Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2017-16-AC2, 2017

Interactive comment on "Uncertainty information in climate data records from Earth observation" *by* Christopher J. Merchant et al.

Christopher J. Merchant et al.

c.j.merchant@reading.ac.uk

Received and published: 19 June 2017

Responses to reviews and comments

Responses to anonymous referee #1

The authors thank the referee for constructive comments on the manuscript. Each point is addressed in turn below.

1. In this review paper, the authors discuss some issues and necessities related to the communication of uncertainty information in climate data records. This is an important topic, which has too often been neglected in the provision of Earth observation data. I hence very much welcome this review, and recommend publication after some revision. Regarding that revision, I have an overarching recommendation, which I am not sure

the authors can fulfill. After reading the paper I felt that it could possibly have an even larger impact if it were to provide more concrete guidance on how to report uncertainties. I am, however, uncertain if such concrete guidance is possible in the scope of this review paper, or whether this would be too large a task. I will leave this for the authors to consider and eventually decide - the paper certainly carries enough significance even if it largely remains in its present form.

We appreciate these comments, and the encouragement to be more definite to readers about how to meet our recommendations 1 to 3 (quantify uncertainty within the dataset, use well defined measures of uncertainty, and discriminate data that are more and less certain per datum.

A fully comprehensive "Earth observation metrology" is not yet established, and so we cannot give comprehensive concrete guidance for every case – in particular for class variables or vector-outline variables.

However, we can strengthen the recommendations in section 8 and reword the recommendations to be clearer at least to give a concrete minimum expectation for future CDRs containing numerical variables, as follows: Pg 14 line 4: Insert a conclusion from the discussion in this section, giving clearer guidance on presenting uncertainty information in climate datasets:

"Producers of CDRs therefore have to reflect on the expected applications of their data, and make a judgement about the balance to strike between conflicting requirements, such as ease-of-use versus completeness of the uncertainty information. Nonetheless, the collective experience across the CCI ECVs represented in Table 1 shows that provision of per-datum standard uncertainty emerged as a rigorous-but-simple approach adopted within most projects (other than products comprising classifications). The standard uncertainty provided is generally the total from all sources of error, although uncertainty components with different error correlation structure are additionally provided in one case. Although not sufficient for every possible application, quantifying the total standard uncertainty per datum in a CDR product therefore emerges as a baseline standard for future good practice."

Pg 15 line 10: Reword recommendations 1. to 3. more concretely:

"1. Include quantitative uncertainty information within the dataset. (Don't expect users to find uncertainty information from reading related papers.) 2. Follow metrological practice for quantifying uncertainty. Baseline good practice is to provide the total standard uncertainty for numerical variables. 3. Uncertainty estimates should be provided per datum in CDRs where uncertainty varies significantly, so that the uncertainty information discriminates which data are more and less certain."

Table 1: where mention uncertainties are specifically standard uncertainties, this has been made explicit.

2. Given the very wide coverage of topics suggested by the long list of authors, it seems possible to suggest a rather concrete best practice on how uncertainty information should be communicated in concrete products. I expect that within the ESA Climate change initiative, quite some discussion must have revolved around this question. For example, many of the records included here will have issues related to re-gridding, sensor drift, physical limitations, etc. Can standard procedures be defined on how these should be reported? Also in chapter 5, for each sub-topic a possible recommendation would be helpful, going beyond the mere description of possible issues. For example, in 5.2, it would be helpful to learn how non-gaussian uncertainties should be reported, and how gaussian uncertainties should be reported. Just by providing standard deviation? As I only see this overarching issue as a recommendation, I ranked this as a minor revision.

The current manuscript reflects the range of topics on which discussions between teams within the CCI project converged in agreement. Essentially, reporting the total standard uncertainty estimate per datum emerged to be an strong baseline for numerical data (non-categorical).

Adding recommendations for each subsection of section 5 turns out to be rather artificial when we tried it. It is agreed that more concrete direction in section 5.2 is useful, and the last few sentences are amended to:

". . . some statistical uncertainty. Cases such as this require numerical representation of error distributions and Monte-Carlo style simulation for the propagation of uncertainty. Where quantization is negligible, which is often the case for contemporary sensors, the Gaussian distribution may realistically describe the signal noise, and should be characterised by the standard deviation of the error distribution, which is the standard uncertainty."

3. Additional issues: - The abstract can and should be improved. It currently remains unclear how the ab- stract relates to the actual paper. The crucial information that this is a review paper is only given towards the end of the introduction, and only then it became clear that what was said in the abstracts refers to this paper rather then to some general previous knowledge. - p.1, l.30: I recommend indicating that these are just examples - p.2, l.21-24: perfect for abstract

Agreed: proposed amendment:

The question of how to derive and present uncertainty information in CDRs has received sustained attention within the European Space Agency Climate Change Initiative (CCI), a programme to generate CDRs addressing a range of essential climate variables (ECVs). Here, we review the nature, mathematics, practicalities and communication of uncertainty information in CDRs from Earth observations. This review paper argues that climate data records (CDRs) derived from Earth observation (EO) should include rigorous uncertainty information, to support application of the data in contexts such as policy, climate modelling and numerical weather prediction reanalysis. Uncertainty, error and quality are distinct concepts, and the case is made that CDR products should follow international metrological norms for presenting quantified uncertainty. As a baseline for good practice, total standard uncertainty should be quantified per datum in a CDR, meaning that uncertainty estimates clearly discriminate more and less certain data. In this case, flags for data quality should not duplicate uncertainty information, but instead describe complementary information (such as the confidence held in the uncertainty estimate provided, or indicators of conditions violating retrieval assumptions). The paper discusses the many sources of error in CDRs, noting that different errors may be correlated across a wide range of time and space scales. Error effects that contribute negligibly to the total uncertainty in a single satellite measurement can be the dominant sources of uncertainty in a CDR on the large space-scales and long time-scales that are highly relevant for some climate applications. For this reason, identifying and characterizing the relevant sources of uncertainty for CDRs is particularly challenging. Characterisation of uncertainty caused by a given error effect involves assessing the magnitude of the effect, the shape of the error distribution, and the propagation of the uncertainty to the geophysical variable in the CDR accounting for its error correlation properties. Uncertainty estimates can and should be validated as part of CDR validation, where possible. These principles are quite general, but the form of uncertainty information appropriate to different essential climate variables (ECVs) is highly variable, as confirmed by a quick review of the different approaches to uncertainty taken across different ECVs in the CCI. User requirements for uncertainty information can conflict with each other, and again a variety of solutions and compromises are possible. The concept of an ensemble CDR as a simple means of communicating rigorous uncertainty information to users is discussed. Our review concludes by providing eight concrete recommendations for good practice in providing and communicating uncertainty in EO-based climate data records.

- p.3, l.3: The term 'climatic' implies a time frame of around 30 years, but artifacts can arise on shorter time scales. I suggest to simply drop "climatic" from this sentence

Agreed.

- p.3, l.8: Bellprat, 2016 is missing from the reference list. I did not check all references but given that apparently the reference list has not been included automatically, the

authors should carefully check its completeness.

Done.

- p.3, l.16: Might be helpful to also be explicit about "The first example" and "The second example", as I wasn't quite sure what the "third" referred to.

Modified.

-p.3, l.16: maybe add "and seasonal or decadal forecasts" after "data assimilation"

Comment is correct, but the paragraph focuses on re-analysis.

- Section 3: Maybe this section should be renamed to "Terminology" to clarify what it actually is about?

Modified.

- Section 4: I also here found the heading a bit too generic, as also the previous section deals with "lessons from metrology" when defining terminology. Would "Traceability of uncertainty" work?

Agreed.

- Section 5: I did not fully follow the logical flow of 5.1 to 5.4 in reference to the scenario described in the first paragraph. It might be helpful to describe in some order the actual steps one has to take to come up with an uncertainty estimate for the concrete scenario. (i.e., 1. Explain how one would estimate uncertainties at each step including issues from quantisation. 2. Explain how one can then propagate uncertainties 3. Explain how things then become difficult when moving to spatial fields). So most of the information is already there but some more guidance and re-structuring might be helpful for the reader. Simply moving the first sentence of section 6 to before section 5.1 could be a very simple step forward.

Agreed – the moving of the summary sentence does improve clarity; back references to the scenario have been added throughout the sections to make the links explicity; and an additional paragraph to the “magnitude of uncertainty” subsection makes the discussion more concrete:

“With reference to the scenario described in subsection 5.1, several sources of uncertainty can be identified whose magnitude must be estimated. For example, at step 4, the combined effect of solid-state detector noise, amplifier noise and digitisation causes an uncertainty in counts. This uncertainty can be estimated, for example, by considering the dispersion of measured values when viewing a constant calibration reference. Another example is the retrieval uncertainty associated with the inverse solution that provides the geophysical retrieval from the satellite radiances (step 6.c). Even with perfect data, the process of retrieval is usually ambiguous (more than one geophysical state can be associated with identical radiances). This component of uncertainty can be quantified by simulation of retrieval outcomes compared to the simulation ‘truth’, if a forward model for the satellite observations is available.”

- Section 6: It would be helpful to better understand why the particular choices summarized in table 2 have been made. While it’s interesting to see the range, it would be helpful to have some understanding of its underlying cause. Could the range have been made narrower if there had been more communication/funding/time?

In fact, the range of types of uncertainty reported in CCI is rather narrow while the means of estimation is varied. Some details in Table 1 have been refined to emphasise that. Nonetheless there is a variety of alternatives to standard uncertainty that are metrologically “respectable”. To clarify this, some text can be moved around. The summary outcome that most CCI projects supply a standard uncertainty estimate is better placed at the end of section 6 after the discussion of the alternatives, and now reads:

“Most projects in the CCI programme adopted total standard uncertainty as the provided uncertainty information (Table 1), a convergence that arose after sustained dis-

cussion across the programme and which is in line with metrological guidance. Exceptions include ECVs where the geophysical data are categorical rather than numerical as discussed above. However, there is a wide range of methods employed to develop this uncertainty information, documented by the varied contents of the ‘Uncertainty Characterisation Reports’ prepared for each CDR. (For these reports and other documentation, refer to www.esa-cci.org.)”

- Section 7: The structure of this section was not fully clear to me. It introduces the issue of validation, but from line 25 onwards there is an unclear logical flow. After introducing an equation, an example is given, then triple collocation techniques are introduced, and finally instrument noise is adressed. Could this be structured more clearly? This would certainly help the reader in understanding how best to validate which kind of uncertainty.

Additional “signposting” of the progression will make the flow clearer.

- Reference list: Kobayashi, 2015 should be before Mahlstein, 2012. The list is not always consistent, for example some journal names are not in italics or are not abbreviated. Please double check, or switch to an automated system to create reference lists :-)

Double-checked and made consistent.

In any case, I am grateful to the authors for compiling this review, and hope that it will help to raise awareness of these important issues.

We agree that these issues do need to be aired to raise awareness. Thank you for the comments that have improved the manuscript significantly, and time and attention given to our paper.

Responses to Reviewer 2

1. Careful, thorough and thoughtful paper. Not sure the fit to this particular journal but evidently editors have allowed it so far and - not clear to this reader - where else could

it go?

We enquired with the editor prior to submission and got the go-ahead to submit.

2. Positives: The linguistic and mathematical / statistical connections to metrology, very important reminders for our community. The positive and detailed description of a SSM maturity matrix (which so far, for the NOAA / Bates version, seems to mostly have circulated as a Powerpoint presentation). The very appropriate note of caution about the SSM matrix, and the overall sense of a rigorous application in an ad hoc manner according to data stream and user need.

We are glad these aspects come across in the paper.

Suggestions: 3. The authors started very clearly on a general topic of observations, climate-relevant observations, and then formal CDRs. They mentioned the very strong influence and impact of reanalyses. But by the middle and certainly by the end the manuscript they had pretty much focused entirely - and not inappropriately nor unexpectedly given this set of authors and the motivations and support of ESA's CCI - on satellite data. Perhaps a reminder in the conclusions of how the 8 recommendations, and the tactics of per- datum or ensemble approaches, fit or do not fit other data sets, particularly in situ data sets.

We should make clear that the focus is on satellite datasets in the abstract so as not to raise false expectations – see above rewording. We were not convinced about the degree to which generic comments about in situ observations could be usefully made by us.

4. The authors mentioned the error generation possibilities inherent in the upscaling pro- cess but they missed, at least for this reader, mention of the interpolation processes by which irregular - in space and time - observations get converted to grid spacing. This seems a very prominent process in our community - strongly evident in other ESSD papers for example (c.f. the original and gridded versions of SOCAT both

in ESSD) - and one that perhaps needs explicit discussion in a document like this?

This is a fair point, and is relevant to the impact of error correlation. Therefore, a further paragraph can be added to the end of section 5 as follows:

"Local correlations and correlation of errors from systematic effects need to be properly accounted for when creating 'level 3' versions of CDRs, i.e., gridded products involving averaging of full resolution data. If the correlated nature of errors is neglected, the uncertainty estimate for the gridded data will be poor (usually an underestimate). In averaging data subject only to independent random errors, it is well known that the effect of the errors on the average decreases with the square root of the number of contributing data, but local correlation decreases the averaging-out of errors. In the extreme of pixel uncertainty dominated by an error source that is fully in common across a grid cell, there is no reduction in uncertainty from averaging many pixels. These impacts of error correlation in an average can be evaluated using eq. 1 with the required off-diagonal terms in Ux. Where a grid cell is not completely sampled by the full resolution data, there is an additional uncertainty not quantified by eq. 1 associated with the unobserved part of the cell; Reuter et al. (2010) or Bulgin et al. (2016a) provide an example of a parameterisation of sub-sampling uncertainty."

5. This reader notes the almost complete absence of any discussion of precipitation. Not a current CDR, as I understand (but soil moisture is?) but perhaps the most challeng- ing and troublesome of all our climate-relevant data. The very useful examples in this manuscript relate mostly to radiances: aerosols and sea surface temperatures. We read almost nothing about about the combined morass of satellite, radar, rain gauge, stream gauge, etc. data streams, often with wildly uncertain uncertainties, that con- stitute the basis for any of several global precipitation products. Precip represents perhaps the extreme uncertainty challenge? We should read a mention of it, even if only a subject for future work or as a reminder of the many real-world difficulties?

We agree that precip is a good example of the challenges involved in this business.

But it seems a little artificial to introduce a discussion of a variable that has not been addressed by any of this particular set of authors or used as an example of any of the issues discussed in the text. We would thus prefer to not introduce precipitation in our overview.

6. Finally, as an advocate of open access for data and a frequent reviewer for this journal, this reader wonders if the 8 recommendations might have some impact or appropriate- ness as guidelines both for data providers submitting to this journal and for reviewers struggling to assess the quality of the submitted products? Recommendations 1, 2, 5 and 6 seem highly relevant for example. Some of us have pushed the editors for more detailed standards (at the same time recognising as this paper clearly points out that one standard will never fit all) but I understand that mostly they (or somebody at Copernicus) check the presence absence of functioning access links. Perhaps even listing some or all of these recommendations as guidance on the ESSD web site would assist providers and reviewers (and do a nice job of promoting this paper)?

We are glad that the wide applicability of some of the recommendations suggests this possibility to the reviewer. This proposal is one for the ESSD editors to consider.

7. Also, one wonders if and how the recommendations pertain in an open access environment as promoted by ESSD (among others)? Most of the discussion in the paper seems to refer to a within-house exchange between CCI and reanalysis or climate modelling centres, but for some of these data (clouds, for example) one could image a larger group of less familiar users. One wonders if or how the recommendations might change with those non-specialist users in mind?

The primary audience for the recommendations we see is really producers of data with making uncertainty information as a result more accessible to their users, who in turn are supported by that to make more informed use of data themselves. Producers of data are often themselves scientific users too, and will benefit both their own science as well as the applications of other users if the recommendations are followed. Recom-

mendations 3 to 7 are principally aimed at clarifying and making directly accessible in forms meaningful to non-specialists in products and documentation the understanding of uncertainty developed by the data producer.

Thank you for the comments that have improved the manuscript significantly, and time and attention given to our paper.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-16, 2017.