Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-55-RC1, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Gaussian Process regression model for dynamically calibrating a wireless low-cost particulate matter sensor network in Delhi" *by* Tongshu Zheng et al.

Anonymous Referee #1

Received and published: 30 May 2019

General Comments:

Overall, the paper presents an interesting approach to in-field low-cost sensor calibration using Gaussian Process models. I have identified several points below which should be addressed before publication.

First, a better description of the algorithm should be supplied. The equations are not matched well with their descriptions in the text, and multiple steps of the process (e.g. the linear regression and Gaussian Process hyperparameter optimization) are described simultaneously. The diagrams of figure 3 are helpful, but not sufficient to clarify the entire process. A complete step-by-step breakdown of an example run of

the algorithm could be provided.

Second, it appears that both the Gaussian process hyperparameter calibration and the linear regression calibration of the low-cost nodes are carried out over an approximately 60-day period, using all data collected during this period. This would seem to preclude the use of your methods for on-line calibration. You may want to examine how this technique could be used in an on-line fashion, but designating a "current time" within the dataset and only using data collected prior to that time to calibrate the Gaussian Process hyperparameters and linear regression coefficients which are used to correct the data for that time. You could then also examine the effect of time history on your model, analyzing how the performance changes as more or less past data is included in the calibration process. As it currently is, if I am understanding your approach correctly, it can only be applied retroactively to a designated period of time for which all sensor data are available.

Third, when analyzing possible failure modes of sensors to determine if the algorithm can detect these modes, only two modes are considered: linear drift over time and replacement of the sensor signal with random noise. Other common failure modes should also be examined. These should include a "random walk" baseline drift (rather than simple linear drift), flatlining of the sensor (either at zero or at a non-zero value), and noisy corruption of a true signal (i.e. adding a random noise to the original signal, rather than completely replacing the true signal with random noise).

Finally, while the body of the paper presents a good discussion of the limitations of the proposed approach (mainly its need for spatial homogeneity in the true concentrations to be fully effective), this discussion is missing from the abstract. I believe that this observation is an important result of this paper and should be highlighted in the abstract as well.

Specific Comments:

Page 1, Lines 25-29: This is a very long and complex sentence; consider splitting in

into several sentences and/or revising how the information is presented. For example: “Simulations conducted using our algorithm suggest that in addition to dynamic calibration, it can also be adapted to automated monitoring of WLPMSNs. In these simulations, the algorithm was able to differentiate malfunctioning or singular low-cost nodes by identifying aberrant model-generated calibration factors (i.e. slopes close to zero and intercepts close to the global mean of true PM2.5). The algorithm was also able to track the drift of low-cost nodes accurately within 4% error for all the simulation scenarios.”

Page 1, Line 27: I am not clear on what is meant by a “singular” node.

Page 2, Line 14: I assume you mean “since the emergence of low-cost AQ sensors” Rather than “since the emergence of calibration-related issues”. It might be better to state that.

Page 3, Line 10: These coordinates are likely too precise to denote the city of Delhi generally. It is probably sufficient here to just state “Delhi, India”, rather than providing coordinates, unless you are trying to describe a specific location within the city.

Page 3, Line 18: Rather than “drift nodes” I would say “the drift of nodes”.

Page 4: Line 15: Use “the” rather than “our”.

Page 5, Lines 6-7: It is not clear to me why the GPR model would require data from all stations to operate. If it is interpolating between stations then it should be able to fill in for any missing station data as well.

Page 5, Line 20: The meaning of “with that of after missing data imputation” is not clear.

Page 5, Line 21: I don’t know if “imputed” is the correct word to use here.

Page 5, Line 26: Should be “while outliers have scores significantly larger than 1”.

Page 6, Line 19 to Page 7, Line 5: This description could be improved. In particular,

it is not clear how the alpha and beta parameters of Equation 3 are determined. The description seems to combine a linear regression and a calibration of the hyperparameters of the Gaussian Process. These two steps should be described separately.

Page 6, Lines 23-24: The process of “standardization” is not clear to me. If this is done separately for each node, wouldn’t this eliminate any systematic differences between measurement locations? If this step is only done to the data which are to be used for calibrating the model hyperparameters, then that should be stated. Even so, it is not clear that this is an appropriate step; for example, two node may be systematically higher than other locations, and so should have a mutual correlation, while if the means are subtracted, the data from the nodes would no longer be correlated (in other words, two variables can be made similar in a GP model either by giving them a high mutual correlation or by giving them a smaller prior variance and the same prior mean).

Equation 4: What is Gamma?

Page 7, line 7: What does the bold-face Theta denote? Are these the hyperparameters of the GP model as described in Equation 2?

Page 7, Lines 10-11: It is not clear what it means to re-calibrate a node based on its posterior mean. I am assuming this involves adjusting the alpha and beta parameters, but this is not clear.

Page 7, Line 29: What criteria are used for convergence?

Page 7, Line 31: This relates to a previous comment, I believe, but it should be described how the predictions are transformed back into the original PM scale.

Page 9, Lines 10-12: This sentence can be better written as “. . .the reference node mapping accuracy follows a pattern, with relatively high quality prediction for those nodes whose means are close to the global mean (e.g., global mean $\pm$ SD as highlighted with shading in Table 2) and relatively poor prediction for those nodes whose means differ substantially from the global mean (particularly on the lower end)”.

Page 9, Line 21: It is unclear what the “scale of 10” refers to.

Page 12, Line 4: It is unclear what “quality drift estimation” is.

Page 12, Line 11: This should be “Questions which remain unsolved”.

Page 13, Lines 24-27: The end of this sentence may be incomplete.

Page 13, Line 31: Again, it is not clear what is meant by a singular node.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-55, 2019.