Rebuttal- Reviewer 2

Dear Reviewer,

Thank you for giving us valuable suggestions to improve the presented research work. As per the given suggestions, we have incorporated all the changes and highlighted the changes in the red color and figure, and table-caption changes have been represented using yellow color.

Q-1 In general, the introduction to air pollution and its impact on human health as well as the introduction to the gap in the field of predicting air pollution concentrations that their paper is trying to fill feel scattershot and uninformative. It needs to be reworked so that the disparate ideas within the text are more cohesively tied together and more accurately explained, and so that the reader immediately understands what the author is trying to convey -- air pollution is bad because X, Y, Z; AQI has been used to communicate the health risk of air pollution concentrations but is imperfect; new advanced statistical methods can help; for example, we present the Pollution Weather System. I have provided a few specific comments to this effect, but altogether I suggest a full reconsideration and revision of content and presentation.

Answer: Thank you for giving valuable suggestions. Thank you for giving us valuable suggestions. As per the given suggestions, we have revised the introduction section by highlighting the suggested comments.

Q-2 The English grammar needs a great deal of revision to improve the clarity of the authors' ideas and analyses and their relevance to the topic at hand.

Answer: Thank you for giving valuable suggestions. As per the given suggestions, we have proof-read the whole manuscript and made the required changes throughout the manuscript including the introduction section. We have also highlighted the changes in the red color.

Q-3 What are the details of the 90-day exposure study? More context is needed for understanding the behavior of pollutants and sensors and the appropriateness of specific prediction methods employed in this paper. Was a human subjects review board involved in data collection?

Thank you for giving us valuable advice. As per the given suggestions, we have added the required details.

Q-4 The strengths, importance, or usefulness of AQI as a pollution risk metric should be more clearly explained – why did you chose this metric to predict?

Thank you for giving us valuable advice. As per the given suggestions, we have included an AQI risk metric table 6 and describe the same in section 2.2.3.

Q-5 A brief summary description of the device that includes all of the sensors together (listed in Table 1a and Section 2.1) is needed. Perhaps a few sentences before section 2.1.1.

Thank you for giving us valuable comments. As per the given comments, we have added the required configuration details as described in Table 1 to Table 5 and section 2.1, subsection 2.1.1 to 2.1.5. We have also added a Figure 3 which represents the system architecture of a PWP system.

Q-6 Optical PM sensors – and perhaps the other gaseous sensors – often require considerable calibration as the optical properties of aerosols change over the course an hour, day, week, season, etc. How is this effect accounted for? If not, a paragraph pointing out this weakness in low-cost sensors is needed -- perhaps in the Discussion

Thank you for giving us valuable comments. As per the given comments, we have added the required configuration details as described in Table 1 to Table 5 and section 2.1, subsection 2.1.1 to 2.1.5. We have also added a Figure 3 which represents the system architecture of a PWP system.

Q-7 A flow diagram describing the data sample sizes from measurement, to training, to testing (etc.) along with data missingness would be very useful.

Thank you for giving us valuable suggestions. As per the given suggestions, we have added the required details in the manuscript as shown in Table 7.

Q-8 Could you explain why or how you decided on the ANN parameters you chose (e.g., relu, 7 layers, the hidden layer sizes, learning rate of 0.2, etc.)?

Thank you for giving us valuable suggestions. As per the given suggestions, we have added the required details in the manuscript.

Line-specific commentary:

Q-9 Lines 37-42 (ending at "… mentioned above.") may not be necessary; removal would make the document more succinct.

Thank you for giving us valuable suggestions. As per the given suggestions, we have removed the required text.

Q-10 Lines 44, 58: "WHO Reports" is not a specific enough resource, and the reference items "1" and "2" could be more rigorous and primary. Cite the specific reports from which the information came or, perhaps, one or more of the recent Lancet articles on IHME's global burden of disease study (this is the other major reoccurring global GBD study): https://www.thelancet.com/gbd, especially https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)32225-6/fulltext. Also see: https://www.stateofglobalair.org/report

Thank you for giving us valuable suggestions. As per the given suggestions, we have added the required references in the introduction and reference sections.

Q-11 Line 49: "80% of urban breath air" is unclear. Do you mean "80% of urban populations breathe air"?

Thank you for giving valuable suggestions. As per the given suggestions, we have the text to 80% of urban populations breathe air.

Q-12 Line 49: WHO "limits" should be more appropriately referred to as "guidelines."

Thank you for giving valuable suggestions. As per the given suggestions, we have changed limits keyword proof-read the whole manuscript and made the required changes throughout the manuscript including the introduction section. We have also highlighted the changes in the red color.

Q-13 Line 51 (And elsewhere): define all acronyms the first time they are used, e.g. "AQI". Also, AQI is focused on in this paragraph as if it were a primary air pollutant rather than a mathematical

construct created to provide information on the relative dangers of the concentrations of a variety of specific primary pollutants.

Thank you for giving us valuable suggestions. As per the given suggestions, we have the text to more meaningful text.

Q-14 Line 53: be more descriptive with "a fair amount" – does this mean a lot, a little, a moderate amount, etc.?

Thank you for giving us important suggestions. As per the given suggestions, we have changed the word fair to higher and poor.

Q-15 Line 55: Rework "can reach the human system with ease" with more specificity.

Thank you for giving us valuable suggestion. As per the given suggestions, we have made the required changes.

Q-16

Line 60: I believe we know how PM2.5 is formed quite well under many circumstances. Do you mean that concentrations are difficult to predict?

Thank you for giving us valuable suggestion. As per the given suggestions, we have made the required changes.

Q-17 Line 61: The point of Annenberg et al. is that there are pollutants that affect both climate and health ("co-benefits"), and that BC and methane are two great examples of this. It is not that BC and methane are, specifically, the worst pollutants out there for health.

Thank you for giving us valuable suggestions. As per the given suggestions, we have made the required changes.

Q-18 Line 80: "far better" at what?

Thank you for giving us valuable suggestions. As per the given suggestions, we have made the required changes.

Q-19

Lines 99-155: This is a confusing section. Are you pointing to work that the Pollution Weather System improves upon, or are you demonstrating examples of other methods of air quality prediction, or both? It is unclear, and seems like it could be made into multiple paragraphs that connect an entire narrative, rather than jumping around between failures in the field vs. successes in the field upon which your work builds vs. approaches that are similar to yours. Moreover, it reads more like a Discussion Section than an Introduction. Perhaps some of this content is better suited for the Discussion?

Thank you for giving us valuable suggestions. As per the given suggestions, we have made the required changes in the introduction and related work section.

Q-20 Line 163: What is variability research? It does not seem to be a known field of study. Do you mean that your work focuses mainly on measurement (or some other type of) variability?

Thank you for giving us valuable suggestions. We interpret variability as a measurement of air pollutant AQI values and its variations as described in section 2.2.3. However, to avoid confusions, we have made the required changes in section 2.

Q-21 Line 173: "plagued with issues" should have at least a few high-level examples.

Thank you for giving us valuable comments. As per the given comments, we have made the required changes.

Q-22 Section 2.1: all sensors should have listed their lower and upper detection thresholds, measurement precision and resolution; name of measurement tech (e.g. optical particle counter, electrochemical cell, etc). Measurement accuracy as shown by third party tests where possible (e.g. http://www.aqmd.gov/aq-spec) would be very useful, as well.

Thank you for giving us valuable comments. As per the given comments, we have added the required configuration details as described in Table 1 to Table 5 and section 2.1, subsection 2.1.1 to 2.1.5.

Q-23 Line 242: temperature is not a pollutant. It is an environmental factor

Thank you for giving us valuable comments. As per the given comments, we have removed the word temperature which was written by mistake.

Q-24 Figure 3: not all components of this figure seem to be defined or addressed in the paper's text

Thank you for giving us valuable comments. As per the given comments, we have added the details of all the components and its interface. The Figure 3 has been changed to Figure 4.

Q-25 Line 268: "AQI indexes" is redundant, as AQI stands for Air Quality Index.

Thank you for giving us valuable comments. As per the given comments, we have removed the word 'indexes' and changed it to "values".

Q-26 Lines 271 and 389: Is "1,89,648" a typo or the Indian convention for "189,648"? I believe the 1,89,648 convention is specific to India and may confuse non-Indian readers. 189,648

Thank you for giving us valuable comments. As per the given comments, we have made the required changes.

Q-27 Line 275: confusing. Are "training examples" further divided into training and testing "phases"?

Thank you for giving us valuable suggestions. As per the given comments, we have made the required changes.

Q-28 Line 276: Which specific dimensionality reduction technique?

Thank you for giving us valuable suggestions. As per the given comments, we have included the feature selection dimensionality reduction technique and also given more details in Table 7.

Q-29 Section 2.2.3: a table identifying the specific model features used is needed

Thank you for giving us valuable suggestions. As per the given comments, we have included Table 7 to highlight model features and techniques.

Q-30 Line 283: Which scaling technique, and how are the data already organized in a way that makes the scaling technique unnecessary?

Thank you for giving us valuable suggestions. We have made the required changes. In the conducted experiments standard-scaler scaling technique has been used.

Q-31 Line 287, 312, 319: What does it variability in "µm" supposed to be a mass concentration like µg/m3? If so, was the output from sensors that report in ppm and ppb converted to the appropriate corresponding mass concentration? The "100" does not seem to correspond to a number on the graphics. Generally, do you mean categories are created of size 100 µg/m3 and 25 µg/m3 (or something along these lines)?

Thank you for giving us valuable comments. As per the given comments, we have added the required conversion formula as described in eq(2) and eq(3) and section 2.2.3.

Q-32 Lines 289-305: The point of this visualization as well as the data that underly it. Are these measurements taken in the field [if so, which measurements] vs. predicted AQI? Are they just graphical representations of the mathematical AQI formula?

Thank you for giving us valuable comments. As per the given comments, we have added the required formula of AQI as described in eq(1) and section 2.2.3.

Q-33 Lines 370-373 seem like they should be Results, if these are the parameter values you arrived at using the gradient descent methods you describe.

Thank you for giving us valuable suggestions. Yes, the parameter values have been calculated using gradient descent methods as described in 2.2.4a.

Q-34 Line 436: Linear "regression"?

Thank you for giving us typo corrections. We have made the required changes.

Q-35 Line 516: "Particulate" matter?

Thank you for giving us typo corrections. We have made the required changes.

Q-36 Line 541-542 (and again on 579-581) is far too strong of a generalization and very likely untrue in many cases.

Thank you for giving us valuable suggestions. As per the given comments, we have made the required changes. We have done re-experimentation and recorded better accuracy with the customized ANN regression methodology.

Q-37 Figure 11 suggests the system is predicting pollution concentrations and not AQI, which is antithetical to how the rest of the paper reads. Moreover, are these predictions retrospective or forecasts? Overall, the presentation of methods and results is quite confusing.

Thank you for giving us valuable suggestions. As per the given comments, we have updated the required images as shown in Figure 10.