General comments on the paper entitled "Reduction in vehicular emissions attributable to the Covid-19 lockdown in Shanghai: insights from 5-year monitoring-based machine learning":

This work presents the results of a 5-year monitoring of EC and NO_x in a traffic site in Shanghai, China. The authors used Random Forest to estimate a business-as-usual scenario during the COVID-19 lockdown period. The authors have validated their results using Satellite data of NO_2.

This manuscript can be interesting and is generally well-written. However, there are a lot of important clarifications that need to be addressed especially in the methodology. It is necessary to improve the manuscript through more elaborate discussions and concise take-aways in line with the results presented in this manuscript.

General comments:

- Line 86: Please put a period after the sentence.
- Line 95 to 101: This paragraph is more appropriate in Line 1. Please bear in mind the cohesiveness on the next paragraphs.
- Introduction: Please include more recent studies similar to your work. There are quite a few on COVID-19 lockdown implications and traffic and some have similarly estimated a business-as-usual scenario. Here are a few of those: https://doi.org/10.1039/D3EA00013C, https://doi.org/10.1007/s11869-023-01330-3, https://doi.org/10.1080/02786826.2023.2193237.
- Line 122: Please include the units of each of the meteorological variables. What are the resolution of these observations?
- Line 124: What is the expected seasonality of the air quality parameters considered in the study area? Please elaborate on the seasonal trends of both EC and NO_x.
- Line 132: Was it Random Forest regression that was specifically used in this study? A modelling workflow would be useful to clearly present the Random Forest methodology.
- Line 138: Did you test on any other meteorological variable aside from the mentioned variables? Were all these variables/features included in the final Random Forest model? If yes, what was the criteria used to include/exclude features in the model?
- Line 138: Please elaborate on the units used for the wind direction feature. Did you perform any specific data processing on this variable or in any of the other features?
- Line 132: What is the size of the dataframe worked on this study? How many data points in total was used to train and test the model? How many data points were available for both target and features in the Random Forest? How many data points are available in the 3 periods that were compared in this study (during, transition, lockdown)?
- Line 144: How did you determine that this was the appropriate architecture for the Random Forest model in this study? Was the training and test sets partitioned randomly?
- Line 144: Please check if there is an existing decreasing trend for both EC and NO_x in the years prior to 2020. If there is, it could be useful to use the first 80% of the dataset as the training set and the last 20% as the testing set. In this case, the existing trends can be taken into consideration in the building of the Random Forest model.
- Line 148: What are the model performance measures used? In this paragraph, it was mentioned that a validation step was performed by comparing the predicted and measured ratio of NO_x and EC in the training and testing step. Please elaborate on why this methodology was used to validate the results. There are many different ways to validate a machine learning model. Have you performed an out-of-bag model validation

(using a subset of your dataset not included in both training and testing sets)? The validation step of any machine learning model has to be clear and elaborate to support that the model optimization was reached and substantial. This was not apparent in the methodology of this study.

- Line 157: I suggest the use the term "estimated" than "counterfactual".
- Line 165: Please add a period after the sentence.
- Line172: How did you calculate the differences representing the magnitude of reduction? In the results and discussion section, this difference is in a percentage form. Please be clear on how this was calculated? Is this percentage difference, percentage change? Was this difference calculated on a daily/hourly comparison or using the entire lockdown period? Please explain.
- Line 204: Please add the standard deviation on the mean concentration. Please apply this on all other concentrations mentioned in this section.
- Line 207: This sentence is confusing, please rephrase. Did you mean a 65% reduction was found and that is equivalent to a 43 $\mu g\ m^{-3}$ reduction?
- Line 223: I understand that the effect of the holidays can indeed affect the overall reduction of pollution levels, hence it does make sense to only use the last two weeks of the lockdown period. However, since the holiday is a yearly event, it would still be interesting to see the reduction of levels during the entire lockdown period even if it does include the holidays. This should be included in the analysis as well.
- Line 258: What is the correlation coefficient mentioned here? Is this Pearson or Spearman or something else? Please include the RSQ value as done in Line 265.
- Line 260: The term failed can be subjective as some studies have referred to this level of correlation as moderate. Please rephrase.
- Line 261: This sentence is quite redundant as this has been mentioned several times in the prior sections already. Please check other redundant statements in the manuscript.
- Line 271: The RSQ value is missing. Is this also 80% same as for NO_x?
- Line 277 to 290: Does the feature importance (SHAP values) make sense? In theory, temperature should probably have been a strong feature for NO_x, but it doesn't seem to be. Why do you think that is? What is the wind speed considered in this study, is this wind speed close to the ground? There is a need to elaborate the seasonal/climate conditions during the lockdown restrictions in the study area? What are the specific (for example) seasonal influence on the levels of NO_x and EC?
- Line 283: In this sentence, it was mentioned that low wind speed contributes up to 100 $\mu g\ m^{-3}$ in NO_x and high wind speed contributes negatively. However, Figure 4 presents that wind speed did not vary a lot. How did the authors categorize low and high wind speed in this analysis? A lot of the SHAP values are resting in zero value, which could mean very little impact any of the features. How about the temporal variables considered What is the impact of these variables?
- Section 3.5: The use of Satellite data to validate the results does not add much information in the analysis. Usually ground measurements are used to validate Satellite data. Hence, this analysis does not make sense and seems out of place.
- Line 341: Please check the tenses of the verbs used in the manuscript. Check for the rest of the text.
- Line 358: There are already existing studies focusing on traffic emissions, which also uses long-term datasets and using Random Forest to estimate BAU levels. There is a need to discuss/compare the results in this study with existing literature.
- Line 367: Please explain the geochemical meaning behind high importance of wind speed and wind direction as features in a Random Forest model.

- Section 4: The discussion section lacks elaborate discussions about the results of this study. How does the results of this study compared to other studies using a more traditional/simplistic approach on evaluating the effects of the lockdown restrictions on air pollution? Is there an under- or over-estimation in other approaches compared to using a Random Forest model? Since the measurements are in an hourly resolution, was there any diurnal or hourly variation during the lockdown period?
- Section 5: The conclusions section read like a summary. This section needs to be improved. I suggest using bullet points to enumerate the main take-aways of the study.

Additional comments:

- Was it only vehicle/transportation that was restricted in Shanghai? How about the industrial sector? There are other sources of NO_x and EC apart from vehicular emissions. Are there no other known anthropogenic sources in the area?
- Was there an already existing trend that needs to be considered apart from the impact of COVID-19 restrictions? A long-term dataset should be able to check this.
- There is a lot about the % reduction but the analysis and discussions lack on the model optimization. Can we use SHAP to improve the Random Forest model by choosing the most appropriate features to be used in the model?
- Figure 1a and Figure 2a both show time-series of NO_x and EC. Is this a daily average? If it is, please add the standard deviation for year 2020. This might show that in fact the change is not as apparent as it seems.
- It is essential that a dependence scatter plot be provided to show the effect of a single feature across the whole dataset. It will also be useful to just take the mean absolute value of the SHAP values for each feature to get a standard bar plot that can be easily interpreted.