Interactive comment on "How to improve your metocean datasets" *by* Erik Quaeghebeur and Michiel B. Zaaijer

Erik Quaeghebeur and Michiel B. Zaaijer

e.r.g.quaeghebeur@tudelft.nl

This response to reviews is done per reviewer and then follows the general sectioning structure of the paper, but first the general comments are treated. The format is as follows:

- The reviewer comment is shown in an upright font.

  Our response follows in blue and slanted font.

  The planned changes are written in red and slanted smaller font.

1 Rémi Gandoin (RC1)

General comments

- (Comment 1.2) Add project owner next to 'producers and users' as a stakeholder; they define the scope of work, where requirements can be specified and have an effect on the actual measurement campaign. Currently the term 'dataset producer' is used ambiguously. Relevant locations in the text: P2L20, P20L18, P22L4–5.

  It is a good suggestion to include 'project owner', as this allows us to improve the global argument and indeed would make responsibilities clearer. Also, we should indeed be more precise about what we mean with 'dataset producer' and not unnecessarily mix in 'dataset provider', as this does not add a useful distinction for the paper's goals.

  We will rewrite the second paragraph of the introduction to explicitly introduce the 'project owner', 'dataset producer' (including its role as provider), and 'dataset user'. We will include 'project owner' as a recommendation recipient (cf. P2L20). We will change all occurrences of 'dataset provider' to 'dataset producer', to consistently use the terminology introduced for the different parties as delineated in the introduction. We will add recommendations for the 'project owner' as well, making it explicit that they are the ones that can actually enforce things (cf. P20L18). In the conclusions, we will make it clear that the project owner is the one that gives the dataset producer duties that would be beneficial to the dataset users.

- (Comments 1.1, 2.1 first point) There would be value added by discussing more uses for the datasets, including a deduced ranking of importance of specific quantities (e.g., wind speed being more important than humidity). At the very least a clear justification should be given if no such ranking is provided. The ranking

could be initially given in the Introduction and worked out and used to guide the argumentation in Sect. 2.1.

We discuss the context and generality of our work in the fifth paragraph of the introduction. We state that it is relevant outside wind energy as well, but did not explicitly discuss whether we discriminate between the measured quantities. We should make it explicit that our choice is to not focus on any specific subset of quantities. We feel that not doing so would make the paper more complex by adding another consideration and deviate from the actual approach we took during the research. It would force us to discuss matters that we have not sufficiently investigated and which would need far more work than is reasonable in the context of the revision of this paper.

We will add a sentence to that fifth paragraph to make our approach explicit. (We will, however, add the technical report of Beeken et al. mentioned by the reviewer at the end of Appendix A1.2.)

- The tone of the paper is sometimes inconsiderate towards data creators. Examples: P7L13 'more intelligent handling', P20L23–24 '..., which ...the data', P21L15 'especially with ECN' may be read to imply the opposite for other parties, P21L16 'do the effort'.

While a lack of consideration was not intended, we agree that the formulations come across as such and that this is not appropriate.

P7L13: 'more intelligent handling' will be replaced by 'more elaborate handling'. P20L23–24: We will removed 'which is most likely already available in your data management systems'. We will remove P21L15 'especially with ECN' (we express our gratitude to ECN in the acknowledgments). P21L16: 'do the effort' will essentially replaced by 'invest in'.

- (Comment 4.1) Take into account FAIR data principles.

C3

See RC2.

See RC2.

- (Rating) The reviewer finds the paper to be a bit long and that it should be shortened where possible.

We agree that a shorter text is more accessible. But of course there is a trade-off between conciseness and amount of content. Our choice of content and what was placed in an appendix was deliberate. We do not see easy opportunities for shortening the length in a meaningful way. Therefore, without concrete pointers and arguments, we are not inclined to work on reducing the paper's length.

No changes will be made.

- Both 'off-shore' and 'offshore' are used; be consistent. (The reviewer suggests 'offshore'.)

We agree a single spelling should be used; 'offshore' seems preferred also by dictionaries.

We will change all occurrences of 'off-shore' to 'offshore'.

- Make sure the meaning of italic text is clear when used.

It is correct that we use italic text in different meanings: for emphasis, for foreign language names, and when introducing some concept/terminology. Of these the journal's style guidelines *only allows the first two. We should of course follow the style guide. We assume that if we do, our usage may be considered sufficiently clear.*

We will follow the style guide and will remove italics for concept introduction, but sometimes add 'called' in front of the term introduced or quotes around it.

C4

Frontmatter

- The title should make it clear the paper treats measurement datasets. It is unclear who 'your' in the title refers to.

  Making the title clearer and more specific is a good idea. (At the expense of being less catchy, perhaps.)

  We will replace "How to improve your metocean datasets" by "How to improve the state-of-the-art in metocean measurement datasets".

1.1 Introduction

- Make it clear that 'data' can refer to both measurement as model data and that this paper discusses measurement data.

  It is true that talking about 'measurement data' makes things clearer. We do not think it necessary to mention and discuss 'model data' if it is clear we the paper is about measurement data.

  P1L1: We will replace 'metocean datasets of 10-minute statistics' by 'datasets of 10-minute metocean measurement statistics'. P1L14: We will add a 'measurements' keyword. P1L16: We will replace 'data' by 'measurement data'. P20L27: We will replace 'metocean statistics datasets' by 'metocean measurement statistics datasets'. P21L20,P22L18: We will replace 'metocean datasets' by 'metocean measurement datasets'.

- P1L17–18: Tower & substructure design and installation planning also need wave data. P1L19: Instruments are also placed on fixed offshore platforms.

  We agree.

  We will integrate the suggested additions into the corresponding sentences.

- P1L24: Integrate parenthetical in preceding sentence or remove parentheses. (Similar cases pointed out: P6L18–19, P8L26–27, P13L14–15.)

  The sentences on P1L24, P6L18–19, P7L8–9, P8L26–27, P13L14–15, P13L17, P15L17–18, P16T3caption, P16L1–2, P19L10–12, and P27L16–17 are made parenthetical to de-emphasize them. This is appropriate according to the style advice we found online. Removing parentheses would remove this intentional de-emphasis. Shifting the parentheses into the preceding sentence would add a whole sentence as a parenthetical within that preceding sentence, which is stylistically strange (though appropriate for parenthetical material that is not a full sentence). We prefer to keep our current, intentional stylistic choice.

  No change will be made.

- P2L5–6: Another reviewer (RC2) feels this sentence belongs in conclusions. This reviewer explicitly mentions the surrounding P1L23–P2L9 is well-formulated.

  See RC2.

  See RC2.

- P2L14,18: Replace 'instructive' with 'instructional'.

  Indeed, 'instructional' expresses intent, whereas 'instructive' expresses an effect.

  We will replace P2L14 "To achieve the informative and instructive goals of this paper, we [. . .]" with "We [. . .]" (so we will drop the first part of this sentence). We will replace P2L18 'instructive' with 'instructional'.

- P2L17–18: Rephrase sentence (missing comma?).

  Indeed, the sentence is a bit confusing as it is.

  We will add comma after 'described'.

1.2   The datasets and their analysis

- P2L23: Is 'qua' a typo?

  No, but it is uncommon and may be incorrect as used.

  We will replace it by 'in terms of'.

1.2.1   A first look at the datasets

- Use 'FINO1' as used by BSH instead of 'FINO 1'.

  It is correct that BSH uses 'FINO1'.

  We will use 'FINO1' throughout, including in the transformation scripts, except in Figure 1, for which the effort to change this would be disproportionate (for us).

- Details about instruments (make, type) may provide added value.

  This may indeed be useful to some readers. However, we feel that providing this information in the existing tables in the paper would make them too cluttered. Adding extra tables in an appendix is an option. However, this information is already available in the metadata we include in the transformed datasets and so in the transformation scripts. That is not the ideal location to reference, but if we move that out to separate metadata files, not only this, but also other metadata we have chosen not to include in the paper is made available conveniently.

  We will separate out the metadata from the scripts into separate human-readable and machine-readable YAML files included in the script bundle. We will mention this in the 'Code and data availability section' and at the end of the introductory paragraphs of Section 2.1.

- It may provide added value to add the logger to the information provided about the measurement setups.

  This may indeed be useful to some readers. The information available varies greatly between the three datasets. We feel that going into this in the paper would introduce too much detail that is not of interest to many readers. It is easy and possible to provide some information in metadata files in the script bundle.

  We will add files with some information about the loggers (make, type, number, reference) to the script bundle. We will mention this at the end of the introductory paragraphs of Section 2.1.

- Make explicit what the provenance is of the uncertainty and range values given in the instrument & quantity tables.

  That can indeed be of interest to some readers.

  We will add this information (for all datasets) to the metadata YAML files as comments.

- P4T1: Specify orientation in degN instead of 'NE'. Kouwenhoven (2007) states a sampling frequency of 39 Hz for the ultrasonic anemometer instead of the listed 4 Hz. Perhaps make it clear that the thermometer and hygrometer are integrated into one instrument. Perhaps 'thermometer' is more correctly called 'temperature sensor'?

  We agree that exact angles should be given.
  The sampling rate of the ultrasonic anemometer is 39 Hz, but this raw data is not sent to the logger; the output rate is 1 or 4 Hz (as can be seen in the spec sheet in Kouwenhoven's report) and I have seen the raw data files (a colleague had obtained access) and there were 2400 samples per 10 minutes, so 4 Hz.
  The thermometer and hygrometer are indeed two sensors integrated into one package.
  We feel that 'thermometer' is the correct term to use (in this case it is a resistance thermometer, but such detail is left for the metadata files); 'temperature sensor' may be interpreted to mean just the sensor part of the instrument and exclude the part that converts the quantity sensed to a numerical value.

We will give exact angles in the footnote 'ao' describing the orientation.

We will add a footnote 'i' indicating that the thermometer and hygrometer are two sensors integrated into one package to this table and also the others, where the same remark applies.

We will keep 'thermometer'.

- Information from the original data file headers may provide value in the data file descriptions.

  This information is available in the instrument overview tables and in the metadata in the transformation scripts.

  We will make no changes because of this comment. However, our separation of the metadata into separate YAML files will make all the information not in the tables more accessible, sufficiently so, we feel.

- Use 'specification sheets' instead of 'spec sheets'. Use 'met. mast' instead of 'met-mast'.

  We have looked to actual usage (on-line) and think 'specification sheet' and 'met mast' are the most common ways of writing (although 'spec sheet' is also common).

  We will change 'spec sheets' to 'specification sheets' and 'met-mast' to 'met mast' throughout the paper.

- P5L23: FINO1 generated data after 2016.

  Indeed. This may be of interest to the readers.

  We will add a phrase "measurements are still ongoing" to the corresponding appendix (A.1.2).

1.2.2 Dataset issues

- (Comment 4.1 b) Discuss the possible reasons for data quality issues in the datasets. (This is also relevant for Sect. 4.)

  See RC2.

  See RC2.

- P7L1: What is the 'normal range' for the series?

  Our language use was sloppy here. We meant 'instrument's range'.

  We will replace 'normal range' by 'instrument's range'.

- P7L6–7: Perhaps give numerical examples to illustrate the standard deviation bound.

  Given that this is a purely mathematical, rough bound that serves as a sanity check, specific examples are not really of interest. We had a look at the empirical distribution of $2s_x/|\hat{x} - \breve{x}|$ for a variable (MMIJ 'TrueWs' at 92 m), but also that did not show anything interesting, i.e., almost all samples lie far below the bound (mean is 0.35, standard deviation is 0.05). We really think a numerical example for this will not provide added value.

  No changes will be made.

- P8L1–5: 'Max' is strange for this categorical variable; is the same issue present for 'Value'?

  The MMIJ dataset has all four statistics calculated from the raw values even for categorical variables. In Section 2.2.3 "Statistic Selection" we already comment on this. Therefore the 'avg/Value' column really contains averaged values and is not really useful (but that was not the focus of this part of the paper). The 'Max'

column should only contain real samples and therefore it should not contain non-existent codes, even if the concept of maximum is not really applicable (strange). So it is still a useful column in the context of this part of the paper, i.e., checks for faulty (categorical) data.

No change will be made.

- P11L2: Move footnote superscript before comma.

  The footnote refers to all material delimited by the comma the footnotemark is attached to and the preceding comma, not just to the parenthetical or a specific word. Therefore, placing the footnotemark there is appropriate according to style guides.

  No changes will be made.

- P11L7–12: Another, uncommon issue is drift of the logger clock.

  Yes, we have heard colleagues discuss this in the context of SCADA data. We have no indication this was an issue for these datasets.

  No changes will be made.

- P11L13–17: Descriptions and drawings do not always reflect actual placement; pictures or videos of the mounted instruments are useful in this regard.

  We can agree with that.

  We will add '(Pictures or video footage would of course further increase confidence in the accuracy of the drawings.)' after the first sentence of this paragraph.

- P11L19: Isn't the precipitation detector mentioned on page 56 of ECN-Wind-Memo-12-010?

  No, that is the precipitation monitor, a different instrument, which is well-documented.

No changes will be made.

- P12L4: Refer to Sect. 2.2.5 to make it clear why the sampling frequency is important.

  That would be helpful indeed.

  We will add a reference to Sect. 2.2.5.

- P12L29: Make explicit relative to what in 'relatively little effort'.

  Making that explicit would indeed be helpful.

  We will use 'little effort relative to the whole of the measurement campaign'.

- (Comment 2.1, second point) Sensor and quantity-specific treatment is missing (cf. comment about ranking of quantities). For the discussion of uncertainty, references to and comparisons with existing work (including industry standards) are lacking.

  As stated in the discussion of the 'General comments', we choose not to do a quantity-specific treatment. We also choose not to do an instrument-specific treatment. However, it would be useful to explicitly inform the reader that our treatment is generic and about the existence of such specific treatments. The reviewer suggests focusing on cup anemometers. He also provides texts about cup anemometers that can be used in the comparison.

  - *Of these, Kristensen's paper discusses and quantifies biases in the 10-minute averaged wind speed and suggests an approach to remove (much of) that bias using wind direction measurements; if we understand the results correctly, these biases are supplementary to the uncertainties we derive.*

  - *Pindado et al.'s review presents dynamical models and empirical data, but as far as I can see, no explicit expressions for uncertainty.*

    – *The standard IEC61400-12 discusses how per-wind speed bin absolute uncertainties should be calculated (Appendix F.8); these can form the basis for the absolute and relative uncertainties used in our procedure. It also discusses the uncertainty of wind speed (Appendix E.5.3) as a combination of component uncertainties, not all of which may be of interest to be included in the statistics datasets. I find the discussion in the standard to not be very clear about the impact the time interval the wind speed is averaged over; this makes positioning its procedure relative to ours difficult.*

The first and last texts provide sufficient material to create a paragraph to inform the readers.

We will add an extra paragraph at the end of Sect. 2.2.5: "Before closing this Section, it is important to stress that the expressions for propagated uncertainties and biases above are generic. Namely, their derivation does not depend on the specific quantity considered or instrument used. Detailed knowledge of the measuring instrument's properties may allow for better uncertainty estimates or additional uncertainty and bias terms. For example, for cup anemometers, it is known that there is a positive bias of 0.5%–8% in the mean wind speed, but that this bias can be greatly reduced using wind direction variance estimates (Kristensen, 1999). Also, the IEC 61400-12-1 standard prescribes how the wind speed uncertainty should be calculated for calibrated cup anemometers (IEC, 2017, App. F), which may lead to high-quality estimates for ε_{a} and ε_{r}."

- P14L31: $\varepsilon_{\bar{x}}^2$ instead of $\varepsilon_{\bar{x}^2}$.

It should indeed be $\varepsilon_{\bar{x}}^2$.

We will correct this.

- P15L17–18: A numerical example would add value for understanding the magnitude of the bias's effect. P16T3: More generally, numerical examples can clarify how the derivation of uncertainties is done. Perhaps such examples can or can also be provided as Python code.

<center>C13</center>

Making the bias's effect more concrete would indeed add value. The derivation of the uncertainties is done as per the equations in this section, of course, and the code for generating the table should be made available to make it clear that effectively just that is done. It is not clear to us whether the reviewer would prefer more steps to be put in the paper's text (we do not think this would have sufficient added value).

The code for generating the table's values will be included in the script bundle. To allow for easier interpretation, we will add an extra column to the table, for the relative value of the bias-corrected standard deviation. We will furthermore make our comment about the impact of the bias on turbulence intensity more concrete and add a comment about bias and uncertainty for ambient temperature.

- Choose more formal or precise alternative word for 'bunch' (P10L1) and 'quite a lot' (P12L27). 'Timestamp' instead of 'Time stamp'.

We disagreed amongst ourselves about 'bunch', so we will follow the reviewer's preference to replace it. Given its context, we think 'quite a lot' is fine here, i.e., the discussion above makes it explicit what we mean. Nevertheless, we think 'a good amount' would be a better formulation (not that it is more formal). Both 'Timestamp' and 'Time stamp' are in use, but we have no objections to your preference.

We will replace 'a bunch' by 'several' and 'a cluster', respectively. We will change 'quite a lot' to 'a good amount'. We will change 'Time stamp' to 'Timestamp'.

1.3   Dataset formatting

1.3.1   A comparison of data file formats

- P17L7: Clarify what is called useless and why.

<center>C14</center>

We meant to say that using a text editor for analysis is useless. This comment in the text is not essential and apparently not clear.

We will remove 'and useless for analysis'.

- P17L27: Rephrase to avoid quotes around 'knows'.

We agree that this formulation is not that good.

We will change 'knows' to 'has access to'.

1.4   Recommendations

- (Comment 4.1 a) Give reasons for dataset creators to follow the recommendations. (Comments 4.1 a, c) Sketch opportunities, barriers to change, and means to resolve them.

The main reason why dataset producers should follow the recommendations is because it would improve the usefulness for users of the datasets they deliver. This is already clear in the paper. But of course, even if we think this would cost relatively little effort, this costs time and therefore money. By introducing the project owner as a stakeholder (see discussion earlier), we can make it clear how following the recommendations can be fit into the agreed-on duties of the dataset producer. Also, the value of improved datasets to project owners as input to future measurement can then be mentioned. We do not think a wider discussion of opportunities and barriers to change falls within the scope of our paper. We have the perspective of the dataset user and gaining the necessary insight for such a discussion would for us be a project unto itself; we certainly do not wish to speculate on this.

We will add the following sentences to the last paragraph of the Conclusions: "This effort can be seen by the project owner as necessary for getting the most value out of the

raw data collected. Such a well-documented dataset with uncertainty and quality information included creates the possibility for consciously making possibly different choices (trade-offs) when setting up future measurement campaigns.". No further changes will be made apart for those related to other comments involving the stakeholders (see "General comments") and their shared responsibility (see below).

- (Comment 4.1) The reviewer states that solving the issues discussed in the paper is a shared responsibility and that more normative requirements are not realistic.

We agree that this is a shared responsibility. The earlier suggested introduction of the project owner makes it possible to sketch the responsibilities of the stakeholders in the paper. We do not argue explicitly for more normative requirements, but would recommend project owners make certain concrete requirements for dataset producers. Our assessment is that the benefits of these outweigh their costs.

In the recommendations, we will add some parentheticals specifically aimed at making the shared responsibility and each stakeholder's role clearer.

- Keep the recommendations impersonal; avoid 'you'.

We have no clear preference here, so we will follow yours.

We will reformulate the recommendations to remove 'you(r)'.

- P20L26: Why the parentheses around 'also'?

To express that providing the dataset binary format can be done next to CSV file (or some such), but that we do not think providing the latter is necessary. But actually, this shouldn't be our concern and focus. Getting binary format files is, so we should just leave out the 'also'.

We will drop '(also)'.

- P21L4: Rounding to the expressed uncertainty would lead to a loss of information in case the uncertainty is revised downward.

  We think the metadata and data should be consistent. So if the information (metadata) used to determine the uncertainty of values is revised, then a revised dataset should be published, based on the reprocessed raw or intermediate data. If conservative estimates (lower bounds) of the uncertainties are used (e.g., as proposed in this paper), the revised datasets should in general not include uncertainty reduction. Because the difference between possible precision and actual precision is large in general, binary-rounding also leads to substantial space savings (after compression; non-significant digits are essentially random and do not compress well). Given all these reasons (and some others that would lead too far), we stick to our current recommendation of binary-rounding values.

  We will add a recommendation "Use clear version identification in dataset files, to avoid confusion when updated or extended datasets are released.".

- P21L6: Provide original sample standard deviation next to the bias-corrected version.

  In principle, we think this should not be done, for the same reasons as mentioned in the reply just above. However, we understand this feels like a more invasive change than rounding to uncertainty, even if there is a real error that is corrected by this procedure. Because of that, we mentioned the alternative option of not correcting but just including the bias values.

  No changes will be made.

1.5 Appendices

- P25T2: Wasn't there a statistic labeled 'variance' in the FINO1 datasets?

  No, not in the version we downloaded. But BSH may have changed the files they make available; I think that the current version may even be different from the one I downloaded and analyzed.

  No changes will be made.

Backmatter

- In the list of references there are stale URLs and missing version and techreport numbers.

  See RC3.

  See RC3.

2 Nikola Vasiljevic (RC2)

General comments

- The tone of the paper is sometimes inconsiderate towards data creators.

  See RC1.

  See RC1.

- Take into account FAIR data principles. Try to score the FAIRness of the datasets before and after implementation of the recommendations of the paper.

  Given the current interest in FAIRness, this may indeed be of interest to many readers. We can perform a FAIRness analysis, although more qualitative rather than quantitative in nature. (So without really scoring each dataset.)

We will add a whole new section that looks at the current status, then moves to what role the recommendations of this paper play in changing that status, to finally evaluate the role of the non-user stakeholders. We will refer to that Appendix at the end of the second introductory paragraph of Sect. 2.1.

- Review the FAIRness analysis of DTU Wind Energy.

  This is an interesting overview of FAIRness activities at DTU. However, we feel that our paper is not the appropriate place to provide a review of this material.

  No changes will be made.

- Add EERA-JP wind energy metadata to the created binary datasets. Review the EERA-JP wind energy taxonomies, metadata, and vocabularies in the paper.

  We looked at the datasets on https://data.dtu.dk/DTU_Wind_Energy *to find the proper way to add EERA-JP metadata to netCDF files (in a single attribute? in multiple attributes?), but found no example. We feel that our paper is not the appropriate place to provide a review of this material, just as we do not review, e.g., the CF Conventions.*

  We decided that we will add the metadata in a '`EERA_JPWind`' attribute, with one descriptor per line:

  ```
  Activities:Measurements:Field Experiment
  External Conditions:Location:Offshore:Offshore
  External Conditions:Water Depth Category:Shallow Water
  Data Categories:Meteorological
  ```

2.1 Introduction

- P1L24: Integrate parenthetical in preceding sentence or remove parentheses.

  See RC1.

  See RC1.

- P2L5–6: This reviewer feels this sentence belongs in conclusions. Another (RC1) explicitly mentions the surrounding P1L23–P2L9 is well-formulated.

  The content of the sentence is also present in the conclusions. We think that this 'preview' is useful for readers and feel supported concerning this by one reviewer.

  No change will be made.

2.2 The datasets and their analysis

2.2.1 A first look at the datasets

- Download URLs should be provided for the datasets analyzed.

  See RC3.

  See RC3.

- P3L1–2: Sentence unclear.

  It is not clear to us in what way the sentence is unclear.

  No change will be made.

2.2.2 Dataset issues

- Discuss the possible reasons for data quality issues in the datasets. (This is also relevant for Sect. 4.)

  We write from the perspective of the dataset user, as indicated in the third paragraph of the introduction. Of course we have some ideas about some of the

possible causes for data quality issues, but do not have the insight necessary to usefully discuss this. This would be very interesting, but should probably be done from the perspective of the dataset producer, so by a dataset producer. What we do is, however, provide tools (including code) that can help the dataset producer identify, explain, and eliminate some classes of quality issues. It is also important to make it clear that we realize that it is inevitable for faulty data to be present in the raw measurements. Our aim is to improve the processing of that data into datasets such as the ones studied.

We will modify the first paragraph of Sect. 2.2.1 to make the last point made above explicit: "It is normal that the measured signals (raw data) contain faulty data. [. . .] The dataset producers deal with such faulty data, e.g., by removing it, when creating the datasets of statistics series we study. Nevertheless, each of the three datasets presented above contained remaining faulty data."

- P5L30: Faulty data and quality flags are interrelated.

  It is true that quality flags can be used to indicate possibly faulty data. This is a good idea to include.

  We will add a paragraph at the end of Sect. 2.2.4: "Of course other information next to missingness mechanisms can be included in the quality flag bit field, also for non-missing values, as is done for FINO1. For example, this can be used to indicate possibly faulty data (cf. Sect. 2.2.1) that has not been removed (made missing)."

- P6L5: Clarify which datasets were converted to which formats.

  That can indeed be of interest to the readers and it can be convenient that they do not need to look this up in the referenced script bundle.

  We will change '(HDF5 or netCDF4)' to '(HDF5 format for OWEZ and netCDF4 format for MMIJ and FINO1)'.

- P6L7: Be more concrete regarding the automatic data issues detection.

C21

We elaborate on this in the list below this sentence. For even more concrete information, the scripts themselves are available. We think this is enough, but perhaps the reviewer had something else in mind.

No change will be made.

- P7L1: What is the 'normal range' for the series?

  Our language use was sloppy here. We meant 'instrument's range'.

  We will replace 'normal range' by 'instrument's range'.

- Axis labels are missing in Figs. 3–6. These figures are hard to understand without extra information in the captions.

  The axis labels were omitted consciously to not overburden the plots. The information necessary for understanding the figures (including axis meanings) used to be closer to the plots themselves, but we were requested to move this in-text to better adhere to the WES style. However, we understand that such information may be useful.

  We will not add any axis labels. We will add the parenthetical "(cf. pages x–y for an explanation)" to each caption to point the reader to the explanation. (There is too much explanation to put a summary in each caption.)

- P11L26: The term 'accuracy' is used instead of 'uncertainty'.

  We consciously chose to use 'accuracy' here, trying to follow the usage described by the JCGM and mentioned in footnote 1. As we understand JCGM's definitions, 'accuracy' is the qualitative counterpart to the quantitative 'uncertainty' and so uncertainty provides accuracy information. We do not know whether the reviewer means to say that he disagrees with our interpretation or is following another definition. We assume the latter for now.

  We will add a footnotemark to 'Accuracy' that refers to footnote 1, to clarify our usage of the term.

C22

2.3 Dataset formatting

2.3.1 A comparison of data file formats

- P17L29–31: Refer to EERA-JP wind energy taxonomies etc. here.

  If we include terms from the taxonomy in our datasets, we should indeed also cite reference material.

  We will cite

    – *https://github.com/wind-energy/taxonomies-and-vocabularies*
    – *https://doi.org/10.5281/zenodo.1199488*

2.3.2 Practical experiences with binary formats

- P19L27: How did you help fix the buggy Python netCDF4 code?

  We filed a bug report and actively assisted in getting it fixed. Actually, we did the same with another issue. However, I now think this paper is not the place to try and get credit for that.

  We will remove the remark about helping to fix the buggy code.

2.4 Recommendations

- Give reasons for dataset creators to follow the recommendations.

  See RC1.

  See RC1.

<div align="center">C23</div>

3 Hans Verhoef (RC3)

General comments

- The reviewer feels that the paper lacks a research question.

  We can understand that judgment by the reviewer. The paper is not a classical research paper and that is reflected in the formulation of its goals. These goals are given in the Introduction (P2L4–6) as a pair of questions and brief answers that; we feel these are appropriate as a formulation of the research objectives.

  No changes will be made.

- Avoid uncommon words such as 'non-onerous', 'relegate', 'gleaned', 'tuple'.

  Judging what and what words would interfere with readers' reading is difficult. We will not make that judgment, but will change all but one of the words mentioned. (We prefer 'tuple' over alternatives such as 'set' or 'vector' for precisely expressing what we want.)

  We will change 'non-onerous' to 'non-burdensome', 'relegate to' to 'put in', and 'gleaned' to 'learned'.

- ECN is mentioned in various places in the paper; this organization is in an unfinished state of name change.

  Thank you for reminding us; it is useful for the readers to mention this.

  We will add a footnote mentioning that ECN is now part of TNO and that its name will change.

<div align="center">C24</div>

3.2   The datasets and their analysis

3.2.1   A first look at the datasets

- Download URLs should be provided for the datasets analyzed.

  Download URLs were included in the references, but this was not clear from the citations, as no year was included for these references.

  We will add the (URL-visiting) year to to make download references stand out when citing.

- A comparison of the metocean climates at the datasets' sites would have added value.

  Such a comparison would indeed have value, but falls outside of the scope of this paper.

  No changes will be made.

3.2.2   Dataset issues

- The colors in Fig. 2 must be explained in the caption or the legend of that figure.

  This information used to be closer to the plots themselves, but we were requested to move this in-text to better adhere to the WES style. However, we understand that such information may be useful.

  We will add a summary of the in-text explanation: "(Mean in black; mean $\pm$ standard deviation in blue; minimum and maximum in red.)".

- Axis labels are missing in Figs. 3–6.

  See RC2.

See RC2.

- 'North' is mentioned without clarifying whether it is geomagnetic or geographic.

  North is only mentioned in the context of the MMIJ dataset. The reason is that the boom designation is offset from typical direction angles. In the documentation it is not mentioned whether this is geographic or magnetic North (or even a grid North), but we assume it is not magnetic, because that would be atypical.

  We will preface '(geographic)' in front of the two occurrences of 'North'.

3.4   Recommendations

- The reviewer contrasts the current acceptance of binary formats in academia with a preference for text-based formats in the commercial sector.

  The last recommendation for users is relevant in this context: our experience shows that binary formats are much more efficient to work with and we have become convinced that this would be the case for almost any party, be it commercial or academic. We do think this is actually the most relevant recommendation in the list and needs to be made more prominent and forceful.

  We will move that last recommendation to the front and reformulate it to: "Invest in learning to work with format like HDF5 or netCDF4, as this will allow working more efficiently with datasets (cf. Sect. 3).".

- The reviewer indicates that the recommendations to users are more reminders and that the main benefit for them would be the generalized use of a standardized time and effort-saving format.

  We think the argument that binary formats can reduce time and effort spent by users is sufficiently made in the paper. The recommendations for users are indeed not as strong as those for producers. However, even the obvious recommendation about not trusting the data blindly must, we feel, be kept: in our own

project there were mathematician/computer science researchers that use such datasets in a purely instrumental fashion, without an inclination to perform checks first.

The change that will be made due to the point above this one already makes the recommendations a bit stronger. We will make a minor further improvement by also switching the order of the other two recommendations and by making it clear that we realize not trusting the data blindly would be obvious for many readers.

3.5 Appendices

- A reference is needed in support of the statistics-heavy material.

  We understand that this part is not as accessible, but it is a bit hard for us to judge which statements require referencing.

  We will add three more specific citations to the standard text by Cramér, to support statements that may not be as well-known as we assumed. (The first author has a background in probability theory...)

- The first three equations in line 28 on page 26 have the same right-hand sides.

  Indeed, and that is correct.

  No changes will be applied.

Backmatter

- In the list of references there are stale URLs and missing version and techreport numbers.

  The reviewers are correct about the stale URLs and missing numbers.

  The changes to be made to the reference entries are therefore:

  - *fix stale URLs,*
  - *remove unnecessary stale URLs,*
  - *update all software entries to the currently used version,*
  - *move all version numbers to title field,*
  - *make sure BibTeX entry types are chosen such as to expose the number field once compiled.*