

Review of manuscript essd-2020-276 “The winter 2019 air pollution (PM<sub>2.5</sub>) measurement campaign in Christchurch, New Zealand” by Dale et al.

### **General comments**

The manuscript presents data from an intensive campaign to measure spatio-temporal variations in airborne particulate matter (PM) concentrations across a large urban area in New Zealand through the winter months of 2019. PM measurements were made with 2 types of instrumentation at ~50 sites and were complemented with additional surface meteorological stations and boundary layer profiles from balloons and ground-based remote sensing. The measurement campaign aimed to generate the data required to generate hourly emissions maps using inverse methods.

Datasets of this scale are rare, especially in New Zealand where the dominant PM source (solid fuel domestic heating) differs from many cities internationally. The scale, along with the addition of atmospheric boundary measurements make this data unique. The data will undoubtedly be useful for the stated purpose as well as for direct data analysis and model validation exercises. The data is mostly well described and complete. Specific comments address areas where further description would be useful, particularly regarding the methods used to calculate uncertainties and a description of fraction of data flagged by each quality control step for each instrument and location. Also, the data from the inter-comparison periods is notably absent from the datasets, so it is impossible to test the author assertions about derived calibration coefficients and uncertainties.

Regarding data quality, the data are available at the stated DOI's. Given that the authors do not intend for the 'version 2' PM correction algorithm to supersede the 'version 1' the datasets should be contained in the same file. Currently the website lists each version as an 'update' to the previous versions, so it appears as if the 'raw' and 'version 1' data are redundant. The data from the meteorological analysis are not available (for copyright reasons), neither is a list of sites used for meteorological data (only points on a map), so the readers cannot retrieve the data from the respective repositories. No analysis of which data is removed from the full AWS dataset presented, so the quality control analysis of meteorological data presented in the manuscript is largely redundant. These issues could be addressed by listing the sites used (along with metadata) in table and providing a timeseries of quality flags for each site for each 10-min period in the netCDF dataset. In addition, a much fuller description of the miniMPL and ceilometer output variables is needed for this data to be useful.

Regarding the presentation quality – the manuscript is well written but contains several sections and paragraphs that are extraneous and distract from a concise description of the data and processing (see specific comments for suggested edits). The methods are mostly well described with the exception of the method used to calculate uncertainties where a more detailed description of the procedure and equations used is needed, along with further tables and figures describing the uncertainties. The lack of a clear treatment of uncertainties is evident in the conflicting statements regarding the uncertainty in the results and conclusions.

I am sure that with some clarification of methods and addition of the inter-comparison and meteorological quality control data to the repository that the manuscript will present a make a good contribution to the scientific literature and provide a significant and useful dataset to the community.

**Specific comments** (line numbers in **bold** and suggested rewording in *italics*)

**5.** Data from only 46 to 49 ODIN instruments is presented here. Please revise.

**13-15.** “We find that for while for the ODINs a correction based on environmental conditions is beneficial, this results in over-fitting and increased uncertainties when applied to the measurements obtained using the more sophisticated ES-642s.” This sentence is ambiguous, please reword (e.g. *A correction based on environmental conditions improves the quality of data retrieved from ODIN instruments but results in over-fitting and increased uncertainties when applied to the more sophisticated ES-642 instruments*)

**13-15.** This statement does not appear to be supported by analyses in the paper – both versions of the correction algorithm are shown to reduce errors with respect to the reference. No analysis of the effect of the correction on the uncertainties is presented. Please revise to be in line with results presented in the manuscript.

**57.** “data required to test and validate the MAPM methodology”. Some additional context is needed here to elaborate on what the requirements for the inverse method are (e.g. spatial and temporal resolution and extent) and whether these data exist in NZ or elsewhere. This will help justify the contribution of the dataset to the literature.

**70.** Some discussion of similar PM field campaigns in other locations (especially internationally) would give useful context for this dataset.

**76.** “additional 10 low cost nephelometers units were”. Please give details of these units in the text.

**105.** You mention the second most common wind, which begs these question – what is the most common wind? And how do these typical meteorological conditions relate to the occurrence of poor air quality?

**113-125.** It is not clear how these statements are relevant to the measurement campaign - consider removing for brevity.

**135.** A clearer description of the campaign timeframe including designation of the 'deployment' period is needed here. Also consider changing title of section to include description of time periods for easier reference.

**138.** “i.e. around 2 to 3 m above the surface”. This is quite an important piece of information – can a description of the variation in height be given?

**170.** A short description of how the ODIN determines concentrations in different size fractions would be useful here.

**244.** No planetary boundary height data is provided in files. Please revise and/or provide the reader with links to software than can be used to derive PBL height.

**245.** Data is only provided till mid-July – please revise text or provide data for the whole campaign online.

**297.** It would be useful if the authors refer specifically to the structure of different versions of the datasets stored online and the various flags contained within them. e.g. “ three version of the PM data were produced, the first “raw” version has pre-screening applied. Two corrected versions...” This would guide the reader and remove ambiguities in nomenclature between the paper and dataset.

**303.** please describe how much data from each instrument was flagged as missing or out of range. A table or some summary statistics for each instrument type would be very useful context for readers evaluating the dataset for their intended future uses.

**310-334** – you do not present the data from the smoke barrel tests or use it in corrections, so I would suggest removing this section.

**340.** “several large sections of data” – please be more specific and provide a summary of how much data from how many sites became unusable because of timing errors.

**347.** “was flagged in the netCDF files as such”. Please be more specific – what is the flag called.

**376.** “These steps were repeated...”. what was the motivation for introducing a new basis function? perhaps better to introduce both together, rather than the second appearing as an afterthought. At present it reads like an additional step required for the version 1 processing, rather than a separate method.

**377.** Please give some more explanation of why these particular formulations were chosen. i.e. what is the justification for a polynomial term for PM<sub>2.5</sub>?

**413.** How do you know that 3 passes is sufficient or that 12.5 % of the data is an appropriate level of data to remove? Based on the data presented in Figure 5, there is no obvious change in the histogram of differences at 87.5% that would suggest a break between good and poor quality data?

**416.** A table describing the sites used along with how much data has been removed from each site would be a useful as a reference here.

**424.** “recommended values’ what do you mean here? are you referring to the proxy time series? Or do you mean the time periods recommended for use? Also, do you provide these for just the three sites you operated or all 30 sites? Please revise for clarity.

**444.** “in the absence of further co-location data’ did you consider using the co-located ES-672 and ODIN? Especially given that i) corrections for ODIN instruments had similar movement between the two co-location periods, ii) the ES-462 had fairly stable corrections. These would provide timeseries of ODIN correction against the ES-642, which would allow you to test your assumption that a change in emissions and/or environmental conditions were driving these differences. Perhaps by accounting for changes in the corrections on a daily or hourly basis would reduce the ODIN errors shown in Table 1.

**466** “very weak correlations with temperature or relative humidity”. Figure 10 appears to be at odds with this statement – showing clearly larger scatter with higher RH, implying that there is larger uncertainty at higher RH. This seems a critical aspect of the uncertainty given the environment that the instruments are being deployed in. Please present and discuss the analyses used to justify the uncertainty parameterisation here. At the very least the final equations used are needed.

**467** “uncertainty estimates were parameterised’ – please provide the equations used to parameterise the error here.

**482.** “There was no strong correlation in the instrument type accuracy of either the ODINs or ES-642s with either hourly mean temperature or relative humidity”. Again Figure 10 shows larger scatter with higher RH - this seems at odds with your statement here. Please explain.

**492.** “temperature inversion forming below 250 m”. Is there a reason that the ceilometer measurements and retrieval of PBL height was not presented? You note earlier it is a critical

element, and as you only have two nights of validation, it would be worth presenting this information here.

**493.** Some comment on the miniMPL backscatter measurements needed here as they are not mentioned in the text. It appears that the concentration is increasing towards the ground, so perhaps there are higher concentrations that are not being captured by the miniMPL that you could comment on?

**499-513.** This section is an application of the data and would be better suited to the end of the manuscript after the corrections and uncertainties have been discussed.

**520.** If the coefficients are changing because emission sources are changing with the season then how does this impact the ability of the measurements to infer emissions sources? Also how do you know that a linear fit between the two periods is robust? Have you tried withholding part of the co-location timeseries from each co-location to determine the quality of the correction without including the training data in the results?

**536.** The change in ODIN correction does raise concerns about the ability of the ODINS to report absolute values of PM, particularly if the emission sources varies. This may not be an issue if the objective is the measure the relative concentrations of PM from a similar source under similar environmental conditions (which seems to be the stated purpose), but this should be explicitly discussed here.

**Figure 9.** Coefficients from only 46 ODINS are presented here – please provide more detail on why units have been excluded.

**537.** Why are uncertainties calculated from the inter-comparison not presented here?

**555.** Overfitting and increased uncertainties are mentioned in the abstract, but it is not clear what results support this statement - please explain?

**557** “The relation between bias and relative humidity is very different from that of the ODIN due to the inlet heater, built into an ES-642.” Could this be why the version 2 is worse than version 1? Worth discussing.

**566** “The intra-instrument variability was found to have little dependence on environmental factors and a constant value was used” This does not agree with 468 “[Intra-instrument variability was] parameterised in terms only of the PM2.5” - please revise.

**567** “On the other hand the instrument type accuracy was found to vary with environmental factors.” This contradicts statements at line 481 "There was no strong correlation in the instrument type accuracy of either the ODINs or ES-642s with either hourly mean temperature or relative humidity, nor was there any correlation of the uncertainty estimates with higher measured concentrations." Please revise.

In order to assess the overall data quality without having to download and process the entire dataset, it would be very useful to include tables containing: Lat/lon, elevation, sensor height, fraction of data with missing flag/out of range flag/interpolated time stamp etc, for both the PM data and the AWS data.

## Dataset comments

ODIN/ES-642 netCDF:

- The intercomparison data should be included
- The paper states 50 ODIN were deployed, but data for only 49, 47 and 46 instruments is provided for the raw, version 1 and version 2 datasets, respectively. The manuscript should state this along with the reasons that sites were excluded in the processing.
- For version 1, no coefficients are given for ODIN-SD0167
- Please state in the netCDF which geoid was used as a reference for elevations (e.g. NZGeoid2016), or are the heights given with respect to the reference ellipsoid from the GPS (i.e. GRS1980).
- Please correct the units for inlet height, which appear to be in mm.
- The use of a 16-bit integer to represent the quality control flags makes this data hard to use in many programs. Given that there are only 2 or 3 flags used, why not just store the flags separately as Boolean variables?

Meteorological data files:

- Readme files are lacking for each instrument. These should describe the naming conventions, processing steps, sources etc, including describing the meaning of 'recommended' variables.
- The AWS and radiosonde data files are fairly self-describing, but the variables in miniMPL and Ceilometer files need further description including specifying units in some cases.
- Ceilometer data is only provided till mid-July – this should be noted in the manuscript
- Why not include the proxy data in the AWS file to provide a continuous time series?
- The file is very large 8.2BG. It would be better to split this up into multiple files, perhaps by instrument type.

## Technical corrections

**1** ...infer *airborne* particulate...

**34.** ...PM *primarily* resulting from...

**39.** "is one of the most polluted cities in New Zealand". This statement needs qualifying – e.g. *experiences poorer air quality than most cities in NZ.*

**41.** "regional councils". Need to give context for role of regional councils for readers outside of New Zealand – e.g. *the regional government responsible for managing emissions of PM*

**49.** "up to" – suggest change to *in line with*

**67-69.** this is a repeat of lines 38-40. Please revise or remove

**70.** remove comma between "sites" and "that"

**71.** "area" >> *areas*

**72.** please remove "(measuring both PM10 and PM2.5)" from parentheses.

**78.** "differences with the" >> *differences to the*

83. "(Sect. 3)" >> "(refer to Sect. 3)"

163. "Nine Dust Motes and five Dust Met Motes". Please use the acronyms you have defined or don't define them.

185. "appendix" >> *Appendix*

189. "and" >> *an*

200. as for line 163 "the Weather Observation Website (WOW) maintained by the United Kingdom Met Office." >> *WOW*

213. "During night-time, radiative cooling at the surface of the atmosphere causes temperature inversions to form in the lower layers of the atmosphere. These regions of stable air 215 prevent mixing of aerosol above the boundary layer. Therefore," this detail is not needed. Consider removing.

219. remove duplicate "mini"

222. "and a ceilometer" >> *and ceilometer*

226. as for line 163 "A Sigma Space mini micro pulse lidar" >> *The miniMPL*

241. as for line 163 "A Lufft CHM 15k ceilometer" >> *The ceilometer*

248-250. A generic description of a radiosondes is not required. Perhaps move line 250 to the end of the following paragraph.

269. "This philosophy" >> "This *design* philosophy"

280. 'producing' >> *produced*

296. please add reference to Figure 2 here.

459. "question of" >> *question*

468. why is deployment in italics?

489. "The boundary layer is of specific interest as its stability influences the concentration of pollutants such as PM2.5 at the ground level." >> consider removing this theory for brevity

493. "Inversion layers such as this cause the air to have a strong static stability. This prevents vertical mixing of air, constraining pollutants to the lower layer of the atmosphere. Thus, inversions play a large role in enhanced PM levels at the ground." >> consider removing this theory for brevity

499. "measurements as two ODIN sites is compared" >> *measurements at two ODIN sites are compared*

**Figure 7.** Is the timestamp UTC or NZST? The x-axis and caption do not agree.

**Figure 8.** Are these hourly or 1-minute averages? Please add this info to caption.

**Table 1.** The naming of the ODIN and ES-642 in the caption is not given anywhere in the manuscript so these references are confusing. Either remove names or add tables listing the names and metadata for each site. Same for captions of Figure 10 and 11.

**Table 1.** What are the units?

**537-541.** These sentences repeat – please reword.

**566** “was found to found the have” >> *was found to have*

**Appendix A** – missing negative sign for lower limit of Air Temperature for AWS