UPPER HUNTER MINING DIALOGUE AIR QUALITY MONITORING DATA ANALYSIS PROJECT

CSIRO Review Comments Register - 20200917 (Prepared by ERM)

#	Comment	Response
1	When discussing differences in annual mean PM10 and PM2.5 concentrations between sites and groups of sites determine if the differences are statistically significant at a certain confidence level (generally 95% confidence is used but you could report significance at other levels). This will determine whether the differences are due to variability within the dataset used to calculate the mean concentrations or due to differences in the mean. Reporting/plotting standard deviations with mean concentrations will also indicate this.	We haven't been able to apply confidence intervals to annual data for station groups on the limitation that the several regional groups (and the UHAQMN background group) include insufficient data points to establish a standard deviation within the group/year. As raised in later comments, a statistical analysis of mean differences has been added for consideration (see Responses #41, #42, #43).
2	When referring to correlation, perform a regression between the variables and determine if the correlation coefficient is significant at a certain confidence level, ensuring to account for outliers.	Regression analyses have been added for correlations reviewed in Section 6. Analyses provided in Appendix A, and include significance tests at a 95% confidence level with consideration of outliers. References to Appendix A have been added to the relevant parts of Section 6.
3	Page 1 last paragraph- use of the word whilst-while not incorrect I think that the use of whilst instead of while detracts from the readability of the report.	Whilst' has been replaced with 'while' across all instances in the report.
4	Page 5 Paragraph 1 It's not clear what the point of this paragraph is, it doesn't present any new information except the reference to the standards. I suggest just including:

The health-based assessment criteria used by NSW EPA have, to a large extent, been developed by reference to epidemiological studies undertaken in urban areas with large populations where the primary pollutants are the products of combustion (National Environment Protection Council [NEPC], 1998a; NEPC, 1998b). | Text has been revised in accordance with suggestion. |
| 5 | Page 6 Table 4.1 A map show the station locations would be very helpful | Map of station locations has been added (Figure 4.1) |

	6	Page 7 Para 2 (starting at line 7) the description of the colour gradients needs to be clearer and should match the information given in the legend for Table 5.2 etc. The legend indicates that red is the 100% (maximum). The text on page 7 states Results have been shaded using a green to red colour relative gradient scheme with lowest values shown in green, and highest values shown in red, with the median value shown in yellow – is this gradient across all the sites and all the years or is the gradient applied to each station across the years or to each year across the stations? The discussion on page 9 indicates the gradient is across all years for all stations however it will be clearer to specify this here.	Clarifications made to text (Section 5, Paragraph 2 - final line). Text added to legend within figures to reinforce grouping.
	7	Page 7 Table 5.1 Note before this section (maybe in section 4?) that you are using data from X number of (how many?) DPIE AQMS, otherwise it's not clear here why you are excluding these stations listed in Table 5.1 (at this stage the reader is not even aware that using these stations in the first place).	Context added to beginning of Section 5, including the total number of DPIE AQMS across NSW (i.e. in addition to UHAQMN).
	8	Page 8 Table 5.2 Define BG, DF, LP and SC in the UHAQM cells in column 1	Table footer added to define abbreviations.
	9	Page 9 Para 1 Suggest changing *The second highest concentrations are measured at Camberwell, as associated with mining emissions and other sources such as wood smoke, as identified within UHAQMN reporting (OEH, 2016).* To: *The second highest concentrations are measured at Camberwell, where PM10 concentrations have been associated with mining emissions and other sources such as wood smoke (OEH, 2016).*	Text modified as per suggestion.
	10	Page 9 Table 5.3 here and elsewhere – when referring to regions use Title case (i.e. the first letter of each word in the region is captilised). For example, Lower Hunter and Central Coast region. Also use Diagnostic station instead of diagnostic station in the text.	Text modified as per suggestion.
	11	Page 9 Para 3 here and elsewhere where the term 'average' is used, include the period being averaged e.g. annual average	Original text "Across NSW, the UHAQMN diagnostic stations show the highest average (24 µg/m³) across all years" retained. This text is explicit in defining that it is an average across all years.

	12	Page 9 – Para 4 This difference is expected to be primarily due to the varied influence of bushfire and dry conditions (inclusive of dust storms) in late-2019 between regions. – can you add some information that shows there were more bushfires, smoke and dust events in 2019 compared to other years?	We don't have a concise reference for the influence of these events on ambient air quality across NSW (given their recent nature). Given the statement is providing context, it's considered appropriate in its current form.
	13	Page 9 Final paragraph Whilst these data are not weighted spatially or by population, they are useful in considering interannual variability at a broad scale -why would weighting the data provide any better indication of interannual variability?	This text is stating that irrespective of selection bias implicit in the varying spatial density of regions, it is still useful to look at an overall average. Text amended to reinforce that looking at a region-based average is useful.
	14	Page 9 Final paragraph reference to progression of drought conditions- while more discussion about rainfall etc does occur later in the report, I suggest you also include a reference to a BoM statement or other evidence about the region being officially in drought.	Reference to 2019 DPI drought synopsis added.
	15	Page 13-16 Figures 5.2-5.4 These figures are not referred to in the text. Include standard deviations as error bars on the bars? Move to the appendix?	Text added to refer to these figures. Inclusion of standard deviations / error bars would not be consistent with the objective of these charts being to present the data. In addition, the standard deviations (of annual variability) by station would be of limited use, as we are not seeking to make inferences on future annual variability, or to establish the significance of the study period relative to a broader period. We feel that this aspect is best covered by presenting the data.
	16	Page 19 Section 6 Use of the term correlation – correlation or visual relationship? As stated above when referring to correlation, perform a regression between the variables and determine if the correlation coefficient is significant at a certain confidence level, ensuring to account for outliers.	'Correlation' is being used in a plain english context and is considered fit for purpose. Regression analysis has been added for correlations reviewed in Section 6. Analysis is provided in Appendix A, and references to Appendix A have been added to the relevant parts of Section 6.
	17	Page 19 Section 6.1 *Raw coal production is considered a general indicator of the intensity of mining operations within the Hunter Valley, and hence provides context on the potential contribution of mining operations to changes in ambient air quality.* It would be good to include a brief discussion about other metrics that could be used as an indicator of the intensity of mining operations and why raw coal production was chosen over others—could be a practical reason like the most accessible and complete data set.	Original text retained. Additional context on other potential metrics such as volume of material moved or diesel consumption would compromise the continuity of this section - i.e. we are looking for indicators of PM10 emissions - given that we progress to NPI PM10 emission quantities from here, examination of other variables isn't considered warranted.

18	Page 19 Footnote Table 6.1 – there isn't an * in the table	Footnote removed from table.
19	Page 19 Table 6.1 State why was the 2013 coal production value used as the reference to compare other years to?	Context added to text below Table 6.1.
20	Page 19 2nd Para below Table 6.1 *The ratio of particulate matter emissions to raw coal production would be influenced by several factors including the following:* The ratio of particulate matter emissions to raw coal production implies a new metric, which is not presented in the table or figure. Figure 6.1 compares the PM and production but does not present a ratio. Perhaps you could change the text to t*he relationship between PM concentrations* (since you are not actually presenting emission data here) and *raw coal production.*	Wording amended as per suggestion.
21	Page 19 5th Para below Table 6.1 *Noting the progressive nature in which the majority of mines have developed within the Hunter Valley, as well as variety of mines operating across different points of their lifecycle, coal production data is considered generally relevant to this analysis.* This paragraph suggests coal production isn't a good metric.... perhaps you could say that coal production data is ONLY considered generally relevant.	This text is suggesting that coal production is a useful metric, as the factors raised in the 3 points above will average out across all of the mines. Wording amended to reinforce this.
22	Page 20 Figure 6.1 Does the plot actually show a correlation analysis? It shows that there doesn't appear to be much of relationship between the variables but to show whether or not there is a correlation you should include a statistical test e.g. a regression. I think this plot is open to interpretation depending on your point of view e.g. the lowest PM concentration in 2015 and 2016 happen when the coal production rates are lowest.	Text amended. Regression analysis has also been added in Appendix A to reinforce correlation. Reference to Appendix A have been added to text.
23	Page 20 Para 1 As shown in these data, there is no clear correlation between raw coal production and ambient PM10 concentrations measure by the UHAQMN, an inconsistency in the scale of changes across 2013 – 2019. As a further supplement to this analysis, a comparison against NPI mining emissions is provided in Section 6.2. What does an inconsistency in the scale of changes across 2013 – 2019 mean? Please clarify	Text amended to clarify the inconsistency in the scale of variability.

	24	Page 20 Section 6.2 Paragraph 1 The NPI does not track pollution it just collates data on emissions.	Text retained as it is referenced directly from the NPI website .
	25	Page 20 Paragraph 4 Facility operators determine their own emissions and transfers, and diffuse emissions from households and other sources like motor vehicles are estimated by government agencies. Suggest changing determine to calculate and report It would be good to include a brief discussion about what is done in the NPI to verify the emissions being reported. This may address concerns around under reporting.	Clarification is required on the verification programs alluded to.
	26	Page 21 -Para 1 state why 2013/14 was chosen as the reference time	Text added below figure 6.2. It is noted that 2012/13 is the reference year.
	27	Page 22 Table 6.2 2018/19 data are also shown in the table	Table heading amended.
	28	Page 23 Fig 6.3 2018/19 data are also shown in the figure	Figure heading amended.
	29	Page 24 Figure 6.4 again Does the plot actually show a correlation analysis? It shows that there doesn't appear to be much of relationship between the variables but to show whether or not there is a correlation you should include a statistical test e.g. a regression. Again, I think this plot is open to interpretation depending on your point of view- e,g, the increase in PM concentrations and emissions in 2018?	Text amended. Regression analysis has also been added in Appendix A to reinforce correlation. Reference to Appendix A have been added to text.
	30	Page 24 Section 6.3.1 Paragraph 1 Suggest changing the text to make it clearer: *Here we consider the trends in PM concentrations observed in the UHAQMN in the context of changes in PM concentrations measured at other locations in the NSW monitoring network. An increase in the influence from mining in the UHAQMN would lead to increases in annual average PM10 concentrations that are not observed elsewhere in NSW. Conversely, consistency with the remainder of NSW would indicate the regional factors influence the PM concentrations rather than emissions from mining operations.*	Text amended incorporating suggestion. Minor amendments to suggested text to remain consistent with reporting conventions - e.g. the report has not used 1st person.

#	Comment	Response
31	Page 25 Table 6.3 caption -Please define % of average? Is this average of all years? If so why is that a useful metric to compare? In other parts of the analysis (e.g. coal production values in Table 6.1) you have compared to the 2013 data. You may need to include a brief discussion about how variability can be measured in general and why you have chosen this method- particularly since the Figure 6.5 shows the point that UHAQM shows the same trend as the rest of NSW and this is a key message from the analysis.	Text amended. Footnote added to clarify the choice of average.
32	Page 26 Para1 This paragraph is one of the key messages - however it needs to be rewritten to clearly bring out the point. I suggest: *Consistency in the pattern of changes in PM10 concentrations as function of time between the UHAQMS and the remainder of NSW stations suggests that all sites are influenced by regional PM sources and that the contribution of mining operations on the UHAQMS is not discernible. This finding was not sensitive to the inclusion of Stockton data, with 'Remainder of NSW' values showing 1% or less variance with the exclusion of Stockton data.*	Text revised in accordance with suggested wording. Reference to regression analysis is also provided.
33	Page 27 Section 6.3.2 Para 1 suggest replacing the UHAQMN with the UHAQMN dataset and widening the gap with greater difference.	Suggested revisions incorporated.
34	Page 28 Figure 6.6, Table 6.4 and discussion in Para 1 as discussed above suggest including a statistical analysis of whether the annual differences are significantly different. What is the variance in each of the averages? This is particularly important since one of your key message is that the PM10 concentrations measured at the Diagnostic stations is greater than at the other stations.	Appendix A2 added to assess the statistical significance of annual differences.

	35	Page 28 Paragraph 2 This paragraph needs to be reworded so that it is very clear what is being compared as you have introduced a new type of analysis here. Also, I am not sure the table is needed as the plot shows the point very effectively. Suggested rewording: *Table 6.5 shows the difference between the annual PM10 concentrations for the 2013 – 2019 and the annual averaged concentration for each year for each of the different station groups. The same data are shown graphically Figure 6.7.*	Text expanded to improve clarity. Table retained to provide transparency / assist the reader.
	36	Page 29 Para 1 This paragraph repeats information on page 27 so could be deleted	Paragraph deleted.
	37	Page 29 Section 9.4 Define how the rainfall anomaly is calculated. Also is it the BoM analysis that determines 2019 is the driest on record or your interpretation?	Footnote (6) added to define rainfall anomaly. The figure shows that 2019 features the highest negative rainfall anomaly, hence this is not interpretation, but a feature of the record presented.
	38	Page 30 Fig 6.8 Is this Figure reproduced from the BoM? if so, include a link to the original figure and acknowledge the BoM in the caption.	Image source is shown in figure. Image source reference has now also been added to base of figure as well. The link for this chart is included in the reference for these data (BoM, 2020).
	39	Page 30 Table 6.6 It's not clear that the data presented in this table are differences, they look like annual values. where in the table are the differences against the 2013-2019 average for each of these parameters?	Text referring to differences deleted.
	40	Page 31 Paragraph 1 Again not actually a correlation but a visual inverse relationship. Suggest carrying out a regression analysis as discussed above.	Negative correlation nominated to be retained. Regression analysis has also been added in Appendix A to reinforce correlation. Reference to Appendix A has been added to text.

	41	Page 32 Conclusions - Finding 1: *Annual average PM10 concentrations within the Upper Hunter are broadly consistent with Lower Hunter, but also higher than a range of other regions within NSW. The difference between the Upper Hunter and average concentrations in NSW is small relative to the variability observed between years.* This statement needs a statistical analysis to understand the significance of small relative to variability between years.	Appendix A2 added to assess the statistical significance of annual differences as per CSIRO comments.
	42	Finding 2: *Across the UHAQMN, the difference between background, diagnostic and population-based station groups is indicative of an influence from mining as well as other anthropogenic sources. Lower concentrations are observed at background stations, and higher concentrations are observed at diagnostic stations. Concentrations observed at population-based stations fall between these two monitor groups.* Was there a statistical difference between the Background and Diagnostic stations?	Appendix A2 added to assess the statistical significance of annual differences as per CSIRO comments.
	43	Finding 3: *A further analysis of the trends in station group values was conducted for the review period. This analysis has shown a consistent difference between station groups (within each year), indicating that changes in Upper Hunter PM10 concentrations over time are associated with regional conditions such as rainfall and are indicative of a minimal change in the contribution from local emission sources inclusive of mining.* Is this a statistically significant difference? Was the relationship with rainfall significant? Rather than state indicative of a minimal change in the contribution from local emission sources including mining you could state that the contribution of local emissions including mining may have been less than the influence of regional conditions.	Suggested to retain original wording, as the physical plausibility of the noted relationship (decreasing concentration with distance from source) supports the use of the word 'indicative'. A regression analysis against rainfall has been added. Appendix A2 added to assess the statistical significance of annual differences.

	44	**Finding 4:** *Significant increases were observed in annual average UHAQMN PM10 concentrations between 2017 and 2019. These increases have been found to be generally consistent with trends observed across the remainder of NSW, which show a correlation with the progressive decrease in annual rainfall and increased prevalence of drought conditions.* Why is the difference between 2017 and 2019 and not 2013 and 2019 as presented in the rest of the analysis? Is the increase statistically significant or within the noise of variability?	We have reviewed 2013-2019 in the report. The increase we are referring to occurs in the period 2017 to 2019. Hence the statement is appropriate as it is referring to something that occurred during a portion of the study period. Regression analyses have been added (Annual average PM vs Mean Annual Rainfall, and Annual Average PM: UHAQMN Vs Remainder of NSW) to assess the significance of the quoted correlations. These regressions support the original statements, hence they have been retained as is.
	45	**Findings 5:** *A review of trends in mining operations has not found a correlation between ambient PM10 concentrations and coal production, or ambient PM10 concentrations and NPI reported PM10 emissions.* Is this a correlation or a visually identified relationship?	Text retained - A correlation has not been found. Regression analysis has been added to support.
	46	**Finding 6:** *A review of NSW/ACT average rainfall has shown a correlation between below average rainfall and above average UHAQMN PM10 concentrations.* Is this a correlation or a visually identified relationship?	Text retained - this is a correlation. Regression analysis has been added to support.
	47	**Finding 7:** *As consistent with PM10 monitoring results, PM2.5 concentrations are elevated across 2018, and 2019, with highest concentrations measured at the UHAQMN large populations station group, which likely contains a significant influence from wood smoke (CSIRO, 2013). Trends in annual average concentrations were also found to be consistent with the remainder of NSW.* Qualify that woodsmoke is important during winter months.	This comment was inadvertently missed. Text has been revised to clarify that wood smoke is associated with wood heating (conclusion and executive summary) should be: *As consistent with PM10 monitoring results, PM2.5 concentrations are elevated across 2018, and 2019, with highest concentrations measured at the UHAQMN large populations station group, which likely contains a significant influence from wood smoke **during winter months** (CSIRO, 2013). Trends in annual average concentrations were also found to be consistent with the remainder of NSW.*

| 48 | Page 32 and Page 33 -Answers to the objectives- as stated throughout this review, support the conclusions about correlations and significant variation etc with objective statistical analysis. | Report has been amended to provide a statement about the statistical significance of the correlations reviewed. We have not amended the conclusions as they still stand as is. Supporting material has been added as Appendix A, with references to this analysis within the relevant sections of the report. |

| | = Revision made since 20200828 draft. |