

Comments on “DUACS 1 DT2014 : the new multi-mission altimeter dataset reprocessed over 20 years”

by M.-I. Pujol, Y. Faugere, G. Taburet, S. Dupuy, . Pelloquin, M. Ablain and N. Picot
Ocean Science Discussions Manuscript #os-2015-110

Reviewer: Dudley Chelton

This paper is a very valuable contribution to the altimeter data user community and is likely to be heavily cited because of its comprehensive summary of the improvements made to the latest version of the DUACS altimeter datasets (more commonly referred to colloquially as the “AVISO dataset”). In addition to a general overview of all of the changes that were made between the old altimeter dataset (named DT2010 in this paper) and the new DT2014 dataset, the two most important contributions of this paper are the documentation of the estimated errors of the gridded products in Sec. 3.2.3, and the extensive validation effort by comparisons with in situ data in Secs. 3.2 and 3.4 (monthly mean tide gauge sea level; dynamic height derived from temperature-salinity profiles; and surface drifter data).

While the value of this paper is indisputable, the present version needs a lot of work. I list below two major issues, followed by numerous specific comments, questions and suggestions and, lastly, a list of minor comments, mostly editorial in nature. My comments are mostly focused on the gridded products, which are what I have used almost exclusively for my own research.

My comments that follow are embarrassingly extensive, especially for a journal that posts reviews in their entirety in a Discussion section. The level of detail of my comments reflects the amount of time I’ve devoted in the past to thinking about these issues with the old DT2010 dataset. This paper is an opportunity to put all of those issues to rest once and for all for any future researchers using the DT2014 altimeter dataset.

Major Comments

1. Users have long been frustrated that it is so difficult to track down the details of the OI procedure used by DUACS to generate the gridded SSH fields. The complicated paper trail that users must claw their way through to understand all of these details is evident from lines 10-12 on page 2 of this manuscript in which the authors list a sequence of 7 publications (to which I would add Le Traon and Ogor, 1998) that describe the evolution of the DUACS system and associated products. My own frustration with this is summarized near the beginning of Appendix A2 of Chelton et al (2011, *Progress in Oceanography*) where it is stated that

[...] The details of this procedure have evolved somewhat over the years and an up-to-date and comprehensive documentation of the procedure is not available. Many of

the details are described in a sequence of published papers: Le Traon et al. (1995, 1998, 2003), Le Traon and Ogor (1998), and Ducet et al. (2000). [...] we summarize here our understanding of the procedure as it has most recently been implemented. In addition to the above publications, this summary is based on personal communication in December 2010 with G. Dibarboure at CLS in Toulouse, France who presently oversees the AVISO processing.

Rather than include the subsequent description of all of the details, I would have much preferred to point to a single publication in which users would find all of the details in a single place. Not including the details in this paper would perpetuate the difficulty for users to understand how the parameters of the OI procedure impact the interpretation of the data. One consequence of this is that many (probably even most) users believe that the new DT2014 dataset resolves features with spatial scales of $\frac{1}{4}^{\circ} \times \frac{1}{4}^{\circ}$ and temporal scales of 1 day. This is, of course, not even vaguely close to the true spatial and temporal scales of the features that can be resolved. But without easy access to the specifications of the spatial and temporal autocorrelation functions used in the OI analysis, it is understandable that users will make such gross misinterpretation of the resolution capabilities of the DUACS/AVISO dataset. This paper is an opportunity to provide users with “one-stop shopping” for everything they need to know about the dataset.

2. The English grammar and spelling are in need of extensive work by a copy editor. There are also a lot of instances where the authors used incorrect words. These English problems are understandable since English is not the native language of any of the authors. And their English is infinitely better than my French. But the extent of the problem is quite extreme. I did not carry out a careful statistical analysis, but I have a sense that more than half of the sentences have an English problem of one sort or another, albeit many times only a minor problem. Much of the time, the reader can guess what the authors are trying to say. But there are quite a few cases where I just couldn't figure it out. I draw attention to some of those in the Minor Comments below.

Specific Comments

1. Page 1, lines 14-15: It is stated here and in several other places in the paper (e.g., page 3, lines 12-13, and somewhat less forcefully on Page 21, lines 23-25) that the main improvement in the DT2014 dataset compared with DT2010 is in the use of a 20-year reference period rather than the 7-year reference period used previously. I think this is actually a relatively minor point. Indeed, the authors explain how to change the reference period in a very short appendix at the end of the paper. The main improvements in the DT2014 dataset are the reduced smoothing of the along-track data, improved orbit accuracy, improved ionospheric and atmospheric corrections and the refined scales of the correlation functions used in the OI procedure (the details of which the authors relegate to a reference to Dibarboure et al., 2011, instead of

including them in this paper).

2. Page 6, lines 13-14: “Reference mission” is defined on lines 2-4 on this page, but “secondary mission” is not clearly defined anywhere. The informed reader will know that the sequence of secondary missions is ERS-1, ERS-2, ENVISAT, Cryosat, AltiKA and HY-2A, but this should be spelled out here to avoid ambiguity.
3. Page 8, lines 2-3: I’m not certain I understand what is meant by “The wavelengths ranging nearly 200 to 65 km are filtered”. I think that the short end of this may be imposed by the along-track smoothing applied to the track data. But from page 11, lines 11-12, I think that 65 km is a global average for that, rather than a fixed number everywhere. And furthermore, the wavelengths shorter than 65 km are also attenuated, so there should not be a lower range of 65 km for attenuation by filtering. Is the 200 km upper end of the range something that is imposed explicitly in the OI procedure? In any case, the wording is awkward. Perhaps say something like “SSH variability with wavelengths shorter than about 200 km is attenuated.”
4. Page 8, lines 8-15: This is the place to include the specifics of the parameters for the OI mapping procedure. In my opinion, it is not enough to simply say that the parameters used previously for DT2010 were refined for DT2014 and then send the readers to Dibarboure et al. (2011) to find those details.
5. Page 8, lines 16-18: I would follow this first sentence of this paragraph with a qualifying sentence like, “But note that the time scales of the variability that are resolved in the DT2014 dataset are not substantially different from DT2010; these time scales are imposed by the temporal correlation function used in the OI mapping procedure.”
6. Page 8, lines 21-22: Change “an improved resolution” to “a higher grid resolution” so that readers don’t draw the incorrect conclusion that the smaller grid spacing using essentially the same OI correlation parameters somehow magically resolves smaller-scale variability.
7. Page 8, line 22: Prior to the sentence that begins “These latitudes include the main part...”, insert a qualifying sentence like “Note, however, that the spatial scales of the features that are resolved in the DT2014 fields are about the same (perhaps slightly smaller) than in the DT2010 fields; these spatial scales are imposed by the spatial correlation function used in the OI mapping procedure.”
8. Page 8, lines 25-26: I’m not sure I agree with the statement that the new gridding “reduces the capability of the gridded products to accurately represent the mesoscale signal in high latitude areas.” The $\frac{1}{4}^{\circ} \times \frac{1}{4}^{\circ}$ is much finer than the spatial correlation scales used in the OI mapping procedure. I therefore do not expect there to be much

difference in the feature resolution capability of the new $1/4^\circ \times 1/4^\circ$ gridded product compared with the old $1/3^\circ \times 1/3^\circ$ Mercator gridded product.

9. Page 8, line 29: I am wondering what happens at the highest latitudes in the DT2014 dataset at times when data are not available “from the recent altimeters like C2”. Are the gridded values flagged as missing, or are they filled with some other value such as the mean?
10. Page 9, line 8: After this first paragraph summarizing the processing for SSH, it would be helpful to include a summary of the details of the calculation of geostrophic velocity. For example, it is my understanding that the DT2014 processing used a wider stencil as suggested by Arbic et al. (2012) to compute the SSH derivatives. Arbic et al. (2012) is included in the reference list, but I could not find it cited anywhere in the paper. The authors apparently intended to discuss the changes in the derivative calculations but either forgot to include that text or inadvertently deleted it from an earlier draft of the paper.
11. Page 9, line 12: The phrase “The mesoscale signal is indeed better reconstructed...” in the discussion of the all-sat-merged product compared with the two-sat-merged product is ambiguous. The meaning of “better” should be defined. In lines 20-21 on this same page, it is stated that the OI parameters are the same for both products. The all-sat product therefore doesn’t have any better space-time resolution than the two-sat product. So “better” evidently means “more accurate”. This is an important point because virtually all users believe that the all-sat product has higher spatial resolution, which is not the case since it uses the same correlation parameters in the OI mapping procedure. But the all-sat product is presumably more accurate.
12. Page 11, lines 5-7: Defining the resolution of the along-track data as the point where the spectra of signal and noise intersect is too liberal. With a signal-to-noise ratio of 1, it is impossible to distinguish between signal and noise. In my experience, based on visual assessment of various fields with simulated noise added, a more reasonable choice is a S/N variance ratio of 10, which corresponds to a signal standard deviation that is about 3 times larger than the noise standard deviation. At a minimum, a S/N standard deviation ratio of 2 is needed to distinguish signal from noise (i.e., a S/N variance ratio of 4).
13. Page 11, line 7: It should be clarified that the 55 km mesoscale resolution capability corresponds to wavelength resolution. And the apparent contradiction between 55 km on line 7 and 65 km on line 12 should be clarified.
14. Page 14, lines 18 and 25: I am very surprised, and indeed skeptical, that changing the gridding from the original $1/3^\circ \times 1/3^\circ$ to $1/4^\circ \times 1/4^\circ$ is a bigger factor than the improved processing in determining the increased SSH variance of DT2014 compared with

DT2010. I would have thought that the primary explanation for the increased variance in DT2014 is the reduced smoothing applied to the along-track data that are used in the OI procedure. I wonder if the authors can somehow assess the impact of the different along-track smoothing independently of other changes to the gridded products.

15. Page 15, lines 11-21: I find the discussion in this paragraph to be rather confusing. But in any case, I disagree with the statement on line 14 that wavelength scales of 80-100 km are “fully observable”. With DT2010, SSH is fully resolved only for scales longer than 200 km (Chelton et al., 2011, *Progress in Oceanography*). It is asserted later (Page 24, line 3) that DT2014 has a wavelength resolution of 150 km (but see comment #21 below). The statement on lines 11-21 that scales of 80-100 km are fully resolved is therefore incorrect.
16. Page 15, lines 29-30: It is not intuitive that the variance in DT2014 can be 20% smaller than the variance in DT2010 in low variability areas and near the eastern boundaries. It would be good to include some discussion of why this happens.
17. Page 16, lines 2-5: Same comment as #14 above for SSH variance: I am very surprised that the finer gridding has a bigger impact on geostrophic current variance than do all of the other changes in the processing.
18. Page 16, lines 27-31: The error estimates are stated as 4.9 cm^2 in low variability areas, 32.5 cm^2 in high variability areas and 1.4 cm^2 in very low-variability areas. These numerical values seem to differ from the numbers given elsewhere in the manuscript. For example, the three numbers that are stated on lines 21-23 of the abstract on page 1. There are perhaps similar inconsistencies elsewhere in the paper.
19. Page 18, lines 1-15: It would probably be easier to interpret comparisons of speeds rather than velocity components since the velocity fields are very anisotropic so that analysis of each velocity component separately is difficult to interpret.
20. Page 22, lines 5-7: Same comment as #14 and #17 above.
21. Page 24, lines 3-8: This discussion of the resolution capability of DT2014 versus DT2010 is important but confusing. On line 3, the authors assert that the resolution of DT2014 is 150 km, which can be compared with the 200 km resolution estimated for DT2010 by Chelton et al. (2011, *Progress in Oceanography*). The authors cite the OSTST presentation by Chelton et al (2014) for evidence of the 150 km resolution for DT2014. If they want to include this level of detail in this paper, then I think they should include a figure showing this result, rather than pointing the readers to an abstract for the OSTST meeting. On line 6, the authors again state the resolution of the along-track data to be 65 km, but as I have noted above, I believe this is a global

average number. So they should refer to the cutoff as less than about 65 km, rather than a rigid < 65 km. In line 8, I can't figure out where the 300 km end of the range 300-65 km comes from.

Minor Comments (mostly editorial in nature)

1. The word “restitution” on page 1, line 26 and on page 4, line 10 does not seem to be the correct word. My dictionary defines “restitution” as “the restoration of something to its original state.” I think “representation” is a better choice.
2. The word “homogeneous” is used frequently throughout the paper. While I believe I know what the authors have in mind by this, it should be defined unambiguously the first time this description is used. In most instances, I think the word “consistency” is better than “homogeneous”.
3. Page 2, line 26: The phrase “large panel of ocean signals” is unclear.
4. Page 3, line 2: Chelton et al. (2011) is missing from the list of references.
5. Page 3, line 7: The phrase “large panel the AVISO’s users” is unclear.
6. Page 3, line 16: I think the authors mean “derived products” rather than “derivated products”. My dictionary doesn’t include the word “derivated”. This shows up also on page 4, line 26, and on page 24, line 19.
7. Page 4, line 12: Here and at various places later in the manuscript, “ERA Interim” should be replaced with “ERA Interim Reanalysis”.
8. Page 5, line 24: The word “homogene” should be “homogeneous”. But see minor comment #2 above.
9. Page 5, lines 15-16: I am confused about how close to land the various altimeter measurements are retained. Here it is stated that “detection of erroneous measurements was strongly restricted in coastal areas”, but I can’t find a clear description of what these strong restrictions are. On line 17, it is stated that non-repeat data are omitted within 20 km of land. But then on page 7, line 14, it is stated that the proximity to can range anywhere from 0 to 15 km. I guess that this is for exact repeat data. What are the details on how it is decided whether an along-track measurement near land is erroneous?
10. Page 8, line 4: “Keep one point over two” is awkward wording. I think the authors mean “keep every second point along track”.

11. Page 12, line 21: The use of the word “traduce” here and in other places in the manuscript is not correct. I had never seen this word before. My dictionary defines it as “to speak badly of or tell lies about”. I don’t think this is what the authors have in mind, but I’m not quite sure what word to substitute. Perhaps “introduces” works, at least in some cases.
12. Page 13, line 10: What reference level was used for the dynamic height calculations?
13. Page 13, line 24: As noted above from Page 11, lines 11-12, I believe the 65 km is a global average, rather than a fixed value everywhere.
14. Page 14, lines 1-2: Define “short wavelength signal”.
15. Page 15, line 11: I’m not sure what “twice more important” means here. I think the authors mean to say “twice as energetic”. Same comment on page 15, line 26 where it says “EKE is more important”.
16. Page 15, line 30: The phrase “can represent below 80%” is awkward. I think this should be “can represent less than 80%”.
17. Page 16, line 23: The along-track data are not entirely independent of the gridded SSH fields since the mapping procedure is based on spatial and temporal smoothing of the along-track data from the various altimeters.
18. Page 18, line 5: Change “meridian” to “meridional”.
19. Page 18, lines 25-26: I can’t figure out what is intended by the phrase “a thinner data selection”.
20. Page 19, line 31: Define GIA.
21. Page 20, lines 15 and 16: The use of the word “regional MSL trend” and “at such hemispheric scales” seems contradictory. Regional usually applies to much smaller than hemispheric.
22. Page 20, lines 23-29: The word “underline” appears three times here and in many other places in the manuscript. This is not the correct word, but I don’t think a single word can capture every instance of “underline”. In some places, “emphasize” can work. In other places, “underscore” can work.
23. Page 21, line 4: The acronym LWE is defined on page 6, line 16, but I had long forgotten its definition by the time I got to page 21. Since it only appears three times in the whole manuscript, two of them on page 21, maybe it doesn’t need an acronym

that will require readers to go back 15 pages to find the definition.

24. Page 21, line 13: The phrase “was quite two times more stronger” is unclear. I think the authors mean “was not quite two times stronger”.
25. Page 21, line 18: Change “the last altimeter standards” to “the most up-to-date altimeter standards”. The standards used in this paper are most assuredly not the last standards that will ever be established.
26. Page 22, line 1: Describing the dataset as more energetic is not correct. It is the SSH variability in the DT2014 dataset that is more energetic than in the DT2010 dataset.
27. Page 22, line 16: The phrase “is quite 10 times less important than” is unclear. I'm not sure what the authors mean. Perhaps “is not quite 10 times as important as”?
28. Page 23, line 26: I believe that “between eastern and western basin” should be changed to “between eastern and western hemispheres.”
29. Page 24, line 9: I think that “liked” should be “linked”.