Oceanographic Application of Temporal MAF Analysis

Jacob S. Vestergaard and Allan A. Nielsen

Technical University of Denmark

Introduction

Application of multivariate statistical methods to oceanographic datasets originate in the desire to find interesting patterns that can be used to generate new – or confirm existing – hypotheses about changes in global climate.

Features that can be extracted from geophysical datasets include periodical signals, unique phenomena, and trends. Statistical analysis can be used to determine the periodicity of the signals, or predict an event.

Trends are of particular interest, as they represent a property evolving over time. When applied to temperature-, precipitation-, or sea surface height (SSH) data, strong trends may indicate a change in our global climate.

El Niño is an example of a large scale phenomenon that will appear prominent in most statistical analyses. El Niño builds up over a period of 6-7 years, where masses of hot water is pressed against Indonesia and Northern Australia. At some point the warm water floods along Equator to Southern America. This instantly affects weather and animal life globally, and therefore has tremendous socio-economic impacts over the entire planet.

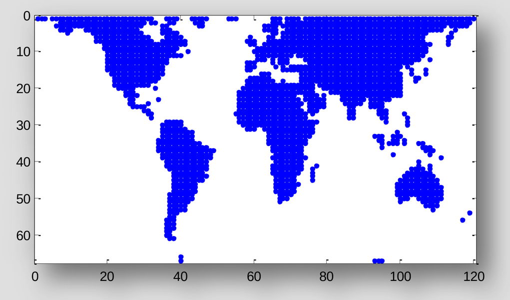
Santa Cruz, February 1998

Data

The data analyzed is sea surface height (SSH) anomalies from the period October 1992 through February 2009. It consists of combined data from the TOPEX/Poseidon and Jason-1 satellites and is the longest consecutive SSH dataset existing.

The data is sampled once a month, resulting in 201 months of data gridded in a 67 x 120 grid. In the latitudinal direction, the observations are spaced by 2 degrees, and 3 degrees in the longitudinal direction.

The northern- and southern bounds of the dataset, are limited by the satellite lanes, as their maximum and minimum latitude is at the Arctic and Antarctic Circle respectively.



A mask has been created to distinguish land- from sea observations. It is defined as the locations where observations are 0 throughout the entire dataset.

Methods

The data is analyzed using the temporal variant of Maximum Autocorrelation Factor (MAF) analysis.

MAF analysis aims to find linear combinations of the zero mean original variables that maximize autocorrelation between the neighboring observations in time.

The autocorrelation is in MAF analysis expressed as:

$$\rho = 1 - \frac{1}{2}\frac{a^T\Sigma_\Delta a}{a^T\Sigma a} .$$

- where a is the linear combinations, and $\sum_\Delta$ is the dispersion matrix of the difference between the original *X(x)* and the temporally shifted variables *X(x+Δ)*

To maximize this expression, the Rayleigh coefficient is minimized by solving the generalized eigenvalue problem:

$$\Sigma a = \lambda \Sigma_\Delta a$$

The SSH data is analyzed with $\Delta = 1$ month and results in 200 spatial representations of data, sorted with decreasing temporal autocorrelation.

Corresponding correlations between these components and original data are used to determine where – in time – each component's behavior is prominent.

Results

The temporal MAF analysis has been applied to the SSH data, projecting the data into 200 components, each orthogonal to each other.

By inspecting the correlations of the first component with the original data, a noticeably linear trend, combined with some El Niño related behavior, can be seen.

Different areas are highlighted by the temporal MAF analysis on the spatial representation of the component; dark areas west of Southern America, and light areas south of Greenland, and near Indonesia.

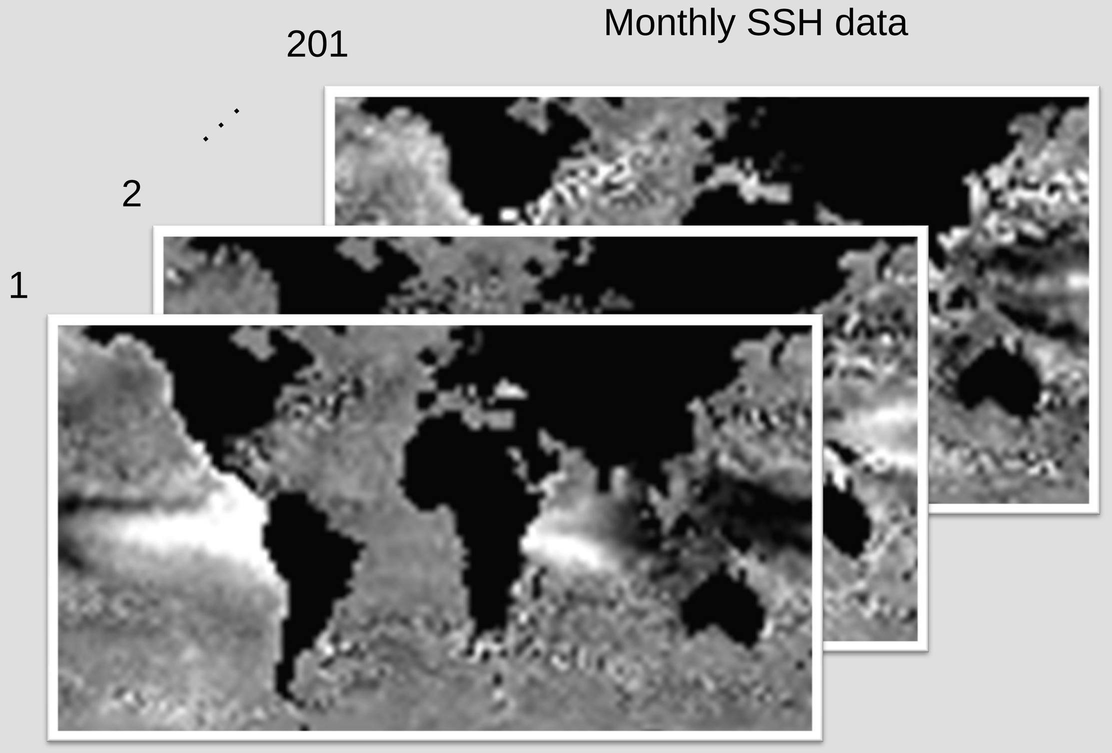
These locations are selected manually, and the SSH anomalies are plotted together with randomly selected pixels in non-significant areas (green) for reference.

The areas in the pacific (blue, cyan) exhibit strong oscillations at the great El Niño of 1997-98, and also a strong positive trend around Indonesia.

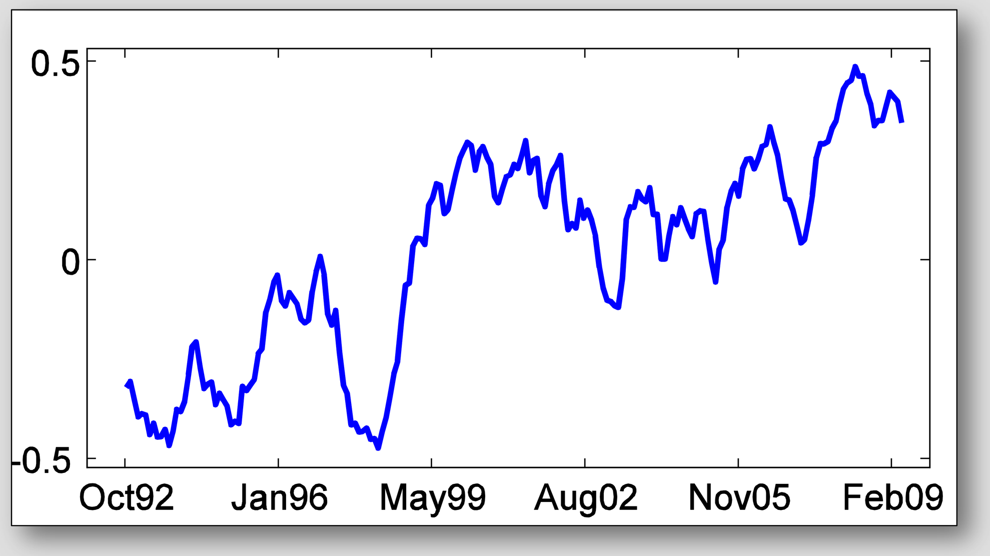
The observations from Greenland (red), however, show no relation to the El Niño at all, and compared to the reference observations (green), the trend is much stronger. This indicates that **the waters around Greenland, are rising with a higher rate, than the global mean**.

Melting ice sheet in Greenland

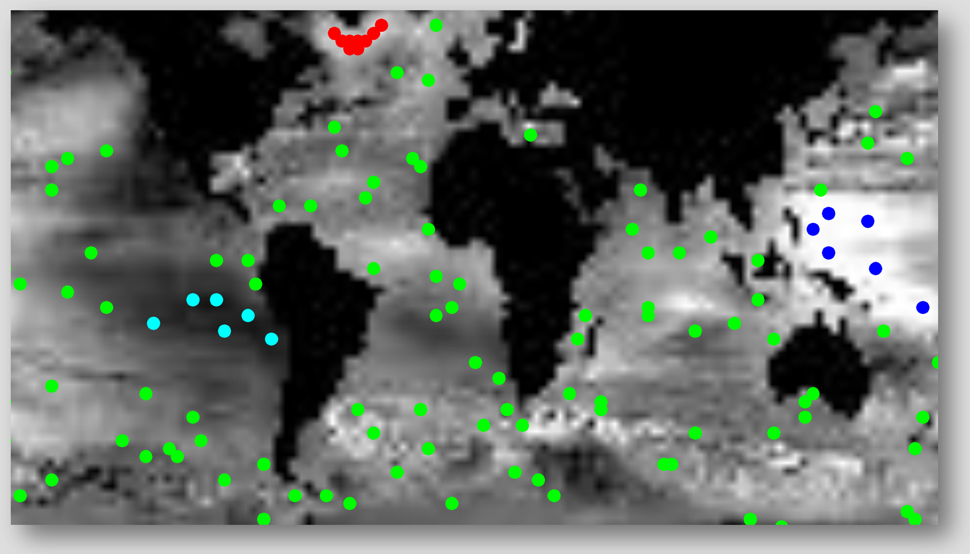
Monthly SSH data



First temporal MAF component's correlations with data



Identification of highlighted areas



SSH values at identified points