

Impact of COVID-19 pandemic on pollution concentration

Małgorzata Łazęcka, Katarzyna Woźnica

1 Introduction

Air pollutants have a significant impact on human health and are indicated as a major cause of premature deaths. For this reason, more and more attention is being paid to studying what regulations should be introduced to improve air quality. This aspect is particularly important in Poland, where the air quality is one of the worst in European Union according to a report from 2020 published by the European Environment Agency (EEA). Therefore, it is important to identify on which factors the level of each pollutant depends.

It is widely believed that transport has significantly impact on the concentration of particulate matter in urban area. Reduced road traffic has resulted in changes in the emission of harmful substances into the atmosphere and has provided an opportunity to quantify how much air pollution depends on vehicle exhaust and personal transportation.

Next to vehicle traffic, weather conditions are strongly correlated with pollution measurement. The relationship between these readings can be direct - temperature, pressure or humidity affect the concentration of chemicals in the air. For nitrogen dioxide, it has been shown that wind speed and cloud height have the greatest effect on the concentration of this pollutant of all weather-related components.

Since the introduction of lockdown, it has provided unprecedented opportunities in the industrial era to disable or reduce factors associated with human activity and movement. Researchers in various countries are conducting studies to verify hypotheses about air quality and the relationship with vehicle transport.

The article is organized as follows: Section 2 lists the data used in further analysis, Sections 3 and 4 describes the methods and the application with results presented in the end of Section 4. Section 5 concludes.

2 Data

The main purpose of the analysis presented here is to examine the impact of imposed restrictions on resident activity and mobility in Warsaw. We focused most of our attention on obtaining data for the Polish city. Importantly, the data acquisition methodology presented below can also be extended to other cities in Poland and it is possible to extend the conclusions to other cities in Poland. In order to verify the obtained conclusions, data concerning foreign cities were also used: Wuhan and Paris. Wuhan was chosen as the city where COVID-19 hit the hardest and restrictions were introduced earliest. Paris has a similar climate to Warsaw and the restrictions were introduced at a similar time and were not as severe as in Wuhan. For every location we gathered data from January 2016 to May 2020. Data on pollution levels and data describing weather are time series. Due to their nature, it was important to collect as much data as possible before implementing lockdown, as this allowed for robust analysis and estimation of potential trends present in the data.

2.1 Pollution

For Warsaw location we collected air pollutant data from air quality monitoring sites - Chief Inspectorate of Environmental Protection¹. For the remaining cities, however, we used data provided on the repos-

¹<http://www.gios.gov.pl>

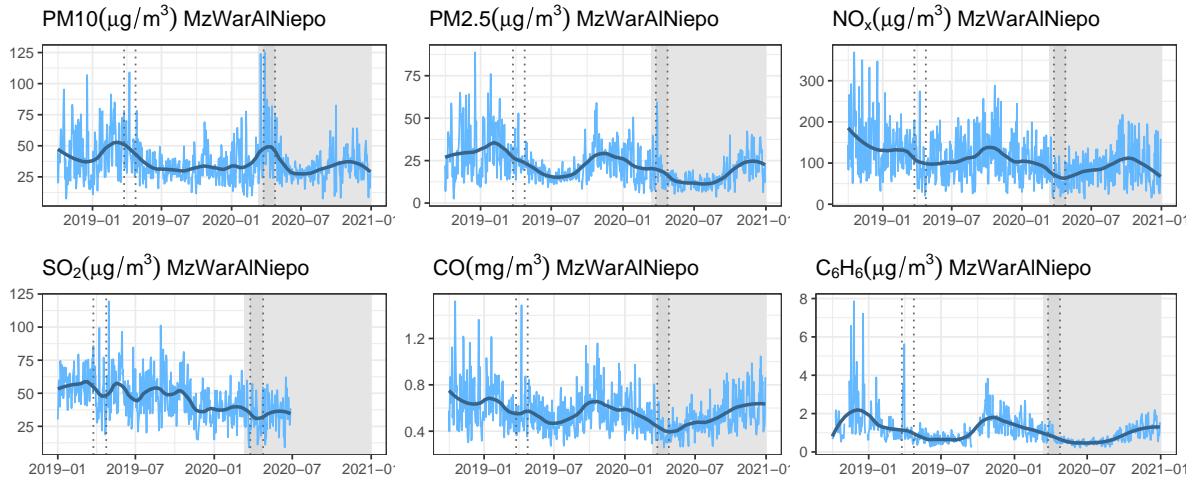


Figure 1: Pollution time series with the lockdown (gray background), strict lockdown (dark gray background and vertical lines) and strict lockdown a year earlier (vertical lines) marked. We present pollutants from **MzWarA1Niepo** station. A smooth curve is obtained with a local regression technique called loess.

itory by authors of [1]. They downloaded data from websites maintained by national (Paris) and local organizations (Wuhan). Source of Wuhan data is unavailable so updating is impossible.

For Warsaw we collected pollutant concentration for 3 locations: **MzWarA1Niepo**, **MzWarKondrat** and **MzWarWokalna**. Every location is in the urban quarter. In [1], for Wuhan an Paris are provided data form urban, suburban and rural areas so for comparability and consistent validation we select urban stations. For every location there are available measurements for PM10, PM2.5, NO₂, SO₂ and CO (with one exception: in Warsaw we use NO_x instead of NO₂). For every location there are available some auxiliary pollutions but we selected ones available for all of three stations.

In Warsaw original pollutant measurements are reported daily but for Wuhan and Paris data has hourly resolution. So we aggregated readings from these foreign locations to one record per day with average. In Figure 1 there are presented considered pollutant concentrations in Warsaw station **MzWarA1Niepo**.

2.2 Weather

Weather data for temperature, relative humidity, atmospheric pressure, and wind speed and direction were downloaded from the National Oceanic and Atmospheric Administration (NOAA) website. The R package **worldometer** was used. For each station with a pollution measurement, the nearest available meteorological station was selected. Data are reported hourly, but to match the frequency of the data with the frequency of the pollutants they were aggregated to one measurement per day.

2.3 Lockdown

The movement restrictions in different countries were introduced at different times. The province in which Wuhan is located was the first to introduce such restrictions on January 23, 2020; they were later extended to all of China. The French government introduced the lockdown on March 16, while Poland introduced it on March 25. In addition to the different dates of introduction of the restrictions, the scope of restriction of human activity varied. All governments recommended working remotely, avoiding contact with other people, keeping social distance and going out only to satisfy the basic needs of life. In France it was only possible to leave the place of residence for an hour a day and at a distance of no more than 1km and a failure to comply the restrictions was punishable by a fine. In Poland, there was no such firm approach to enforcing restrictions on leaving the place of residence, but it was highly recommended to stay at home and leave for essential purposes only. Many workplaces, schools and universities were

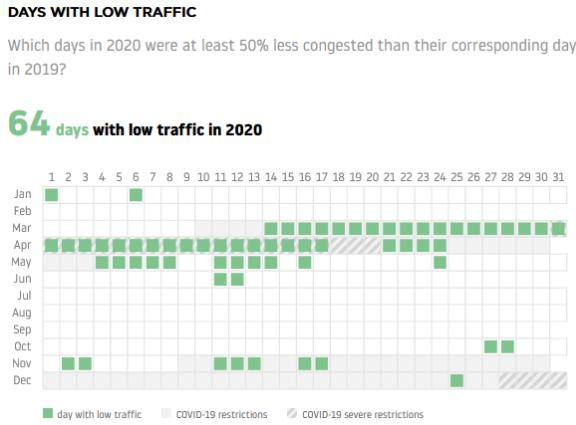


Figure 2: The plot shows the days in which vehicle traffic decreased the most compared to the same day year earlier in Warsaw. The traffic is estimated using travel times (TomTom Traffic Index).

transitioned hybrid or a majority-remote regime. Moreover, wearing masks covering both mouth and nose was mandatory and any outdoors gatherings were banned.

These restrictions, coupled with the public's sense of responsibility and fear of contagion, have resulted in a significant reduction in traffic, both individual and mass passenger transport as well as transport linked to industry. In many countries production in industrial plants and international trade have been restricted. A significant drop in car traffic has been recorded by the car navigation company TomTom². In Figure 2 we see that during strict lockdown time in Warsaw most days has significant lower traffic than one year earlier.

Concurrent with the movement restrictions and other restrictions associated with the COVID-19 pandemic, more states began reporting significant improvements in air quality. This has been observed in China, India, USA as well as European countries. Figure 3 shows the difference in nitrogen dioxide concentration on the satellite photos³ in Wuhan Province before and after the lockdown. The presented maps were one of the first indications of a positive impact of lockdown on the pollution levels.

3 Meteorological normalisation

The first step in our approach is excluding the weather impact on pollution concentrations. Controlling for such changes supports robust trend analysis because there is more certainty that the observed trends are due to changes in emissions or chemistry, not changes in meteorology. Previous research suggests that the effect of weather on changes in pollution levels can be crucial and sometimes far greater than external interventions to control particulate matter release [2]. The variability and trend that we observe in pollutant readings to a large extent may be correlated with external factors that are not influenced by air quality control strategies. To avoid erroneous conclusions about the impact of lockdown-induced limitations, we should address the problem of how meteorological factors affect observed pollutants. The importance of controlling for meteorology over time in air quality time series is shown in [1]. Decrease of nitrogen dioxide concentration measurements reported in [3] was significantly higher than deweathered values reported in [1].

We used the meteorological normalisation presented in [4] and use random forest machine learning models to account for changes in meteorology over time in an air quality time series. The principle of this method is reduce variability by training a machine learning model which can find dependency between pollutant concentrations and a number of independent variables. The explanatory features used

²TomTom page about Warsaw

³Satellite photos of China

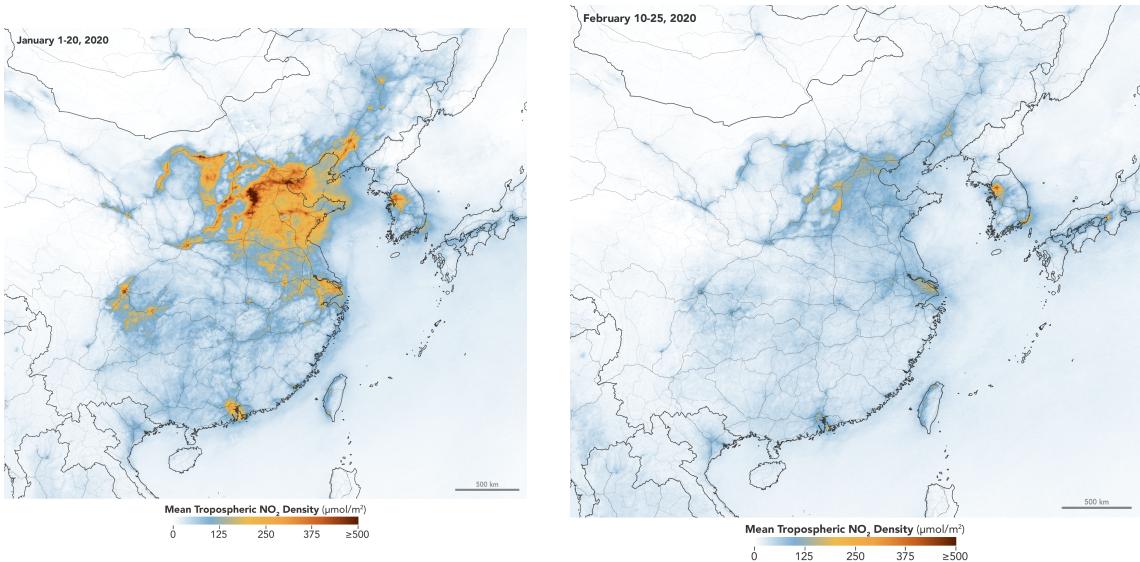


Figure 3: NASA and European Space Agency (ESA) pollution monitoring satellites have detected significant decreases in nitrogen dioxide (NO_2) over China. There is one of the evidences that limiting population's mobility is at least partly related to rapid improvement of air quality.

are typically surface-based meteorological observations and time variables which act as proxies for regular emission patterns such as hour of day and season.

For each city, this process was conducted independently; statistical models were learned on the entire available data. In the method used, the independent variable is the concentration level of the selected pollutant. As meteorological variables whose influence we want to reduce we used all available data: daily temperature, pressure, humidity, wind direction and speed. Meteorological variables in the preprocessing were scaled to the interval [0,1]. The time variables to help detect trend is number of days starts from '1970-01-01' (`date_unix`). To model seasonality we include features such as month, day of year (`date_julian`), and day of week. At this stage, we treat all observations as independent and neglect information about autocorrelation of observed variables. We conducted meteorological normalisation for every pollution independently.

The data normalization method is a heuristic. We briefly describe the principles of its operation (see Figure 4). In the first step, a random forest model [5] is trained for a selected pollution as a dependent

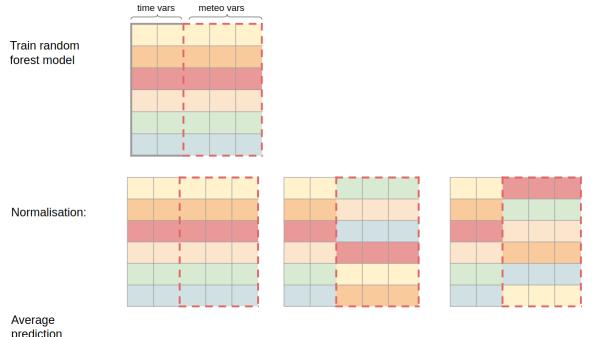


Figure 4: Scheme of meteorological normalisation. Firstly, random forest is trained on original data to model dependency between time along with weather features and pollutant. Normalisation consists in averaging prediction for replicated data with permuted meteorological columns.

variable and a selected set of predictors. This model finds relationships between variables and learns what the dependencies between weather conditions, day of the week, time of the year and observed pollution levels are. In the next step to mimic the exclusion of the effect of weather changes, new modified data sets are created in which the meteorological variables are permuted (jointly). For such changed data, a previously trained model makes a prediction of the pollution level. This operation is repeated 1000 times, so that for the same time data we check what are the model predictions under changing weather conditions. To summarize these results, they are aggregated as an average and this we take as a time series with the effect of meteorological changes removed. In Figure 5 we can compare observed pollutant versus meteorological normalized values.

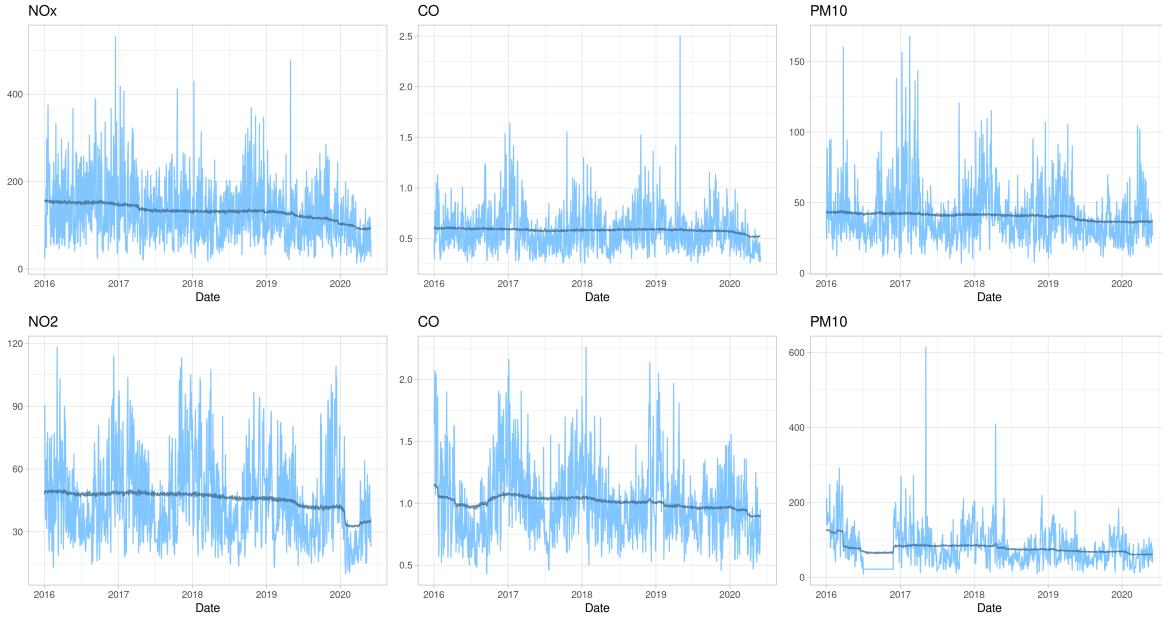


Figure 5: Observed (blue lines) and deweathered (navy lines) concentrations for Warsaw (upper panel) and Wuhan (lower panel).

3.1 Analysis of meteorological normalisation models

As validation criteria for trained random forest models we use Index of Agreement (IOA) values [6]. This accuracy measure is commonly used to evaluate numerical models of climatic, hydrologic, and environmental systems. We trained random forest models on 70% of data and check the prediction goodness-of-fit on the remaining 30% observations. We repeated this process 10 times for every pollutant and site. In Table 1 we provide mean value of IOA along with standard deviation.

Pollution	Warsaw	Paris	Wuhan
CO	0.6 (0.02)	0.56 (0.04)	0.46 (0.02)
NO ₂ /NO _x	0.71 (0.01)	0.7 (0.02)	0.59 (0.02)
PM10	0.51 (0.02)	0.54 (0.03)	0.6 (0.03)
PM2.5	0.66 (0.02)	0.6 (0.02)	0.65 (0.02)
SO ₂	0.61 (0.03)	0.4 (0.04)	0.69 (0.02)

Table 1: Mean value of Index of Agreement (IOA) for meteorological normalisation models for every site and pollutant. Every model was trained 10 times and IOA values are averaged, in brackets there is standard deviation.

Random forest models are composed of series of independent decision tree so it is difficult to summarise the dependency between single explanatory variable and dependent variables. To insight into the structure of these models we employed explainable model analysis methods (XAI) [7]. In Figure 6 we present permutation-based variable importance for random forest models for CO, PM10 and PM2.5 for all locations. Importance of features varies across meteorological normalisation models for different locations and different pollutants. In Wuhan, regardless of considered pollution, day of year and date are in the top 2 the most relevant features. Crucial is that `weekday` variable is insignificant in every model.

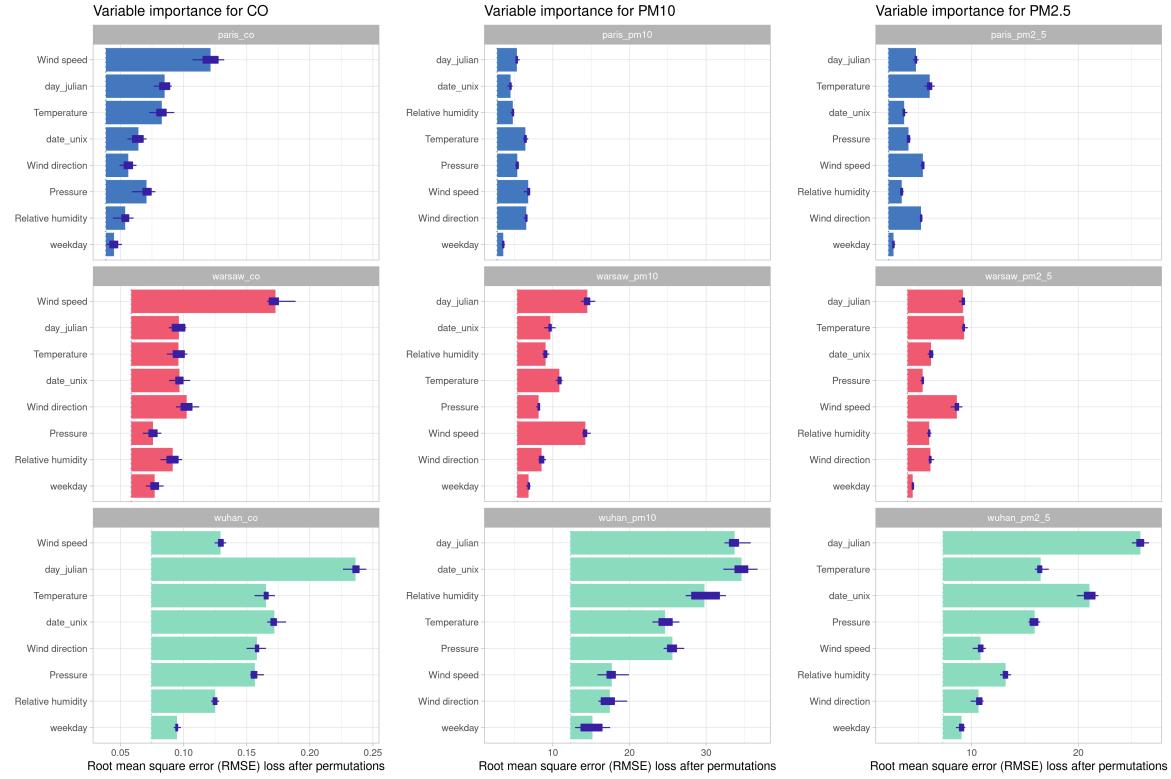


Figure 6: Permutation based variable importance for selected pollutants: *CO*, *PM10*, *PM2.5*

To better understand what the dependence of particulate air pollution is we used partial dependence plots [8]. In addition to evaluating which variables are most important, we can tell how the prediction depends on meteorological variables. In Figure 7 we show PDP profiles for selected pollutants and weathers features. Considered stations are located in various climate zone so next to PDP we provide histograms to assess the density of variables.

4 Changepoint detection

The main goal of that section is to detect multiple changepoints in the deweathered data using PELT algorithm (packages: `changepoint` [9], [10] and `EnvCpt` [11]), identify (if exists) the changepoint that occurred up to a month after the strict lockdown was imposed and test if there was a change in the trend or the mean (package `strucchange` [12]). First, we describe used methods, then we present the results of the deweathered pollution time series analysis.

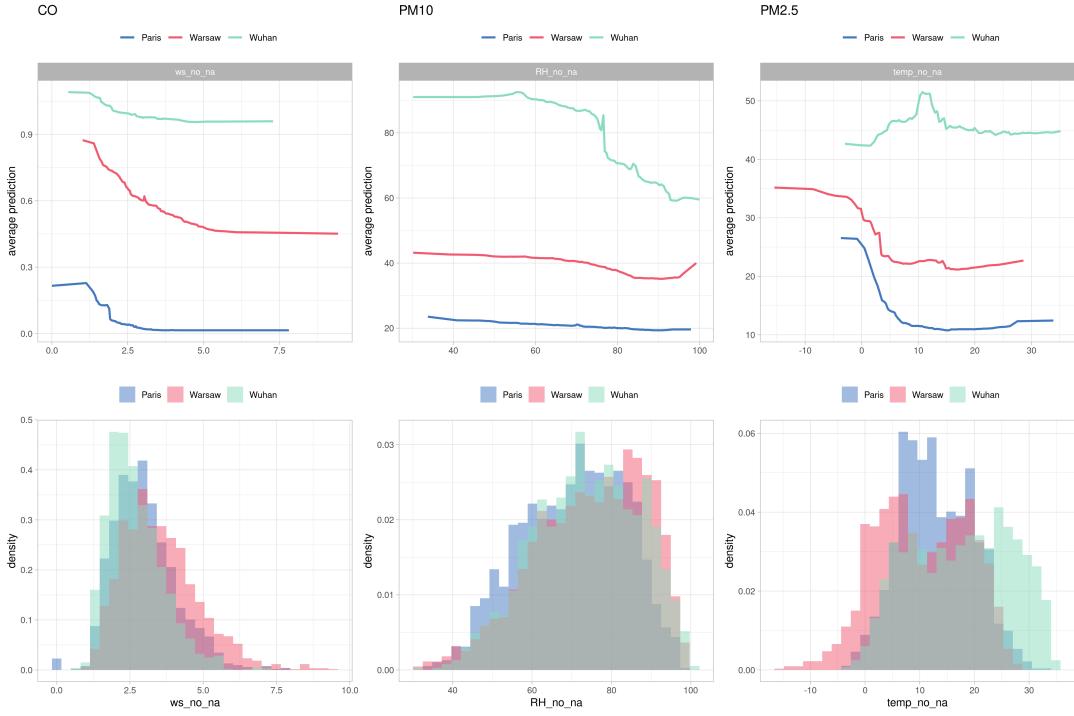


Figure 7: Partial dependence plots for selected pollutants: CO, PM10, PM2.5

4.1 PELT algorithm

Assume we have a time-ordered data: y_1, y_2, \dots, y_n and a set of m changepoints $\tau = \{\tau_{1:m}, \tau_{2:m}, \dots, \tau_{m:m}\}$. By $\tau_{i:m}$ we denote i th earliest changepoint and by $y_{(\tau_{(i-1):n}+1):(\tau_{i:n})}$ we denote the segment between $i-1$ th and i th changepoint i.e. $y_{(\tau_{(i-1):n}+1):(\tau_{i:n})} = (y_{\tau_{(i-1):n}+1}, y_{\tau_{(i-1):n}+2}, \dots, y_{\tau_{i:n}})$. The i th changepoint means that the statistical properties of $y_{(\tau_{(i-1):n}+1):(\tau_{i:n})}$ and $y_{(\tau_{i:n}+1):(\tau_{(i+1):n})}$ differ. The most common approach is to minimize the following function

$$\sum_{i=0}^m \mathcal{C}(y_{(\tau_{i:n}+1):(\tau_{(i+1):n})}) + \beta f(m), \quad (1)$$

where \mathcal{C} is a cost function of a segment and $\beta f(m)$ is a penalty for introducing changepoints. The most common choices of a cost function is twice the negative log-likelihood or quadratic loss, whereas the penalty usually is a linear function in the number of changepoints, namely $\beta f(m) = \beta m$. Two basic methods of optimizing (1) are binary segmentation ([13], [14]), which delivers only approximate results, but is fast, and the segment neighborhood algorithm [15], which gives exact results if the maximum number of the breakpoints is known, but its computational complexity is high. PELT (Pruned Exact Linear Time) [16] is the method based on optimal partitioning [17] with an addition of pruning. The optimal partitioning is based on solving the recursion

$$F(s) = \min_{\tau \in \mathcal{T}_s} \left[\sum_{i=0}^m (\mathcal{C}(y_{(\tau_{i:n}+1):(\tau_{(i+1):n})}) + \beta) \right] = \min_t [F(t) + (\mathcal{C}(y_{(t+1):n}) + \beta)], \quad (2)$$

where $\mathcal{T}_s = \{\tau : 0 = \tau_0 < \tau_1 < \dots < \tau_{m+1} = s\}$ and $F(0) = -\beta$. We note that the recursive formula holds for a linear penalty. The optimal partitioning algorithm improves the computational efficiency of the segment neighborhood algorithm, bus its complexity is still higher than the binary segmentation. The PELT algorithm introduces pruning to the optimal partitioning, which consists in removing redundant values of τ (cf. Theorem 3.1 in [16]) and that way the computational cost of the procedure is reduced.

4.2 Detecting significant changepoints

Concerning testing for changepoints, we distinguish two cases:

- the first, in which we assume that two segments and the changepoint are recognized and the task is to test if the change is significant,
- the second, in which we have no knowledge about any changepoints, thus we assume that parameters in the considered period do not change and we want to check if the claim is valid.

In both cases, we consider the standard linear regression model $y_i = x'_i \beta + \varepsilon_i$ for $i \in 1, 2, \dots, n$, where y_i is the dependent variable, $x_i = (1, x_{i1}, x_{i2}, \dots, x_{ip})'$ is a vector of observation at time i and ε_i is an error term with mean equals 0 and a constant variance σ^2 independent of ε_j for $i \neq j$. The vector of coefficients $\beta = (\beta_0, \beta_1, \dots, \beta_p)$ might varies over time $\beta = \beta(i)$ and the main difference between the two cases is what we assume about that dependence.

4.2.1 F test

In the first case we assume that we exactly know, where the change point is, thus both hypothesis can be formulated in the following way:

$$\begin{aligned} H_0 : \quad & \beta^{(0)} = \beta^{(1)} \\ H_1 : \quad & \beta^{(0)} \neq \beta^{(1)} \end{aligned}$$

where $\beta^{(0)} = \beta(i)$ for $1 \leq i \leq \tau$ and $\beta^{(1)} = \beta(i)$ for $i_0 < i \leq n$ and τ is the change point. The first to suggest a test for that situation was [18].

4.2.2 Generalized fluctuation tests

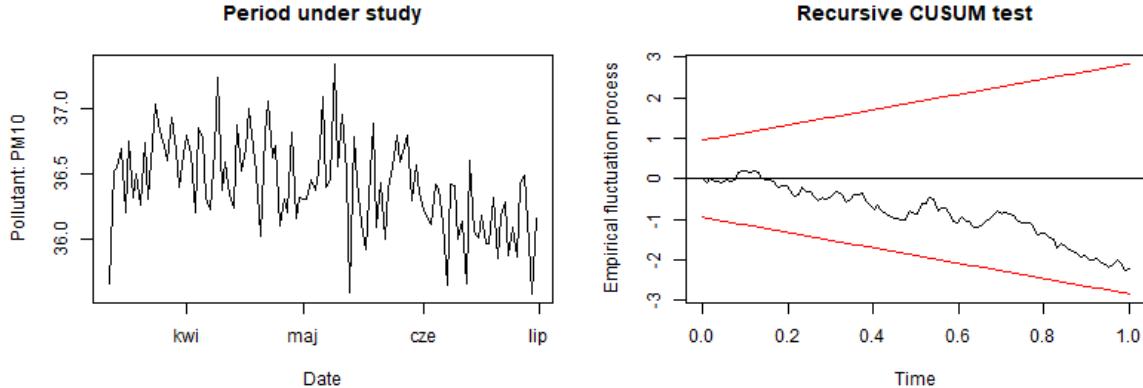


Figure 8: Deweathered PM10 in Warsaw (left panel) and EFP for the data shown in the left panel (right panel).

If the change point is not known, we apply the procedure, which detects the structural changes regardless of when the change has occurred. We are going to test

$$H_0 : \quad \beta(i) = \beta(1) \quad \text{for } 2 \leq i \leq n$$

against the alternative hypothesis, that the vector β varies over time. The test we apply uses CUSUM processes (more generally this type of processes are called empirical fluctuation processes (EFP)), which

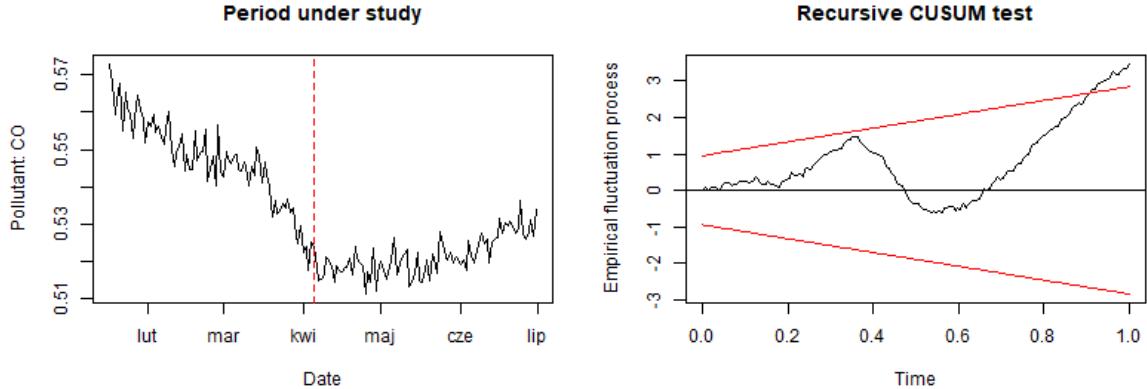


Figure 9: Deweathered CO in Warsaw with detected chang point (left panel) and EFP for the data shown in the left panel (right panel).

is based on cumulative sums of recursive residuals \tilde{e} ($\tilde{\sigma}^2$ is a recursive residuals variance):

$$W_n(t) = \frac{1}{\tilde{\sigma}n} \sum_{i=p+1}^{k+\lfloor tn \rfloor} \tilde{e}_i, \text{ for } 0 \leq t \leq 1, \quad (3)$$

where $\eta = n - p$ is the number of \tilde{e}_i and the recursive residuals are defined in the following way

$$\tilde{e}_i = \frac{y_i - x_i' \hat{\beta}^{(1:(i-1))}}{\sqrt{1 + x_i' (X^{(1:(i-1))'} X^{(1:(i-1))})^{-1} x_i}} \text{ for } i = p+1, p+2, \dots, n,$$

where $\beta^{(1:(i-1))}$ is the common OLS estimate based on observations up to i and similarly $X^{(1:(i-1))}$ is the experiment matrix containing observation up to i . The variance is $\tilde{\sigma}^2 = \frac{1}{n-p} \sum_{i=p+1}^n (\tilde{e}_i - \bar{\tilde{e}})^2$. The CUSUM process firstly appeared in [19]. The test is based on the fact, that under the null hypothesis (and additional assumptions, which we there omit as this is not the focus of this work; for more details or further references see e.g. [12]) the empirical fluctuation process $W_n(t)$ asymptotically is the standard Brownian Motion $W(t)$.

The illustration of the EFP process for the data with and without detected change point (see Section 4.3) is presented in Figures 8 and 9 on the example of Warsaw's PM10 and CO pollutants. In the left panels deweathered pollutants are shown (in case of CO the break point is marked). In the right panel the process $W_n(t)$ is presented with appropriate boundaries that process $W(t)$ crosses with the probability $\alpha = 0.05$. There is no evidence that there is a change in structure for PM10 pollutant, whereas for CO the boundary is crossed, thus the null hypothesis is rejected. We note that Figure 9 is just to visualize the situation, in which the null is rejected, because as we know, where the break point is, we use in that case F test. The chosen period corresponds to the segment between two break points detected through a procedure described in Section 4.3. In both cases in the tests we use as regressors lagged observations x_{i-1} .

4.3 Procedure

To detect changepoints in our data, we use a package `EnvCpt` [11] for automatic model selection between a variety of models including trend and autocorrelation models. The models with and without changepoints (changepoints are detected using PELT algorithm described in Section 4.1) are being compared using AIC (Akaike information criterion), which deals with the trade-off between a goodness of fit and overfitting as it combines likelihood and a penalty term for the number of parameters in the model. For the pollution

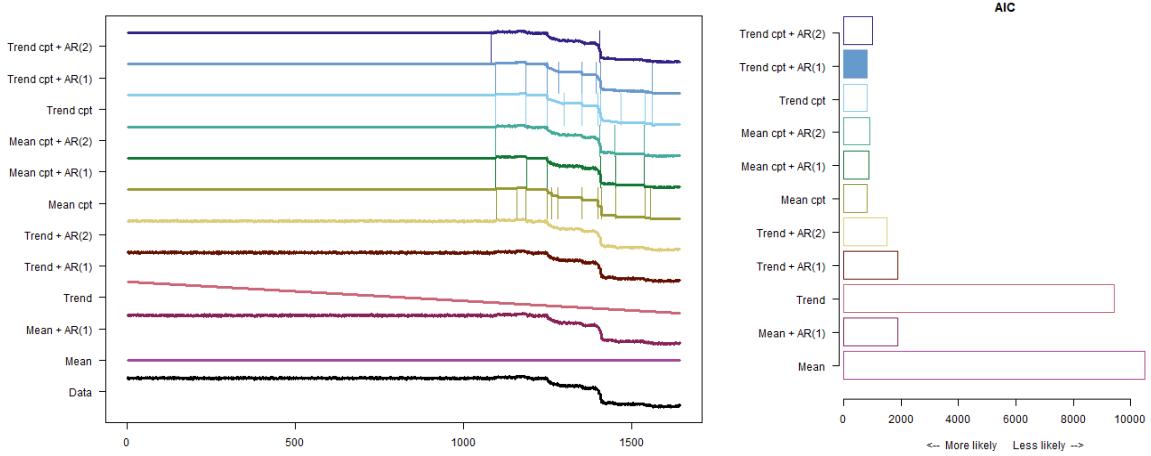


Figure 10: The 12 models fitted by `envcpt` function for the deweathered SO_2 pollution in Warsaw. The first plot shows fitted values by the models (in the first six graphs the predictions for models with changepoints are presented). The values are sorted by the date. In the second plot the AIC value for each model is shown and the model with the smallest value is chosen.

datasets models containing changing trend (sometimes with additional autoregressive AR(1) or AR(2) terms) work the best (in Figures 10 and 11 we an example for Warsaw and SO_2 pollutant).

Next, changepoints that occur up to 30 days after lockdown was announced (tabela z datami) and segments before and after that point in time are selected. Then the F test is run. If there is no breakpoint meeting the requirements, the test based on recursive residuals and a CUSUM process is performed on the segment that contains the date of the beginning of lockdown.

4.4 Results

In Figures 12 and 13 we present results for Warsaw for CO, SO_2 , PM10, PM2.5 and for Warsaw for NO_x , Wuhan and Paris for NO_2 (results for other pollutants in Wuhan and Paris can be found in the Appendix - Figures 14 and 15). Table 2 shows results of the tests described in Section ???. We report test statistics values and p-values without any adjustment for multiple testing. We point out that the results of the considered statistical tests highly dependent on previously used procedures: weather normalisation and detection of homogeneous segments.

We note that all listed figures show downward trend of pollutants' concentrations in recent years, which make analysis more challenging. Some of detected changepoints might be explained by the introduction of emission restrictions or car-free zones. For example, the visible decrease in all pollutants in China in the beginning of 2016 might be an effect of Air Pollution Prevention and Control Law (2015) (English translation is available at cleanairchina.org), which entered into force on 1st of January, 2016. In Paris, Low Emission Zones, which limits vehicle access to a defined area, are being introduced. The LEZ was implemented on 1st of July, 2016, the subsequent phases started on July of 2017 and 2019. The details are described in *Report: Impacts of the Paris low-emission zone and implications for other cities* by Bernard Y. et al. The effects of the LEZ might be spread over time, as LEZs apply restrictions to certain vehicle types and the engines might be replaced gradually.

The obtained results of the tests suggest, that in the cases, in which the change point connected with lockdown is detected, the change is significant (see the results of the F tests), otherwise it is not (there is only one EFP-based test, which rejects the null - for Wuhan and SO_2). In all three sites the change in nitrogen oxides appears to be significant. Nitrogen oxides are considered one of the critical pollutants found in exhaust gas from all types of combustion engines. In Warsaw the changes in CO and SO_2 concentrations are also considered significant, whereas particulate matter concentrations remain stable in the run-up to and during lockdown. This might be caused by the relatively small effect of the lockdown

Deweathered data vs fitted values of the chosen model (trend + AR(1))

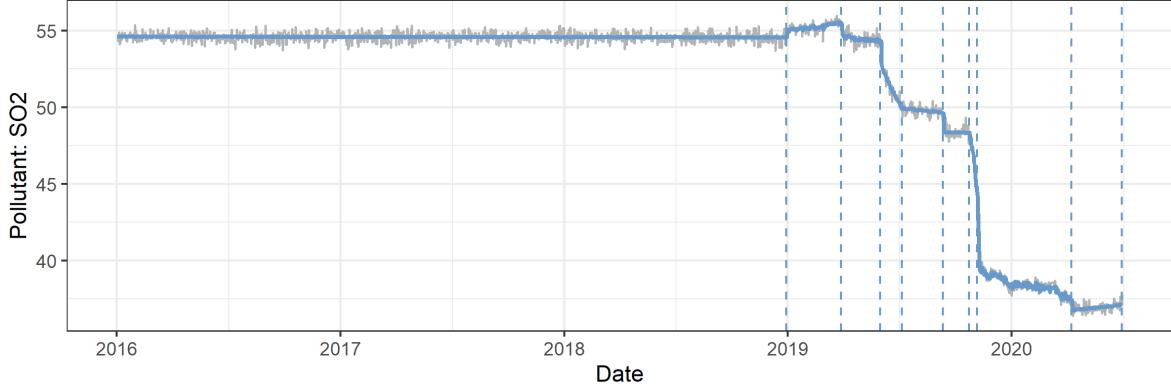


Figure 11: Comparison of the deweathered time series for SO₂ in Warsaw (gray points) and fitted values of the best model (blue points) chosen from 12 models (the best model there is a model with trend and an autoregressive term AR(1), cf. Figure 10).

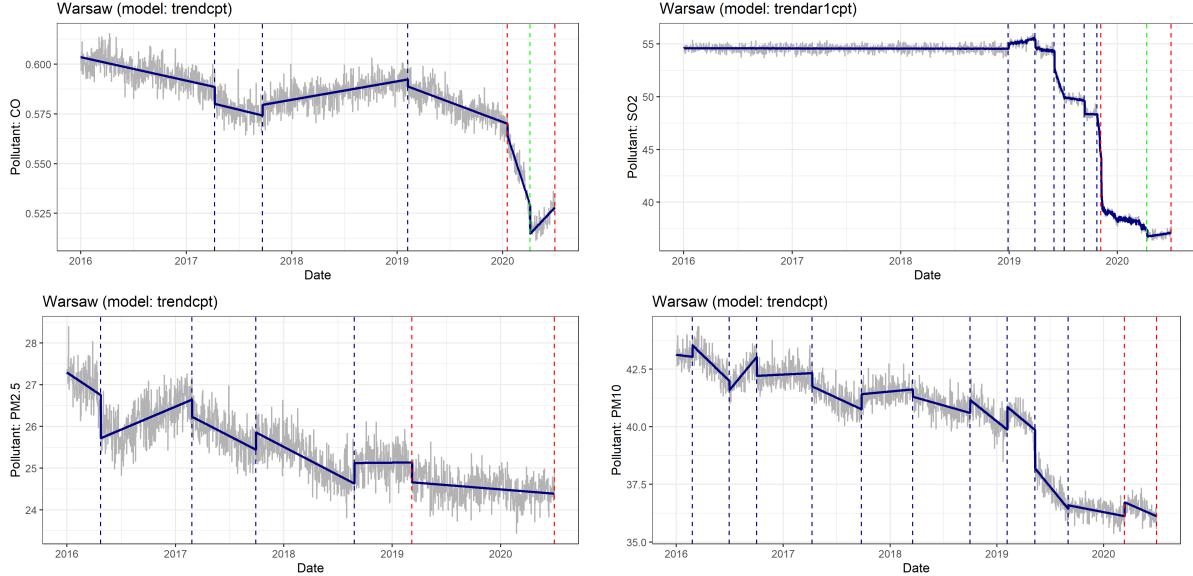


Figure 12: Deweather concentrations of pollutants in Warsaw with fitted values by the most suitable models (see the titles of the plots). Change points are marked with vertical lines, a green vertical line denotes a change point that occurs up to 30 days after lockdown has been announced, the red lines, which are also detected change points, show the beginning and the end of the period under study.

compared to usual seasonal variations in PM levels throughout the year.

5 Conclusions

The tested methods of detecting changepoints in time series are subject to the selection error of the data preparation technique. It is possible that we are too conservative in the meteorological normalization process and lose some of the trend effect. In conditions of unrestricted mobility, worsening weather affects people's choices, and it is possible that they are more likely to use cars to commute to work. The chosen normalization method is not robust to such correlation and this procedure will remove such effect of

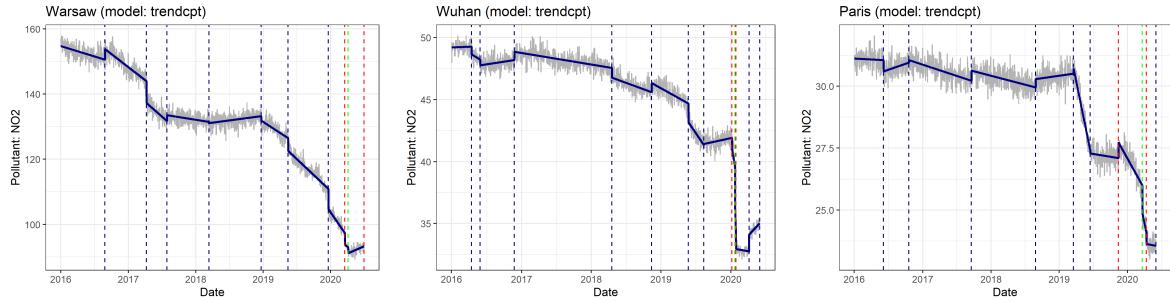


Figure 13: Deweather concentrations of NO_x in Warsaw and NO_2 in Paris and Wuhan. For more information see a caption of Figure 12.

	Warsaw			Paris			Wuhan		
	Method	Statistic	p-value	Method	Statistic	p-value	Method	Statistic	p-value
NO_2/NO_x	F test	30.47	<1e-10	F test	103.34	<1e-10	F test	60.73	9.95e-09
CO	F test	297.56	<1e-10	EFP	0.74	2.00e-01	F test	144.08	<1e-10
SO_2	F test	17.09	1.19e-07	F test	62.27	<1e-10	EFP	1.33	1.53e-03
PM10	EFP	0.76	1.79e-01	EFP	0.89	7.86e-02	F test	13.41	4.43e-05
PM2.5	EFP	0.81	1.33e-01	EFP	0.59	4.40e-01	F test	523.72	<1e-10

Table 2: Test statistics and p-values for all pollutants in Warsaw, Paris and Wuhan. EFP denotes a test based on EFP (Rec-CUSUM).

increased pollution.

The procedure of meteorological standardization is crucial to properly evaluate the effect of the imposed regulations. However, it requires careful validation because its quality significantly affects the subsequent steps of the analysis. XAI methods may be applied to audit model and collate domain knowledge and intuitions with relationship found by model. On the example of the normalization performed for the city of Wuhan, we can suspect that the statistical model was overfitted.

The conclusions we can draw from this analysis concern the overall effect of limitations introduced during the lockdown. The conclusions that we can draw from the analysis presented here relate to the overall effect of the lockdown restrictions on mobility, but also to reduced trade-related transport, reduced production and reduced air traffic.

References

- [1] Zongbo Shi, Congbo Song, Bowen Liu, Gongda Lu, Jingsha Xu, Tuan Van Vu, Robert JR Elliott, Weijun Li, William J Bloss, and Roy M Harrison. Abrupt but smaller than expected changes in surface air quality attributable to covid-19 lockdowns. *Science Advances*, 7(3):eabd6696, 2021.
- [2] Vo Anh, Hiep Duc, and Merched Azzi. Modeling anthropogenic trends in air quality data. *Journal of the Air & Waste Management Association*, 47(1):66–71, 1997.
- [3] Tianhao Le, Yuan Wang, Lang Liu, Jian Yang, Yuk L Yung, Guohui Li, and John H Seinfeld. Unexpected air pollution with marked emission reductions during the covid-19 outbreak in china. *Science*, 369(6504):702–706, 2020.
- [4] Stuart K Grange, David C Carslaw, Alastair C Lewis, Eirini Boleti, and Christoph Hueglin. Random forest meteorological normalisation models for swiss pm 10 trend analysis. *Atmospheric Chemistry and Physics*, 18(9):6223–6239, 2018.
- [5] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

- [6] Cort J Willmott, Scott M Robeson, and Kenji Matsuura. A refined index of model performance. *International Journal of climatology*, 32(13):2088–2094, 2012.
- [7] Przemyslaw Biecek and Tomasz Burzykowski. *Explanatory Model Analysis*. Chapman and Hall/CRC, New York, 2021.
- [8] Jerome H Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.
- [9] Rebecca Killick and Idris Eckley. changepoint: An R package for changepoint analysis. *Journal of statistical software*, 58(3):1–19, 2014.
- [10] Rebecca Killick, Kaylea Haynes, and Idris A. Eckley. *changepoint: An R package for changepoint analysis*, 2016. R package version 2.2.2.
- [11] Rebecca Killick, Claudie Beaulieu, Simon Taylor, and Harjit Hullait. *EnvCpt: Detection of Structural Changes in Climate and Environment Time Series*, 2021. R package version 1.1.3.
- [12] Achim Zeileis, Friedrich Leisch, Kurt Hornik, and Christian Kleiber. strucchange: An r package for testing for structural change in linear regression models. *Journal of Statistical Software*, 7(2):1–38, 2002.
- [13] Andrew Jhon Scott and M Knott. A cluster analysis method for grouping means in the analysis of variance. *Biometrics*, pages 507–512, 1974.
- [14] Ashish Sen and Muni S Srivastava. On tests for detecting change in mean. *The Annals of statistics*, pages 98–108, 1975.
- [15] Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, pages 47–78, 1998.
- [16] Rebecca Killick, Paul Fearnhead, and Idris A Eckley. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500):1590–1598, 2012.
- [17] Brad Jackson, Jeffrey D Scargle, David Barnes, Sundararajan Arabhi, Alina Alt, Peter Gioumousis, Elyus Gwin, Paungkaew Sangtrakulcharoen, Linda Tan, and Tun Tao Tsai. An algorithm for optimal partitioning of data on an interval. *IEEE Signal Processing Letters*, 12(2):105–108, 2005.
- [18] Gregory C Chow. Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, pages 591–605, 1960.
- [19] Robert L Brown, James Durbin, and James M Evans. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2):149–163, 1975.

6 Appendix

6.1 Breakpoint detection for Wuhan and Paris

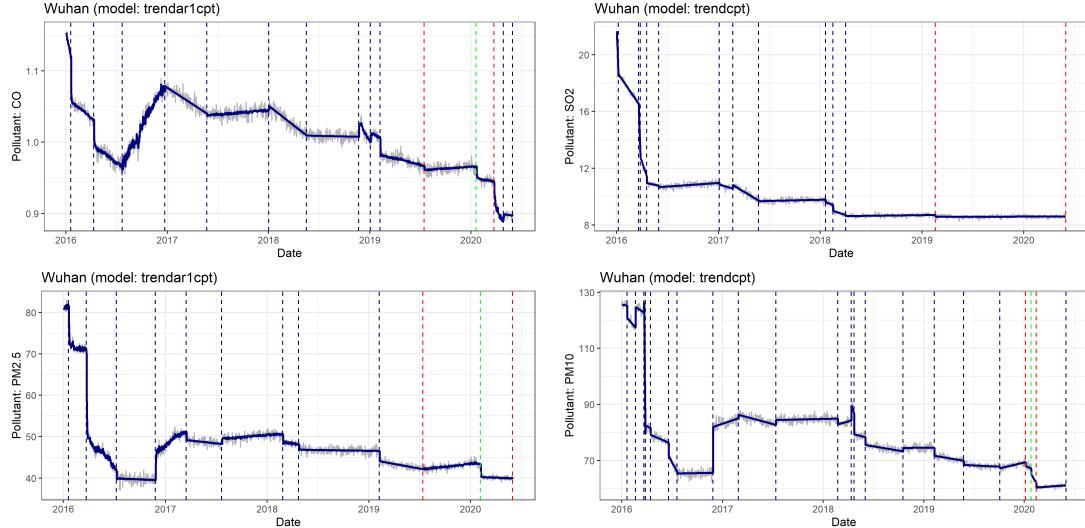


Figure 14: Deweather concentrations of pollutants in Wuhan. For more information see Figure 12.

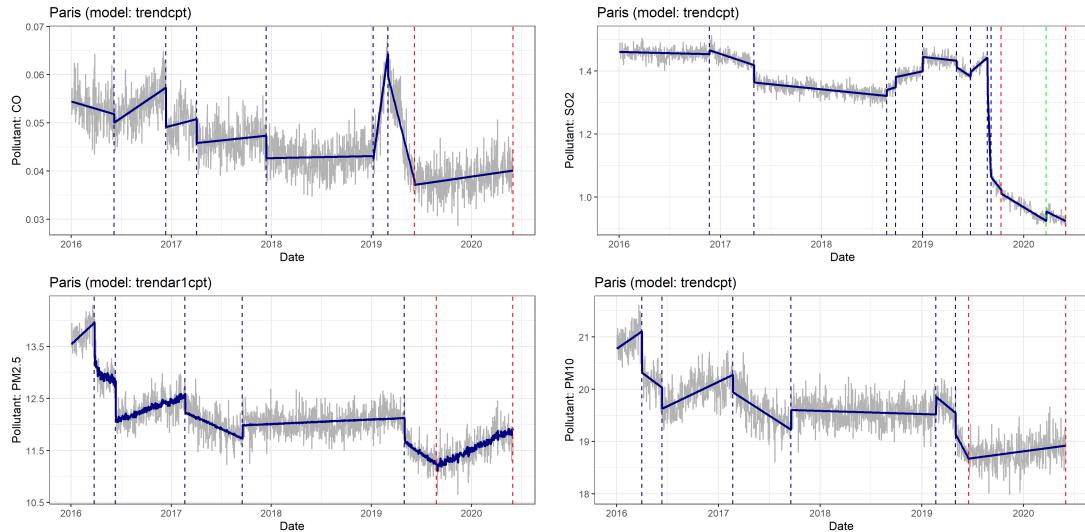


Figure 15: Deweather concentrations of pollutants in Paris. For more information see Figure 12.