

Interactive comment on “Forecast and analysis assessment through skill scores” by M. Tonani et al.

M. Tonani et al.

Received and published: 26 July 2009

We thank the referee for pointing out missing elements in our study. We think we have answered all the comments and the paper is improved. We start by addressing the general comment and then move on to the minor points and editorial remarks. A revised manuscript has been submitted to OS, a paragraph has been added to ch. 4 and all the figures have been redrawn.

Referee: I suggest the authors to choose a more suitable set of scores to carry out their analysis. In addition, no observations are included in the assessment, so everything is rated using analyses as "sea truth", and this hide the actual skill of the model forecast. Answer: Thanks for this remark. We have re-formulated the skill score used in this work and we have also taken into account the available observations in order to evaluate the analysis and to explain the choice of assessing the forecast against the analysis.

GENERAL REMARKS Referee: The authors claim to design this new score, the SSP, following Murphy's papers. Briefly, Murphy's MESS is $MESS = 1 - MS(\text{forecast-observation}) / MS(\text{reference forecast} - \text{observation})$ In this paper, analyses are considered to be the "truth" instead of observations (by the way, that may be a severe assumption, and not really justified in the manuscript), and persistence is considered as the reference forecast. Following Murphy's theory, MESS should then be read: $MESS = 1 - MS(\text{forecast-analysis}) / MS(\text{persistence} - \text{analysis})$ The reason for that is: $MESS > 0$, then forecast is better than persistence, i.e., add more information, $MESS < 0$ means the forecast is worse than persistence (generally you don't want this). Instead the author choose the score in the following way (I write it disregarding the percentage): $SS = 1 - RMS(\text{forecast} - \text{analysis}) / RMS(\text{forecast} - \text{persistence})$ or, equivalently, $SS = 1 - RMS(\text{analysis} - \text{forecast}) / RMS(\text{persistence} - \text{forecast})$ So, the authors are rating analysis vs. persistence considering the forecast as it was the true state. It doesn't sound that close to Murphy's approach. If $SS > 0$, it means that analyses are closer than persistence to forecasts, which is not a useful information. If you show that forecast are closer than persistence to analyses it may help more. Using Murphy's MESS, you can judge if the forecast is good/bad compared to a reference forecast. Using the score presented in this paper you can't. So what are you assessing then? I'm confused. Answer: We thank the referee for pointing out this error. We have reformulated the SSP following your suggestion. Referee: I don't agree with the interpretation of the relative weight between FA and FP (and therefore the resulting considerations throughout the manuscript), which affects also SSP estimates (by definition). For example, in section 4, para 3 and 4, the authors state that forecast is better than persistence since $SS > 0$ (or in other words, $FA < FP$). Now pretend that in one location the analysis (supposed to be the sea truth) is 10 degC. The forecast is 12 degC. Persistence is 9 degC. Then $FA = 2$ degC, $FP = 3$ degC. In this example, persistence is more accurate than forecast, but $FA < FP$ and consequently $SS > 0$. On the opposite, if persistence is 11 degC, then it is still more accurate than the forecast, but now the $FP = 1$ and therefore $SS < 0$. . . I don't understand why you

want such a score! If you want to rate the forecast compared to persistence, I suggest to use $\text{RMS}(\text{forecast-analysis})$ vs. $\text{RMS}(\text{persistence} - \text{analysis})$. . . and not to use what you call "persistence error", $\text{RMS}(\text{forecast-persistence})$. Answer: We have modified the formulation of SSP and we consider the $\text{rms}(\text{forecast-analysis})$ vs. $\text{rms}(\text{persistence-analysis})$ and not $\text{rms}(\text{forecast-analysis})$ vs $\text{rms}(\text{forecast-persistence})$.

MINOR REMARKS (1) in the manuscript actually MFS analyses are not assessed (as stated in the title). Answer We have added a paragraph at the beginning of Section 4 and we have modified Figure 2 in order to consider the analysis assessment.

(2) Sometimes it seems that the authors actually computed time/space averages of SSP scores (for example, the statement in the first para of 4.2). I recommend accumulating separately the numerator and the denominator first and then to compute the final resulting SSP. Otherwise the median is preferable as measure of central tendency. Answer Thanks for this remark. We have checked the computation of the SSP; the time/space average is computed accumulating separately the numerator and the denominator as first step and then the final result is computed.

(3) about the SSP, I don't see positive outcomes in taking the root of MSE, instead of just sticking with the MSE. But I see a drawback, since RMSE skill score decomposition would not be as straightforward as MSE skill score decomposition is. Answer We consider the rmse instead of the mse in order to keep the physical dimension of the considered variables. We believe that this information could be useful and makes the result more understandable from an oceanographic and physical point of view.

(4) add reference for FGAT Answer We have reformulated that sentence and added the proper reference.

(5) ch.3, 3rd para: what do you mean with "best" analysis? How did you rate them? Answer The best analyses are the analyses computed every J from J-15 to J-9. The analyses from J-8 to J-1 will be recomputed the next week. The differences between the analyses for the same day computed at Tuesday (J) and Tuesday (J+7) are in the

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

quality of the assimilated data.

(6) I would drop figure 1, it doesn't seem relevant. If not, please state in the figure captions what's "J". It is explained in the text but well below your first pointing to figure 1. Answer We have created a new Figure 1 instead of dropping it, in order to explain the MFSTEP forecast production cycle and let the reader understand how the weekly ten-day forecast cycle is performed. We have explained in the caption of the figure what J is.

(7) ch 4.3 second para. different scores from summer 2005 to winter 2006 may be associated to a seasonal signal ? and not necessarily an improvement due to data assimilation?. Answer We have rephrased most of ch 4.3 because we have extended the period of study to a full year. We have rephrased that consideration and provided a more detailed explanation.

EDITORIAL REMARKS

(1) font-size of the labels in figure 7-8 is too small Answer We have redrawn all the figures and now the labels are bigger.

(2) ch.4 first para. It is not figure 4, it is 3. 2nd para: it is not figure 3 it is 4. Answer We have taken this remark into account

(3) references are not in alphabetical order Answer We have taken this remark into account and the references are now in alphabetical order.

Interactive comment on Ocean Sci. Discuss., 4, 189, 2007.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

