Ocean Science
Discussions

Open Access

OSD

Interactive comment

Interactive comment on "GEM: A Dynamic Tracking Model for Mesoscale Eddies in the Ocean" *by* Qiu-Yang Li et al.

Anonymous Referee #1

Received and published: 31 August 2016

General Comments

The authors present a novel method to assess merging and splitting of ocean mesoscale eddies. The input for the associated program code are sea level anomalies. First, eddies are identified a specific way (labelled "segmentation") to allow for the assessment of merging and splitting. Second, eddies are tracked over time based on the extent of areal overlap of eddies in consecutive time steps, labelled "similarity vector". A look-ahead procedure is part of the code which improves the tracking results.

The newly developed method is unique in that it is to my knowledge the first one to address eddy merging/splitting events of eddies in a semi-Lagrangian manner and very valuable in that merging/splitting of eddies (i) poses major challenges for existing conventional eddy tracking algorithms (ii) is scientifically very interesting to investigate,

[Printer-friendly version]

[Discussion paper]

and last but not least (iii) the method is promised by the authors to be efficient, hence allowing for the processing of large datasets. I strongly recommend the publication of the paper. Given its potential great value I suggest a major effort to improve its perceptibility (see major comments below). I have been quite detailed with comments on typos/sentence structure etc- this may look like nit-picking, though my motivation was to help improve the readability of the paper to make its value easier to appreciate.

Major comments:
- The English is difficult to follow in some places (e.g. pay attention to sentence structure tenses), but also the technical description itself is not easy to follow in some places (noted in more detail below). Please keep in mind also that you submitted to an ocean focused journal/readership (e.g. maybe try to make the headings less technical, to use less abbreviations, to sufficiently describe the figures in the captions and to generally try to further streamline the paper).
- The first step of the algorithm, the eddy identification has been published previously as the authors say (Li et al., 2014, Li and Sun, 2015). Please state very clearly what code/section/figure you have taken straight from there, where you summarize the previous papers and where you potentially add a new facet for the purpose of the merging/splitting assessment in the present paper.
- Given that the paper presents a method and the associated code, my suggestion would be to (i) state in the text overly clearly what the input of the code is, what programming language it is written in, how efficient it is (based on the North Pacific example) and (ii) publish it along with the paper (or provide a link to a repository where you uploaded the code). Obviously, it is left to authors if they like to follow this suggestion but I believe it could greatly enhance the impact of the paper.
- I was wondering if it was useful at all to apply your tracking algorithm to eddies identified not with the watershed? Would you recommend to use the two algorithms you present to identify and track eddies "in conjunction" or does it make sense to test your tracking algorithm separately e.g. with eddies identified with OW. In the text it appears not conclusive to me, on the one hand you say that segmentation is necessary in the

identification process, on the other hand you mention in future research that one could test the algorithm with eddies identified based on other methods.

Specific Comments:

Text:

L14/15: isn't it mainly the look-ahead approach which solves the missing eddy problem, and the similarity approach helps to reduce the number of missing eddies?

L15 "parents and children": unclear at this point; if you want to mention this in the abstract I suggest to explain the terms.

L54 "related to sampling errors and measurement noise": I suggest to make this more general (as you do in L79) and add sth like "but also due to limitations of the eddy detection step"; a missing eddy problem can occur also in model data (e.g. if an eddy is not well defined in a time step or too weak/small etc).

L66 "larger computational complexity": I suggest to define computational complexity.

L93 "If the algorithm works well": don't you show that it works well?

L96 "to limit the size of the study area": why do you do this? Because of computational costs?

L99 "computation complex": unclear, please rephrase.

L105 "mainly include": unclear; "consists of"?

L126-129: I suggest to order the constraints the same way as you apply them in section 2.4.

L127 "the SLA value": the average SLA value?

L128: do you need the 1 cm constraint? It would be nicer without as you may discard long-lived weak eddies.

L129 "large enough": large enough for what, any reason for the 16 pixels? The "potential usefulness" noted in L136 is not obvious to me.

L138 "The above criteria...": isn't it mainly criterion (1) which helps to get rid of the huge features?

L141 "territory": any reason not to call this "area" as you have done before? In general, I suggest to stay either with "territory" or "area" unless you have a reason to use the two (e.g. also L147 "area"). The nomenclature is not always clear to me. Similarly, I would stay either with "eddy detection" or "eddy identification" if you talk about "finding" individual eddies in an SLA map (e.g. you call the section 2 "Eddy detection" and sub-section 2.2 "Eddy identification", and if you do so say early on in the paper that you will refer with "Eddy identification" always to XXX). To keep the nomenclature consistent and as simple as possible will help the reader to keep track of what you are doing. Maybe you can come up with an expression for the "final" eddies also, which includes identified eddies and their tracks (e.g. "tracked eddies")? Along these lines, what is the difference between "link" and "connection" (e.g. Fig. 4)? And, I suggest to either use "intersection" or "overlap" for the territories.

L143 ff "Necessity of segmentation": if this is taken straight from the previous Li papers I am not sure if you really need 2 Figures explaining it- I like Fig 2 to get the point across, but is Fig 3 really necessary? Also, it is somewhat unclear what is taken straight from Li (previous papers) and what is new here in terms of the eddy identification, I suggest to be very clear about it and start section 2.2 with sth like "we have taken/adopted the eddy detection step from Li et al XXX which provides us with the necessary input for the tracking routines, namely eddy territories and boundaries", followed by a summary of the method.

L143 subsection title: as you submitted to an oceanographic journal I suggest a less technical and more purpose based section title, sth like "Determination of merging and splitting events" (..."based on segmentation" if you wish).

L149 "above-identified": unclear, what do you refer to here?

L154 ff: A side note: the watershed segmentation appears elegant but the so defined territories cut across SLA contours- which appears somewhat non-physical. You mention this briefly but I would make it clearer in a sentence. If I think about trapping of mass and material properties by eddies, "stuff" in the outer area (closed SLA contours which enclose both extrema) intuitively may end up in either of the two enclosed eddies

if the split in a subsequent time step, i.e. this outer area technically does not belong to either of the two but is a separate area.

L154: it is irritating presently that 2.3 and 2.4 are named very similarly. I suggest to insert the text of 2.4 right in the beginning of 2.2 (as it provides kind of an overview), and then have subsections for the segmentation step if necessary.

L156: when do you apply constraint (4)?

L168 "which is less dependent on physical parameters": can you say why this should be a good thing? Not necessarily intuitive to me as an oceanographer.

L173 ff: this section didn't help me all that much when reading it the first time as I wasn't familiar with the technical terms used in there. I suggest to either try to make it more general/easier to understand or to get rid of it.

L177 "first": you have a "first" and "second" part here but 4 subsections which follow thereafter. I suggest to either refer first, second, third and forth here or to reduce to two subsections, if needed with subsubsections.

L178 "link between eddies in different snapshots are saved": could you define "link" here in "plain English", e.g. "link of an eddy from one temporal snapshot to the next, namely living, missing, death, birth, and the associated dynamical processes of merging and splitting?". I suggest this as after your elaboration of "map link" I still didn't get what it exactly represented.

L183 ff: I don't find it straightforward to get here the differences of "links", "branch" and "tree".

L193 "Similarity vector": once more I suggest to phrase the sections less technical and more objective driven, e.g. "Eddy similarity" or "Recognizing eddies in subsequent time steps" or so. Similarly, section 3.3 – 3.5 are titled very technically, too.

L193 ff: the section is rather long and sounds rather complicated. In my understanding, you simply define similarity based on the overlapping area of eddies in consecutive time steps. Subsequently, the overlapping area which is closest to the one of the original eddy is defined to be the successor of the original eddy (if the threshold is met). Can you summarize the section in plain English in the beginning of the section?

C5

L195 "evaluates the similarity of these eddies": "evaluates the similarity of these eddies which is defined here based on the overlap of the territory of an eddy in two consecutive time steps." or so.

L203 "rectangular comparison region": I cannot find this in the figure.

L219 "The last type...": unclear sentence, please rephrase (e.g. "can also be identified" to "is prescribed"?).

L239 ff: I find 3.3, the look-ahead method, difficult to follow, especially paragraph 2 and 3. Partly this is due to the phrasing using T0 to T3 (rather than "unrelated" etc). Can you try to rephrase and make it clearer? Also, doesn't belong the first section more to the previous section?

L255 "closed day": I don't understand the concept of closest day, please try to elaborate.

L264: do I understand correctly that an eddy branch involves 2 time steps, and track tree is then the result of concatenating all time steps? If so, please say so, if not, please clarify. It is not entirely clear to me what makes an eddy branch.

L277 "we could not decide": why not in a technical sense? Because the similarity was not sufficiently high for either eddy?

L279 "increases": why increase, doesn't it reduce the number of detected eddies (as one has one eddy birth less as a result)?

L279 "other tracking problems": such as?

L293 "two generations": "three generations"? Parent, child and grandchild?

L294 "due to the complexity of the output": is it only that or does also the computational complexity and cost increase with more generations?

L296 "Computation complexity": "computational cost" or so? This is what I, as user of the code applying it to my data, would be interested in.

L302 This might be the fastest method possible": possible to do what? Why "might"? Change to sth like "The look-ahead method can hardly be made any faster/more efficient".

L321 "NECC": the NECC is rather close to the equator and at the boundary of your

detection domain, hence I would be careful with interpretations here.

L330 "The long-lifetime eddy trajectories imply that the quality of the tracking results is reasonable": I am not sure if this is true. You can easily connect many tracks which do not belong to each other and get a very long, but "bad" track.

L331: you look at only one example here. Can you elaborate on a few more examples and/or if you have done a visual evaluation of your tracks (e.g. animations)? One example appears not sufficient to justify the conclusion that the tracking algorithm "works".

L335 "puzzle": has this happened only once? Why did you come across it? Have you checked this somewhat systematically in your eddy data set?

L343 "A similarity vector": is it obvious if the number of missing eddies is reduced due to the fact that it is a vector or your approach of using the overlapping of territories?

L345 "as scale": "a scalar"? Btw, oftentimes, people previously have used not only distance but in addition a similarity parameter as well, but a scalar defined based on eddy properties such as amplitude/size/vorticity etc.

L364 "that cyclonic": "a pair of cyclonic"?

L365 "for cyclones": "for atmospheric cyclones"?

L369: I would be interested in the number of merging/splitting events per lifetime, could you mention it? And/or if you showed Fig 9 in terms of eddy tracks passing a grid box per year (instead of eddy extrema), one could easily compare Fig 12 to Fig 9 to roughly estimate such a number.

L388: I suggest to highlight the regions you mention in the Figure, with a box or so.

L407 "false": I don't like "false" as it is not obvious what extrema are true and what are false. The original data may have errors and hence false extrema, the smoothed data likewise.

L411 "This is one of the reasons why we need look-ahead": "The ambiguity of the eddy detection procedure and the potential errors in the input data strongly suggest the application of a look-ahead approach. The resulting eddy tracks are largely insensitive for instance to a filtering method applied to the input SLA data." Or so.

L413: in this section, it would be nice to quantify a bit more the uncertainty due to vari-

ations of N and variations of rc. For N you have partly done it (L424 where you provide the

L418 "The numbers of eddies seldom change": "The number of eddies does not change substantially". Can you quantify this a bit more? E.g. provide the

L421 "The numbers of merging and splitting events seem to converge for rc >0.5 as N increases.": "The sensitivity of the number of merging and splitting events seem to converge to 0 for rc >0.5". Can you briefly comment on why this is the case?

L425 "missing eddies, which may also reduce both total eddy numbers and dynamic events.": I would have expected the opposite (unless you consider only eddies with a certain minimum lifespan, e.g. 30 days. I guess this is the case here?)

L427 "0.6-0.7": why is this optimal?

L445 "Thus, it is difficult to directly compare the influences of eddy territory using different tracking algorithms": I don't see how this follows from what you said in the previous sentences. Also, you originally point out that you need specifically identified eddies which allow you to find the merging and splitting events. Which is why you use the watershed algorithm. Hence, overall I don't really see the point of this paragraph.

L447: The paragraph is difficult to follow except for the last two sentences. I like the sensitivity test with respect to the minimum amplitude, only an issue is that you don't strictly test the sensitivity of the eddy boundary as the modification of the amplitude threshold also changes the number of identified eddies (as you state yourself). The whole point of section 5.3 appears to be that the sensitivity of the eddy boundary is not straightforward to determine but is estimated to be small.

L457: does it make sense at all to test the tracking algorithms with the input from other detection algorithms which do not include what you refer to with "segmentation"?

L465 "A better way to obtain these parameters might be to use a nonlinear fitting of the flow field": unclear, please rephrase. The point of the whole paragraph is unclear to me.

L472 "computation": is this referring to the computation time of the tracking procedure only or including the identification? This is great. Could you repeat the detail here on

the model domain and other constraints (eddy size, amplitude, N etc) that the reader can get an impression what it would mean for his/her own application?

L475 ff: I would stress this, this is the really exiting part for me as an oceanographer.

Figures:

General:

-If you show snapshots of eddies, e.g. Fig 11 or 13, I suggest to roughly maintain an aspect ratio which doesn't distort the eddies that much (the extreme case is Fig 13!).

-please describe everything you show in your Figures, e.g. "white dots mark eddy centers" etc.; I mention some of the missing references below.

-the resolution could be increased in some of the figure and the labels increased (e.g. longitudes/latitudes are sometimes difficult to read).

Fig. 1: -mention that the background field shows SLA, and white dots mark eddy centers.

Fig. 2: -say what the letters h, A, t etc in the figure refer to.

Fig. 3: -cryptic caption, please expand (explain letters, what quantity do you illustrate in panel a etc).

Fig. 5: -missing colorbar, say what A1 etc refer to.

Fig. 6a: I suggest to shade only the overlap/intersection of eddy territories.

Fig. 6b:

-change "their are four types" to "there are four similarity types".

-in the text it says you used rc=2/3; you may want to illustrate this threshold here.

-it would help if you noted the "types of similarity" you write about in the text in the Figure in plain English, i.e. "unrelated" etc. You could also insert E2 to E4 in the respective types/quadrants to continue the illustration of 6a.

Fig. 7:

-why are the eddies labelled now Ec1 etc instead of E1 etc? If possible, make it consistent.
-cryptic caption, please expand.

Fig. 8b:
-"of the eddies" or of an example eddy?

Fig. 9:
-are all extrema counted here or only the ones belonging to eddies which existed at least 30 days?
-is the scale logarithmic or just nonlinear?
-mention the red box.
-instead of number of eddy extrema you could show number of eddy tracks per year (i.e. count an eddy which passes through only once – or have you done this? This is unclear) to make the numbers interpretable together with Fig 12.
-9c: I am a bit surprised that the eddy numbers are distributed that evenly.

Fig. 12:
-adjust colorbar so that the figure does not appear only in blueish colors.
-I cannot see the spatial pattern you describe in the text very clearly.
-12d: you could show the ratio of merging and splitting to highlight the difference (as you have done in Fig 9 for cyclones/anticyclones).

Fig. 13:
-note in the caption what are the box, letters, white dots.

Fig. 14a: I like figure 14 a lot, it is great that you tested the sensitivity of the results to the parameters.
-why do you have the spike at rc=0.6?

Technical Corrections:
L31 "passive": delete as you refer not only to passive tracers if I understand correctly.

L32 "has": "may have".

L39 "Recently": delete as automated ocean eddy tracking/identification has been carried out for over a decade (e.g. Isern-Fontanet et al 2003, http://dx.doi.org/10.1175/1520-0426(2003)20<772:IOMEFA>2.0.CO;2)

L42 "in any given global SLA map": insert "in any given global SLA map and they frequently interact" to connect to the following sentence.

L47 "Various": delete.

L48 "closest eddy strategy": "nearest neighbour strategy"?

L50 "searching for": "assessing" or "estimating" or so.

L58 "eddy to another": "eddy track to another".

L58 "as a result, it was abandoned, even though this look ahead feature is": "As a result, the look-ahead approach was abandoned, even though it is".

L80 "in details": "in detail".

L81 "may be more than": "may arise from more than" or so.

L82 "are merged": "subsequently merged"

L82 "it may": "Ei may". Also: be consistent with using subscript or not (Ei, E1 etc).

L82 "at time step k+1": "at time step k+1 if it splits".

L90 "as daughters": "as children"?

L93 "L the pixels": "L denotes the pixels"

L94 "research field": "research fields".

L99 "examples of merging and splitting events in a sample area in the North Pacific".

L100 and others "noises": "noise".

L100 "parameters": "parameter uncertainties".

L108/109 the sentence "The data were ...": I suggest to delete this sentence. The SLA estimates are not perfect/errorless.

L111 "of gridded products": "of the gridded Aviso products".

L114 "in related studies": unclear; "depending on the purpose of a study"?

L115 "smoothing is required to remove data errors": "smoothing may be required".

L131 "especially for the small ones": delete as this is obvious with the 16pixel size

constraint.

L137 "its parameters": which parameters do you refer to?

L138 "remove": "make obsolete"?

L139 "So, they are simpler and more consistent.": unclear; delete?

L141 "to the": "to as the".

L144 "using": "based on".

L151 "seldom": "only marginally" or so.

L152 "shortened": "reduced"?

L153: I much like this approach and description!

L156 "Second...": I suggest you don't need the "second", any closed SLA contour has at least one extremum unless I am missing sth.

L156 "Third...": I suggest to say sth like "Next we apply the constraints (1) to (4)".

L158 "However, we need...": "We used a segmentation approach when applying constraint (1) [only one extremum]". Or else say why you "need" it, sth like "for the subsequent tracking procedure it is necessary to obtain eddies detected based on a segmentation approach" or so.

L161 "watershed": define/explain watershed somewhere.

L162 "slide": "tend to slide" or "would slide" or so.

L166 "to produce": I am not sure what you want to say here: I presume that you want your code to be able to do this so that the code may be more generally/widely applicable, say e.g. "and include in our code the calculation of amplitudes/etc which may be of use for oceanographic and other applications/analysis."

L167 "other": "future"?

L168 "The GEM...": I find it confusing that GEM shows up here all of a sudden; I suggest to connect to the previous paragraph with sth like "the output eddy "territory" and "boundary" from MEI is then used as input for our novel tracking algorithm GEM."

L174 "connecting": sth is missing in this sentence; connecting what?

L177 "input": "as input".

L177 "eddy data": "eddy data obtained in the prior eddy identification step".

C12

L177 "with": delete.
L178 "different snapshots": "subsequent snapshots"?
L178 "in this part": "in this part of the work flow".
L188 "roles": do you refer with roles to parent and child or death, birth etc?
L191 "it includes": in addition to eddy territory and boundary, it needs two parameters as input".
L192 "impact of these parameter choices": "impact of varying these parameters".
L195 "As shown in Fig 5a...": to help the reader follow sth like "We first illustrate this with an example (Fig 5a) before proceeding to the mathematical expressions."
L195 "there were eddies": "there were three eddies".
L198 "territories on the first day, and on the second day": unclear; please rephrase.
L198 "the intersection was": "the intersections were"?
L198 "both": "to their respective"?
L202 "logical": "mathematically logical"?
L206 "days 0": "day 0.
L206 "we overlap the eddy territories into a single map": "we intersect the territories of day 0 and day 1"?
L211 "possibility E2": "possibility that E2".
L213 "the overlap of the same eddy territory should be large enough": unclear, please rephrase; sth like "we expect an overlap to always exist."?
L213 "one grid": "one grid box".
L213 "should be large enough.": insert a transition to the next sentence, sth like "this is one of the parameters to set" or so.
L214 "for this study": "for this study as threshold to assign E2 to E1".
L223 "way for dimensionless": "to scalar" (as oceanographer I think "unit" if I read "dimension").
L225 "good": "high".
L226 "a connection": "an evolution"?
L226 "... eddy branch procedure...": I suggest to delete this sentence, it is confusing

that you refer here to the following procedure.

L228 "However": delete.

L232 "direction": unclear, can you define/rephrase?

L236 "measure neighboring eddies": "to assess the relationship of eddies in subsequent time steps", as you refer to neighboring in time not space, don't you?

L36 "overlap in": "overlap of".

L241 "some possibilities for a give eddy": "an example for a give eddy".

L243 and others "Ec1": why do you change from simple E0, E1 etc to Ec1? Unless you want to express sth with it (e.g. it is a child of E1) I would stick with E0, E1 etc.

L243 "the eddy on day 1": "E3"?

L249 "with N daily": "with N=3"?

L250 "it is similar but slightly different than": "it is similar to the MHA method with important modifications [or differences]".

L256 "potential one": potential for what? Successor of E1?

L265 "look-ahead day and the following eddy": unclear; can you rephrase?

L266 "braches": "branches".

L280 "simply": delete.

L286 "should be": "is"?

L287 "basically": delete? Or alternatively, explain why basically?

L302: redefine M here as it is a while ago that you have done so and you define again also all the other variables.

L307 "effective": "effective compared to the previously suggested methods".

L314 "more": "more frequent".

L321 "the ratio of difference in number of": "the ratio of the difference of the numbers of".

L331 "long-life": "long-lived".

L341 "in eddy tracking used in recent studies": "applied in standard eddy tracking routines".

L342 "we should connect": "that it was consistent with our approach to connect" or "that

it was dynamically meaningful to connect" or so.

L343 "in tracking": "in the tracking procedure".

L344 "usage": unclear; "cases where the look-ahead procedure is activated" or so.

L347 "also": delete.

L355 "also occurred": "also occurred in the box the same time".

L363 "evolution process": "evolutionary process".

L371 "ubiquitous": do you mean that the events are homogeneously distributed in space? Ubiquitous and "very few times each year" sounds slightly contradictory to me.

L374 "merging frequency": "merging frequencies".

L377 "Although merging and splitting events may be ubiquitous in the ocean (Fig 12c,d), there are several types of special regions where merging and splitting events occur more frequently": the first and second part appear contradictory, how about "Although merging and splitting events may occur anywhere in the ocean there is spatial variation in the number of events." or so.

L384 "seldom": replace, with "not"? Or, if you know a study, cite the study and say "only noted by".

L388 "few": "relatively few".

L388 "The existence of "eddy desert" may be due to the fact that the eddy was too small to be detected": "One hypothesis for the existence of the eddy desert has been that eddies there were too small or weak to be detected".

L390 "However, Figures 9 and 12 suggest that fewer eddies accompanied by frequent dynamic (merging and splitting) events caused the "eddy desert."": "Figures 9 and 12 suggest that merging and splitting events may be a major contributor to the "eddy desert.""

L395 "DT14": "the Aviso product DT14".

L399 "As some parameters are used": "As certain parameters need to be chosen".

L401 "Alternatively, we can simply use a five-point quadratic smoothing to remove the noises in SLA data": "As sensitivity test we apply a simple five-point quadratic smooth-

ing to the SLA data."

L402 "C^2-smooth": "C^2-smoothed", also please explain.

L405 "extrema": "extrema (not shown)".

L406 "These imply that the noises in DT14 data are very small.": "This implies that the noise in the DT14 data is sufficiently small for our purpose."

L413 "Impact of parameters": "Impact of variations of GEM parameters" or so.

L414 "these": which do you refer to?

L414 "apply the GEM to the eddies detected in the NPO.": "carry out a sensitivity study in the north Pacific".

L415 "the numbers": "the number".

L420 "are a little more": "occurs slightly more frequently".

L426 "the tracking results should be insensitive": "one would like the tracking results to be insensitive".

L427 "is a potential choice": "appears to be a choice with relatively robust results".

L431 "representative period": please clarify, do you refer to the temporal resolution? In this case I would say N should be larger than 4.

L432 "N=6": N=6 is not shown on the Figure.

L432 "bias": bias compared to what? E.g. "bias compared to N=XXX".

L433 "for these regimes": unclear, please rephrase.

L435 "which might to though complex evolution process": unclear, please correct sentence.

L447 "of eddy boundary": "of the eddy boundary".

L458 "GEM requires": "GEM only requires".

L460 "when considering more complex conditions.": unclear, please rephrase.

L462 "eddy parameters": I suggest to generally replace with "eddy properties" or "eddy characteristics" as you earlier refer with "parameters" to input of you model code.

L471 "increases linearly": "increases only linearly".

L473 "accompanied by other (e.g., velocity-based, O-W-based) identification methods,": delete.

L480 "for tracking": "for the tracking of the".
L484 "Each eddy track was very smooth": "Eddy tracks were smooth".
L484 "NPO": I suggest to not use this abbreviation but just say north Pacific.
L487 "observation data": "observational data".
L487 "for numerical simulation outputs": "for the output of numerical simulations".

———————————————————————

OSD

Interactive comment

[Printer-friendly version]

[Discussion paper]