Atlantic Multidecadal Oscillation
An Overview

Nikolaus Koopmann

Abstract

On the basis of monthly SST-data, inherent limitations to signal analysis of short data sets are illustrated. A Northern Atlantic multi-decadal SST-variability signal, short AMO, is investigated and theories to its nature, mechanism and connections to other climate phenomena are discussed.

1 SST Analysis

The freely available *KAPLAN EXTENDED v2 ssta* [NOAA, 2012] SST anomaly data set was used to investigate the major low-frequency periodicities of Atlantic SST. This analysis has obviously been done before with greater rigor and expertise [Delworth et al., 1993, Ghil and Vautard, 1991, Schlesinger and Ramankutty, 1994]. This re-analyzation is simply meant as a demonstration tool of high educational value to the author. Besides, modern desktop-PC computing power and high end numerical computing software make it possible to analyze large data sets quickly and with ease. Furthermore it should be considered that SST data sets[NOAA, 2012] have improved since the early nineties [Delworth et al., 1993, Ghil and Vautard, 1991] and that approximately a quarter of a century has passed since. The later is vital since a major problem of investigating low-frequency SST-signals is the shortness of data compared to the period of the signal (see 1.1.4).

1.1 Singular Spectrum Analysis (SSA)

SSA is designed to extract information from short and noisy time series and thus provide insight into the unknown or only partially known dynamics of the underlying system that generated the series... [Ghil et al., 2002]

The implementation method used for this paper will be explained briefly in the following. For more detailed information on SSA-theory see [Ghil et al., 2002, Ghil and Vautard, 1991].

Let $\mathbf{S}$ be the array of SST's with dimensions $(lat, lon, time)$.

1.1.1 Choice of Spatial Window

Instead of simply setting the entire Northern Atlantic as the data domain to be analyzed, a window of the Atlantic where the low-frequency signals are strongest was chosen. This was done by calculating the first two Northern Atlantic $\lambda_{i,j}^{1,2}$ (see section 1.1.5) for all Northern Atlantic grid-points separately (see Figure 1). The effect is that the low-frequency component is dominating strongest over for example an ENSO teleconnection-signal, which is rather strong when the window is chosen to $40° - 60°N$ and $50° - 10°W$ as was done in [Latif et al., 2004].

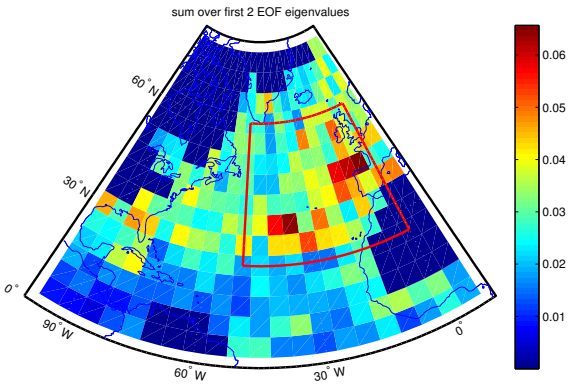

Figure 1: First 2 EOF eigenvalues for entire time span. The window spans from $25° - 60°N$ and from $50° - 0°W$

.

1.1.2 Mean North Atlantic SST Anomaly

Once the desired geographical window is cut out, the matrix is averaged in space.
Since $\mathbf{S}$ is gridded in longitude and latitude, its cells are not equidistant in meridional direction. The averaging must hence be performed in an area-weighted fashion:

- Calculate total number of values, i.e. sum of all weights

$$N = \sum_{i=i_1}^{I} \sum_{j=j_1}^{J} W_{i,j} = \sum_{i=i_1}^{I} \sum_{j=j_1}^{J} cos(lat_{i,j}) \quad (1)$$

- Calculate global mean via

$$= \mathbf{m}(t) = \frac{1}{N} \sum_{i=i_1}^{I} \sum_{j=j_1}^{J} \mathbf{S}_{i,j,t} \qquad (2)$$

The effect is that, for example 10 grid cells combined at $\cos(lat) = 0.1$ have the same effect on the domain mean as one grid cell at the equator. This technique is obviously not exact, especially so for poor resolutions. The correct area of a grid cell at some latitude lat would easiest be calculated in spherical coordinates. Yet, since the coordinates refer to the center of the cell, the applied method is assumed to be of sufficient accuracy for the experimental purposes of this paper.

1.1.3 Detrending

A crucial step is to detrend $\mathbf{m}$. If this step was skipped (as was done in Ghil and Vautard 1991), the strongest variance in the signal would be a non-oscillatory trend stemming from external forcing e.g. global warming. An investigation of inherent low-frequency modes via SSA would therefore be unsatisfactory as the strongest EOF's would have to account for the trend. Yet whether or not a simple linear detrending solves the issue entirely is questionable. As [Zhang and Delworth, 2006] pointed out: *The idealized linear detrending [...] does not necessarily remove all the influence of changing external forcings. Fluctuations of the external forcings about a linear trend may have contributed to some of the observed multidecadal variability.*

1.1.4 Choosing the Lag

The next step is to perform an auto-correlation $\mathbf{c}$ on $\mathbf{m}$ with arbitrary lag $l - 1$. In this paper l was chosen as $l = 76a \cdot 12\frac{months}{a}$. 76 years was chosen as it is the longest period within the results of any of the reference papers [see Schlesinger and Ramankutty, 1994]. A ratio $\frac{n}{l} \approx 2$ [1] is mathematically hard to justify. The whole record is only 156 years long which suggests a maximum lag of around 50 years in order to satisfy at least 3 repetitions of the phenomenon within the data [Ghil and Vautard, 1991]. Ghil et al. 2002 even called an $\frac{n}{l} = 11$ only *fairly* safe. The power density spectra of the reconstructed components (see 1.1.6) will always show highest energy at around $50a$[2]. It is only the $\mathbf{r}_i$ themselves that, if l is chosen too short, will fail to grasp the fundamental signal and mix it into components of higher frequency. Choosing a longer lag will push short frequency signals further down into weaker $\mathbf{r}_i$ but would not alter the shape of the low frequency components. Hence if it is the lowest frequency component that is sought, the lag should better be chosen too long than too short. It must be kept in mind however that at such small ratio of data length over signal period, there is no reason to assume the signal component to be of continuous sinusoidal nature. Due to its high heat capacity, oceanic heat content is strongly auto-correlated. The apparent pattern might therefore as well be a purely undeterministic red noise walk that happens to feature a rather constant period for the time of record. (see also sections 4, 2.1)

1.1.5 Empirical Orthogonal Functions

The EOFs ρ_i of the spatially 0-dimensional time series are nothing other than the eigenvectors of the covariance matrix i.e. the Toeplitz[3] matrix $\mathbf{C}$ of $\mathbf{c}$. The corresponding eigenvalues λ_i (see Figure 2) are crucial in that they show how much of the total variance each principal component carries. The sum over all, i.e. the trace of $\mathbf{C}$, gives the total variance of m.

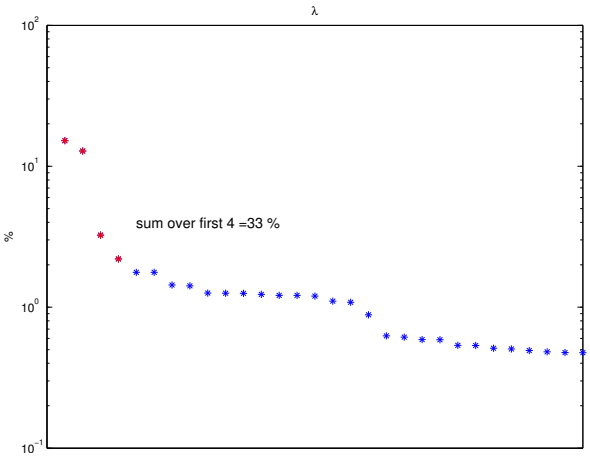

Figure 2: λ

[1] where n is the length of $\mathbf{m}$

[2] ballpark figure

[3] $\mathbf{C}_{i,j} = \mathbf{C}_{i+1,j+1} = \mathbf{c}_{i-j}$, where negative indices count backwards from end to beginning and index counting starts with 0.

1.1.6 Reconstructed Components

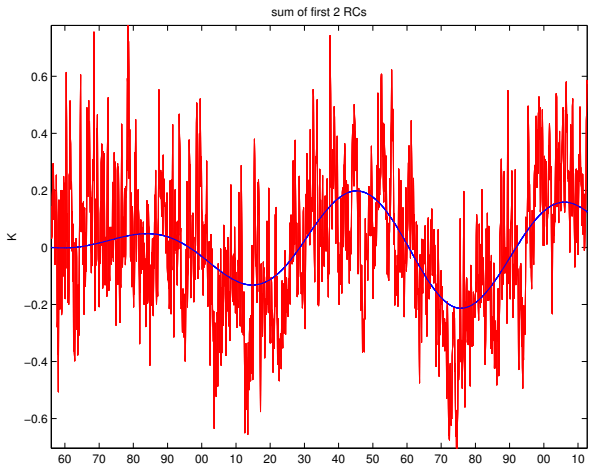

Figure 3: Red is $\mathbf{m}$ (detrended), and blue is the sum of the first two $\mathbf{r}$. Numbers on the x-axis are the last two digits of the year.

Projecting $\mathbf{m}$ onto the eigenvectors ρ_i yields the principal components of $\mathbf{m}$:

$$\mathbf{p}_i(t) = \sum_{j=1}^{l} \mathbf{m}(t+j-1)\rho_i(j) \tag{3}$$

These can in turn be used to reconstruct the i'th component of the actual time series $\mathbf{m}$ corresponding to the i'th EOF ρ_i by

$$\mathbf{r}_i(t) = \frac{1}{M} \sum_{j=L}^{U} \mathbf{p}_i(t-j+1)\rho_k(j) \tag{4}$$

where

$$(M, L, U) = \begin{cases} \left(\frac{1}{t}, 1, t\right) & \text{if } 1 \le t \le l-1; \\ \left(\frac{1}{l}, 1, l\right) & \text{if } l \le t \le N. \end{cases}$$

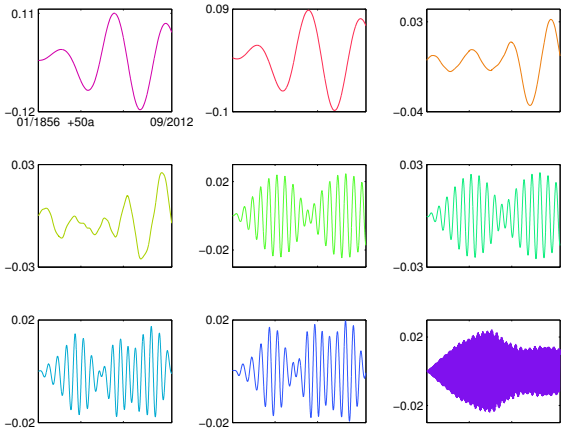

Figure 4: The first 9 SSA reconstructed components of $\mathbf{m}$.

This *back-projection* onto the eigenvectors transforms the principal components $\mathbf{p}_i$ back into the original coordinate system of $\mathbf{m}$, which makes the reconstructed components $\mathbf{r}_i$ very useful. They represent a decomposition of $\mathbf{m}$ into its statistically orthogonal subassemblies so that their sum is exactly equal to $\mathbf{m}$. Ghil et al. 2002 put it in a nutshell: *SSA allows one to unravel the information embedded in the delay-coordinate phase space by decomposing the sequence of augmented vectors thus obtained into elementary patterns of behavior. It does so by providing data-adaptive filters that help separate the time series into components that are statistically independent, at zero lag, in the augmented vector space of interest. These components can be classified essentially into trends, oscillatory patterns, and noise.*

1.1.7 Results

As can be seen from Figure 2 most of the variance stems from the first two components. This becomes also apparent by looking at the first 2 $\mathbf{r}$ in Figure 4 respective Figure 5 and their corresponding power density plots in Figure 7. The first four $\mathbf{r}$ represent the sought-after low-frequency signal, where $\mathbf{r}_{3,4}$ also feature energy at a period of approximately 23 years. $\mathbf{r}_{5,6}$ are almost identical and represent an underlying frequency of approx. 10 years. $\mathbf{r}_{7,8,9}$ feature energy in the ENSO-range and particularly $\mathbf{r}_9$ shows a sharp annual signal. See Figure 6 for further comments on the different peaks. The slopes in Figure 7 indicate red-noise character, as to be expected for SST-data (see section 1.1.4).

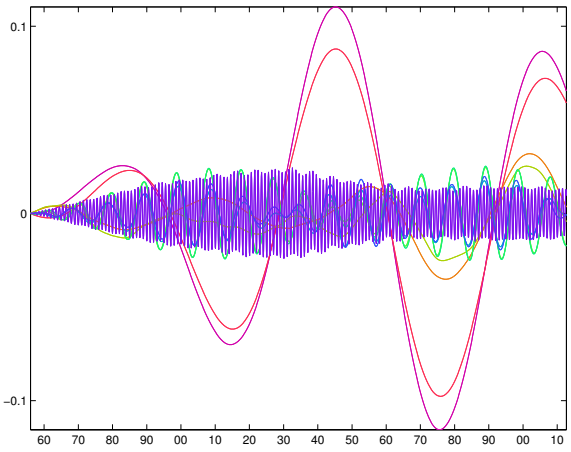

Figure 5: The same as Figure 4.

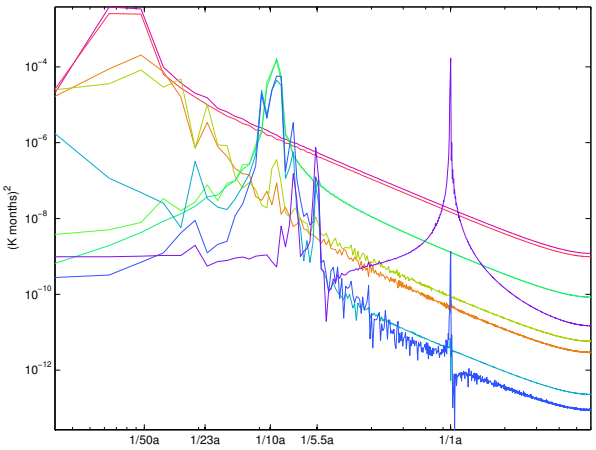

Figure 7: Power density spectra of the first 9 **r**. Colors correspond to those in Figures 4 and 5. A sum over all reconstructs a noise-filtered version of the signal in Figure 6. Note the different types of vertical scaling between Figures 6 and 7.

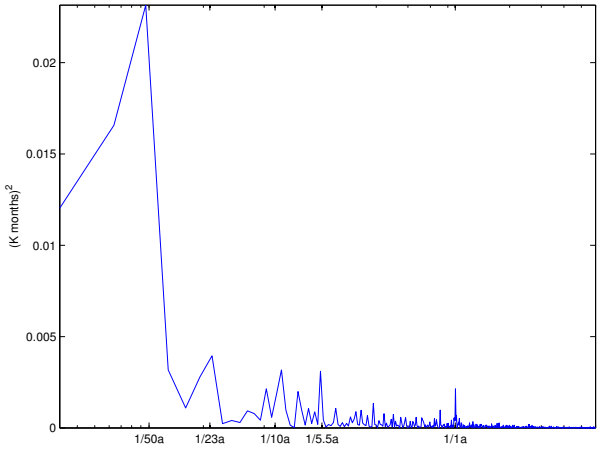

Figure 6: Power density spectrum of the entire data set. 5 distinct peaks emerge: A peak at an annual period stemming from some inter-annual dependency of SST-anomalies, a peak at approx. 5 to 6 years, probably corresponding to ENSO-tele-connections, a peak at a period of approximately 10 years, assumed to be atmospherically forced via NAO [BLACKMON, 1993], a peak at around 23 years, assumed to represent a global fundamental signal found in many climatic variables [Ghil and Vautard, 1991] and a strong poorly resolved peak at say 70 ± 20 years.

[4] trial and error

2 AMO Theories

2.1 Red Noise

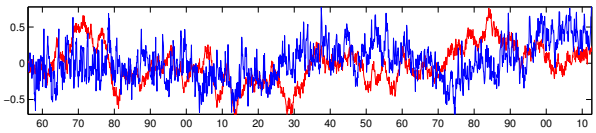

Figure 8: Red: artificially generated red-noise signal as outlined in 2.1. Blue: **m**.

In 1977 Lemke 1977 already proposed the possibility that the low frequency SST oscillations in an ocean could just be red noise forced by atmospheric white noise. In other words, that due to the auto-correlatedness of SSTA, i.e. the ocean's *thermal memory*, the SSTA undergoes a Brownian walk, continuously and asymptotically pushed back towards a long time mean via negative feedback, without which the SSTA would wander randomly along the temperature scale, like a dust particle in a pool of gas molecules [Lucretius, 1744]. The forcing being entirely random, but biased on limited time windows like a small ensemble of thrown dice. Once such bias forces the SST, the sea surface tends to retain this anomaly for some time (depending on its auto-correlation) and then slowly drift back to its mean state.
The process can easily be illustrated:

Consider the second-order auto-regressive time series

$$x(t) = \varphi_1 x(t-1) + \varphi_2 x(t-2) + \varepsilon \qquad (5)$$

I.e. a signal in which the new value $x(t)$ depends on the two preceding values but on top of that is forced by a random value ε. Note that $\sum \varphi < 1$. Otherwise the process would not be auto-regressive and hence not stationary anymore, but instead explode sooner or later. In the example the parameters[4] and the implementation were set as follows:

```
x=zeros(size(sstm));
phi=.53; phi2=.45;
for t=3:length(x)
    x(t)=phi*x(t-1)+...
         phi2*x(t-2)+...
         (rand-.5)*mean(abs(sstm));
end
```

[5] This code obviously has a different outcome on each realization.

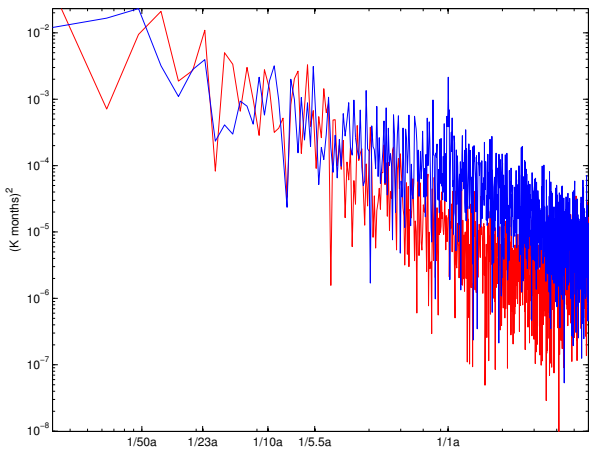

Figure 9: Power density spectra of **m** (blue) and **x**.

[5] `sstm` is identical to m

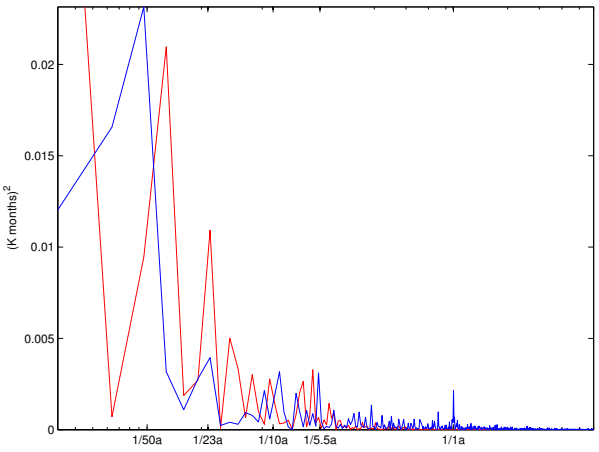

Figure 10: Same as Figure 9, with linear y-axis.

One interesting realization can be seen in Figure 8. The density power spectrum resembles that of **m** remarkably well. At first sight the two graphs in Figure 9 seem to be clearly of the same nature. With a linear vertical axis the graph of **x** reveals relatively narrow peaks between periods of $50 - 5$ years, completely annihilating any grounds on which the assumptions in the description of Figure 6 regarding the three intermediate signals were made, noting that **x** has peaks of equal or larger magnitude within this band. This means that all those peaks in **m** could perfectly well be entirely random.
The same argument can be made via SSA (see Figure 11).

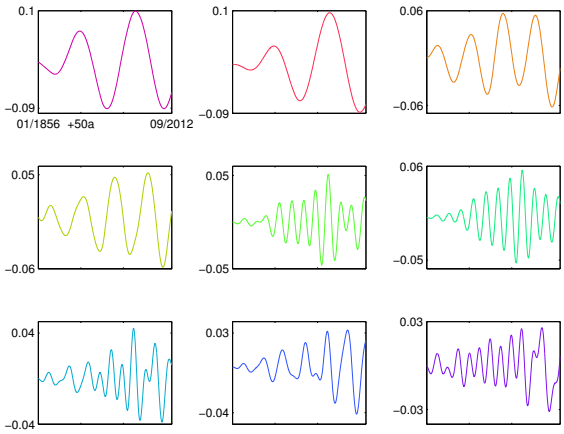

Figure 11: Same as Figure 4, yet with **x** as underlying data instead of **m**.

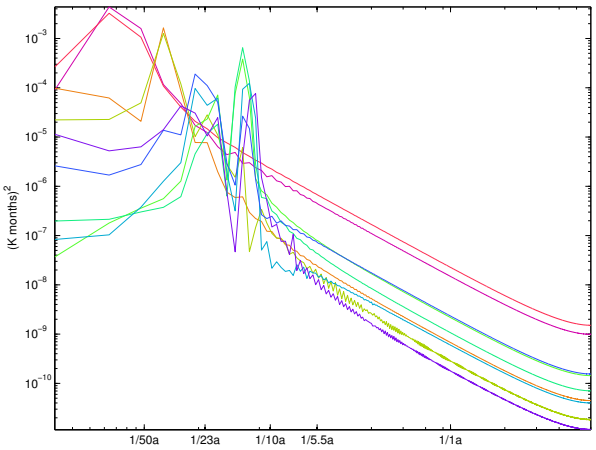

Figure 12: Density power spectra corresponding to Figure 11.

2.2 Internal Modes

The red noise time-series via equation (5) is not meant as an actual model of oceanic SST. The idea was merely to investigate potential pitfalls of stochastically forced signals.
Timmermann et al. 1998 did practically the same, only that they compared their SSTA spectrum to a more appropriate spectrum of red noise by calculating the first-order maximum entropy spectrum of the THC index. Further, they also created confidence levels on the basis of probability considerations. In other words, they can detect parts of the spectrum, that are almost certainly[6] not red noise. The strongest peak of their spectrum penetrates the $> 95\%$-confidence zone, indicating that *other processes than just the integration of synoptic weather noise are involved* [Timmermann et al., 1998]. If not stochastic, one would naturally assume that there be some form of Atlantic eigenmode. Timmermann et al. 1998 though extend this assumption immediately to that of a Pacific-Atlantic coupled mode. They do so because they find the aforementioned Atlantic spectral peak mirrored in the Pacific SSTA spectrum. This is unlikely to be a coincidence as the geometries of the two basins are very different, suggesting different time-scales. Furthermore, cross-spectral analysis of the two oceans revealed both SSTA's to be in phase at that particular frequency, suggesting that the coupling is via the atmosphere and not of ocean-internal nature, as tele-connections via the world ocean would create a lag of at least several years due to the slowness of planetary waves [Olbers et al., 2012].

[6] confidence of 95% or more

2.3 Thermo-Haline Circulation

Delworth et al. 1993 already mentioned that coupled ocean-atmosphere models show irregular THC oscillations creating spatial patterns in SSTA resembling those observed on interdecadal scales in the North Atlantic. Multiple later studies with more sophisticated models manifested the connection between THC and SST further. Latif et al. 2004 for example ran 400 year MPI Global Climate Model runs and found strong covariance between SST and THC. *Specifically, the strength of the meridional overturning at* 30° *correlates almost perfectly with the North Atlantic SST index.* The fact that in their result the northward ocean heat transport consistently leads SSTA, whilst SSTA and surface heat flux are strongly anti-correlated i.e. that SST is damped by the atmosphere, leads the authors to conclude that long time scale SST fluctuations are driven by ocean dynamics instead of atmospheric forcing. Cross spectral analysis found the wind driven part of the northward heat transport to be relevant only on annual time-scales, whilst the thermohaline part dominated strongly on decadal scales.

2.4 Arctic Sea Ice

Dima and Lohmann 2007 found a hemispheric wavenumber-1-like structure as a fourth EOF of SLP data, that turns out to vary on time scales of the AMO and is significantly correlated with the THC index. This atmospheric mode is also found by Grosfeld et al. 2007 and attributed to the SST gradient between Atlantic and Pacific. The pattern has a strong gradient across Fram Strait and hence strong influence on along Fram Strait winds, which in turn are the driving mechanism of Fram Strait sea ice export [Dima and Lohmann, 2007]. They proposed hence, that FSSIE [7] be an important link between atmospheric- and THC-variability.

3 AMO Mechanism

Attempts to sketch a causal cycle of the AMO were done several times [see Timmermann et al., 1998] before. The most recent of which is from Dima and Lohmann 2007 and goes as follows:
Assume a positive THCA state, causing positive SSTA in the Atlantic. The warm Atlantic creates a negative SLPA over the Atlantic extending towards Eurasia, which via tele-connections (see 2.4) causes a weakened Aleutian low, i.e. a positive Pacific-SLPA. This strengthened SLPA discrepancy manifests as a strong

[7] Fram Strait sea ice export

pressure-gradient across Fram Strait causing an intensified FSSIE i.e. and increase of fresh-water input to the North Atlantic. The decreased density weakens the MOC and we are in a state of negative THC, causing negative Atlantic SSTA, hence negative Atlantic SLPA, eventually strengthening the Aleutian low. The resulting weakened southward wind stress over Fram Strait weakens FSSIE/fresh water export and allows for stronger MOC, closing the cycle.

4 Discussion

Section 1 elucidates the complications of interpreting insufficiently resolved signals. Merely three, or even less, repetitions of a full AMO cycle within recorded SST data are simply not enough to even proof the existence of the AMO. As appealing as the first SSA modes may look and as convincing as the spectral peaks on semi-logarithmic axes[8] may seem at first sight, it turned out that same such picture can quickly be drawn by replacing SST data by a trivial auto-regressive time series. Whether or not the signal be just that, a manifestation of Brownian noise, forced by a random atmosphere, or the result of an actual distinct underlying dynamic oscillation system, can only be answered by means of probability.

[...] purely random processes, particularly those that have even mildly red spectra, have a behavior that comes as a surprise to many, and there is great risk of misinterpretation. That is, the purely random behavior of a rigorously stationary process often appears visually interesting, particularly over brief time intervals, and creates the temptation to interpret it as arising from specific and exciting deterministic causes. [Wunsch, 1999]

Multidecadal Oscillations are therefore a prime example of what coupled ocean-atmosphere models are useful for. Parameters can be altered, fluxes can be integrated, SST can be prescribed or the model can be unforced, runs can be arbitrarily long etc. There is no limit to creativity and the possibilities of sensitivity studies to help find the mechanistic indicators of the physical system and their causal interconnections are endless. Proxy data from e.g. tree rings can also be exploited to lengthen the physical data, as was done by Gray et al. 2004. Their results of 400 year tree ring proxies are also indicative of a quasi-periodic AMO cycle at 65-80 year period. Crucial to a description of any poorly resolved data is hence to use all information and forms of interpretation at hand and then to compare the robustness of apparent underlying signals among all types of derivation.

The interesting aspect of climatic variabilities on scales as long as a human life is that it is beyond our timeframe of intuitive perception. While humanity has naturally grasped the ENSO cycle as a *come and go* phenomenon, effects of the AMO become apparent only by stepping back and looking at the larger picture in the form of time series. The AMO period is simply too long for the short data set that is our life. It could be argued that the effects are only subtle and thus elude immediate perception. This depends on where you are. An AMO linked low-pressure anomaly over the southern United States for example has been *associated with precipitation reductions of up to 20% [...] Over western Europe, there is enhanced precipitation([...] 5% to 15% of the mean summer value). This multidecadal change in European precipitation has been previously documented [...] but had not been linked with the AMO.* [Grosfeld et al., 2007]. Further regional multidecadal climate variabilities that have been related to the AMO are North Eastern Brazilian and African Sahel rainfall [Knight et al., 2006] and even glacier mass budget in the Swiss Alps varying in phase with the AMO [Huss et al., 2010].

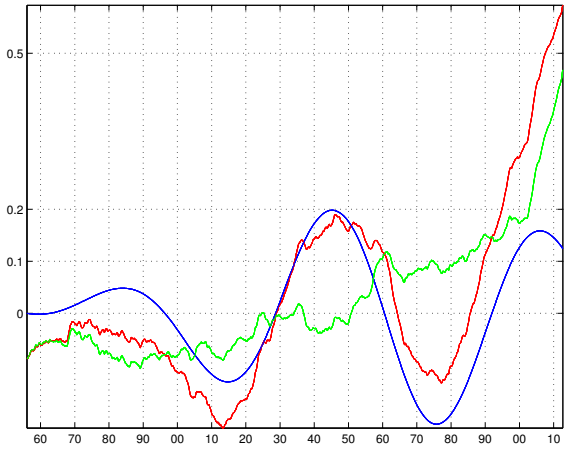

Figure 14: Red: 20a-running mean of $\mathbf{m}$ (non-detrended), blue: sum of first two $\mathbf{r}$, green: their difference.

Slow changes in climate as such tend to be interpreted as trends instead of oscillations. Somebody born in the 70's for example has experienced only one direction of the AMO throughout all his life; up. Somebody from Northern Germany for instance might remember extremely cold and snowy winters and sunny summers of the late 70's and early 80's in contrast to the rainy summers and mild winters of the 00's. Not to suggest that there be any connection to the AMO, for lack of prove. Needless to say that personal perception as such has many other much stronger, psychological pitfalls e.g. selective perception/memory.

[8] A density power spectrum, depicted with logarithmic axis scaling in f-direction can be deceiving in that it hides the integrated power of high frequencies.

The point is that low-frequency signals like the AMO have great potential to mask real climate trends. Figure 14 shows the difference between the actual monthly mean SSTA data and the sum of the first two reconstructed components. I.e. it shows the Atlantic SST trend with the AMO removed. The AMO could possibly have made climate change appear more dramatic than it actually was in the last three decades. But what is more important is that it might mask a continuously increasing upwards temperature trend in the AMO downhill phase, thereby partially compensating climate change effects only to turn the compensation into an intensification one generation later.
A very fortunate fact about the quasi-periodicity of the AMO is that it should be possible to extrapolate its signal and thereby make predictions about future climate [Latif et al., 2004].

References

M.L. BLACKMON. Surface climate variations over the north atlantic ocean during winter: 1900-1989. *J. Climate*, 6:1743–1753, 1993.

T. Delworth, S. Manabe, and RJ Stouffer. Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *Journal of Climate*, 6(11): 1993–2011, 1993.

M. Dima and G. Lohmann. A hemispheric mechanism for the atlantic multidecadal oscillation. *Journal of Climate*, 20(11):2706–2719, 2007.

M. Ghil and R. Vautard. Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, 350(6316):324–327, 1991.

M. Ghil, MR Allen, MD Dettinger, K. Ide, D. Kondrashov, ME Mann, AW Robertson, A. Saunders, Y. Tian, F. Varadi, et al. Advanced spectral methods for climatic time series. *Reviews of Geophysics*, 40(1):1003, 2002.

S.T. Gray, L.J. Graumlich, J.L. Betancourt, and G.T. Pederson. A tree-ring based reconstruction of the atlantic multidecadal oscillation since 1567 ad. *Geophysical Research Letters*, 31(12):L12205, 2004.

K. Grosfeld, G. Lohmann, N. Rimbu, K. Fraedrich, and F. Lunkeit. Atmospheric multidecadal variations in the north atlantic realm: proxy data, observations, and atmospheric circulation model studies. *Climate of the Past, 3.{www. clim-past. net/3/39/2007/}*, 39, 2007.

M. Huss, R. Hock, A. Bauder, and M. Funk. 100-year mass changes in the swiss alps linked to the atlantic multidecadal oscillation. *Geophysical research letters*, 37(10): L10501, 2010.

J.R. Knight, C.K. Folland, and A.A. Scaife. Climate impacts of the atlantic multidecadal oscillation. *Geophysical Research Letters*, 33(17):L17706, 2006.

M. Latif, E. Roeckner, M. Botzet, M. Esch, H. Haak, S. Hagemann, J. Jungclaus, S. Legutke, S. Marsland, U. Mikolajewicz, et al. Reconstructing, monitoring, and predicting multidecadal-scale changes in the north atlantic thermohaline circulation with sea surface temperature. *Journal of Climate*, 17(7):1605–1614, 2004.

P. Lemke. Stochastic climate models, part 3. application to zonally averaged energy models. *Tellus*, 29(5):385–392, 1977.

T.C. Lucretius. *De rerum natura*. 1744.

NOAA. Kaplan sea surface temperature (SST). http://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html, 2012. URL http://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html.

D. Olbers, J. Willebrand, and C. Eden. Ocean dynamics. 2012.

M.E. Schlesinger and N. Ramankutty. An oscillation in the global climate system of period 65-70 years. *Nature*, 367 (6465):723–726, 1994.

A. Timmermann, M. Latif, R. Voss, and A. Grötzner. Northern hemispheric interdecadal variability: a coupled air-sea mode. *Journal of Climate*, 11(8):1906–1931, 1998.

C. Wunsch. The interpretation of short climate records, with comments on the north atlantic and southern oscillations. *BULLETIN-AMERICAN METEOROLOGICAL SOCIETY*, 80:245–256, 1999.

R. Zhang and T.L. Delworth. Impact of atlantic multidecadal oscillations on india/sahel rainfall and atlantic hurricanes. *Geophysical Research Letters*, 33(17):L17712, 2006.

Appendix - Code

```
% NK 2012
close all
clear all
addpath('./sub_routines')

%% ########################INPUT#########################################
%netcdf data
NC='./sst.mon.anom.nc';
%variable key
var='sst';
%lat/long keys
lat='lat';
lon='lon';
%time key
time='time';
%days since 'YYYYMMDD'
ds='18000101';
%chosen dimensions file
input='./input.txt';
%#######################################################################
dims=dlmread(input);

%% read data
ncdisp(NC)
lon=double(ncread(NC,'lon'));
lat=double(ncread(NC,'lat'));
time=ncread(NC,'time');
sst=ncread(NC,'sst');
[J,K,T]=size(sst);
%make flag nan
sst=flag2nan(sst,max(sst(:)));

%% show data snapshot
fig=snap(sst,lat,lon);

%% get time
[time,from,till]=gettime(time,ds,dims);

%% get space
[south,north,west,east,lon]=getspace(lat,lon,dims);

%% cut piece
[sstc,latc,lonc]=cutpiece(sst,lat,lon,south,north,west,east,from,till);

%% create mean
[sstm]=lat_weighted_mean(sstc,latc);

%% build lag shifet matrix Y, toeplitz covariance matrix C and get
%EOFs (rho) and eigenvalues
[Y,C,RHO,lambda]=Toep_cov(detrend(sstm),dims);

%% build principal components
PC=Y*RHO;

%% build reconstructed components
[RC]=RCs(PC,RHO,dims);

%% autoregressive stuff
x=zeros(length(sstm)*1,1);
phi=.56;
phi2=.45;
for t=3:length(x)
    x(t)=phi*x(t-1)+phi2*x(t-2)+(rand-.5)*mean(abs(sstm));
end


```

```
68
69
70
71
72
73  %% plot stuff
74  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
75
76  ... not shown ...
77
78  printall2eps
79
80
81
82
83
84  %prints all figures into current dir as .eps
85  function []=printall2eps()
86  figs = findobj('Type','figure');
87  for l=1:length(figs)
88    %  pos=get(figure(figs(l)),'Position');
89    %  set(gcf,'PaperPosition',pos/100); %no idea why that works... the /100 i mean
90      eval(['print -f', num2str(figs(l)),' -r600 -depsc figure_',num2str(figs(l)),'.eps;'])
91  end
92
93
94
95  % cut piece out of in
96  function [out,lato,lono]=cutpiece(in,lat,lon,s,n,w,e,f,t)
97  % make in from 0:360 lon to -180:180 lon
98  in=[in(37:end,:,:); in(1:36,:,:)];
99  out=in(w:e,s:n,f:t);
100 lato=lat(s:n);
101 lono=lon(w:e);%turn 3D-data flag values into nan
102 function in=flag2nan(in,fl)
103 disp('making flags nan..')
104 [X,Y,Z]=size(in);
105 for x=1:X
106     for y=1:Y
107         for z=1:Z
108             if in(x,y,z)==fl
109                 in(x,y,z)=nan;
110             end
111         end
112     end
113 end
114
115
116 % mean lat/lon gridded data in
117 function [sstm]=lat_weighted_mean(in,lat)
118 [X,Y,T]=size(in);
119 inw=nan(size(in));
120 %% create weight
121 weight=nan(X,Y);
122 for y=1:Y
123     weight(:,y)=cos(deg2rad(lat(y)));
124 end
125 %% "total number" of values
126 N=nansum(weight(:));
127 %% mean
128 for t=1:T
129     % kill land
130     for i=1:X
131         for j=1:Y
132             if isnan(in(i,j,t))
133                 weight(i,j)=nan ;
134             end
135         end
136     end
137     % weighted in
```

```
    inw(:,:,t)=in(:,:,t).*weight;
    % build mean
    sstm(t)=nansum(nansum(inw(:,:,t)))/N;
end
sstm=sstm';% get window on globe


% lat(s,n) and lon(w,e)
function [s,n,w,e,lon]=getspace(lat,lon,dims)
disp('organizing space stuff..')
lon=lon-180;
south=lat(1); north=lat(end); west=lon(1); east=lon(end);
disp(['data spans from ',num2str(south),' south to ',num2str(north),...
    ' north and from ',num2str(west),' west to ',num2str(east),' east ',...
    'at a resolution of ', num2str(length(lat)), ' x ',num2str(length(lon))])
disp('reading space coordinates from input file..')
so=dims(2,1);
no=dims(2,2);
we=dims(3,1);
ea=dims(3,2);
disp('finding best fit..')
[~,s]=min(abs(lat-so));
[~,n]=min(abs(lat-no));
[~,w]=min(abs(lon-we));
[~,e]=min(abs(lon-ea));
%
windowlat=[linspace(so,no,100),no*ones(1,100),linspace(no,so,100),so*ones(1,100)];
windowlon=[we*ones(1,100),linspace(we,ea,100),ea*ones(1,100),linspace(ea,we,100)];
disp('creating snapshot..')
fig=figure('Color','white');
axesm hatano;
coast = load('coast');
axis on; framem on; gridm on; hold on;
plotm(coast.lat,coast.long)
plotm(windowlat,windowlon,'r')
title('close this depiction of the window you have chosen to continue')
tightmap
waitfor(fig)



% build RCs
function [RC]=RCs(PC,RHO,dims)
[pcl,pcn]=size(PC);
M=pcn;
rcn=dims(4,2);
%% build Z
%Z size of PC times number of eigenvectors
Z=zeros(pcl,M,rcn);
%loop over PCs
for z=1:rcn
    for x=1:M
        Z(1+x-1:end,x,z)=PC(1:end-x+1,M-z+1);  %we need the pcn-lon end part of PC
    end
end
%% build RCs
RC=nan(pcl,rcn);
for r=1:rcn
       RC(:,r)=squeeze(Z(:,:,r))*RHO(:,M-r+1)/M;
end% show snapshot of data at random time interval


% time assumed to be third direction
function [fig]=snap(in,lat,lon)
disp('creating snapshot..')
data=in(:,:,round(rand*length(in(1,1,:))));
fig=figure('Color','white');
axesm hatano;
coast = load('coast');
```

```
axis on; framem on; gridm on; hold on;
plotm(coast.lat,coast.long)
pcolorm(lat,lon,data')
title('close this random snapshot of your data to continue')
tightmap
waitfor(fig);% get time info, ask time window



% time(f:t)
function [time,f,t]=gettime(time,ds,dims)
disp('getting time..')
time=time+datenum(ds,'yyyymmdd');
disp(['data spans ',datestr(time(1)),' through ',datestr(time(end))])
disp('reading time coordinates from input file..')
from=datenum(num2str(dims(1,1)),'yyyymm');
till=datenum(num2str(dims(1,2)),'yyyymm');
disp([datestr(from),' till ',datestr(till)])
[¬,f]=min(abs(time-from));
[¬,t]=min(abs(time-till));


% perform fft on monthly data
function [p,freq]=fourierm(sstm,time)
f = sstm;
N = length(f); %% number of points
T = (time(end)-time(1))/(365.25/12); %% define time of interval, days to months
t = [0:N-1]/N; %% define time
t = t*T; %% define time in seconds
p = abs(fft(f))/(N/2); %% absolute value of the fft
p = p(1:N/2).^2; %% take the power of positve freq. half
freq = [0:N/2-1]/T*12; % find the corresponding frequency in Hz


% colors=[red green blue]/2+.5;
function [colors]=rainbow(R, G, B, l, L)
om=2*pi/L;
red=R*sin(l*om-2*pi*(0/3));
green=G*sin(l*om-2*pi*(1/3));
blue=B*sin(l*om-2*pi*(2/3));
colors=[red green blue]/2+.5;


function [Y,C,rho,lambda]=Toep_cov(in,dims)
%lag
M=dims(4,1);
%autocorrelation
C=autocorr(in,M-1);
%toeplitz form
C=toeplitz(C);
%EOFs
[rho,lambda]=eig(C);
lambda=diag(lambda);
%Y
[I,¬]=size(in);
Y=zeros(I,M);
[K,L]=size(Y);
for x=1:L
    Y(1:end-x+1,x)=in(x:end);
end


% mean lat/lon gridded data in
function [sstm]=lat_weighted_mean(in,lat)
[X,Y,T]=size(in);
inw=nan(size(in));
%% create weight
weight=nan(X,Y);
for y=1:Y
    weight(:,y)=cos(deg2rad(lat(y)));
```

```
278 end
279 %% "total number" of values
280 N=nansum(weight(:));
281 %% mean
282 for t=1:T
283     % kill land
284     for i=1:X
285         for j=1:Y
286             if isnan(in(i,j,t))
287                 weight(i,j)=nan ;
288             end
289         end
290     end
291     % weighted in
292     inw(:,:,t)=in(:,:,t).*weight;
293     % build mean
294     sstm(t)=nansum(nansum(inw(:,:,t)))/N;
295 end
296 sstm=sstm';
```