Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2020-469-AC1, 2020

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Meteorology-driven variability of air pollution (PM1) revealed with explainable machine learning" *by* Roland Stirnberg et al.

Roland Stirnberg et al.

roland.stirnberg@hotmail.com

Received and published: 30 November 2020

"I am afraid that even after reading the article, I do not understand why the authors choose to explain air quality over Paris based on meteorology at the SIRTA location, when regional and local emissions and atmospheric transformations during long range transport are the major drivers of ambient pollution. These major drivers are mentioned towards the end as future work, but studies should start there. For example, even if MLH or wind speed is low, zero emissions = no air pollution."

Answer: Thank you for this comment.

Printer-friendly version

Discussion paper

Indeed we do not inted to explain Paris air quality using the SIRTA site - the focus is on air quality at the SIRTA site, which is near Paris and representative of suburban background concentrations. We have clarified this in the manuscript (see changes in L11, L88, L94, L127, L234, L502). The aim of this study is to quantify how meteorological factors influence pollutant concentrations and thus add to system understanding. It has been shown in previous studies that pollutant concentrations are not solely driven by emissions, and can be exacerbated or weakened substantially by certain meteorological conditions. For SIRTA, this has been described for example by Dupont et al., 2016.

It was not intended to set up a prognostic model to forecast PM1 as accurately as possible in time. Emissions of pollutants or precursor gases undoubtedly constitute a prerequisite to air pollution, but pollutant concentrations are not solely driven by anthropogenic emissions, but strongly affected by varying amounts of natural background emissions (see e.g., Liora et al., 2016, DOI: 10.1016/j.atmosenv.2016.04.040, Jiang et al. 2019, DOI: 10.5194/acp-19-15247-2019). Thus, high concentrations of particulates could also occur during episodes with low anthropogenic emissions. This is particularly the case in summer, when biogenic organic emissions are high (this is mentioned in the introduction, L56). The influence of meteorology can lead to quite different air pollution situations, even if emissions are constant. In winter, meteorological conditions exert great influence on formation pathways, as we describe in chapter 4.2.1. For example, condensation of ammonium nitrate in the aerosol phase is enhanced at low ambient temperatures and high relative humidity (see e.g., Pay et al., 2012; Bressi et al., 2013; Petetin et al., 2014; Petit et al., 2015). Hence, even if emissions would be not above average levels, this formation mechanism would increase the concentrations of pollutants.

Transformation processes are partly covered by meteorological parameters, e.g., through the influence of temperature (please see also answers to Referee #3 comments). Obviously, the model does still not capture all of the occurring variance of PM1

concentrations, but since it was not the intention to set up a predictive framework, the focus is not primarily on accuracy, but on interpretability. Hence, the included parameters are deemed adequate for the analysis

In the updated version of the manuscript, we have taken great care to consider this comment and now more clearly communicate the main goal of this manuscript:

- L3: However, the scientific understanding of the ways by which complex interactions of meteorological factors lead to high pollution episodes is inconclusive, as the effects of meteorological variables are not easy to separate and quantify

- L6: In this study, a novel, data-driven approach based on empirical relationships is used to characterise, quantify and better understand the meteorology-driven component of PM1 variability.

- L8: Changed to “Based on the model, an isolation and quantification of individual meteorological influences for process understanding is achieved using SHapley Additive exPlanation (SHAP) regression values.

- L87: Changed to “Here, the multivariate and highly interconnected nature of the processes determining local PM1 concentrations is analysed in a data-driven way. Therefore, a state-of-the-art explainable machine learning model is set up to reproduce the variability of PM1 concentrations, thereby capturing empirical relationships between PM1 concentrations and meteorological parameters. The goal is to separate and quantify influences of the meteorological variables on PM1 concentrations to advance the process understanding of the complex mechanisms that govern pollution concentrations at the measurement site.”

"The authors frame it as "we should take atmospheric and environmental processes into account during the development of efficient pollution mitigation strategies"/"a basis for future clean air programs", but AirParif can’t exactly change wind conditions or MLH orT/RH."

Answer:

These statements were aimed to show the potential benefits of considering atmospheric and environmental conditions when future measures to prevent air pollution are discussed. This relates mainly to three points:

a) A realistic assessment of the effectiveness of measures against air pollution needs to take atmospheric and environmental processes into account as these processes partly control its variability. For example, if changes in PM concentrations due to traffic restrictions were to be determined, a simple comparison of pre-restriction and post-restriction concentrations would not be sufficient, as meteorological influences would be omitted. Machine learning approaches can be very useful to characterize the efficiency of mitigation policies. Recent lockdown in Spain is an adequate example here (Petetin et al., 2020, DOI: https://doi.org/10.5194/acp-20-11119-2020)

b) Weather conditions which exacerbate pollutant concentrations are identified using the SHAP framework. On this basis, air-pollution measures could be adjusted depending on expected meteorological conditions. For example, warnings could be expressed to the public to remain vigilant or stay at home if possible.

c) In a changing climate, more unfavorable meteorology could trigger and/or exacerbate PM pollution episodes, lowering the role of emission restrictions.

To make this clearer in the manuscript, the following changes were made:

- L47: The sentence “It is therefore crucial to take atmospheric and environmental processes into account during the development of efficient pollution mitigation strategies” was removed.

The explanation relating to future clean air program was shifted to the end of the introduction and expanded; L96-106 now read: “... allowing to infer meteorology-dependent processes driving PM concentrations at high temporal resolution.. Typical situations that lead to high PM1 concentrations are identified, serving as a decision support to

policymakers to issue preventative warnings to the public if these situations are to be expected. In addition, by directly accounting for meteorological effects on PM1 concentrations, such a machine learning-based framework could help in assessing the effectiveness of measures towards better air quality. Furthermore, the proposed ML framework can be viewed as a first step towards a data-driven, prognostic tool in operational air quality forecasting, complementary to CTM approaches.

"Maybe this can be used to forecast periods of bad air quality - but they describe some important events that the model fails to reproduce because it is missing major drivers in the inputs (lines 390, 426-427)."

Answer:

We show several examples where the model is well able to reproduce episodes of high pollutant concentrations (sections 4.4.1-4.4.3). This is encouraging and shows the appropriateness of the approach. There are of course also situations in which the model fails to reproduce high-pollution situations. These situations are shown in detail to stimulate further research in this direction.

As stated earlier, the current model setup was chosen for the purpose of improving the understanding of how meteorological factors influence pollutant concentrations, and to quantify potential influences (which was more clearly stated in L87-105). Setting up a probabilistic forecast model based on top of the presented framework would undoubtedly require many adjustments to include the factors correctly pointed out by the referee..

"So I am not sure this study is an advance over previous knowledge."

Answer:

Extensive changes were made to the manuscript to emphasize new scientific insights (see previous answers and answers to Referee#3 comments)

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2020-469, 2020.

Printer-friendly version

Discussion paper