CANONICAL CORRELATION ANALYSIS TO EXTRACT EXTERNAL INFLUENCES IN SEA LEVEL AROUND SPAIN

J. García Lafuente [1], J. Del Río [1] *, Enrique Alvarez [2], Begoña Perez [2], J.M. Vargas [1], F. Criado [1]

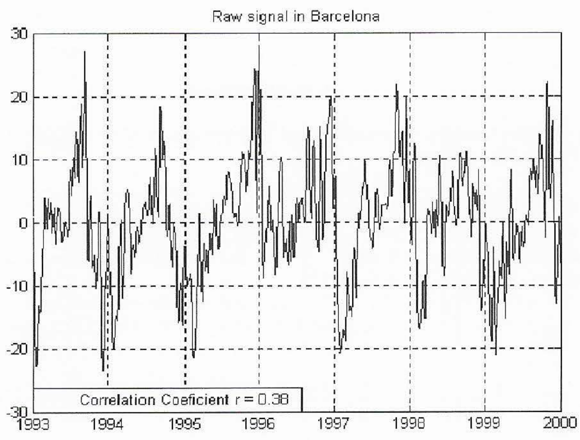
[1] Departamento de FISICA APLICADA II, Universidad de Málaga.
[2] Área del MEDIO FÍSICO, Puertos del Estado, Spain.

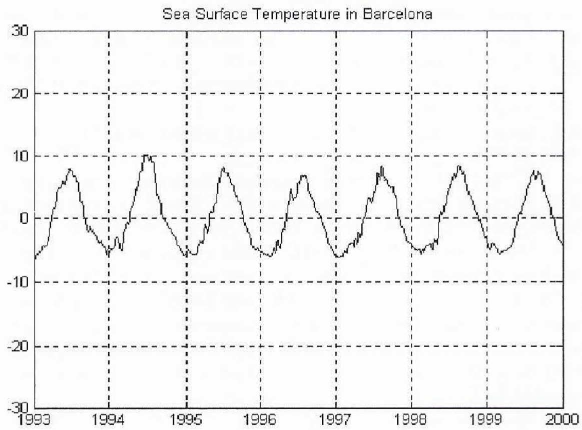
Keywords: CCA, Sea Level, Temperature, Pressure.

Canonical Correlation Analysis (CCA) is a mathematic technique [1,2] that uses two datasets, each one consisting of a given number of time series, in order to look for the canonical functions (or canonical series) that optimize the relationship between a linear combination of variables from the first data set and a linear combination of variables from the second data set. The technique provides ensembles of canonical series, each one accounting for decreasing amount of co-variation between both data sets, in a manner that recalls the decomposition done by Principal Component Analysis (PCA). In other words, the first canonical function will describe what both datasets are sharing most, then the next one will be computed over the residual variance, and so on.

This work uses CCA to extract the variability of sea level in locations around Spain collected by Puertos del Estado, Spain (http://www.puertos.es) induced by external forcing such as atmospheric pressure and sea surface temperature (SST). Atmospheric data have been collected from National Centers for Enviromental prediction (http://wwwt.ncep.noaa.gov/) and SST have been downloaded from Deutsches Zentrum für Luft un Raumfarht (http://www.dlr.de/). The number of canonical series used to reconstruct the signal in each location may be computed through the correlation coefficient between the canonical series of both data sets. If this value is higher than 0.5 we consider that the canonical series contains enough information to be considered as significant. The canonical functions may be used to reconstruct the original signals. If all canonical functions are used, then the original series is recovered (all contributions to the signal are present) but if only a selected and suitable set of canonical functions are chosen (the most significant ones) then the recovered signals mostly contain the effects of the external forcing (parameter) we are considering (pressure or SST for instance) and influences due to other physical variables are rejected.

The recovered signal is better related to the parameter (pressure, SST) than the original series, as a measure of that relationship we use the correlation coefficient. An example of the performance of this technique is shown in the figures, where sea level signal in Barcelona, Sea level recovered in the same location using two canonical functions (the ones that mostly take into account the thermosteric effect) and their respective correlation coefficient with SST are presented, Sea Surface Temperature nearby Barcelona is also shown. Note that correlation coefficient in Barcelona (see figures) between raw sea level signal and SST is 0.38 and the correlation between the recovered signal and SST is 0.61.





References

1 - Barnett T.P., Preisendorfer R., 1987. Origins and levels of monthly and seasonal forecast skill for United States Surface air temperatures determined by canonical correlation analysis. *Monthly Weather Review*, 115: 1825-1850.

2 - Kirby, M. and Anderson, C.W., 2003. Geometric Analysis for the Characterization of Nonstationary Time-Series. *In:* Springer Applied Mathematical Sciences Series Celebratory Volume for the Occasion of the 70th Birthday of Larry Sirovich, ed. by Kaplan, E., Marsden, J., and Sreenivasan, K.R., Springer-Verlag, Chapter 8, pp. 263-292.