Ocean Sci. Discuss., 3, S282–S287, 2006
www.ocean-sci-discuss.net/3/S282/2006/


Ocean Science Discussions



# *Interactive comment on* "DINEOF reconstruction of clouded images including error maps. Application to the Sea-Surface Temperature around Corsican Island" *by* J.-M. Beckers et al.

**J.-M. Beckers et al.**

Received and published: 31 July 2006

Thank you for your careful review and comments that will help us improve the paper presentation and discussion of validities of statements.

*It is well written but I suggest to simplify and limit the number of mathematical expressions*

**Response** The same comment was done by Michel Rixen, so we moved some of the secondary mathematics into the appendix.

*It would also be better to follow the data assimilation notations (B, H, R).*

**Response** We had some discussion when writing the first version of the paper. For readers not used to data assimilation but to standard optimal interpolation in its basic form, it would make the paper much more difficult to read and would force him to master the concept and manipulations of matrix $H$. Because this matrix is so simple here we preferred to explicitly perform the operations involving H. For readers used to data assimilation, we added the footnote in the first version and we think that for them, switching to the B-H-R notation is straightforward while for readers not used to data assimilation jargon, the inverse would be more difficult.

*DINEOF method. DINEOF uses an EOF decomposition and then project the data onto the selected EOFs. DINEOF thus cannot handle non homogeneous or correlated observation errors. If this is true, this should be stated somewhere. DINEOF uses a limited number of EOFs and basically provides a large scale/low frequency interpolation. It cannot map the fine scale structures (or the number of EOFs would be prohibitive). It is not clear also how DINEOF can handle much larger areas (e.g. a global map of SST, see specific comment). In practice, DINEOF, OI and krigging are similar methods to solve an interpolation problem but OI (and krigging) are much more general (as shown by the paper OI can yield the same results as DINEOF depending on the assumption on the covariance matrices).*

**Response** DINEOF can map fine scale structures as long as the clouds do not mask them too often. Increasing the number of EOF is not much of a practical problem since the problem of size $N^3$ even for a large number of EOFs is still feasible. The handling of the multi-scale aspects is not trivial , be it for OI or DINEOF (see also response on specific comments)

OI and kriging are more general only insofar as the specification of covariance matrices is left open for choice. In reality these covariance matrices should be the real ones. The strength of this method, is that the covariance matrix is entirely based on the data (assuming that the error covariance can be represented by a reduced rank matrix). In practice, OI can use either a parametric error covariance function (for which the parameters, like correlation length, need to be determined) or a reduced rank matrix expressed by EOFs as DINEOF. In the later case, however, one has to deal with an incomplete data set in order to derive the EOFs from the observations (see also comment below).

Concerning the covariance matrix of the observational error (R matrix), we only use a diagonal form to derive the error variance of the analyzed field. However, the formalism can be extended to use a non-diagonal form of the R-matrix. The reason why we decided to use a diagonal matrix is the low cost associated with its inversion. For a non-diagonal R matrix one needs to inverse a $m_p$ by $m_p$ matrix where $m_p$ is the number of present observations at a given time. The cost of the error estimation would be comparable to a global OI.

*The "OI version of DINEOF" which is used to estimate the error field is now very different from the DINEOF method itself. The main issue for any OI is to define and invert the data covariance matrix and this is done in an efficient way here using the SVD decomposition of DINEOF. The paper actually deals with the estimation of the inverse of the data covariance matrix and not on an extension of the DINEOF method itself. If the covariance matrix is well represented by the N selected EOFs and for constant and non-correlated observation errors, OI and projection onto EOFs are indeed equivalent.*

**Response** The classical OI approach and DINEOF are similar but we don't think that they are completely equivalent. One way to look at the differences is to recognize that in order to define a data-based error covariance matrix defined by EOFs one needs a gap-free data set, and to fill a gappy data set with OI one needs to know the error covariance matrix. This circular dependency is resolved in DINEOF with an iterative approach. Also the OI version of DINEOF used to derive error maps is not very different in terms of analysis from the original DINEOF, as shown in the second part of section 6.3. Since the method adds error fields to the DINEOF analysis, we can see this as an extension of the method, though the term extension is probably open to interpretation.



OSD

3, S282–S287, 2006

Interactive Comment

*In any OI, one would remove first the seasonal signal and possibly a large scale signals (mean or trend) (as a first guess) and then analyze the residuals. I assume this should be done in the same way here (otherwise the first EOFs probably represent the seasonal signal).*

**Response:** Yes, the first EOF does indeed represent the seasonal cycle. In the present paper, we used the EOF decomposition to isolate this seasonal signal. If this EOF were removed from the data as a preprocessing step, the new 1st, 2nd, ... EOFs modes would be equal to the 2nd, 3rd EOFs mode obtained in this work. But we agree with the reviewer that the reconstruction will not be the same with or without the seasonal cycle. The benefit of including the seasonal cycle is that the reconstruction will be able to represent modulation of the seasonal cycle (like the 2003 heatwave or 2004, the year "without summer" in Europe). But we recognize that there are also drawbacks. If only a few data points are present, the method could produce a winter SST distribution in summer. This is one of the reasons why we do not attempt to reconstruct SST fields with less than 5 % data. After taking this precaution, we didn't observe such problems. In general, the EOF approach, exactly as any OI works on data *anomalies*, where the anomalies should be calculated in the most intelligent way, dealing with trends, seasonal cycles etc. The definition of anomalies is the responsibility of the scientist who is working on a specific problem.

We further will adapt the text to take into acount the previous comments.

*Page 739. 5. OI is often used in a sub-optimal way, i.e. only data that are correlated with the estimation point are kept. However, this does not really degrade the results as only useful data are kept. I assume this is also needed for DINEOF when dealing with larger areas (extracting local EOFs) (otherwise only large scale EOFs will be kept and mesoscale features will be filtered out).*

**Response** We agree with the reviewer that this an important issue. Global EOFs do not honor the multi-scale nature of the ocean. To our knowledge, there is no obvi-



**OSD**

3, S282–S287, 2006

Interactive Comment

ous/elegant solution to this problem. While a global truncated EOF series is unable to represent small-scale variation, local EOFs ignore large-scale correlation (induced by processes such as ENSO, NAO,...). If the data coverage is uniformly dense (as it is the case for satellite data), long -range correlation can be neglected since the large-scale processes are well present in the data and "oversampled". For in-situ data where coverage is highly non-uniform, one needs to include long and small-range correlations but we think that even for satellite data this is desirable.

One possibility to tackle a larger-scale problem would be to reconstruct global SST, e.g., at 1 degree resolution, then reconstruct the SST anomaly at 1/4 degree (observed SST minus reconstructed global SST) for each ocean basin independently, and then reconstruct SST anomaly at 1/20 degree (observed SST minus basin-wide SST) for each sub-basin independently, and so on.... At each level one would introduce more and more small scale features.

This is however a topic well beyond the scope of the present paper and we will add the discussion into the conclusion for futher possible research.

*The assumption that the first N EOF retain the signal and that the remaining ones correspond to noise is a very strong one. Remaining EOFs mainly correspond to mesoscale/submesoscale signals which are filtered out by the method.*

**Response** Right again, see also comment of Damia Gomis and our response. For further information on our approach: we choose a long SST time series (10 years) in order to be able to represent some (recurrent) mesoscale features. However the optimal number of EOFs included only the 11 most dominant EOFs discarding most of the mesoscale signal. We attribute this to the fact the typical clouds are larger than mesoscale features. Therefore mesoscale features covered by clouds are undersampled from the start and cannot be reliably reconstructed.

We will adapt the text and the discussion on noise accordingly

Interactive comment on Ocean Sci. Discuss., 3, 735, 2006.