Response to reviewer's comments on the manuscript

"A multi collocation method for coastal zone observations with applications to SENTINEL-3a altimeter wave height data"

by Johannes Schulz-Stellenfleth and Joanna Staneva

We thank Reviewer 2 for many helpful and constructive comments. We appreciate the time you have obviously invested in this. In the following, You find point by point responses to all comments given in the review. The original comments are given in bold black and the respective responses in green italic. The page and line numbers refer to the original version and do naturally not exactly match with the revised manuscript.

Anonymous Referee #2

The paper deals with relevant methodology to assess measurement and model errors when data are scattered in space. This is important, particularly to help validation of satellite data which due to its nature, it is difficult to obtain in-situ measurements precisely at same geographical location. The paper is presented in an organized manner, where first a standard approach (triple colocation method) is presented, then the extended method is shown and tested first with synthetic data and subsequently with real data. With the rapid increase of available data (in-situ measurements, satellites and models) this method is expected to be helpful on the assessment and identification of error bars. For this reason, I consider the paper is worth of publishing considering some suggestions for discussions and improvements which would help readers to put the paper more in context.

It seems that the authors want to give special focus on "coastal zone" as this is in the title, however the paper is missing more discussions about the method in the coastal zones, for example the implication of the assumptions for distance selected and the type of interpolation. Although the authors mention the heterogeneity of the coastal zone, probably this heterogeneity is not linear and interpolation methods might be difficult to apply if not considered the physical processes involved in the area where the different measurements come from. Within this context a discussion on what is the implication of the footprint of satellite for this method and in the coastal zone. This together with the performance during high sea states. A quantification of "high sea state" should also be given.

We have added a new subsection 2.4, which discusses the implications of collocation distance and system resolutions based on the coastal background statistics presented in Section 2.3. This is supposed to put the focus on the special requirements of the coast, where we can usually expect stronger spatial gradients than in the open ocean. The consequences of such gradients for collocation errors and resolution related errors are discussed.

The expected errors for the triple collocation method are computed for different collocation distances in a coastal area with a wave height gradient (German Bight).

For the same area the subresolution wave height variance is estimated for different resolution cell sizes. It is explained, that this variance becomes part of the data set error in addition to pure instrumental errors.

The statement about the altimeter performance in "higher sea states" in the conclusions was reformulated and is hopefully more clear now.

We also hope that it is more clear now, that the linear interpolation method used in the presented analysis is of course not always realistic, but still a progress compared to the assumption of spatially constant wave heights, that have to be applied in the standard triple collocation method. It is also more emphasized now, that the multi-collocation method is not restricted to linear approaches, but that higher order interpolations require a larger number of data sources.

Specific comments:

Line 17 page missing "s" in "in-situ wave observations"

Replaced "observation" by "observations" in line 17 of the abstract.

Line5 page2, please specify the time resolution of HF radar

Replaced

"A few instruments, like HF radar are able to capture at least 2D surface currents with large coverage and high resolution quite nicely, but most instruments …"

by

"A few instruments, like HF radar are able to capture at least 2D surface currents with large coverage and high resolution quite nicely. Such systems have a typical range of about 100 km, spatial resolutions on a kilometre scale, and about 20 min sampling (Stanev et al., 2015). However, most instruments …

Line 3 page 3, as mention above, the direct application to coastal zone is not completely explored. Please specify what are the requirements considered when saying "special requirements of the coast in mind"

Thanks, we agree that this point should be explored in more detail. First of all, we have extended the text on page 3 as follows:

"In this study the triple collocation approach is extended and adjusted with the special requirements of the coast in mind, where one can usually expect stronger gradients and smaller scale variations than in the open ocean."

Secondly, we hope that with the added Subsection 2.4. this issue is now presented with more clarity. The main point is that the stronger spatial gradients to be expected in near coastal areas have significant consequences both for collocation errors and errors, which are related to the averaging processes involved in the generation of the different data sets. In the new subsection this point is discussed first in more general terms and subsequently the theory is applied to the example of the German Bight. It is shown that the collocation errors for the standard triple collocation method can in fact be very significant. The multi-collocation method can reduce these errors at the cost of a larger required number of data sources.

Line 12, page 3, related to "the question about accuracy of error estimate: : :: : :Sentinel3a.." Is this solved in this paper? A short conclusion and recommendation should be added

We have added the following text in the conclusion around line 14 on page 19:

"For the analysed 16 months data set the estimation errors are significant, in particular if individual geographic locations are analysed. It would therefore be interesting to continue some parts of the analysis at a later stage of the SENTINEL-3a mission, when a larger data set will be available. More robust results are obtained, if averages over different spatially distributed insitu instruments are considered. …."

Line 15, page 3. Add "The interpolation of numerical model data to given observation locations is usually less critical if spatial resolution is appropriate"

This was added as suggested.

Lines 1:4, page 4. The assumption of linear combination might be not applicable in coastal zone.

The multi-collocation method is not limited to linear interpolation approaches in general. However, if higher order approaches are used, a larger number of data sources is required. To make this point more clear we added

"We will concentrate on linear approximations in this study, however the method is able to deal with interpolation approaches of higher order, if a sufficient number of observations is available."

in line 10 on page 5.

Line 12, page 4. The assumption of 10 km might be questionable and will have a strong impact in the coastal zone. As mentioned above more discussions would be beneficial

In the added Subsection 2.4 an analysis is presented for the effects of the collocation distance on the triple collocation results in a coastal area with a spatial wave height gradient (German Bight). The analysis showed that there is a significant error increase going from 10 km allowed distance to 20 km allowed distance. In, general, the collocation distance is a compromise between the minimisation of collocation errors and the maximisation of the sample size. The 10 km distance limit used in this study is still smaller than the values used in previous studies (Janssen et al.,2007; Caires and Sterl, 2003). If the collocation distance was reduced further, the smaller sample size would lead to increased estimation errors as described in Section 2.1.

We hope that with the added discussion in the new Subsection 2.4 this issue is more clear now.

Line 3, page 5. Please define variable "T"

We added

"Here and in the following, the symbol T denotes the transpose operation."

In line 4.

Line 5 page 7, Can you change the sentence to "Therefore the uncertainties of the estimated vector: : :"?

This was modified as suggested.

Line 20, page 8, change “scaling factors” for “scaling parameters” to have a consistent nomenclature (see i.e. line 16 page 8)

“scaling factors” was replaced by “scaling parameters” in line 20 on page 8, in line 16 on page 10, in line 13 page 11, in line 16 on page 14, and in the caption of table 4.

Line 12 page9, change “For the analysis is: : :” by “For the analysis in..”

Sorry, this was corrected.

Line 24, page 9. Please mention the water depth of buoys

We added information on the water depth for both buoys (25 m for “HEL” and 27 m for “ELB”) in line 25 on page 9.

Page 10, related to table 3. The table 3 is not clear. Please describe each column in the table caption. Why first column appears as “stdv” as column 4, 5 and 6 and units are different. By looking at table 3 it should be easy to see the “truth errors” and also the ones obtained by the Monte Carlo simulation

Sorry, for the confusion – it is a little bit complicated, because the last columns refer to error standard deviations of estimates for variances and covariances, and therefore the units have to be m^2. As suggested, we reformulated the caption of table 3 explaining each of the columns in more detail.

“Parameters used for the Monte Carlo simulations in Section 2.4. The first two columns refer to the stochastic wave height error standard deviation (stdv) and variance (var) assumed for the considered data sources. The third column gives the assumed error cross covariance (covar) values for the two altimeter measurements and the two buoy data sets.The fourth column is the error standard deviation of the estimator for the observation error variances obtained by averaging over 1000 estimation experiments (AVEXP approach) . The values in column 5 refer to the same estimation errors, but derived by application of the method described in Section 2.1 (COMAT approach). The last column gives the COMAT and AVEXP standard deviations for the covariance estimation errors.”

Line 6, page 11 Equation 41, and all the equations. Be sure all parameters are defined explicitly. Hs seems not defined.

We added the definition of the symbol H_s after equation 41.

Line2-3 page 12, does this mean that satellite data are not “very” applicable for storm conditions near the coast? Please discuss

Unfortunately, the altimeter data processing, in particular near the coast, is very complicated (e.g., Chelton et al., 2001). Therefore, one has to be very careful with statements about the expected performance in certain conditions. Because we did not find previous publications , about the likely behaviour in high sea state conditions near the coast, and because the data analysed in this study are not sufficient to answer this question, we would prefer to avoid any statements that could mislead the reader. We have added the following text with an additional reference in the paragraph following eq. 41.

“In particular in coastal areas, the altimeter data processing is quite involved (Chelton et al., 2001), and a number of instrument and processing parameters can have a strong impact on the characteristics of the wave height estimates.”

Ref: Chelton, Dudley B and Ries, John C and Haines, Bruce J and Fu, Lee-Lueng and Callahan, Philip S, Satellite altimetry, International geophysics,69,Elsevier,doi:10.1016/S0074-6142(01)80146-7 2001

Line 20, page 12. Please indicate if water level variations are considered in the wave model

We added the following sentence in the first paragraph of Section 3.3 to clarify this point:

“Spatial variations in bathymetry are taken into account, however temporal variations of water depth due to tides are not included in the simulations”

Line 28 page 13. Is “This is an important question” better as “This is and important result”?

This was changed as suggested

Line 30, page 13 Add “stochastic” before “error” to make it clearer. Same in line 26 of page 14

This was modified as suggested

Line 26 page 15. Referring to “(red dots)” , please refer to corresponding figure

We modified this to

“(red dots in Fig. 8a)”

Equation 50 and 51 use nomenclature (e.g. 62150) which should be introduced earlier, maybe in section of measurements if such specific naming convention is relevant.

We added the following sentence at the end of Section 3.2 on page 12:

“Some of the insitu stations shown in Fig.2b, which are referenced in the subsequent analysis, are labelled by either 5 digit numbers (e.g., ``62168''), or three character strings (e.g., ``ELB”).”

Line 7 page 17, introduce naming “north” and “south” to the locations

We reformulated this part as follows:

“Here, we concentrate on two locations covered by the satellite, which appear as two clusters in Fig. 8b. The “North'” group of satellite observations is shown in blue and the “South” group in red.

Line 13 page 19. Replace “: : : was relatively small and allowed to..” by “: : : was relatively small, however it allowed: : :.”

The proposed formulation would change the meaning in a way we had not in mind in the original version. We hope that the following formulation makes the point a little bit more clear.

"The number of available samples was relatively small and estimation errors had therefore to be taken into account. The usefulness of the derived error bars for the interpretation of the data could be demonstrated."

Line 16 page 19, please specify range of "higher sea states" and also with its relation to varying footprint and implications for coastal applications

We agree, that this requires clarification.This comment is partly related to the comment about Line2-3 on page 12, where we tried to make it more clear, that the performance of the altimeter is depending on a larger number of system and processing parameter, which make statements about the expected performance in certain conditions very difficult.

We also noticed that confusion is caused by the formulation

" … slightly biased high, in particular at higher sea states."

because we have actually only estimated one bias value for the entire sea state range, and what we basically wanted to refer to, is the slope above unity observed for the satellite data on average (1.11). Together with the estimated slight high bias (0.07), this means, that the wave heights are overestimated by about 10% for above mean wave conditions. We have reformulated this part as follows:

"The analysis indicates, that on average the altimeter is overestimating wave heights by about 10% for above mean wave conditions."

Caption of table 2 is missing the description of the mean (third column)

We interchanged the second and third column of the table and added the missing information in the caption.

Figure 3. Is it necessary to show 2 symbols in the legends of the subplots?

We removed the first row of symbols in the legends of all subplots of Figure 3

Caption of figure 4. Mention that the red dashed line only indicates the zero

We added

" …in green, and the dashed red line indicating the zero position."

in the caption of figure 4.

A references that is worth to consider to include: Kaighin A. McColl et al. (2014) Extended triple collocation: estimating errors and correlation coefficients with respect to an unknown target. Geophysical Research Letters.

Thanks - we added this reference in the paragraph following equation 3.

Additional Changes

- *Corrected the n_o value for the 1D case in table 1 (from 6 to 5)*

- *Updated reference for Wiese et al, 2018*
- *Completed bibliography information (e.g., doi) for several references.*