Dear S. Williams,

We greatly appreciate your efforts in commenting on the manuscript. We carefully studied the comments and hope our reply could meet with your approval.

1. *The question "what is an eddy?" may not be answerable, at least not in a complete or well-quantified way, and at a certain point you just have to accept that and move on.*

We quite agree with S. Williams's comment on the epistemological issue of eddy. Referee 1# pointed out that we should provide or at least adopt the definition of an eddy before introducing the detection algorithm in the manuscript. And we realized that how important it is to describing the definition for readers in order to under our proposed method. But just as Williams said, "what is an eddy" may not be answerable in a well-quantified way. Especially, current methods usually use different definitions on the concepts about an ocean eddy. We should not wait until the people in this field reach an agreement on this definition problem, then start exploring the detection algorithm. Maybe "We just have to accept it and move on".

Anyway, in the revised manuscript, we made a compromise and presented our understandings of defining an ocean eddy in the Sect. 2.2.1 (P6 L10-27, P7 L1-9).

2. *In a similar vein to the other reviewer, I think that comparisons of your method against a "ground truth" may be overstated, in that there may not be a reliable ground truth.*

We also agree with Williams's viewpoint about "ground truth". Currently, various identification algorithms have been proposed, but no methods can be perfect and guarantee the results are 100% correct. We try to improve the detection performance as better as possible in this study by combining multiple criteria in an algorithm. And we carried out substantial validations and comparisons to verify whether this development does enhance the detection accuracy or not. Soon, we realize the tricky problem that we have no absolute reliable "ground truth" for validating the detection results.

We think all algorithm studies have to face this problem, and there is no easy way to obtain the absolute truth. In previous studies, Chaigneau et al.(2008) and Nencioli et al.(2010) adopted an objective validation protocol which used the manual results from experts to evaluate the algorithm results. Referee 1# pointed out that *this validation should be discarded for it has a lot of human errors, especially with a few experts*. But, considering this validation approach is still an applicable one among the few choices, we also adopted in our study. And in order to lessen the subjective bias as much as possible, we defined that although five experts were invited to do the manual detection, only the eddies identified by at least three experts are counted in the final "ground truth" results for validation.

2. ROC analysis

As Williams suggested, the ROC curve is a useful tool for determining the optimal parameter of a method and comparing different methods. Here, we present the ROC results of the method comparison in the South China Sea (SCS) and the eastern South-Pacific Ocean (ESP) as follows:

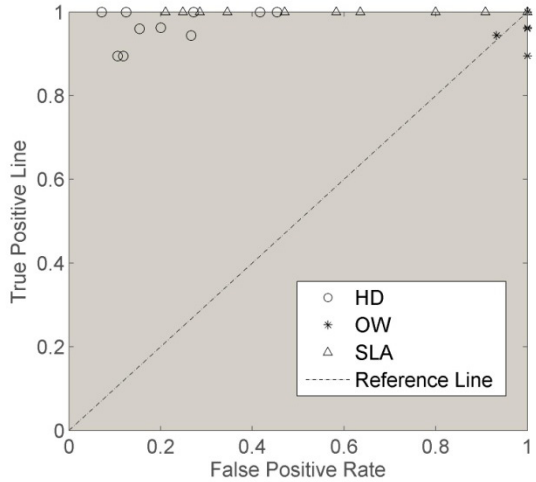

Figure 1. ROC analysis in the SCS

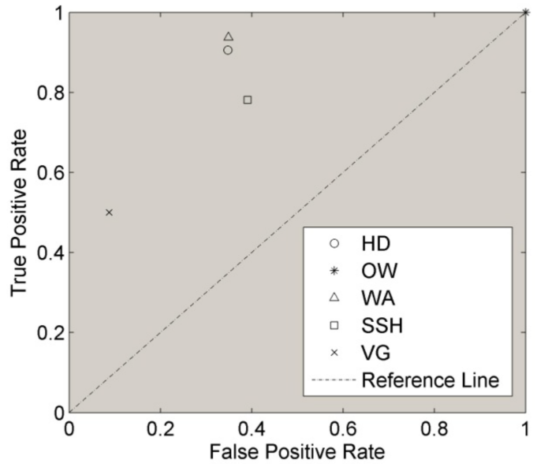

Figure 2. ROC analysis in the ESP

First, we may notice that there are no curves but discrete symbols marking the ROC results of different methods. This is because the main aim of the analysis in this study lies in comparing performance of different methods, not deciding the optimal threshold value. In fact, all the threshold values in the identification criteria of our method are fixed: (1) W<-0.2 σ w, which most research paper adopted; (2) using 3×3 grid moving window to search SLA maxima and minima, which is the smallest scale;

(3) the eddy centers should located over water depths deeper than 100 m, for that the alias from tides and internal waves contained in the SLA data over the shallow shelf area (Yuan et al., 2006); (4) the contours of SLA are generated with 0.5-cm increments, which is finer than the 1-cm ones Chelton et al (2011) used in global study; and (5) unclosed contours or those diameter less than 500 km are discarded, for mesoscale eddies are general about 10-500 km in diameter.

Second, the ROC results are consistent with the accuracy evaluation results using SDR and EDR in our paper. In the ten-sample comparison in the SCS, most ROC points of HD method are located near the top-left corner, which denotes a very good classification, while the ROC points of OW method and SLA-extrema method are not that good. In the ESP, the HD method, though is not the best among the five identification methods, is the second best just next to the WA method. These findings are also evidenced by the EDR/SDR validation protocol. Actually, if we regard the eddies detected by experts as the true eddies, and all the else eddies incorrectly detected by methods as false eddies, then we can derive that:

$$TPR_{(i)} = \frac{TP}{P} = \frac{SDR_{(i)} \times N_e}{N_e} = SDR_{(i)},$$

$$FPR_{(i)} = \frac{FP}{N} = \frac{EDR_{(i)} \times N_e}{EDR_{\max} \times N_e} = \frac{EDR_{(i)}}{EDR_{\max}},$$

Where $TPR_{(i)}$ is the true positive rate of the i^{th} method, TP is the number of true positive, P is the number of positive, $SDR_{(i)}$ is the successive detection rate of the i^{th} method, N_e is the total number of eddies that experts detected, $FPR_{(i)}$ is the false positive rate of the i^{th} method, FP is the number of false positive, N is the number of negative, $EDR_{(i)}$ is the excessive detection rate of the i^{th} method, EDR_{max} is the maximum EDR of different methods.

The ROC analysis is a useful tool, but given the analysis results provide no further findings, we wander if it is very necessary to input this analysis and its background introduction in the revised manuscript in Sect. 3.1 (P13 L19-26) and in Sect. 3.2 (P14 L22-26).

3. “move past detection on to studying eddies functionally”

We quite agree with Williams on this comment. The motivation of working on the detection method and improving the accuracy is to derive more correct dataset of ocean eddies so that further analysis upon these data can help researchers discover more reliable space-time characteristics or valuable underlying patterns of eddies. Currently, we have extracted about 880 eddy tracks since 1992 using the improved detection and companion tracking method. The database records where and when these eddies gave birth and disappeared, and also their spatial trajectories. We are now working hard on discovering possible repeated patterns or spatiotemporal variations, and hoping we could dig up any valuable findings.