The Determination of a Trend in a Multi-Scale Problem

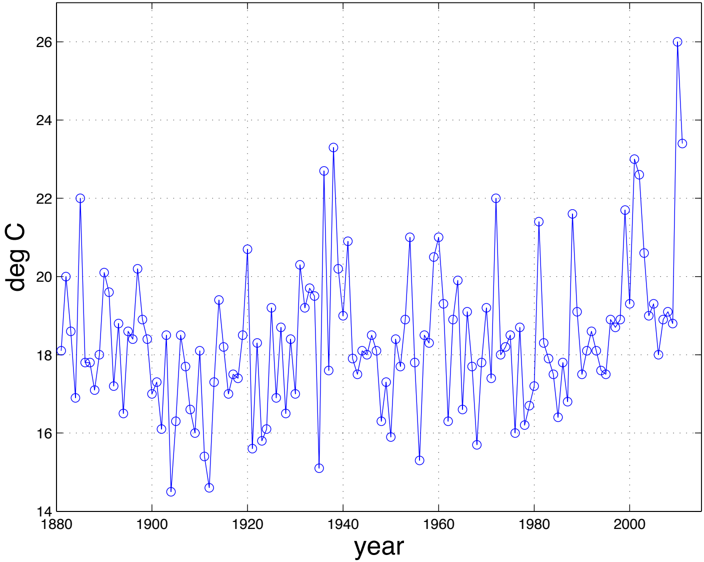
How Warm is it Getting?

JUAN M. RESTREPO

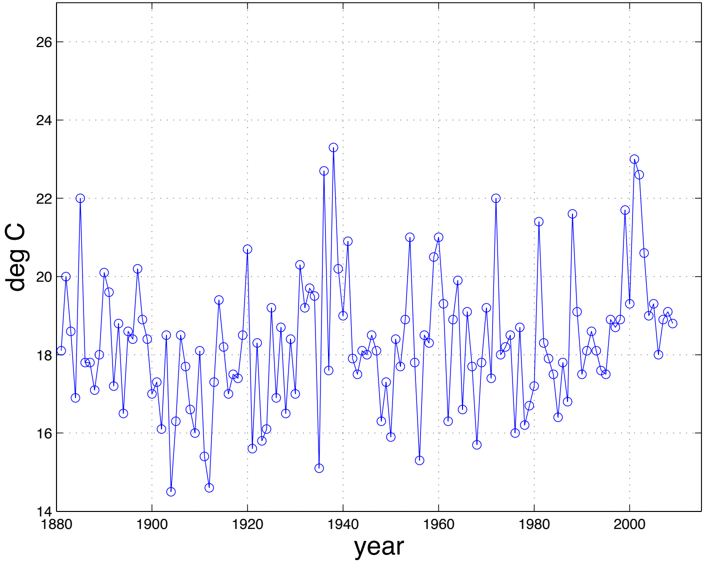
Department of Mathematics, Department of Statistics and Physical Oceanography Oregon State University

May 21, 2018

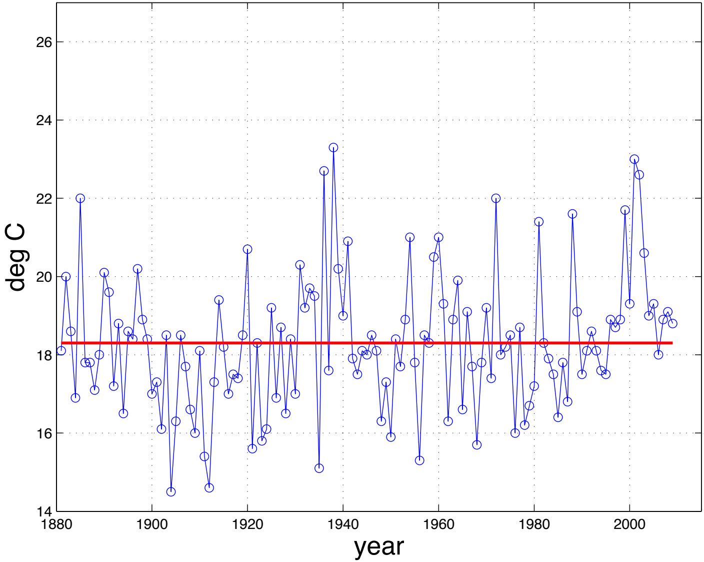
Moscow's Summer Temperatures, 1881-2011



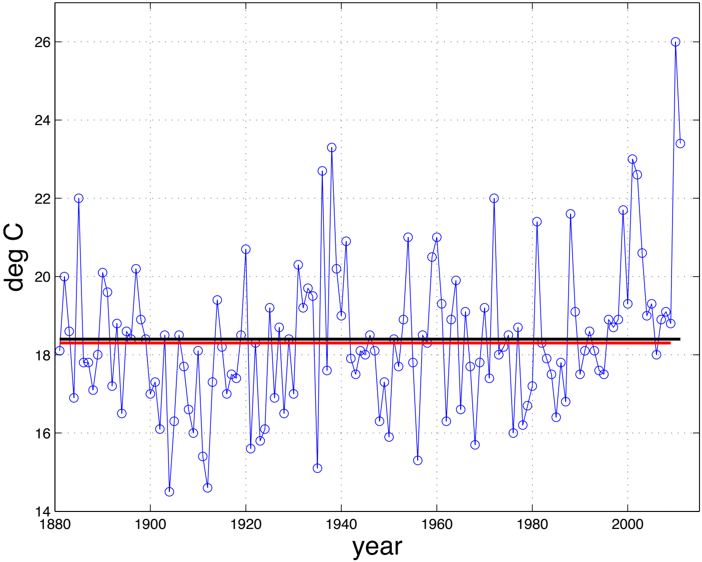
Moscow's Summer Temperatures, 1881-2009



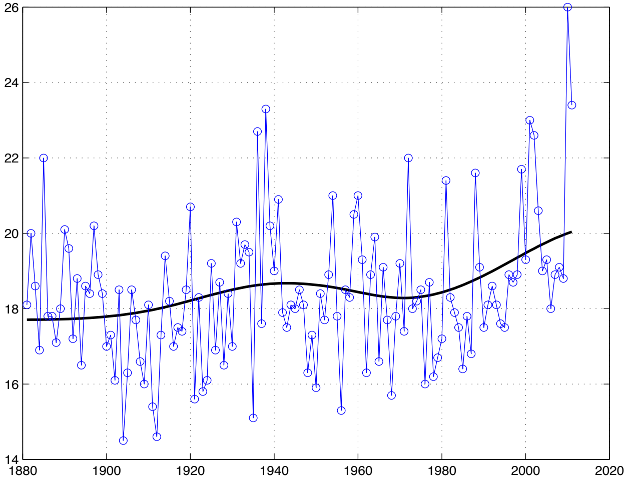
Moscow's Summer Temperatures, 1881-2009



Moscow's Summer Temperatures, 1881-2011



Something Must Account for Changing Mean



Increase of Extreme Events in a Warming World (PNAS **44**, 2011), by Rahmstorf and Coumou.

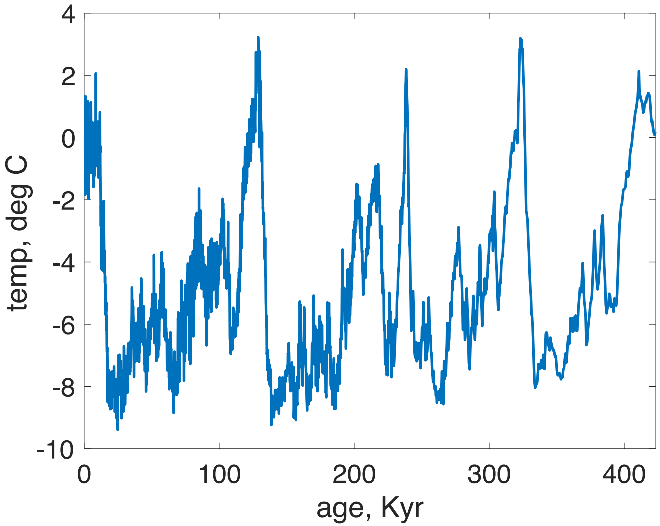
The Trend Problem:

Define a set of simple universal rules with which to compute an underlying *tendency*, given a finite (non-stationary/multi-scale) data set.

Joint work with
Shankar Venkataramani (U. Arizona)
H. Flaschka (U. Arizona) and
D. Comeau (U. Arizona)

GULF OF MEXICO RESEARCH INITIATIVE
NSF National Science Foundation WHERE DISCOVERIES BEGIN

A Climate Signal...



Vostok Ice Core data, Temperature

Given a finite-time time series $Y(i), \quad i = 1, 2, \ldots, N,$

The Tendency $T(i)$ is an *Executive Summary* of $Y(i)$

- Captures essentials of histogram in the abscissa of $Y(i)$; *and*
- Most essential multi scale information, derived from ordinate of $Y(i)$.

The Empirical Uncertainty $U(i) := Y(i) - T(i)$

- is simple entropically,
- The histogram of $U(i)$, is easy to parametrize

Given a finite-time time series $Y(i), \quad i = 1, 2, \ldots, N$,

The Tendency $T(i)$ **is an** ***Executive Summary*** **of** $Y(i)$

- Captures essentials of histogram in the abscissa of $Y(i)$; *and*
- Most essential multi scale information, derived from ordinate of $Y(i)$.

The Empirical Uncertainty $U(i) := Y(i) - T(i)$

- is simple entropically,
- The histogram of $U(i)$, is easy to parametrize

General Procedure:

- Find a decomposition $Y(i) = B^D + \sum_{j=1}^{D} R^j(i)$
- Apply tendency criteria to pick $T(i) := B^D + \{R^j(i)\}_S$, $i = 1, ..., N$.
  - B^D is a constant,
  - $\{R^j(i)\}_S$ is a function made up of a combination of S *rotations*.

The choice of decomposition is motivated by the

- Be non-parametric.
- Ability to handle multi-scale nature of a signal.
- Be lossless.

The Intrinsic Time Decomposition (ITD)

Given a sequence of real numbers $\{Y(i)\}_{i=1}^{N}$,

$$Y(i) = B^D + \sum_{j=1}^{D} R^j(i)$$

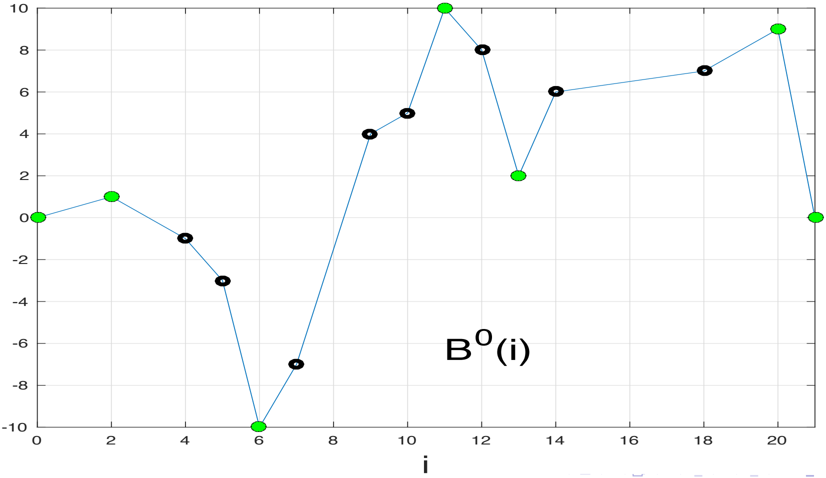
where

$$\begin{aligned} B^j(i) &= B^{j+1}(i) + R^{j+1}(i), \quad j = 0, \ldots, D, \\ &\text{and} \\ B^0(i) : &= Y(i). \end{aligned}$$

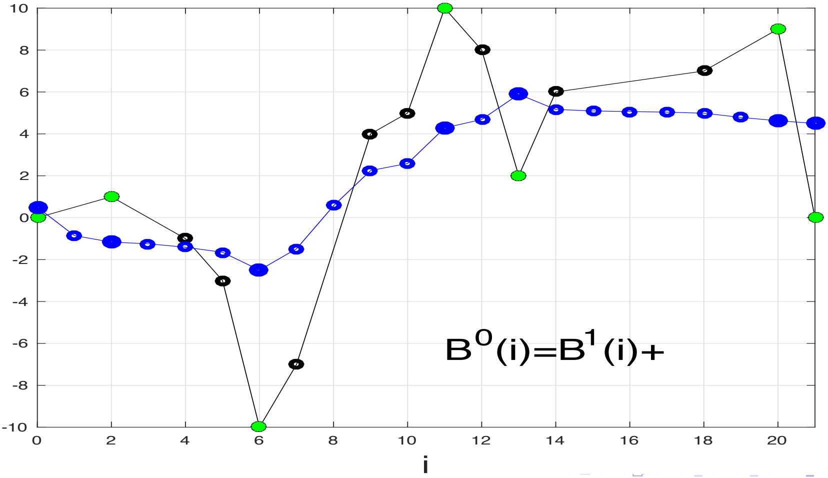
B^j are called *BASELINES*, and R^j are called *ROTATIONS*.

Frei and Osorio, Proc. Roy. Soc. London, (2006).

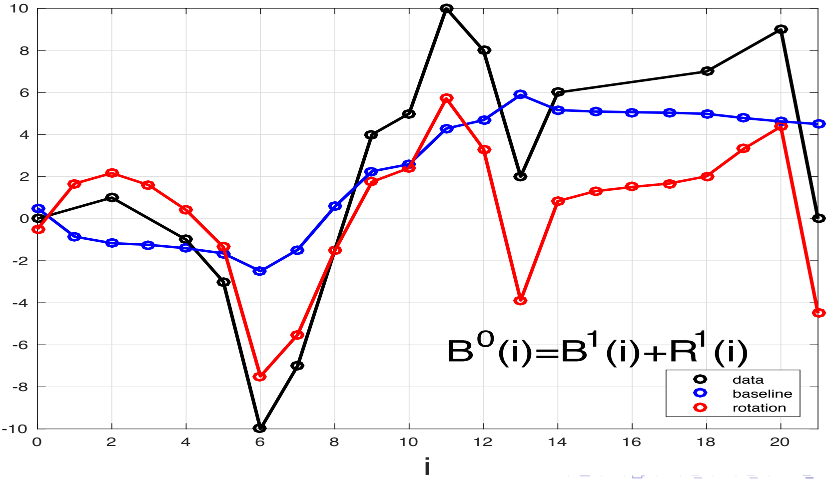
The Intrinsic Time Decomposition (ITD)



The Intrinsic Time Decomposition (ITD)

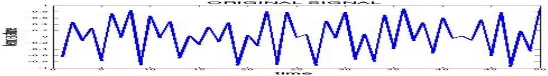


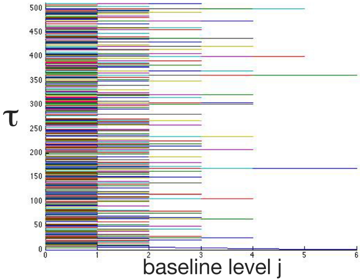
The Intrinsic Time Decomposition (ITD)

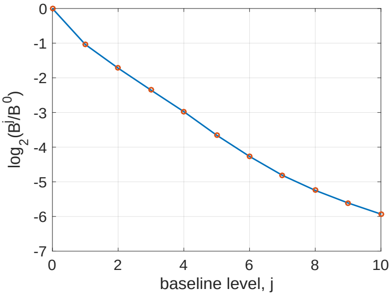


ITD Decomposition: No Low Frequency Slowdown

All Extrema Random Signal $Y = (-1)^i |z_i|, \quad z_i$ from $\mathcal{N}(\sigma = 4)$







Extrema spacing τ^0/τ^j; $\|B^j\|/\|Y\|$

$\tau^j/\tau^{j+1} \approx -0.41$ $\|B^j\|/B^{j+1}\| \approx -0.61.$

How does the ITD (and EMD) work?

$\mathscr{E}[B^j] := \{S^j, b^j\}$.

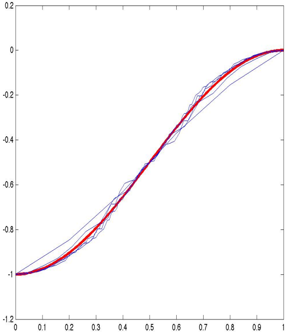
$\{S^j\}_1^{n_j}$ be locations of extrema of baselines, with values b^j.

ITD:
$\{S^{j+1}, b^{j+1}\} = \mathscr{E}[(\mathbb{I} + M^j)b^j]$.

Self Similar Spectrum

ITD: $\{S^{j+1}, b^{j+1}\} = \mathscr{E}[(\mathbb{I} + M^j)b^j]$.

- Spectrum$(\mathbb{I} + M^j)$: $\lambda_k^j = \cos^2(\pi k/n)$,
- e'value 1 corresponding to right e'vector consisting of all 1,
- e'value 0 corresponding to right e'vector consisting of $x_k = (-1)^k$.



Spectrum$(\mathbb{I} + M^j)$, for all levels j

Universality and Decay of the ℓ_2 of the Baselines

$$\{S^{j+1}, b^{j+1}\} = \mathscr{E}[(I+M^j)b^j]$$

at $j+1$, with $S^{j+1} \subseteq S^j$.
$v := M^j b \sim \frac{1}{4}B''(x) + \frac{1}{2}p^j(x)B'(x)$.

$$v_k = \frac{1}{4}(b_{k-1} - 2b_k + b_{k+1}) + \frac{p_k^j}{4}(b_{k+1} - b_{k-1}),$$

$p_k^j \in (-1,1)$ given by

$$p_k^j = \frac{2\tau_k^j - \tau_{k-1}^j - \tau_{k+1}^j}{\tau_{k+1}^j - \tau_{k-1}^j}.$$

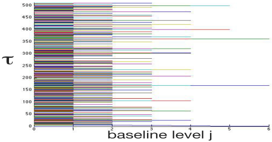
Hence, $(I+M^j)b^j$ is FTCD approximation of

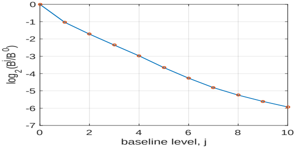
$$\frac{\partial}{\partial t}B = \frac{1}{4}\frac{\partial^2}{\partial x^2}B + \frac{1}{2}p^j(x)\frac{\partial}{\partial x}B = \frac{1}{4w^j(x)}\frac{\partial}{\partial x}\left[w^j(x)\frac{\partial B}{\partial x}\right],$$

$w^j(x) = \exp\left[2\int_0^x p^j(t)dt\right]$.

Distancing of Extremas and Decay of Baselines, with Level j

All Extrema Random Signal $Y = (-1)^i |z_i|, \quad z_i$ from $\mathcal{N}(\sigma = 4)$





Extrema spacing τ^0/τ^j; $\quad\quad \|B^j\|/\|Y\|$

Model: $b^j_k \approx \mu^j (-1)^k + \alpha^j n_k$.

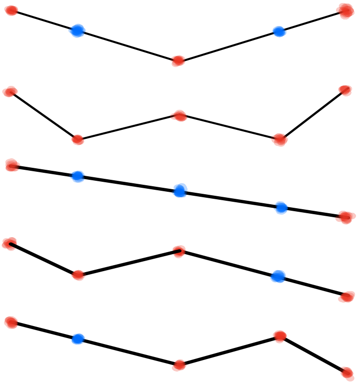
Since $(I + M^j)(-1)^k = 0$ then $b^{j+1} = \alpha^j \mathscr{E}((I + M^j)\mathbf{n})$, where $\mathbf{n} = n_k$ is a vector of independent normal variates.

Estimating μ and α

Model: $b_k^j \approx \mu^j(-1)^k + \alpha^j n_k$.

Let $\mathbf{x} = (x_1, x_2, x_3)$, consecutive entries of the vector $(I + M^j)\mathbf{n}$, then

count ways x_i can be extrema •



Let x_1, x_2 and x_3 consecutive entries of the vector $(I + M^j)\mathbf{n}$, then

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \approx \frac{1}{4} \begin{pmatrix} 1-p_1 & 2 & 1+p_1 & 0 & 0 \\ 0 & 1-p_2 & 2 & 1+p_2 & 0 \\ 0 & 0 & 1-p_3 & 2 & 1+p_3 \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{pmatrix} \equiv A\mathbf{n}.$$

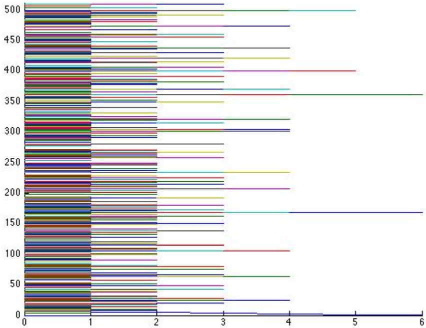
where n_i independent normal, p_i are realizations. So x_1, x_2 and x_3 are jointly Gaussian with mean zero and covariance $\Sigma(p_1, p_2, p_3) = AA^T$ and the (conditional) joint density of x_1, x_2 and x_3 is given by

$$p(x_1, x_2, x_3 | p_1, p_2, p_3) = \frac{1}{\sqrt{8\pi^3 \mathrm{Det}(\Sigma(p_1, p_2, p_3))}} \exp\left[-\frac{1}{2} \mathbf{x}^T \Sigma(p_1, p_2, p_3)^{-1} \mathbf{x} \right].$$

The Joint Density of the Spacing

Let $l_k = \tau_{k+1} - \tau_k$, then $p_k = \frac{l_k - l_{k-1}}{l_k + l_{k-1}}$. Assume every site has a constant probability of being an extremum, and no correlation between neighbors.

Guess l^j is exponentially distributed (mean = 1)



The joint density of p_1, p_2 and p_3:

$$\rho(p_1,p_2,p_3) = \frac{128(1-p_1)(1+p_2)(1-p_2)(1+p_3)}{(3-p_1+p_2+p_1p_2)^3(3-p_2+p_3+p_2p_3)^3}.$$

and

$$\begin{aligned} \Sigma_{av} &= \int_{-1}^{1} dp_2 \int_{-1}^{1} dp_1 \int_{-1}^{1} dp_3 \Sigma(p_1,p_2,p_3)\rho(p_1,p_2,p_3) \\ &= \begin{pmatrix} 0.41666\ldots & 0.25 & 0.058227\ldots \\ 0.25 & 0.41666\ldots & 0.25 \\ 0.058227\ldots & 0.25 & 0.41666\ldots \end{pmatrix}. \end{aligned}$$

We now obtain the joint density of x_1, x_2 and x_3 to get

$$p_{av}(x_1,x_2,x_3) \approx \frac{1}{\sqrt{8\pi^3 \mathrm{Det}(\Sigma_{av})}} \exp\left[-\frac{1}{2}\mathbf{x}^T\Sigma_{av}^{-1}\mathbf{x}\right].$$

We can also compute the mean and the variance of the distribution of the minima (maxima).

$$\mu = \frac{1}{\beta}\int_{-\infty}^{\infty} dx_3 \int_{x_3}^{\infty} dx_2 \int_{-\infty}^{x_2} dx_1\, x_2\, p(x_1, x_2, x_3) = 0.483883\ldots$$

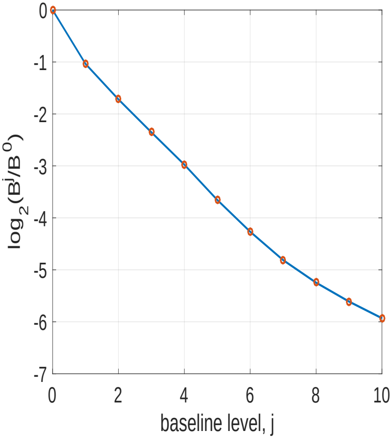
and

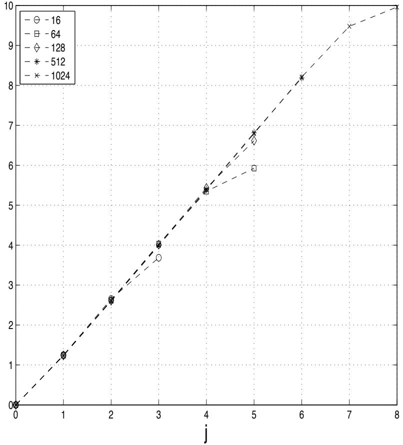
$$\alpha^2 = \frac{1}{\beta}\int_{-\infty}^{\infty} dx_3 \int_{x_3}^{\infty} dx_2 \int_{-\infty}^{x_2} dx_1\, x_2^2\, p(x_1, x_2, x_3) - \mu^2 = 0.302712\ldots$$

From this, we obtain

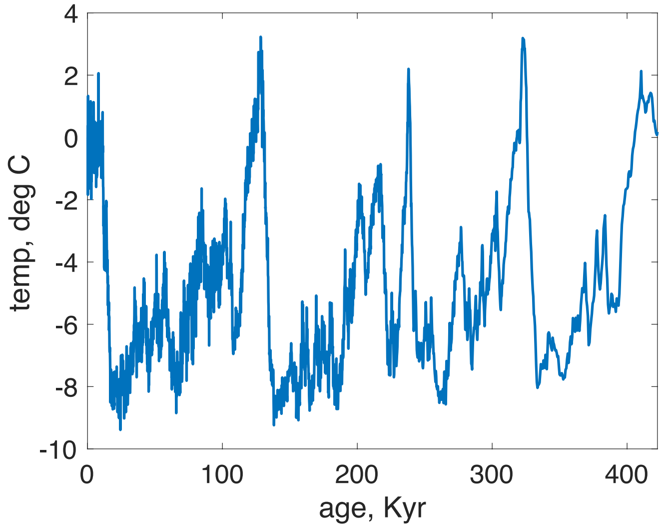
$$(I + M^j)\mathbf{n} \approx \mu(-1)^k + \alpha\mathbf{n}' = 0.483883 \times (-1)^k + 0.550193\mathbf{n}'.$$

	B^j/B^0 DECAY RATE	τ^0/τ^j DECAY RATE
THEORETICAL	-0.55	-0.48
NUMERICAL	-0.61	-0.41



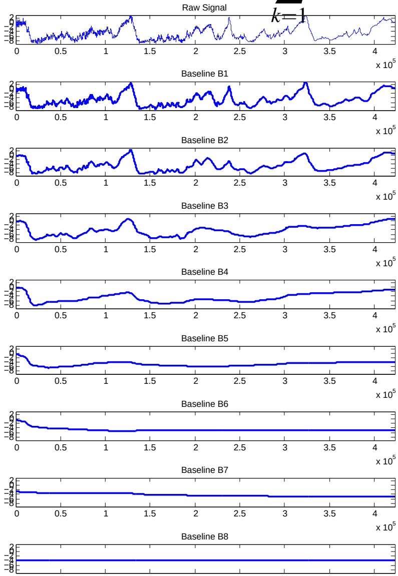


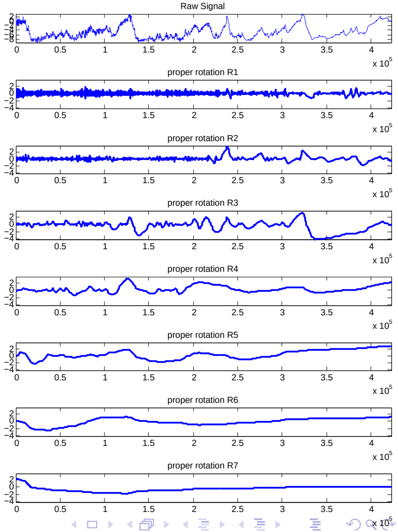
Example Calculation



Vostok Ice Core data, Temperature

$$Y(i) = B^D + \sum_{k=1}^{D} R^k(i), \qquad B^{j+1}(t) + R^{j+1}(i) = B^j(i).$$





Finding the Tendency

- Find ITD:

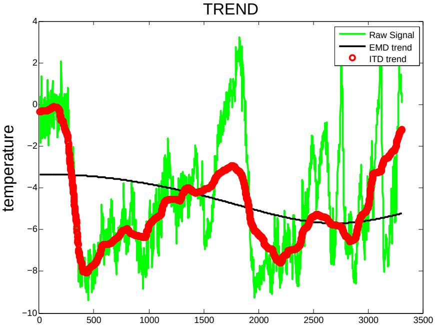
$$Y(i) = B^D + \sum_{j=1}^{D} R^j(i),$$

$$B^j(i) = B^{j+1}(i) + R^{j+1}(i)$$

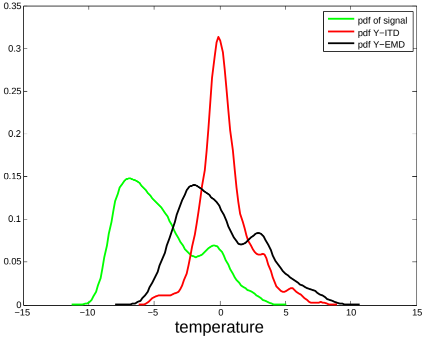
- Find Tendency (picking j^* baseline)

$$T(i) := B^{j^*}(i)$$

The Tendency $T(i)$, the EMD, and the Vostok signal $Y(i)$



Time Series



The Histograms

Find Tendency

Choosing j^* among the baselines $\{B^j(i)\}_{j=1}^D$:

$$T(i) := B^{j^*}(i)$$

The *ABSISSA* information:

- For $j = 1, .., D$ compute $F^j :=$ histogram$[Y(i) - B^j(i)]$
- Determine the Symmetry s^j of F^j via percentiles:

$$s^j := \frac{Pr^j_{75} - 2\,Pr^j_{50} - Pr^j_{25}}{(Pr^j_{75} - Pr^j_{25})}$$

Choosing j^* among the baseline $\{B^j(i)\}_{j=1}^{D}$:

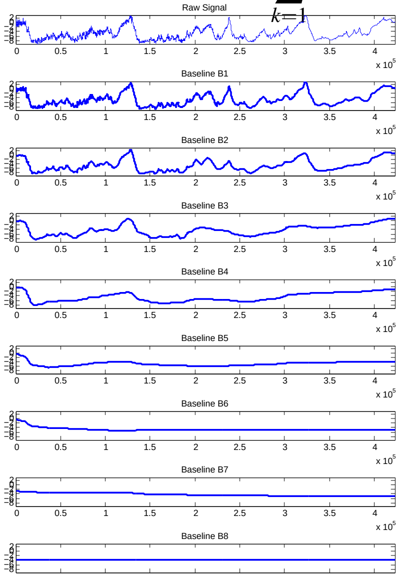
$$T(i) := B^{j^*}(i)$$

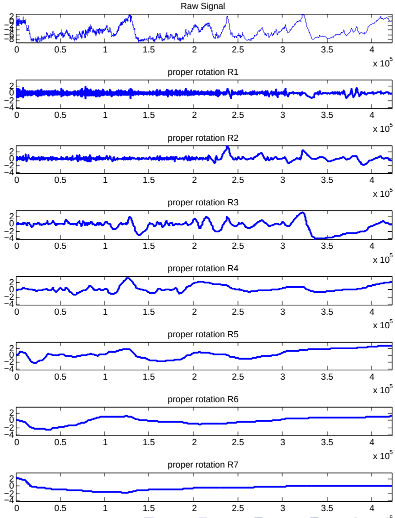
The *ORDINATE* information:

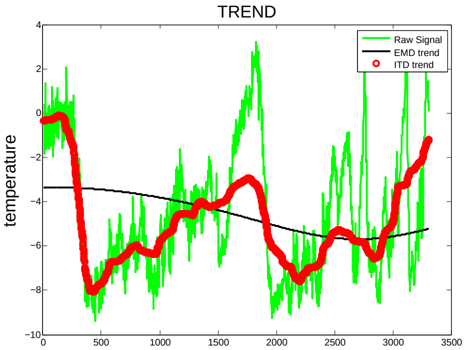
- Compute the Complexity c_j vector

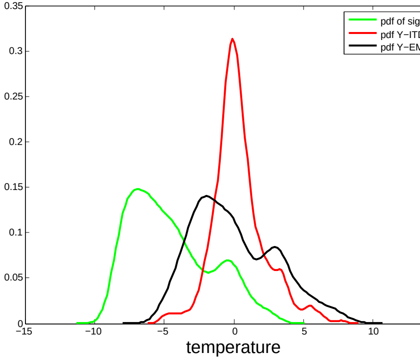
$$c^j := \mathrm{corr}(B^j, R^j)$$

$$Y(i) = B^D + \sum_{k=1}^{D} R^k(i), \qquad B^{j+1}(t) + R^{j+1}(i) = B^j(i).$$







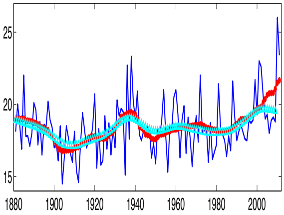


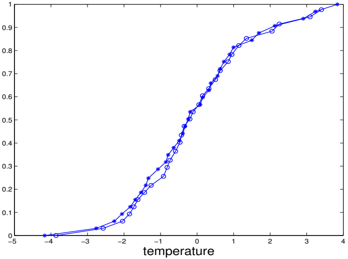
Time Series

The Histograms

Analysis of the Moscow Data

Our analysis confirms Rahmstorf and Coumou's guess: the mean temperature increased, but not its variance:





Further Information

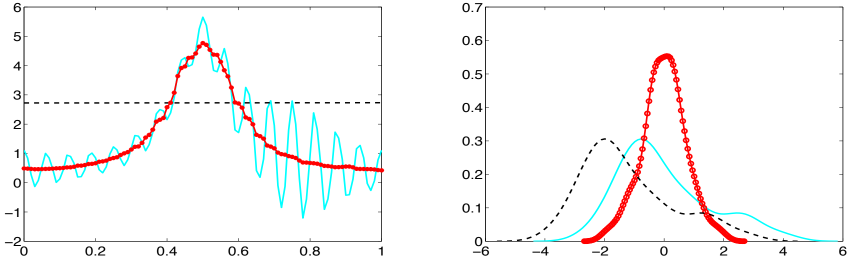
Juan M. Restrepo

http://www.math.oregonstate.edu/∼restrepo

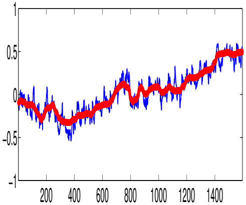
A Multiscale Signal

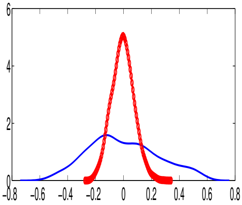
$$Y(i) = \frac{1}{1.5+\sin(2\pi t)}\cos[32\pi t_i + 0.2\cos(64\pi t_i)] + \frac{1}{(1.2+\cos(2\pi t_i))};$$

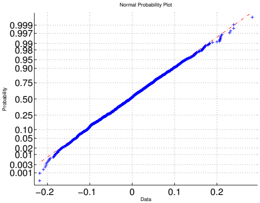
for $t_i \in [0 : 0.0025 : 1]$.

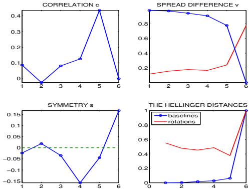


Ocean Temperatures









The Composite Case