Geophysical Research Abstracts, Vol. 7, 10016, 2005
SRef-ID: 1607-7962/gra/EGU05-A-10016
© European Geosciences Union 2005

Spatio-temporal filling of missing data in geophysical datasets

D. Kondrashov (1), M. Ghil (1,2)

(1) University of California, Los Angeles, USA, (2) Ecole Normale Supérieure, Paris, FRANCE (dkondras@atmos.ucla.edu)

Throughout the geosciences researchers have to deal with incomplete datasets. The existence of missing data presents various problems, for example in spectral estimation, specifying boundary conditions in numerical models, etc.

Here we use Singular Spectrum Analysis (SSA) to fill the missing data. SSA is a data-adaptive, non-parametric spectral method based on diagonalizing the lag-covariance matrix of a time series. For a univariate record, our procedure uses only temporal correlations in the data to fill in the missing points. For a multivariate record, multi-channel SSA (MSSA) algorithm takes advantage of both spatial and temporal correlations. We iteratively produce estimates of missing data points, which are then used to compute a self-consistent lag-covariance matrix; and we use cross-validation to optimize the window width and number of dominant SSA (MSSA) modes to fill the gaps. The optimal parameters of our procedure depend on the fraction and pattern of missing data, as well as variance distribution between oscillatory modes and noise.

The algorithm is demonstrated on synthetic examples, as well as various geophysical datasets: global sea-surface temperature, flood-water records of the Nile River, satellite observations of relativistic electrons, and Southern California winds.