General comments:

This paper uses two calibration approaches (one using an approach of calibrating at a monthly interval, one using a test system to calibrate given certain criteria are met) using low cost sensors and nearby “proxy” sites. The approaches build off previous work and seem to provide good performance, although comparisons to other methods are lacking. The article is generally well written with a few sections noted below needing some clarification and attention to units.

Specific comments:

For readability, consider changing the names of the various AQY sensors (e.g. AQY BD-1146) to something more related to their deployment (e.g. AQY RIVR 1) or use a simple number scheme (e.g. AQY 1) for use within the paper. A table with the original names could be provided in the supporting information.

Section 2.3: Can the authors comment on what percentage of the data included fog and was thus discarded from the datasets? Additionally, do other processes that might impact visibility such as wildfires pose a risk for removing data from the calibration periods? Is the main issue with the fog from issues with hygroscopicity (or more generally from high humidity) or with visibility itself? Looking at Figure S1 for the month of November 2021 it seems like a small portion of the dataset was excluded (maybe 10% or less). I recommend that this percentage (either total or broken down by month) be mentioned in the article.

Section 2.3: Could the authors comment on how the drift calibration approach deals with periods dominated by local sources? Are there any checks that are made to determine if the reason for a calibration alarm is due to a local source that would not be picked up by a proxy? For the purposes of this study the collocated reference monitors can be used to verify, but for future deployments is there a protocol?

Section 2.3: What is meant by a “suitable seven-day calibration window” for the monthly calibration approach? Other than removing fog and ensuring data completeness, are there other metrics for choosing the most appropriate window out of the two week period of consideration? Additionally, is this window chosen with the help of the next month’s data or is made completely independently of the next month’s data?

Section 2.4: What are the criteria of classification for roadways used in the analysis? For instance, is the distance of the site from a motorway mean to the nearest highway or does any road at all count?

Section 3.2: How was the most similar land use proxy site determined? I do not see any data related to this (e.g., the metrics discussed in section 2.4 related to roadways). On line 207 the authors state that the nearest proxy was generally more useful than the one with most similar land use (which is unsurprising). Why did the authors choose to use the nearest proxy for all sites rather than the ones which performed the best and have a mix of nearest and similar land use? If it is for simplicity, I would suggest mentioning that.

Section 3.3.1: I would suggest including metrics on the performance statistics discussed on line 224 as “good” can mean different things to different readers. This also applies to line 279.

Section 3.4: Figure 10 needs an explanation in the caption or the legend for the difference between the step changes and the continuous curves.

Section 4: The article could benefit from a table summarizing the performance of the calibration approaches so the reader does not need to use multiple figures to determine the efficiency of each approach (this table could be in the results section and summarized briefly in the conclusions). Additionally, a comparison to other methods of calibration should be discussed. The strengths and weaknesses of the monthly and drift approaches should be discussed relative to these other options.

Technical corrections:

Consider adding units to Figure 1 a and b for latitude and longitude (e.g. [° N]).

Line 41, consider changing “with particle type or properties changes over time” to “with particle type or their properties may change over time”.

Line 59, consider writing out “Los Angeles” for the first time. It is written out on line 74 currently.

Line 70, I believe the non-regulatory air monitors discussed in this line are the AQY systems discussed in section 2.2. Consider clarifying this in section 2.1.

Line 104, consider changing “drive” to “driven”.

Line 167, consider replacing the comma between PM25 and PM10 with “and”.

For Figure 2 consider lining up the various sites vertically and leaving gaps for missing data so that they can be easily compared between the graphs. Not critical but would help the reader compare.

Line 191, consider adding in “respectively” after R^2.

Line 215, consider adding units to the distance columns of Table 1.

Line 264, Figure 5. Some of the equations on the individual panels are cut off. Same for Figure 7.

Line 256, consider changing to “as often is the case”.

Line 275, Figure 6 caption. Add in the metric being plotted to the description of panel a (i.e. MAE).

Line 314, Figure 9. The a0 panel should include units of $\mu g\ m^{-3}$.

Line 319, add in a closing parenthesis at the end of the line.