KERNEL EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS OF 1992-2008 GLOBAL SEA SURFACE HEIGHT ANOMALY DATA

Allan. A. Nielsen, Ole. B. Andersen and Per Knudsen
DTU Space – National Space Institute, Technical University of Denmark
Building 321, Richard Petersens Plads, DK-2800 Lyngby, Denmark
E-mail: {aa,oa,pk}@space.dtu.dk

Linear empirical orthogonal function (EOF) analysis is extended to a kernel version. Results of kernel EOF (kEOF) analysis from the last 17 years have been obtained from analyzing 3 degree longitude and 2 degree latitude, 65 degrees South to 65 degrees North joint TOPEX and JASON-1 monthly sea surface height (SSH) anomaly data over the period 1992-2008. Preliminary analysis shows some interesting features related to large scale ocean currents and particularly to the pulsing of the El Niño/Southern Oscillation (ENSO). Large scale ocean events associated with the ENSO related signals are conveniently concentrated in the first SSH kEOF modes.

1. INTRODUCTION

This paper describes a kernel version of empirical orthogonal function (EOF) analysis and its application to detect patterns of interest in global monthly mean sea surface height (SSH) anomalies from satellite altimetry acquired during the last 17 years. EOF analysis like principal component analysis (PCA) is based on an eigenvalue decomposition of the variance-covariance matrix of the data. The kernel version is based on a dual formulation also termed Q-mode analysis in which the data enter into the analysis via inner products in the so-called Gram matrix only. In the kernel version the inner products are replaced by inner products between nonlinear mappings into higher dimensional feature space of the original data. Via kernel substitution also known as the kernel trick these inner products between the mappings are in turn replaced by a kernel function and all quantities needed in the analysis are expressed in terms of this kernel function. Results from the last 17 years have been obtained from analysing 3 degree longitude and 2 degree latitude, 65 degrees South to 65 degrees North joint TOPEX and JASON-1 monthly sea surface height (SSH) anomaly data over the period 1992–2008.

Similar data but covering a much shorter time span have previously been analyzed in for example (Hilger et al., 2002, Nielsen et al., 2002, Nielsen et al., 2002a and Nielsen et al., 2002b).

2. EOF ANALYSIS

Empirical orthogonal function (EOF) analysis (Preisendorfer, 1988) is principal component analysis (PCA) (Hotelling, 1933). In PCA we construct new uncorrelated components which are linear combinations of the original variables by maximizing the variance of the new components. The task is: find linear combinations $x'a$ of the original (zero mean)

p-variate x such that the variance of $x'a$ is maximized. ' denotes vector or matrix transpose. This leads to an eigenvalue problem.

2.1. *Primal Formulation*

In the primal or R-mode formulation the eigenvalue problem reads

$$X'X/(n\text{-}1)\ u = l\ u$$

where X is the n by p so-called data matrix with rows x', n is the number of observations (pixels), p is the number of variables (months), u is the unit length eigenvector ($u'u = 1$), and l is the eigenvalue which is equal to the maximized variance. $X'X$ is p by p.

2.2. *Dual Formulation*

To get to the dual or Q-mode formulation we multiply from the left with X and get $XX'/(n\text{-}1)\ (Xu) = l\ (Xu)$ or

$$XX'/(n\text{-}1)\ v = l\ v$$

where v is proportional to Xu which is generally not unit length if u is. XX' is n by n. This shows that $X'X$ and XX' have the same eigenvalues. Now multiply from the left with X' to obtain

$$X'X/(n\text{-}1)\ (X'v) = l\ (X'v)$$

which shows that $X'v$ proportional to u is an eigenvector of $X'X/(n\text{-}1)$ with eigenvalue l. Scale u to unit length assuming that v is a unit vector to obtain the relations $u = X'v/\sqrt{(n-1)l}$ and $v = X\,u/\sqrt{(n-1)l}$.

This means that $X'X$ and XX' have the same eigenvalues and that if we have the eigenvectors of the one we can easily obtain the eigenvectors of the other. $G = XX'$ is called the Gram matrix. The elements of the Gram matrix are the inner (dot) products of the rows in X, $x_i'x_j$.

2.3. *Kernel Formulation*

We now replace x with $f(x)$ which maps x nonlinearly into a higher dimensional feature space (Schölkopf et al. 1998, Shawe-Taylor and Cristianini, 2004, Bishop, 2006, and Nielsen and Canty, 2008). As an example 2-D $x = [z_1\ z_2]'$ could be mapped into 5-D $f(x) = [z_1\ z_2\ z_1^2\ z_2^2\ z_1z_2]'$. This mapping makes discrimination between linear as well as quadratic forms including ellipsoids possible. The mapping by f takes X into F with rows $f(x)'$. For primal PCA in this space we get

$$F'F\,u = l\,u$$

where we re-use symbols from above and let l subsume the factor $n\text{-}1$. For the corresponding dual formulation we get

$$F F' v = l v$$

with $u = F'v/\sqrt{l}$ and $v = F\,u/\sqrt{l}$.

2.3.1. Kernel Substitution The elements in FF' above are the inner products of $f(x_i)'f(x_j)$. This is all the information that goes into the eigenvalue problem. This means that to solve this problem we needn't know *f(x)* itself, we need only know the inner products. Applying kernel substitution also known as the kernel trick we replace FF' by some kernel matrix K with elements $k(x_i,x_j)$. K is n by n and must be symmetric and positive (semi)definite. Now the kernel eigenvalue problem reads

$$K v = l\, v.$$

2.3.2. Basic Properties. Apart from the solution to the eigenvalue problem including projection of *f(x)* onto the primal eigenvectors u, $f(x)'u$, several basic properties including the norm in feature space, the distance between observations in feature space, the norm of the mean in feature space, centering to zero mean in feature space, and standardization to unit variance in feature space, may all be expressed in terms of the kernel function without using the mapping by *f* explicitly, see (Schölkopf et al. 1998, Shawe-Taylor and Cristianini, 2004, Bishop, 2006, and Nielsen and Canty, 2008). Again, we don't need *f(x)* itself but but only inner products $k(x_i,x_j) = f(x_i)'f(x_j)$ which are the elements of the kernel matrix.

2.3.3. Popular Kernels. Popular choices for the kernel function are stationary kernels that depend on the vector difference $x_i - x_j$ only (they are therefore invariant under translation in feature space), $k(x_i,x_j) = k(x_i - x_j)$, and homogeneous kernels also known as radial basis functions (RBFs) that depend on the Euclidean distance between x_i and x_j only, $k(x_i,x_j) = k(||x_i - x_j||)$. The Gaussian kernel is used very often: $k(h) = \exp(-½(h/h_0)^2)$ where $h = ||x_i - x_j||$ and h_0 is a typical distance between observations in feature space.

3. DATA

The data applied in this study consist of gridded 3 degree longitude and 2 degree latitude, 65 degrees South to 65 degrees North joint TOPEX and JASON-1 monthly sea surface height (SSH) anomaly data over the period 1992–2008.

4. RESULTS

Preliminary analysis based on a random spatial sub-sample of the data and the Gaussian kernel shows some interesting features related to large scale ocean currents and particularly to the pulsing of the El Niño/Southern Oscillation (ENSO). Large scale ocean events associated with the ENSO related signals are conveniently concentrated in the first SSH EOF modes. A major difference between the classical linear EOF and the kernel EOF analysis is the concentration of patterns associated with large ocean currents in the first two kernel EOF modes. Figures 1 and 2 show the first three classical linear EOF and the kernel EOF modes as red, green and blue. Interpretation of these modes are often carried out by simultaneous inspection of these modes and correlations over time between these modes and the original monthly SSH anomaly data (not shown here).

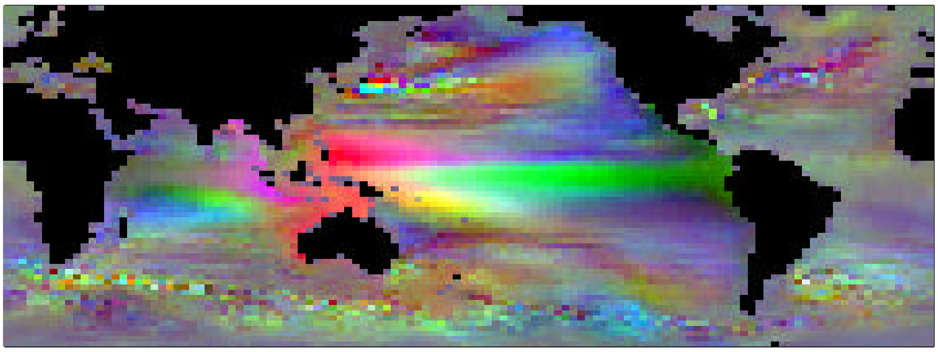

Figure 1. The first three linear EOF modes shown in RGB. The El Niño which built up during the last half of 1997 is very conspicuous. The large ocean currents are seen as very noisy patterns.

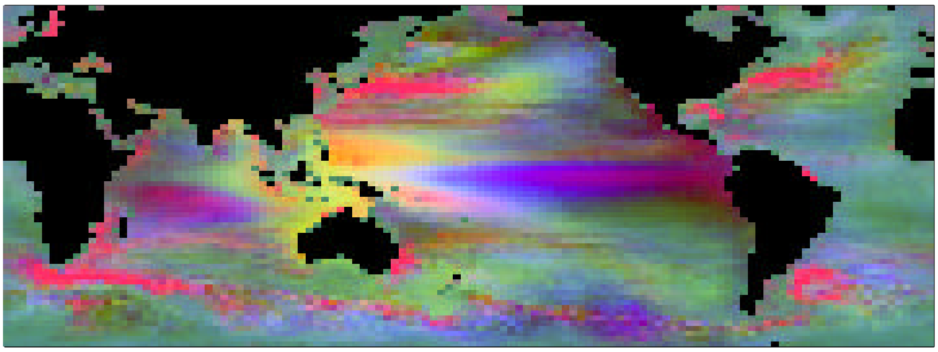

Figure 2. The first three kernel EOF modes shown in RGB. As in Figure 1 the El Niño which built up during the last half of 1997 is very conspicuous. The large ocean currents are here seen as spatially autocorrelated red/magenta patterns characterised by positive score values in kEOF mode 1, negative score values in kEOF mode 2 and zero score values in kEOF mode 3, i.e., strong red, no green and some blue.

REFERENCES

Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*, Springer.

Hilger, K. B., A. A. Nielsen, O. B. Andersen and P. Knudsen, 2002. "An ACE-based nonlinear extension to traditional empirical orthogonal function analysis". In L. Bruzzone and P. Smits (editors) *Analysis of Multi-Temporal Remote Sensing Images*, pp. 171-178, World Scientific. (MultiTemp2001, Trento, Italy, 13-14 September 2001.)

Hotelling, H., 1933. "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology* **24**:417-441 and 498-520.

Nielsen A. A. and M. J. Canty, 2008. "Kernel principal component analysis for change detection". SPIE Europe Remote Sensing Conference, Cardiff, Great Britain, 15-18 September 2008.

Nielsen, A. A., K. Conradsen and O. B. Andersen, 2002. "A change oriented extension of EOF analysis applied to the 1996-1997 AVHRR sea surface temperature Data". *Physics and Chemistry of the Earth* **27**(32-34):1379-1386.

Nielsen, A. A., K. B. Hilger, O. B. Andersen and P. Knudsen, 2002a. "A temporal extension to traditional empirical orthogonal function analysis". In L. Bruzzone and P. Smits (editors) *Analysis of Multi-Temporal Remote Sensing Images*, pp. 164-170, World Scientific. (MultiTemp2001, Trento, Italy, 13-14 September 2001.)

Nielsen, A. A., K. B. Hilger, O. B. Andersen and P. Knudsen, 2002b. "A bivariate extension to traditional empirical orthogonal function analysis". In L. Bruzzone and P. Smits (editors) *Analysis of Multi-Temporal Remote Sensing Images*, pp. 179-185, World Scientific. (MultiTemp2001, Trento, Italy, 13-14 September 2001.)

Preisendorfer, R. W., 1988. *Principal Component Analysis in Meteorology and Oceanography*, posthumously compiled and edited by C. D. Mobley. Developments in Atmospheric Science, 17, Elsevier.

Schölkopf, B., A. Smola, and K.-R. Müller, 1998. "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation* **10**(5):1299-1319.

Shawe-Taylor, J. and N. Cristianini, 2004. *Kernel Methods for Pattern Analysis*, Cambridge University Press.