June 2020).

Results are presented in **table 1** which compares the out-of-sample RMSE for pollution-based (without MIDAS) and PMI-based models, relative to the AR model, across all four quartiles. RMSEs for the pollution-based model are averaged over all 31 days d of the month m. Gains are more substantial – around 40% accuracy gains relatively to the AR model – for the most affected countries in the first two quartiles. For countries in the 4th quartile (for which the maximum fall in industrial production ranges from -6.5% to -0.9%) there are no significant accuracy gains relative to the AR model – but there are however large accuracy gains relatively to the PMI-based model. These results would tend to demonstrate that the contribution of high-frequency data to nowcasting performances is more significant during "crisis" episodes with greater improvement if the crisis is harder.

Table 1. Out-of-sample RMSE relative to the AR model

	1st quartile	2nd quartile	3rd quartile	4th quartile
Pollution-based	0.58	0.55	0.80	0.98
Pollution-based (MIDAS)	0.53	0.51	0.77	1.02
PMI-based	0.67	0.60	0.88	1.18

Dependent variable is the month-on-month growth rate of industrial production. Explanatory variable is the month-on-month growth rate of the moving average of weather-adjusted NO_2 pollution (upper two lines) or the PMI (bottom line). Regressions span over Jan. 2019 to Dec. 2020 covering 33 advanced and emerging countries, with 731 daily observations and 24 monthly observations.

Following Hu and Yao (2019), there might also exist heterogeneities depending on the stage of development of the country. Hu and Yao (2019) – who study night lights – linked this with the type of economic growth based on infrastructure spending that takes place in emerging economies and that produce night lights *vs.* the services-based economic growth in advanced economies. The rationale for NO_2 pollution would be similar with the share of polluting industry being higher in emerging countries. To formally test this, we include an interaction term between the share of manufacturing in total value added[19] and pollution. We also check whether there are additional heterogeneities related to the stage of development by introducing a dummy δ_i if the country is an emerging economy, which we put in a triple

pollution data is still better relatively to the AR model (0.74 on average over the month) and, importantly, this pollution-based model out-performs the PMI-based model which stands at 1.07 for the same metrics.

[19] Source: World Bank; we take the latest data point available (2017 or latter).

interaction with pollution and the share of manufacturing. More formally, the panel regression is now (with additional terms for emerging dummy in brackets):

$$dlog(IPI_{i,m}) = \alpha_0 + \alpha_1 dlog\left(\sum_{k=1}^{31} Pol_{i,d-k}\right) + \alpha_2 manuf_i \cdot dlog\left(\sum_{k=1}^{31} Pol_{i,d-k}\right)$$

$$\left[+ \alpha_3 \delta_i^{EME} \cdot dlog\left(\sum_{k=1}^{31} Pol_{i,d-k}\right) + \alpha_4 \delta_i^{EME} \cdot manuf_i \cdot dlog\left(\sum_{k=1}^{31} Pol_{i,d-k}\right)\right] + f_i + \varepsilon_{i,d}$$

Results are provided in **table 2**. The interaction term is negative and significant (column 2). Thus, the higher the share of manufacturing in a country, the greater the elasticity of pollution to industrial production: this elasticity is indeed the inverse of ($\alpha_1 + \alpha_2 \cdot manuf_i$) which becomes greater when the share of manufacturing is greater. In line with Hu and Yao (2019), if follows that the elasticity is greater in emerging countries than in advanced economies. As such, for a country like France with a share of manufacturing equal to around 10% in value added, this elasticity is 2.8 *vs.* 4.5 for China where this share reaches 27%.[20] Including a triple interaction with dummy δ_i accounting for the stage of development (column 3) does not yield significant results, suggesting that taking into account the importance of manufacturing in the economy is sufficient to account for differences in the means of production across countries. It finally should be noted that the R^2 obtained with pollution data (0.25) is about the same as the one obtained by Henderson *et al.* (2012) with night lights (0.21).

[20] Taking the coefficient estimates from column 2 of **Table 2** since the interactions with the emerging dummy is found non-significant in the column 3.

Table 2. Interactions coefficients

	(1)	(2)	(3)
$constant$	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)
$dlog(\sum_{j=1}^{31} Pol_{i,t-j})$	0.309*** (0.021)	0.431*** (0.069)	0.502*** (0.125)
$dlog(\sum_{j=1}^{31} Pol_{i,t-j}) \cdot manuf_i$		-0.770* (0.439)	-2.048** (0.920)
$dlog(\sum_{j=1}^{31} Pol_{i,t-j}) \cdot \delta_i^{EME}$			0.164 (0.157)
$dlog(\sum_{j=1}^{31} Pol_{i,t-j}) \cdot manuf_i \cdot \delta_i^{EME}$			0.517 (1.064)
Country fixed effects	Yes	Yes	Yes
Adjusted R^2	0.25	0.26	0.30
Number of countries	33	33	33
Number of daily observations	731	731	731

*Dependent variable is the month-on-month growth rate of industrial production. Standard errors in brackets are clustered by country. *, ** and *** denote statistically significant coefficient estimates at respectively 10%, 5% and 1% significance levels. Regressions over Jan. 2019 to Dec. 2020.*

In the end, this relationship provides a nowcasting model for industrial production across all countries, including those without reliable statistics. As shown above, the advantage of our approach for advanced and emerging economies lies in the possibility to deliver an accurate nowcasting of industrial production – well in advance of the publication of official statistics and with greater accuracy relatively to competing models. In a number of developing economies however, official statistics for industrial production are not available at all: the nowcasting provides an informative indicator for the state of industrial production in those regions. This approach might be also be extended at regional level, where it might provide an index for local industrial production.

4.3. Early detecting of turning points

Besides nowcasting, the purpose of such an indicator would also be to detect in real-time the turning points in economic activity. Our identification strategy relies on the detection of structural breaks, as such a break might indicate the transition from one state (normal state) to another (crisis state). The goal is to see whether daily NO_2 pollution allows for a swift and

correct detection of turning points. Following the seminal work of Hamilton (1989),[21] we use first a univariate Markov-switching model that assesses when data appear to be transitioning between two regimes.

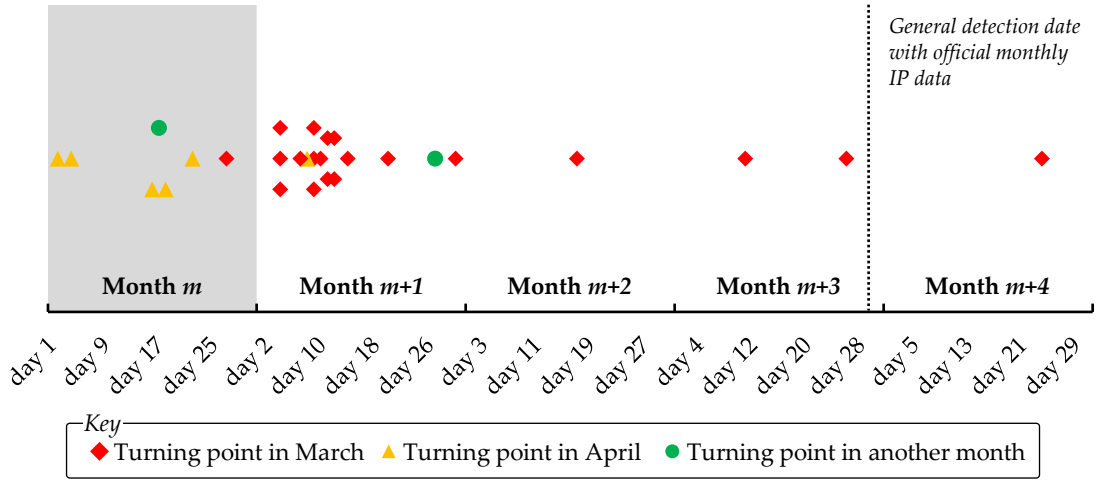
In line with Menezes *et al.* (2006), a key parameter for the real-time detection of turning points is the number *K* of consecutive periods spent in the "crisis" state after which it is appropriate to claim that a structural break has occurred. Fixing an appropriate *K* is important to minimize the risk of false detection – due for example to an outlier. It introduces a trade-off between timeliness and accuracy in the sense that a low (high) *K* will result in swift (lagged) detection but will increase (decrease) the chances of mistaking outliers with a structural break. For monthly industrial production data, we follow Menezes *et al.* (2006) that find $K^m = 3$ months as the most appropriate setting. It means that it would require three consecutive months spent in "crisis" regime by the Markov-switching regression following month *m* to formally identify this month *m* as a turning point. Empirically for daily pollution data, we find that $K^d = 21$ days result in a lower share of false detection while still allowing for swift detection.[22]

We mimic a real-time exercise to detect transition to "crisis" state throughout 2020. We first start by dating *ex-post* the turning point in the business cycle:[23] the corresponding month is represented in grey in **figure 7** and serves as reference for the real-time detection. We then turn to the real-time exercise: going gradually through the year, we assess whether the data available at each day of 2020 would have allowed a univariate Markov-switching regression to detect a transition to the "crisis" state (for at least *K* consecutive periods as defined above). We start using the official monthly industrial production data: for each country and at each data release of official statistics, we assess if the univariate Markov-switching regression would have transited to the "crisis" state for K^m months. When using such monthly data, the crisis is detected around the end of month *m+3* after the "true" start of the crisis – this is represented by the dashed black line. A similar process is performed with the daily NO_2 pollution series: for each day of 2020, we assess whether the univariate Markov-switching regression have detected a structural break (i.e. whether it has transited to "crisis" state for at least K^d consecutive days). The points in **figure 7** represent the date at which the daily NO_2 pollution series would have allowed for a *real-time* detection of this structural break.[24] The detection generally occurs in the beginning of month *m+1* – when official statistics of industrial production are not yet even published. This marks a significant improvement compared to the detection based on official monthly industrial production statistics – which occur only at

[21] In contrast with Hamilton (1989) however, we apply a Markov-switching dynamic model rather than an AR model – as we deal with higher-frequency data (see Guérin and Marcellino, 2013).
[22] We choose empirically K as the number that allows to maximize a signalling ratio defined as the number of true positive over the sum of true positive and false positive.
[23] For this *ex-post* dating, we use a univariate Markov-switching regression on monthly industrial production up to date (as of Feb. 2021).
[24] Note that this is therefore the date d at which the Markov-switching regression first estimates the structural break plus K days.

the end of month *m+3*. For all countries but one,[25] the detection based on daily NO_2 pollution is swifter than the detection based on monthly official data. Finally, the distinction between countries starting their "crisis" state in March (red diamonds), in April (orange triangles), or other months (green circles) show that the detection is even swifter for countries entering in "crisis" state later.

Figure 7. Dates for real-time detection of turning points
Source: authors' calculations



Note: the point indicates at which date the daily NO2 pollution for the country under consideration would have allowed to claim for a real-time detection of this structural break. This is therefore the date d at which the Markov-switching regression first estimates the structural break plus K days.

The main issue with performing Markov-switching regression given our limited timespan – data start only in Dec. 2018 – is that such a framework coerces the existence of two states. Before January 2020, or for countries whose industrial production has been less affected by the Covid-19 crisis, there are however no compelling evidence for the existence of two states. This results in the detection of "false positive" in multiple cases (15 of the 44 countries under consideration) despite fixing a high *K=21*. In addition, such a stringent *K* impairs the detection of turning points in some cases (3 "false negative" out of the 35 countries for which a turning point should have been detected) since NO_2 series are rather noisy, which can prevent the Markov-switching framework to stabilize in the "crisis" state for more than *K=21* days.

We therefore turn to an alternative technique for the real-time detection of turning points by implementing the step-indicator saturation (SIS) algorithm proposed by Castle *et al.* (2015). This algorithm saturates the specification with one step dummy per day – of value 0 before the day and 1 afterwards – and then uses the *Autometrics* software (Doornik, 2009) to select significant step dummies. Those significant step dummies can be interpreted as an indication for regime changes. We combine SIS with an impulse-indicator saturation (IIS) – developed beforehand by Hendry *et al.* (2008) and based on the saturation of time dummy, of value 1 at

[25] Namely Slovakia.

date d and 0 otherwise – with the idea that those time dummies would detect outliers while step dummies would detect regime changes. A number of studies (e.g. Marczak and Proietti, 2016) have shown *Autometrics*' ability to handle such models with much less time observations (T) than variables ($2T+1$) for which it over-performs competing methods (e.g. lasso, stepwise regression), using a combination of expanding / contracting multiple block searches as described in Hendry *et al.* (2008), Doornik (2009) and Hendry and Doornik (2014). Compared to Markov-switching, this algorithm has the advantages of: *(i)* versatility as it does not coerce the regression to a number of states pre-defined by the user; and *(ii)* being able to detect regime changes among noisy data since IIS would detect outliers and distinguish them from regime changes.

We again mimic the real-time exercise. The detection dates are very close to those of the Markov-switching model – and therefore not displayed. The main interest of this alternative approach lies in the fact that the rate of "false positive" (detection of a non-existent turning point) is significantly lower (14% *vs.* 34% in the Markov-switching set-up described above) while the rate of "false negative" (non-detection of turning point) is now 0% *vs.* 9% above.

Conclusion

Our paper shows how satellite data on tropospheric NO_2 pollution can be retrieved, cleaned, interpolated and corrected for weather pattern using machine learning techniques (k-nearest neighbours, random forest). Turning to the econometrics perspective, we show that predictive models for industrial production relying on daily pollution data over-perform comparative standard models based on survey data (PMIs) or an AR term. While the accuracy gain is around 30% *vs.* an AR model when using a simple averaging of daily pollution data at monthly frequency, it can reach around 40% when building on a MIDAS set-up. We finally provide evidence for heterogeneities in the relationship between pollution and industrial production which depends on the share of manufacturing in the value added for a country.

Policy-wise, this paper provides the basis for an alternative daily pollution-based indicator of industrial production. Compared with other potential statistics for industrial production, this indicator has the key advantages of timeliness (daily measurement *vs.* monthly indices for official statistics), global coverage (*vs.* official or alternative statistics often limited to a country of interest and with heterogeneous quality depending on a country's characteristics), granularity, and free access. Evidence suggests that the contribution of high-frequency data is greater during "crisis" periods but only of second order during "normal" times. But evidence also suggests that satellite data allows for a consistent nowcasting of industrial production even in "normal" periods, during which the key advantage of such data might therefore be the ability to provide a measure of industrial production in countries where such official statistics are scarce, or with a shorter delay. Following Henderson *et al.* (2012), it might also

be possible for countries with existing statistics, but of low quality, to combine the official and the pollution-based alternative indices into a single more robust index.

Indeed, with data starting only in December 2018, the limited timespan is arguably a key limit. It impairs most analyses that would focus on a particular country, a limit that we overcome by relying on panel regressions. Another potential key limit might lie in the uniqueness of the sensor. While this feature has advantages for the statistician – measurements are less prone to the arbitrary location of ground sensors or to idiosyncratic failures of some sensors that might derail local measurements, it means that in case of failure, the failure goes global and no reliable NO_2 pollution data can be retrieved. This would remain a threat for practitioners – even though, as the LIBOR indicator has shown, any statistical source may be subject to discontinuity.

Finally, NO_2 pollution represents only a fraction of the data retrieved by satellite. Every day, a number of satellites measure a large range of variables that might be put into use into the economics field. Examples include infrared radiations (which can indicate the heat from industrial sites), volumetric pictures (which can detect construction sites), or even simple pictures (from which the number of ships in ports, or trucks in loading areas can be computed). A number of economic insights can further be derived from these satellite data, for which the advantages highlighted above would remain: high frequency, global coverage with no local bias, granularity, and completely free.

References

d'Agostino A. and Schnatz B. (2012). "Survey-based nowcasting of US growth: a real-time forecast comparison over more than 40 years", *European Central Bank Working Paper*, No 1455

Ahnert H. and Bier W. (2001). "Trade-off between timeliness and accuracy", *Economisch Statistische Berichten*, No 4299

Akdag S., Deran A., and Iskenderoglu O. (2020). "Is PMI a Leading Indicator: Case of Turkey", *Sosyoekonomi Journal*, 28(45)

Alix-Garcia J. and Millimet D. (2021). "Remotely Incorrect? Accounting for Nonclassical Measurement in Satellite Data on Deforestation", *mimeo*
https://people.smu.edu/dmillimet/research/

Anh V., Duc H., and Azzi M. (1997). "Modelling Anthropogenic Trends in Air Quality Data", *Journal of Air & Waste Management Association*, 47, pp. 66-71

Baimatova N., Bukenov B., Ibragimova O. P., Karaca F., Kenessov B., Kerimray A., and Plotitsyn P. (2020). "Assessing air quality changes in large cities during COVID-19 lockdowns: The impacts of traffic-free urban conditions in Almaty, Kazakhstan", *Science of the Total Environment*, 730

Barmpadimos I., Hueglin C., Keller J., Henne S., and Prévôt A. (2011). "Influence of meteorology on PM10 trends and variability in Switzerland from 1991 to 2008", *Atmospheric Chemistry and Physics*, 11(4), pp. 1813-1835

Beyer R., Franco-Bedoya S., and Galdo V. (2021). "Examining the economic impact of COVID-19 in India through daily electricity consumption and nighttime light intensity", *World Development*, 140

Bi J., Belle J., Wang Y., Lyapustin A., Wildani A., and Liu Y. (2019). "Impacts of snow and cloud covers on satellite-derived PM2.5 levels", *Remote Sensing of Environment*, 221, pp. 665-674

Biau G. and Scornet E. (2016). "A random forest guided tour", *TEST*, 25, pp. 197-227

Boersma K. and Castellanos P. (2012). "Reductions in nitrogen oxides over Europe driven by environmental policy and economic recession", *Scientific Reports*, 2(265)

Bricongne J.-C., Coffinet J., Delbos J.-B., Kaiser V., Kien J.-N., Kintzler E., Lestrade A., Meunier B., Mouliom M., and Nicolas T. (2020). "Tracking the economy during the Covid-19 pandemic: The contribution of high-frequency indicators", *Bulletin de la Banque de France*, 231

Bricongne J.-C., Meunier B., and Pouget S. (2021). "Web Scraping Housing Prices in Real-time: the Covid-19 Crisis in the UK", *Banque de France Working papers*, No 827

Bruno G. and Lupi C. (2003). "Forecasting euro-area industrial production using (mostly) business surveys data", *ISAE Working Papers*, No 33

Bulligan G., Golinelli R. and Parigi G. (2010). "Forecasting monthly industrial production in real-time: From single equations to factor-based models", *Empirical Economics*, 39(2), pp. 303–336

Carvalho V., Garcia J., Hansen S., Ortiz Á., Rodrigo T., Rodríguez Mora J., and Ruiz J. (2020). "Tracking the COVID-19 Crisis with High-Resolution Transaction Data", *CEPR Discussion Papers*, No 14642

Castle J., Doornik J., Hendry D., and Pretis F. (2015). "Detecting Location Shifts during Model Selection by Step-Indicator Saturation", *Econometrics*, 3, pp. 240-264

Chen S., Igan D., Pierri N., and Presbitero A. (2020). "Tracking the Economic Impact of COVID-19 and Mitigation Policies in Europe and the United States", *Covid Economics: Vetted and Real-Time Papers*, 36, pp. 1-24

Chetty R., Friedman J., Hendren N., Stepner M., and the Opportunity Insights Team (2020). "How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data", *NBER Working Paper*, No 27431

Chi Y., Wu Z., Liao K., and Ren Y. (2020). "Handling Missing Data in Large-Scale MODIS AOD Products Using a Two-Step Model", *Remote Sensing*, 12(22), 3786

Chodorow-Reich G., Gopinath G., Mishra P., and Narayanan A. (2020). "Cash and the Economy: Evidence from India's Demonetization", *The Quarterly Journal of Economics*, 135(1), pp. 57-103

Cui Y., He S., Jiang L., Kong H., and Zhou H. (2020). "Effects of the socio-economic influencing factors on SO2 pollution in Chinese cities: A spatial econometric analysis based on satellite observed data", *Journal of Environmental Management*, 268

Donaldson D. and Storeygard A. (2016). "The View from Above: Applications of Satellite Data in Economics", *Journal of Economic Perspectives*, 30(4), pp. 171-198

de Ruyter de Wildt M., Eskes H., and Boersma K. (2012). "The global economic cycle and satellite-derived NO2 trends over shipping lanes", *Geophysical Research Letters*, 39, pp. 1-6

Derwent R., Middleton D., Field R., Goldstone M., Lester J., and Perry R. (1995). "Analysis and interpretation of air quality data from an urban roadside location in Central London over the period from July 1991 to July 1992", *Atmospheric Environment*, 29, pp. 923–946

Diamond M. and Wood R. (2020). "Limited regional aerosol changes despite unprecedented decline in Nitrogen Oxide pollution during the February 2020 coronavirus shutdown in China", *unpublished manuscript*

Doornik J. (2009). "Autometrics", In Castle J. and Shephard N. (Eds.), *The Methodology and Practice of Econometrics: Festschrift in Honour of David F. Hendry*, Oxford University Press

Du Y. and Xie Z. (2017). "Global financial crisis making a V-shaped fluctuation in NO2 pollution over the Yangtze River Delta", *Journal of Meteorological Research*, 31, pp. 438-447

Ferrara L., Froidevaux A., and Huynh T-L. (2020). "Macroeconomic nowcasting in times of Covid-19 crisis: On the usefulness of alternative data", *Econbrowser*

Franke K., Richter A., Bovensmann H., Eyring V., Jöckel P., Hoor P., and Burrows J. (2009). "Ship emitted NO2 in the Indian Ocean: Comparison of model results with satellite data", *Atmospheric Chemistry and Physics*, 9, pp. 7289-7301

Fouladgar N. and Främling K. (2020). "A Novel LSTM for Multivariate Time Series with Massive Missingness", *Sensors*, 20(10), 2832

Gardner M. and Dorling S. (2000). "Statistical surface ozone models: an improved methodology to account for non-linear behaviour", *Atmospheric Environment*, 34(1), pp. 21-34

Gerber F., Furrer R., Schaepman-Strub G., de Jong R., and Schaepman M. (2018). "Predicting missing values in spatio-temporal remote sensing data", *IEEE Transactions on Geoscience and Remote Sensing*, 56(5), pp. 2841-2853

Ghysels E., Santa-Clara P., and Valkanov R. (2004). "The MIDAS Touch: Mixed Data Sampling Regression Models", *CIRANO Working Paper*, No 2004-20

Ghysels E., Santa-Clara P., and Valkanov R. (2006). "Predicting volatility: getting the most out of return data sampled at different frequencies", *Journal of Econometrics*, 131(1–2), pp. 59-95

Grange S. and Carslaw D. (2018). "Using meteorological normalisation to detect interventions in air quality time series", *Science of the Total Environment*, 653, pp. 578-588

Grange S., Carslaw D., Lewis A., Boleti E., and Hueglin C. (2018). "Random forest meteorological normalisation models for Swiss PM10 trend analysis", *Atmospheric Chemistry and Physics*, 18, pp. 6223-6239

Griffin D., Zhao X., McLinden C. A., Boersma F., Bourassa A., Dammers E., Degenstein D., Eskes H., Fehr L., Fioletov V., Hayden K., Kharol S. K., Li S.-M., Makar P., Martin R. V., Mihele C., Mittermeier R. L., Krotkov N., Sneep M., Lamsal L. N., ter Linden M., van Geffen J., Veefkind P., Wolde M. and Zhao X. (2019). "High-Resolution Mapping of Nitrogen Dioxide

With TROPOMI: First Results and Validation Over the Canadian Oil Sands", *Geophysical Research Letters*, 46, pp. 1049-1060

Guérin P. and Marcellino M. (2013). "Markov-Switching MIDAS Models", *Journal of Business & Economic Statistics*, 31(1), pp. 45-56

Hamilton J. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57(2), pp. 357-384

Hirabayashi S. and Kroll C. (2017). "Single imputation method of missing air quality data for i-Tree Eco analyses in the conterminous United States", *unpublished manuscript*

Henderson J., Storeygard A., and Weil D. (2012). "Measuring Economic Growth from Outer Space", *American Economic Review*, 102(2), pp. 994-1028

Hendry D., Johansen S., and Santos C. (2008). "Automatic selection of indicators in a fully saturated regression", *Computational Statistics*, 33, pp. 317-335

Hendry D., and Doornik J. (2014). *Empirical Model Discovery and Theory Evaluation*, MIT Press

Henneman L., Holmes H., Mulholland J., and Russell A. (2015). "Meteorological detrending of primary and secondary pollutant concentrations: Method application and evaluation using long-term (2000–2012) data in Atlanta", *Atmospheric Environment*, 119, pp. 201-210

Henneman L., Liu C., Hu Y., Mulholland J., and Russell A. (2017). "Air quality modeling for accountability research: Operational, dynamic, and diagnostic evaluation", *Atmospheric Environment*, 166, pp. 551-565

Herwadkar S. and Ghosh, S. (2020). "Is PMI a good leading indicator of industrial production? Evidence from India", *MPRA Paper*, No 97924

Hu Y. and Yao J. (2019). "Illuminating Economic Growth", *International Monetary Fund Working Paper*, No 19/77

INSEE (2020). "Les données « haute fréquence » sont surtout utiles à la prévision économique en période de crise brutale", *Note de conjoncture de l'INSEE*, June 17th, https://www.insee.fr/fr/statistiques/4513034?sommaire=4473296

Jardet C. and Meunier B. (2020). "Nowcasting world GDP growth with high-frequency data", *Banque de France Working Papers*, No 788

Jean N., Burke M., Xie M., Davis W., Lobel D., and Ermon S. (2016). "Combining satellite imagery and machine learning to predict poverty", *Science*, 353(6301), pp. 790-794

Keola S., Andersson M., and Hall O. (2015). "Monitoring Economic Development from Space: Using Nighttime Light and Land Cover Data to Measure Economic Growth", *World Development*, 66, pp. 322-334

Keola S. and Hayakawa K. (2021). "Do Lockdown Policies Reduce Economic and Social Activities? Evidence from NO2 Emissions", *The Developing Economies*, 59(2), 178-205

Khalaf L., Kichian M., Saunders C., and Voia M. (2021). "Dynamic panels with MIDAS covariates: Nonlinearity, estimation and fit", *Journal of Econometrics*, 220(2), pp. 589-605

Kianian B., Liu Y., and Chang H. (2021). "Imputing Satellite-Derived Aerosol Optical Depth Using a Multi-Resolution Spatial Model and Random Forest for PM2.5 Prediction", *Remote Sensing*, 13(1), 126

Lahiri, K. and Monokroussos, G., (2013). "Nowcasting US GDP: The role of ISM business surveys", *International Journal of Forecasting*, 29(4), pp. 644-658

Lamsal L., Martin R., Padmanabhan A., van Donkelaar A., Zhang Q., Sioris C., Chance K., Kurosu T., and Newchurch M. (2011). "Application of satellite observations for timely updates to global anthropogenic NOx emission inventories", *Geophysical Research Letters*, 38(5)

Laslett G. (1994). "Kriging and splines: An empirical comparison of their predictive performance in some applications", *Journal of the American Statistical Association*, 89, pp. 391-400

Le T., Wang Y., Liu L., Yang J., Yung Y., Li G., and Seinfeld J. (2020). "Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China", *Science*, 369(6504), pp. 702-706

Libiseller C., Grimvall A., Walden J., and Saari H. (2005). "Meteorological normalisation and non-parametric smoothing for quality assessment and trend analysis of tropospheric ozone data", *Environmental Monitoring and Assessment*, 100 (1–3), pp. 33-52

Marczak M. and Proietti T. (2016). "Outlier detection in structural time series models: The indicator saturation approach", *International Journal of Forecasting*, 32(1), pp. 180-202

Menezes Z., McLaren C., Sanden N., Zhanag M., and Black M. (2006). "Timely detection of turning points: Should I use the seasonally adjusted or trend estimates", In *Eurostat Conference on seasonality, seasonal adjustment and their implications for short-term analysis and forecasting*, May 2006

Montgomery A. and Holloway T. (2018). "Assessing the relationship between satellite-derived NO2 and economic growth over the 100 most populous global cities", *Journal of Applied Remote Sensing*, 12(4)

Noor N., Yahayn A., Ralmi N., and Al Bakri Abdullah M. (2006). "The Replacement of Missing Values of Continuous Air Pollution Monitoring Data Using Mean Top Bottom Imputation Technique", *Journal of Engineering Research & Education*, 3, pp. 96-105

Nychka D., Bandyopadhyay S., Hammerling D., Lindgren F., and Sain S. (2015). "A multiresolution Gaussian process model for the analysis of large spatial datasets", *Journal of Computational and Graphical Statistics*, 24, pp. 579-599

Ollech, D. (2018). "Seasonal adjustment of daily time series", *Deutsche Bundesbank Discussion Papers*, No 41/2018

Peng X., Shen H., Zhang L., Zeng C., Yang G., and He Z. (2020). "Spatially continuous mapping of daily global ozone distribution (2004–2014) with the Aura OMI sensor", *Journal of Geophysical Research: Atmospheres*, 121(12), pp. 702-722

Petetin H., Bowdalo D., Soret A., Guevara M., Jorba O., Serradell K., and Pérez García-Pando C. (2020). "Meteorology-normalized impact of COVID-19 lockdown upon NO2 pollution in Spain", *Atmospheric Chemistry and Physics*, 20(18), pp. 11119-11141

Pinkovskiy M. and Sala-i-Martin X. (2016). "Lights, Camera . . . Income! Illuminating the National Accounts-Household Surveys Debate", *The Quarterly Journal of Economics*, 131(2), pp. 579-631

Poloczek J., Treiber N., and Kramer O. (2014). "KNN Regression as Geo-imputation Method for Spatiotemporal wind data", *Advances in Intelligent Systems and Computing*, 299

Rao S. and Zurbenko I. (1994). "Detecting and tracking changes in ozone air quality", *Journal of Air & Waste Management Association*, 44, pp. 1089-1092

Remer L. A., Kaufman Y. J., Tanré D., Mattoo S., Chu D. A., Martins J. V., Li R.-R., Ichoku C., Levy R. C., Kleidman R. G., Eck T. F., Vermote E., and Holben B. N. (2005). "The MODIS Aerosol Algorithm, Products, and Validation", *Journal of the Atmospheric Science*, 62, pp. 947-973

Russell A., Valin L., and Cohen R. (2012). "Trends in OMI NO2 observations over the United States: effects of emission control technology and the economic recession", *Atmospheric Chemistry and Physics*, 12, pp. 12197-12209

Shao Y., Ma Z., Wang J., and Bi J. (2020). "Estimating daily ground-level PM2.5 in China with random-forest-based spatiotemporal kriging", *Science of the Total Environment*, 740

Tadic J., Qiu X., Miller S., and Michalak A. (2017). "Spatio-temporal approach to moving window block kriging of satellite data v1.0", *Geoscientific Model Development*, 10, pp. 709-720

Tobías A., Carnerero C., Reche C., Massagué J., Via M., Minguillón M. C., Alastuey A., and Querol X. (2020). "Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic", *Science of the Total Environment*, 726

Tsuchiya Y. (2014). "Purchasing and supply managers provide early clues on the direction of the US economy: An application of a new market-timing test", *International Review of Economics & Finance*, 29, pp. 599-618

Veronese G., Biancotti C., Rosolia A., Kirchner R., and Mouriaux F. (2020). "Covid-19 and official statistics: A wakeup call", In *8th IMF statistical forum: Measuring the economics of a pandemic*, November 2020

Vu T., Shi Z., Cheng J., Zhang Q., He K., Wang S., and Harrison R. (2019). "Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique", *Atmospheric Chemistry and Physics*, 19, pp. 11303-11314

Weiss D., Atkinson P., Bhatt S., Mappin B., Hay S., and Gething, P. (2014). "An effective approach for gap-filling continental scale remotely sensed time-series", *ISPRS Journal of Photogrammetry and Remote Sensing*, 98, pp. 106-118

Wang C., Wang, T., and Wang, P. (2019). "The Spatial–Temporal Variation of Tropospheric NO2 over China during 2005 to 2018", *Atmosphere*, 10(8), 444

Wang C., Wang T., Wang P. and Rakitin V. (2020). "Comparison and Validation of TROPOMI and OMI NO2 Observations over China", *Atmosphere*, 11(6), 636

Woloszko N. (2020). "Tracking activity in real time with Google Trends", *OECD Economics Department Working Papers*, No 1634

World Bank (2017). "Growth Out of the Blue", *South Asia Economic Focus*, downloadable at: https://openknowledge.worldbank.org/handle/10986/28397#:~:text=At%206.7%20percent%2C%20growth%20is,percent%20over%20the%20medium%20term

Yang J. and Hua M. (2018). "Filling the missing data gaps of daily MODIS AOD using spatiotemporal interpolation", *Science of the Total Environment*, 633

Zhang T., Zeng C., Gong W., Wang L., Sun K., Shen H., Zhu Z., and Zhu Z. (2017). "Improving Spatial Coverage for Aqua MODIS AOD using NDVI-Based Multi-Temporal Regression Analysis", *Remote Sensing*, 9(4), 340

Zhou Z., Zhou Y., and Ge X. (2018). "Nitrogen Oxide Emission, Economic Growth and Urbanization in China: A Spatial Econometric Analysis", *IOP Conference Series Materials Science and Engineering*, 301

Technical appendix

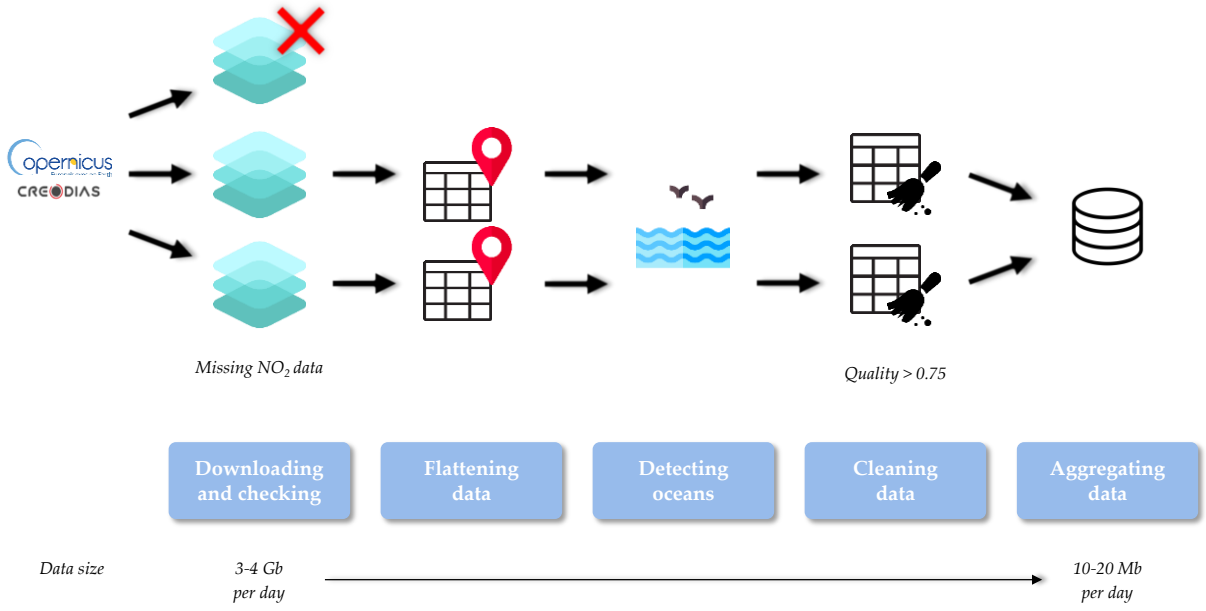
Data availability

The raw data are available on GitHub at https://github.com/thomaspical/Sentinel5_NO2. A complete copy of the GitHub allows obtaining the archived *csv* files – one for each day – of pollution data which form the base data of this study.

In addition to the archived data in a convenient format, the GitHub also includes the script code which would allow updating the database automatically. The program operates the steps defined in **figure A1**: it fetches the data from CREODIAS, flattens it, detects oceans based on latitude and longitude coordinates, eliminates data whose quality falls below the ESA-recommended threshold of 0.75, and aggregates data so that the output file is of a reasonable size (10-20 Mb per day) and in a convenient format (*csv*) for further exploitation.

Figure A1. Steps performed by the script

Source: authors



A daily run of this script would allow keeping the database up to date. If for various reasons the script cannot be executed every day, the program will also recover the missing days. In order to run the script the following Python packages must be installed on your machine: Pandas (v.1.0.3 was used in this project), Numpy (v.1.18.2), Requests (v.2.23.0), json5 (v.0.9.4), zipp (v.3.1.0), netCDF4 (v.1.5.3), and reverse-geocoder (v.1.5.1).

As regards the output *csv* files, observations of NO_2 concentration are aggregated by day and by city at the ZIP code level. While the raw data are provided by the ESA at a 7x3.5 km² level, this aggregation is vital to significantly reduce the size of the *csv* files. Each file (one for each day) is named after the day under consideration (e.g. archived_2019_12_02 relates to the NO_2 pollution data for December, 2nd 2019). **Table A1** provides a description of the variables

included in the *csv* file. Since 7x3.5 km² observations are aggregated at ZIP code level, some of these variables (e.g. longitude, latitude, NO_2) are in fact the average of the constituting observations – this is indicated in the table when this is the case.

Table A1. Variables description

Variable	Description
Year	Year of data measurement
Month	Month of data measurement
Week	Week of data measurement
cc_pays	Country
cc_region	Region
cc_departement	County *(NB: this is not available for some countries, depending on the domestic territorial organization)*
cc_ville	City – ZIP code level
Longitude	Mean longitude of the aggregated observations
Latitude	Mean latitude of the aggregated observations
NO2	Mean nitrogen dioxide concentration per m^3 of the aggregated observations *(NB: this corresponds to the "nitrogendioxide tropospheric column" from ESA data, corresponding to the total air pollution of NO_2 between the surface and the top of the troposphere)*
Quality	Mean quality of the aggregated observations *(NB: before aggregation, individual observations with quality below 0.75 are eliminated, as recommended by the ESA[26])*
Hour_mean	Mean hour of the aggregated observations
Hour_std	Standard deviation of the hour of the aggregated observations
Dayofweek_mean	Mean day of the week of the aggregated observations
Dayofweek_std	Standard deviation of the day of the week of the aggregated observations *(NB: this is a data check, it should be equal to 0)*
Day_mean	Mean day of the aggregated observations
Day_std	Standard deviation of the day of the aggregated observations

[26] See https://sentinels.copernicus.eu/documents/247904/3119978/Sentinel-5P-Level-2-Input-Output-Data-Definition

	(NB: this is a data check, it should be equal to 0)
Counter	Number of individual observations that have been aggregated *(NB: this provides the number of observations aggregated to form data at ZIP code level for this specific day. Due to data availability and quality, the number of observations can be different from one day to the other for the same ZIP code. If the number of constituting points is significantly lower than usual, this is recommended to drop the entire ZIP code to avoid potential composition effects)*

Further details on the raw data are available directly on the ESA website. More specifically, for this study, we relied particularly on:

- https://doi.org/10.5270/S5P-s4ljg54: dataset specification of the TROPOMI Level 2 Nitrogen Dioxide total column products
- https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-NO2-dataproducts: ESA TROPOMI documentation
- https://sentinels.copernicus.eu/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Nitrogen-Dioxide: NO2 product user manual

Individual results

This section illustrates the resulting NO_2 pollution series for some countries, once cleaned, interpolated, corrected for weather and seasonal patterns, and aggregated at country level – as described in **section 3.3**. **Figure A2** shows examples for some key emerging countries (China, India) as well as for some advanced economies (Italy, Spain). The fall in NO_2 pollution is marked following the implementation of lockdown measures across these countries. The slow recovery thereafter is also visible, with tentative evidence that the recovery had been somewhat quicker in China than in other regions, where economic activity apparently remained below pre-crisis level even by the end of 2020 (e.g. in Spain).

Similar pattern can be observed for individual regions. **Figure A3** shows how NO_2 pollution has felt in some regions particularly affected by the Covid-19. The Hubei province (capital: Wuhan; identified as the epicentre of the pandemics) has a coincident fall in pollution as the lockdown is implemented (grey area). The slight lead in the decline of pollution (around mid-January) compared with the implementation of the stringent lockdown (end-January) might be attributed to Chinese New Year falling on the 25 January and inducing a prolonged period of factory shutdowns around this date. The slight lead over the implementation of the lockdowns in other regions (e.g. in India) might be attributed to voluntary distancing spontaneously taken by the population in the context of strong virus circulation. The

Lombardy region in Italy (capital: Milan; where the first large outbreak in Western Europe was detected) presents a similar pattern with bottom-low air pollution during the lockdown period. Even in regions that were severely affected only after the two aforementioned front-runners (e.g. California or São Paulo state), the decline following the implementation of restrictions remains highly visible. For California, the prolonged low pollution might be explained by the re-imposition of stringent lockdown measures throughout 2020 – e.g. most indoor activities (bars, restaurants, museums, gyms) were closed again in the beginning of July 2020 after being able to re-open mid-May 2020.

Figure A2. NO_2 pollution level over 2019-2020 (% change, month-on-month)
Sources: ESA, authors' calculations

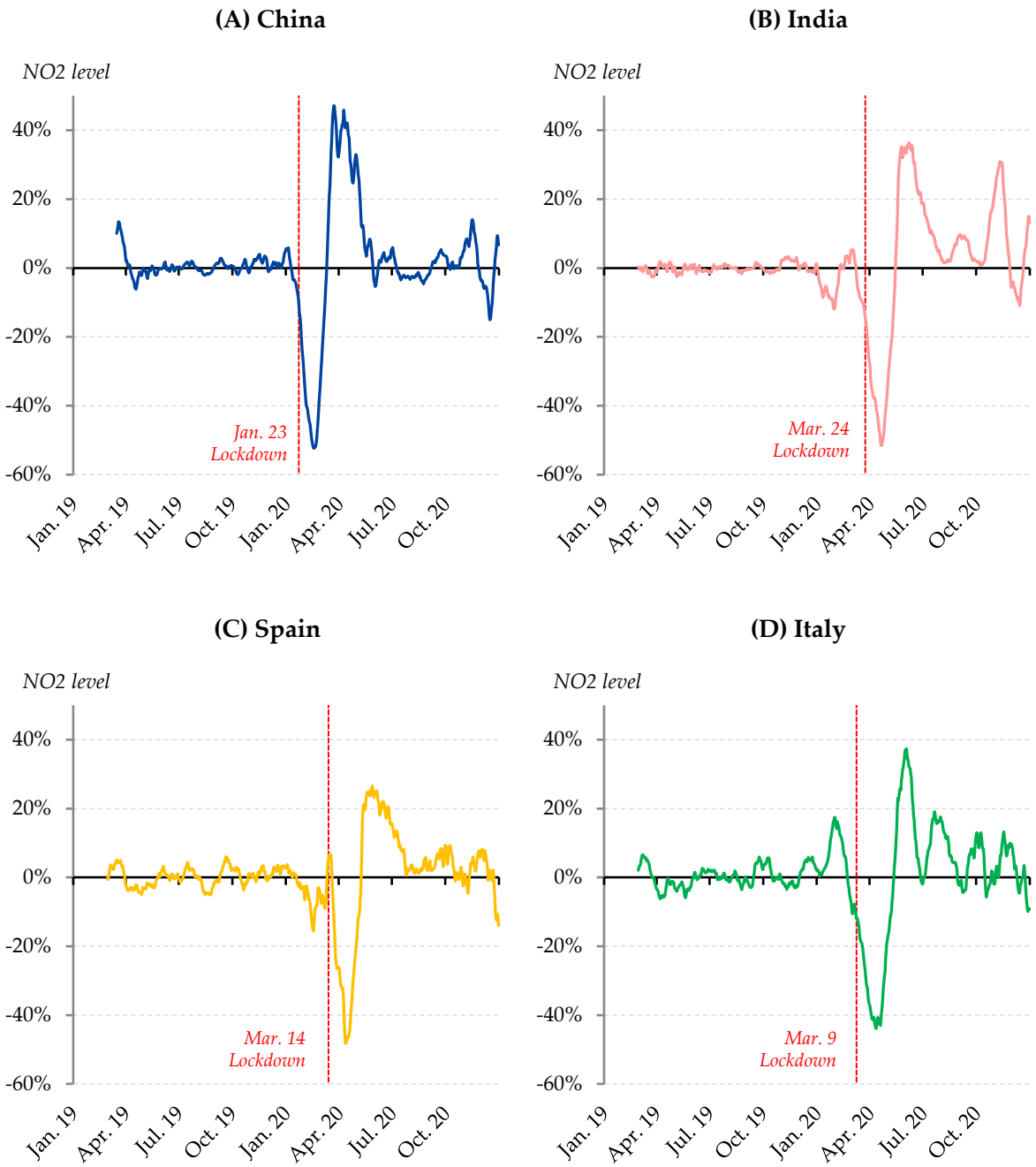
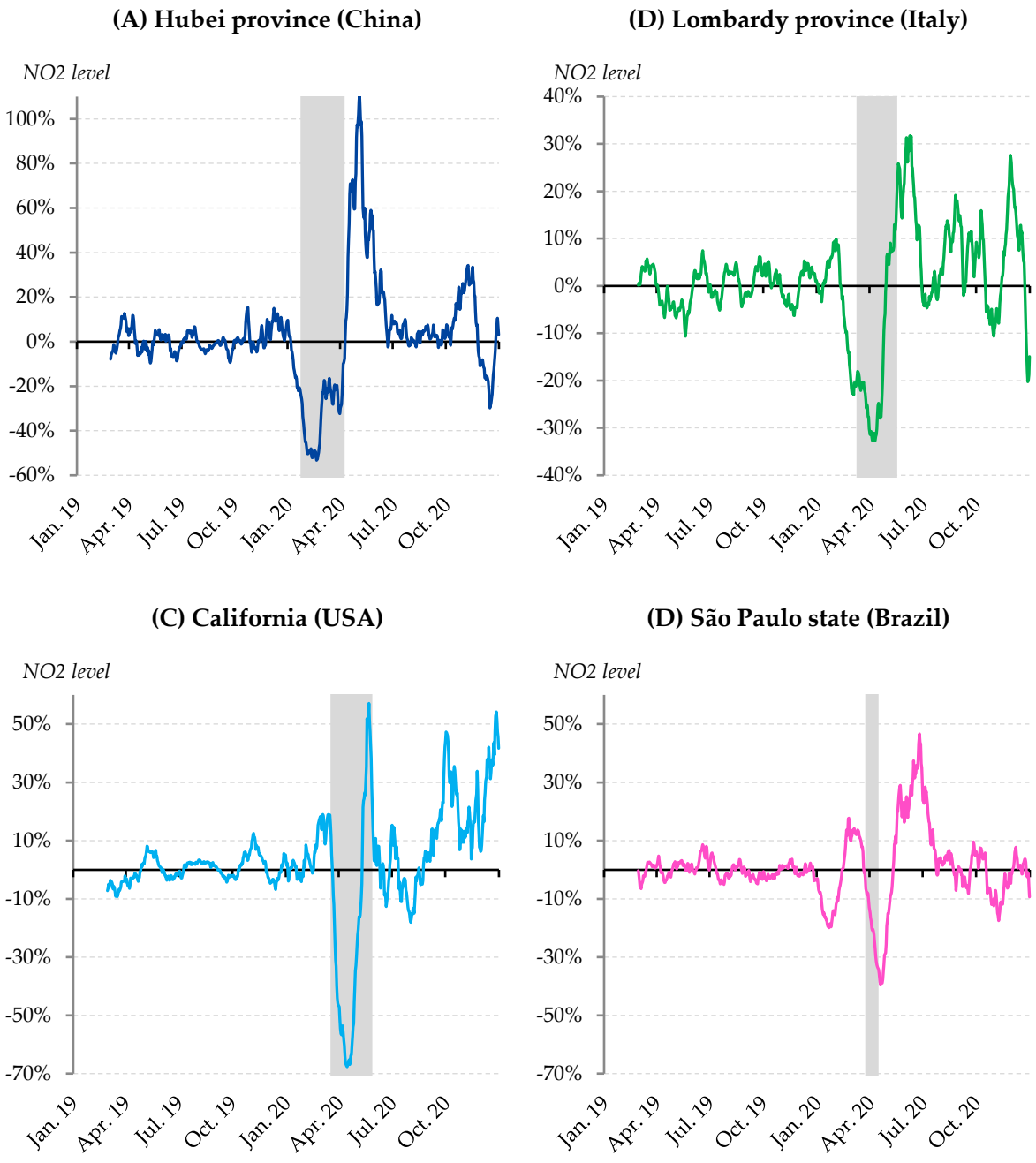


Figure A3. NO_2 pollution level over 2019-2020 (% change, month-on-month)

Sources: ESA, authors' calculations



Note: grey bars represent periods of tight lockdowns (stay-at-home orders in place)