Computational Time Series Analysis (IPAM, Part II, nonstationary methods)

Illia Horenko

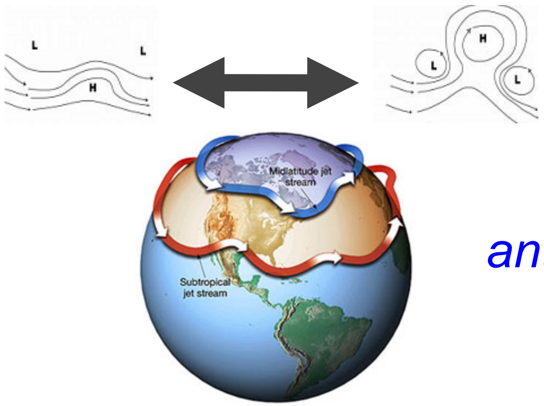
Research Group "*Computational Time Series Analysis* "
Institute of Mathematics
Freie Universität Berlin
Germany

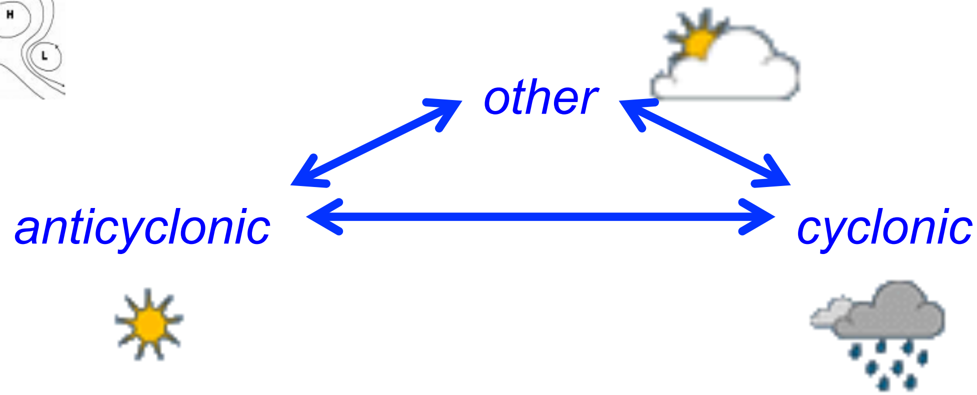
and

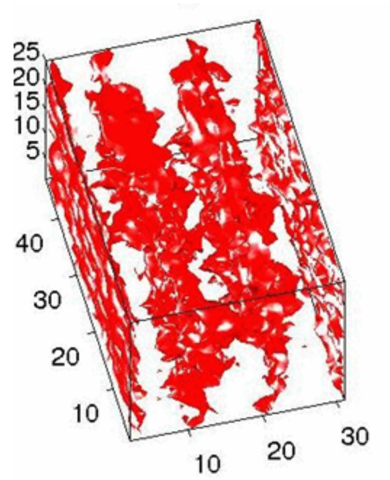
Institute of Computational Science
Universita della Swizzera Italiana, Lugano
Switzerland

Motivation (I): Data

Inverse problem: data-based identification of model parameters







time series of vectors (continous, finite dim.)

time series of regimes (discrete, finite dim.)

time series of functions (continuous, infin. dim.)

Challenges:

VERY LARGE (~Tb size) , non-Markovian (memory), non-stationary (trends), multidimensional, ext. influence

Motivation (II): Models

Inverse problem: data-based identification of model parameters

Dynamical
(f.e., deterministic, stochastic, ...)

Geometrical
(f.e, essential manifolds, ...)

Models

Mixed
(f.e., MTV, PIP/POP, ...)

Challenges:

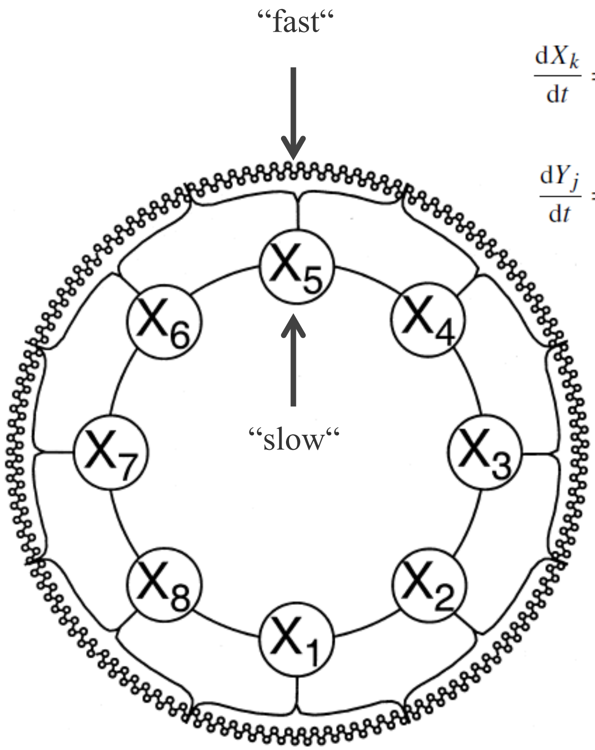
existence and uniqueness, robustness, conditioning

Yesterday (stationary case)

- *“eagle eye“ perspective on **stochastic processes** from the viewpoint of **deterministic dynamical systems***
- *geometric model inference**: EOF/SSA***
- *multivariate dynamical model inference: **VARX***
- *handling the ill-posed problem*
- *motivation for tomorrow : examples where stationarity assumption does not work*

Lorenz96: “order zero“ atmosph. model

Wilks, *Quart.J.Royal Met.Soc., 2005*



$$\frac{\mathrm{d}X_k}{\mathrm{d}t} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b} \sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K$$

$$\frac{\mathrm{d}Y_j}{\mathrm{d}t} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b} X_{\mathrm{int}[(j-1)/J]+1}; \quad j = 1, \ldots, JK.$$

stationary model
(time-independent parameters)

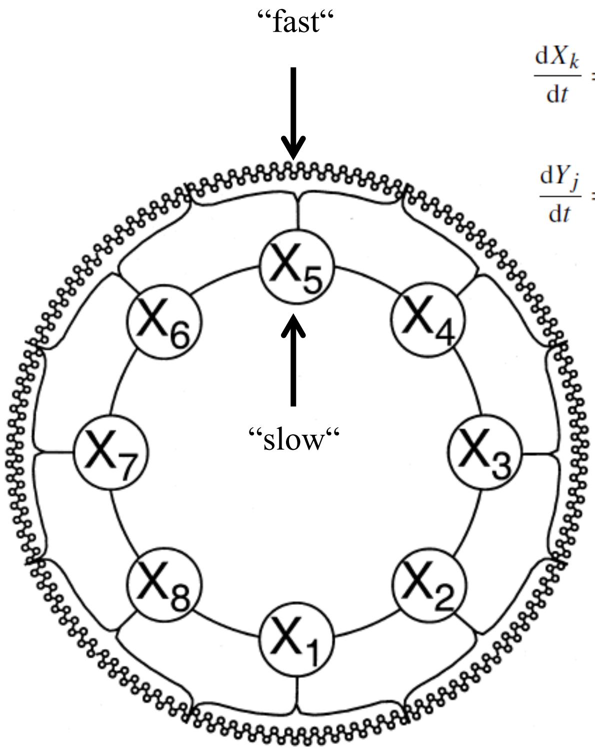
VARX (standard stationary stochastic Model)

$$y_t = \mu + \sum_{q=1}^{m} \mathbf{A}_q y_{t-q\tau} + \mathbf{B}\phi(x_t) + \mathbf{C}\epsilon_t$$

- •Lorenz, *Proc. Of. Sem. On Predict., 1996*
- •Majda/Timoffev/V.-Eijnden, *PNAS, 1999*
- •Orell, *JAS, 2003*
- •Wilks, Quart.J.Royal Met.Soc., 2005
- •Crommelin/V.-Eijnden, *Journal of Atmos. Sci., 2008*

Lorenz96: “order zero“ atmosph. model

Wilks, *Quart.J.Royal Met.Soc., 2005*



$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K$$

$$\frac{dY_j}{dt} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b}X_{\mathrm{int}[(j-1)/J]+1};$$ **+ *F(t)***

+   **=** **non-stationary model (time-dependent parameters)**

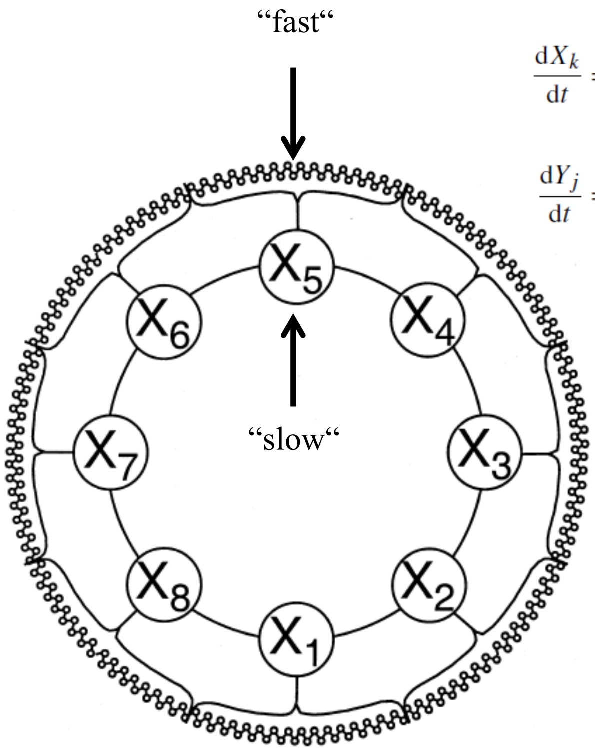
???

VARX (standard stationary stochastic Model)

$$y_t = \mu + \sum_{q=1}^{m} \mathbf{A}_q y_{t-q\tau} + \mathbf{B}\phi(x_t) + \mathbf{C}\epsilon_t$$

•Lorenz, *Proc. Of. Sem. On Predict., 1996*
•Majda/Timoffev/V.-Eijnden, *PNAS, 1999*
•Orell, *JAS, 2003*
•Wilks, Quart.J.Royal Met.Soc., 2005
•Crommelin/V.-Eijnden, *Journal of Atmos. Sci., 2008*

Lorenz96: “order zero“ atmosph. model



Wilks, *Quart.J.Royal Met.Soc., 2005*

$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b} \sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K$$

$$\frac{dY_j}{dt} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b} X_{\mathrm{int}[(j-1)/J]+1};$$

+ *F(t)*

+   = **non-stationary model (time-dependent parameters)**

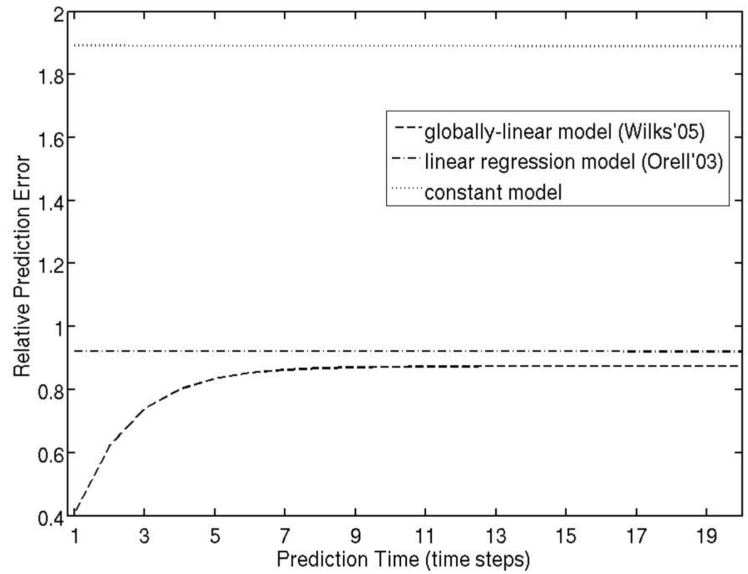
???

~~VARX (standard stationary stochastic Model)~~

$$y_t = \mu + \sum_{q=1}^{m} A_q y_{t-q\tau} + B\phi(x_t) + C\epsilon_t$$

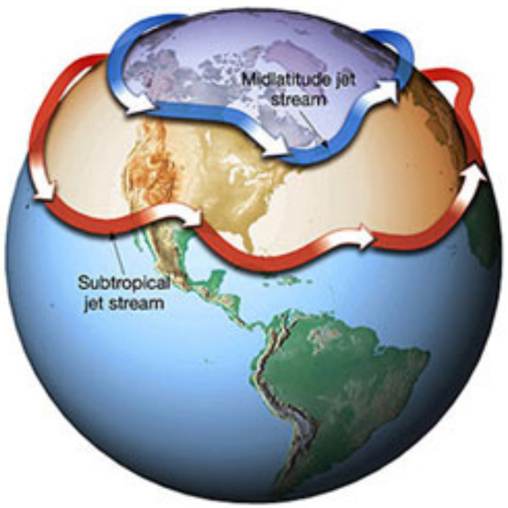
- Lorenz, *Proc. Of. Sem. On Predict., 1996*
- Orell, *JAS, 2003*
- Wilks, Quart.J.Royal Met.Soc., 2005
- Crommelin/V.-Eijnden, *Journal of Atmos. Sci., 2008*

Predictors of Lorenz ‘96-Model

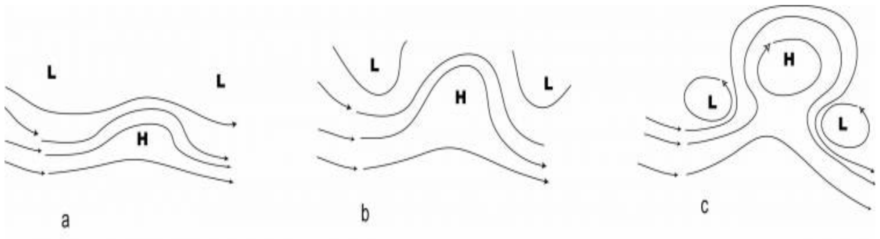


Analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

Wind Jets transport moisture from US to Europe

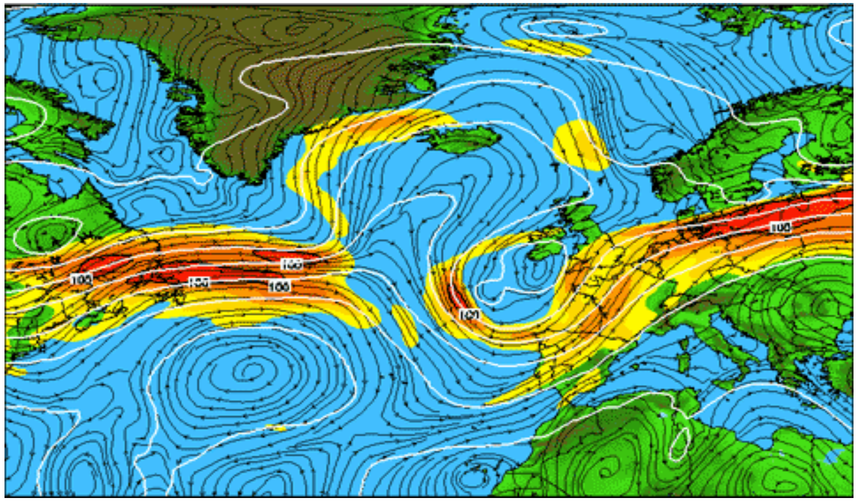


Jet-Blocking



VARX (standard stationary stochastic Model)

$$y_t \quad = \quad \mu + \sum_{q=1}^{m} \mathrm{A}_q y_{t-q\tau} + \mathrm{B}\phi\left(x_t\right) + \mathrm{C}\epsilon_t$$

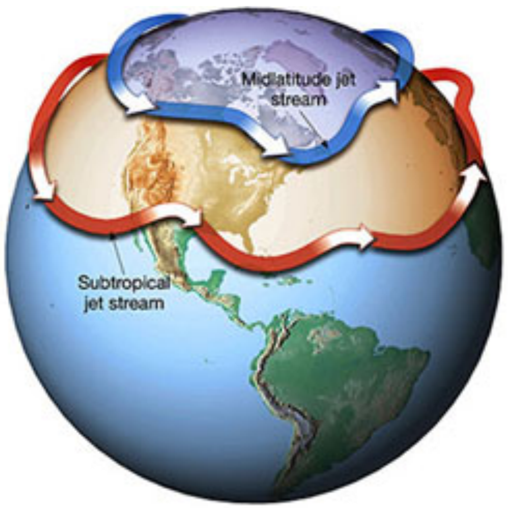


Jet-Blocking results in “indian summer“ and “blackberry cold“

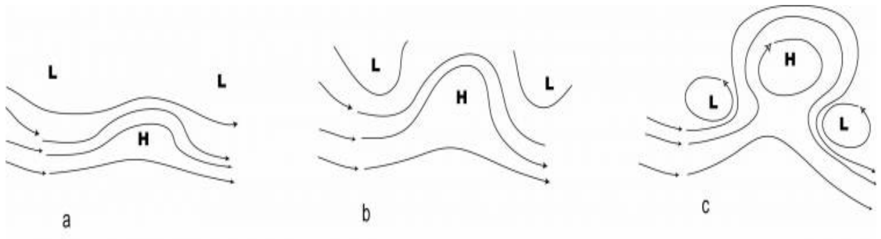
Analysis of ERA40 geopotential time series (PIK, Potsdam)

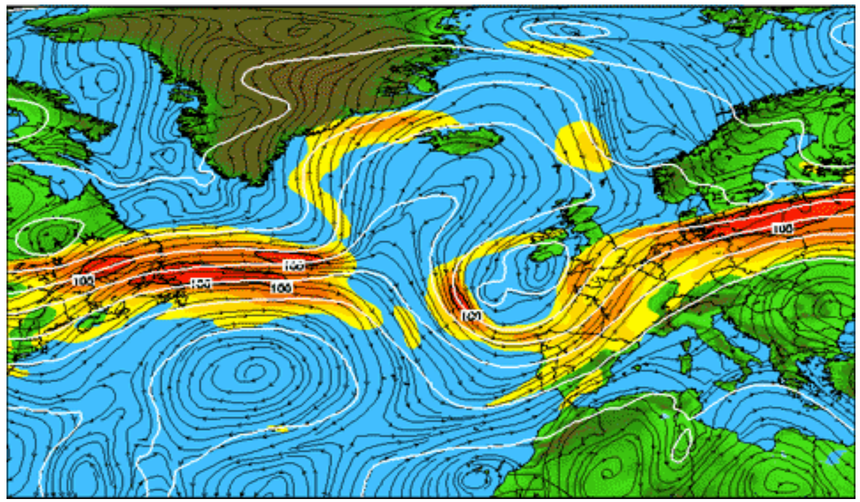
Analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

Wind Jets transport moisture from US to Europe



Jet-Blocking





VARX (standard stationary stochastic Model)

$$y_t = \mu + \sum_{q=1}^{m} \mathbf{A}_q y_{t-q\tau} + \mathbf{B}\phi(x_t) + \mathbf{C}\epsilon_t$$

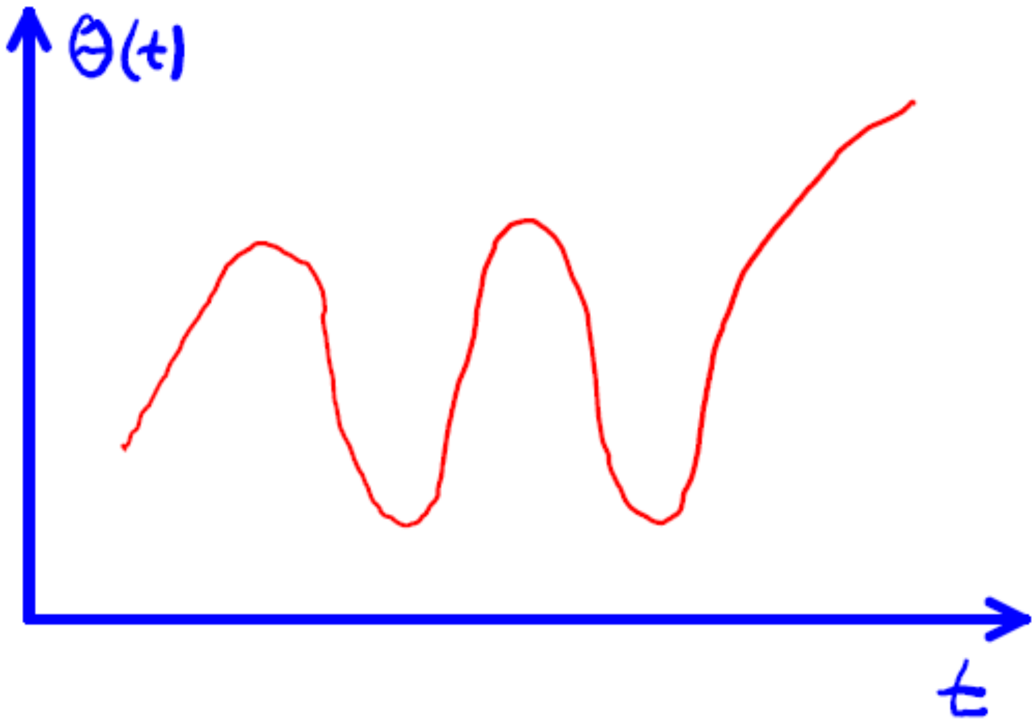
Standard statistics: seasonal factor, greenhouse gases and sun activity are all stat. insignificant!

H. 10, to appear in Journal of Atm. Sci.

Non-stationary Processes

Direct problem: $\mathbf{F}\left(x_t, \ldots, x_{t-m\tau}, \theta(t), t\right) = 0$

Inverse problem: $\int_0^T g(x_s, \theta(s))\,ds \to \min_{\theta(t)}$



<u>Example:</u>

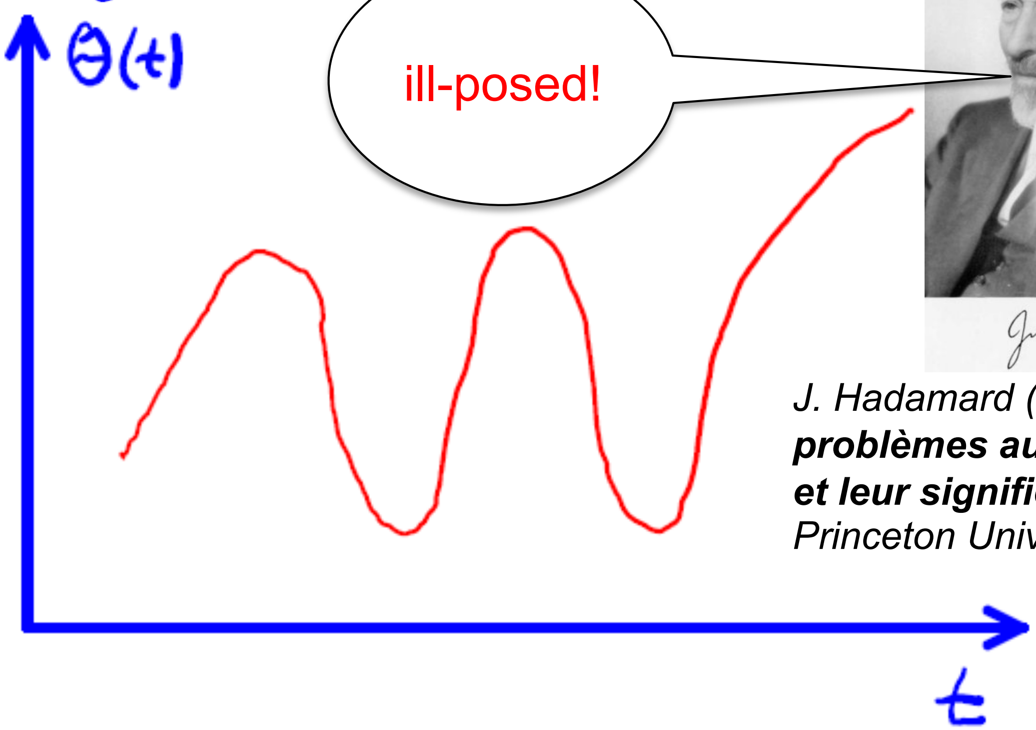
$$x_t = \theta(t) + \epsilon_t \qquad \epsilon_t \text{ (i.i.d.) } \mathbb{E}\left[\epsilon_t\right] = 0$$

$$g\left(x_t, \theta(t)\right) = \|x_t - \theta_t\|_2$$

Non-stationary Processes

Direct problem: $\mathbf{F}\left(x_t, \ldots, x_{t-m\tau}, \theta(t), t\right) = 0$

Inverse problem: $\int_0^T g(x_s, \theta(s))\,ds \rightarrow \min_{\theta(t)}$



J. Hadamard (1902): ***Sur les problèmes aux dérivées partielles et leur signification physique.*** *Princeton University Bulletin, 49--52.*

Ill-posed problem

Example:

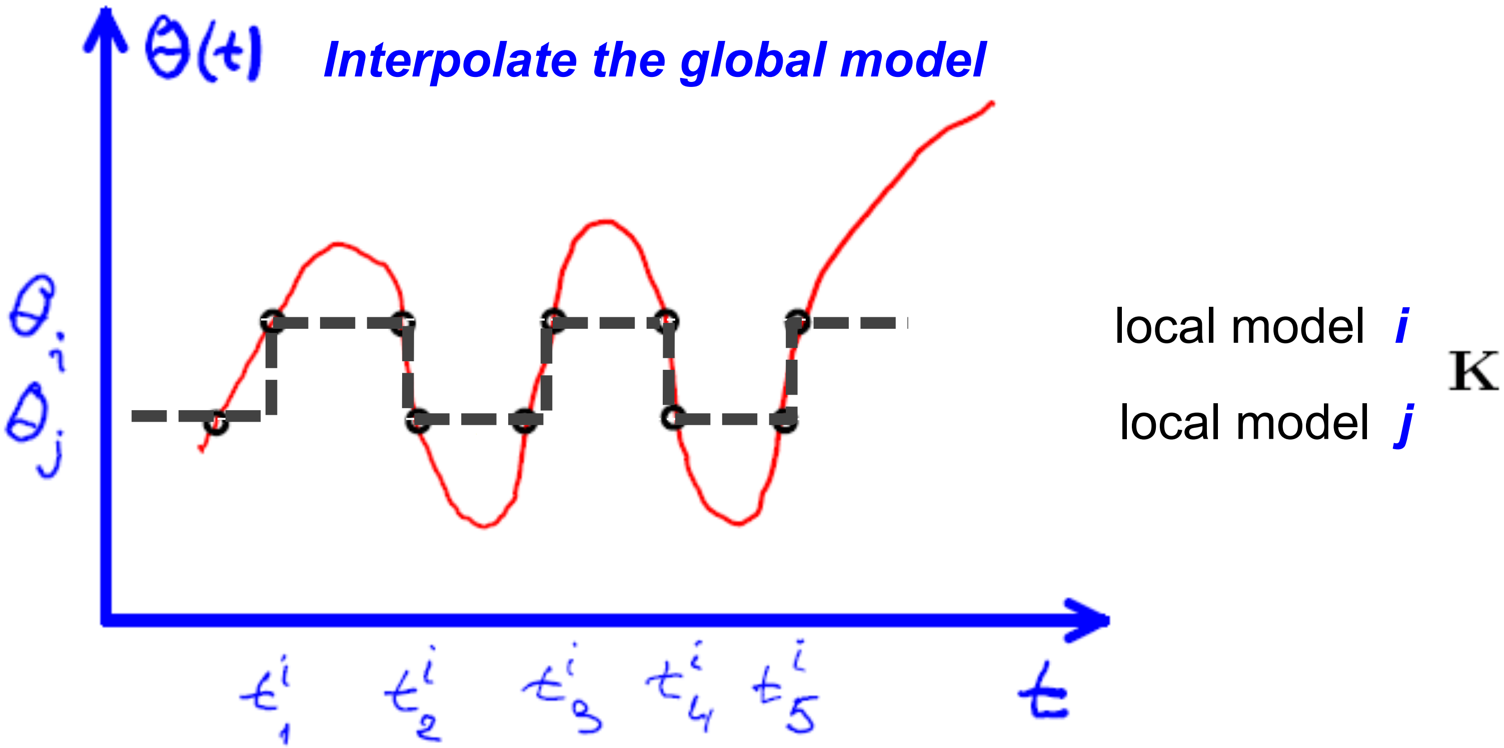
$$x_t = \theta(t) + \epsilon_t \qquad \epsilon_t \text{ (i.i.d.) } \mathbb{E}\left[\epsilon_t\right] = 0$$

$$g\left(x_t, \theta(t)\right) = \|x_t - \theta_t\|_2$$

Non-stationary Processes

$$g(x_t, \Theta(t)) = \sum_{i=1}^{K} \gamma_i(t) g(x_t, \theta_i)$$

Local quality functional $g(x, \theta_i) \quad : \quad \Psi \times \Omega \rightarrow [0, \bar{g}], \quad 0 < \bar{g} < +\infty,$



Averaged Clustering Functional

Find $\Gamma(t) = (\gamma_1(t), \ldots, \gamma_{\mathbf{K}}(t))$ such that for each t :

$$\mathbf{L}(\Theta, \Gamma(t)) \quad = \quad \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t, \theta_i\right) \longrightarrow \min_{\Gamma(t), \Theta},$$

subjected to constraints:

$$\begin{aligned} \sum_{i=1}^{\mathbf{K}} \gamma_i(t) &= 1, \quad \forall t \in [0, T] \\ \gamma_i(t) &\geq 0, \quad \forall t \in [0, T], i = 1, \ldots, \mathbf{K} \end{aligned}$$

Ill-posed problem:
Regularization !

(H. 09, SIAM J. of Sci. Comp.)

Incorporation of Temporal Information

Regularization : $\Theta(t)$ *should be much "slower" than* x_t

$$\mathbf{L}(\Theta,\Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \longrightarrow \min_{\Gamma(t),\Theta},$$

$$\|\gamma_i(\cdot)\|_? \leq \|x_t\|_?$$

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) = 1, \quad \forall t \in [0,T]$$

$$\gamma_i(t) \geq 0, \quad \forall t \in [0,T], i = 1,\dots,\mathbf{K}$$

(H. 09, SIAM J. of Sci. Comp.)

Incorporation of Temporal Information

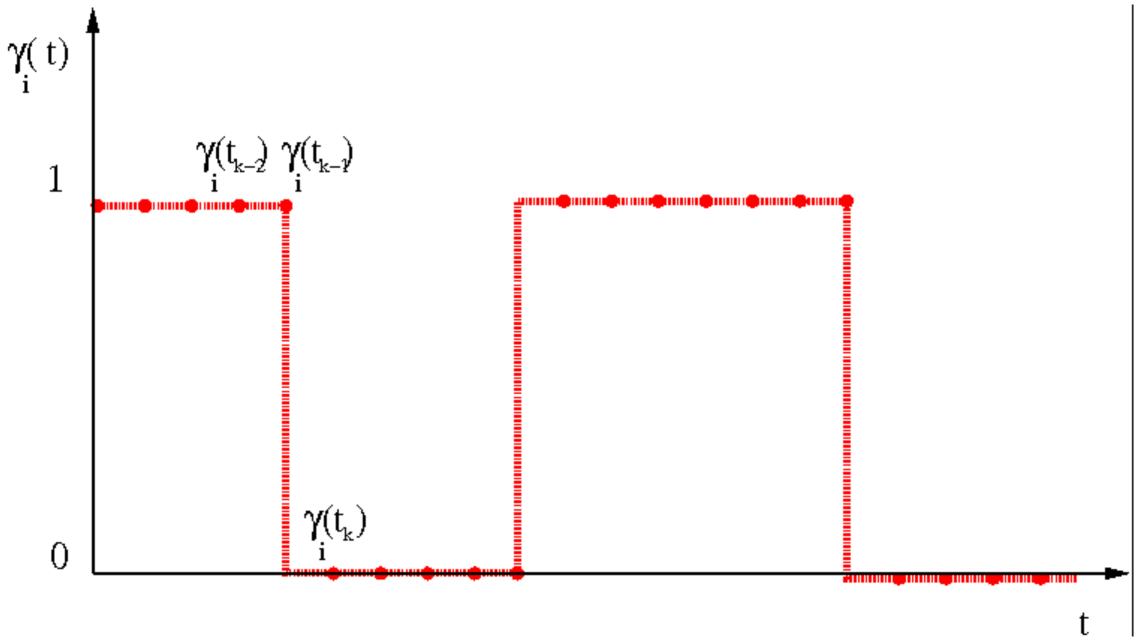
Regularization : $\Theta(t)$ *should be much "slower" than* X_t

$$\mathbf{L}(\Theta,\Gamma(t)) \; = \; \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \rightarrow \min_{\Gamma(t),\Theta},$$

$$\|\gamma_i(\cdot)\|_2 \leq \|X_t\|_2$$

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) \; = \; 1, \quad \forall t \in [0,T]$$

$$\gamma_i(t) \; \geq \; 0, \quad \forall t \in [0,T], i=1,\ldots,\mathbf{K}$$



Number Of Jumps / Time Interval

$$\|\cdot\|_{\mathcal{H}_1} \; : \; \sum_k \frac{\left(\gamma_i(t_{k+1})-\gamma_i(t_k)\right)^2}{\Delta t} \leq C_i$$

$$\|\cdot\|_{BV} \; : \; \sum_k \frac{|\gamma_i(t_{k+1})-\gamma_i(t_k)|}{\Delta t} \leq C_i$$

H. 09, SIAM J. of Sci. Comp.
H. 09, accepted for publ. In JAS

Incorporation of Temporal Information

Regularization : $\Theta(t)$ *should be much "slower" than* x_t

$$\mathbf{L}(\Theta,\Gamma(t)) \;=\; \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \longrightarrow \min_{\Gamma(t),\Theta},$$

$$\begin{aligned} \sum_{i=1}^{\mathbf{K}} \gamma_i(t) &= 1, \quad \forall t \in [0,T] \\ \gamma_i(t) &\geq 0, \quad \forall t \in [0,T], i = 1,\dots,\mathbf{K} \end{aligned}$$

$$|\gamma_i|_{\mathcal{H}^1(0,T)} = \| \partial_t \gamma_i(\cdot) \|_{\mathcal{L}_2(0,T)} = \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \leq C_\epsilon^i < +\infty,$$

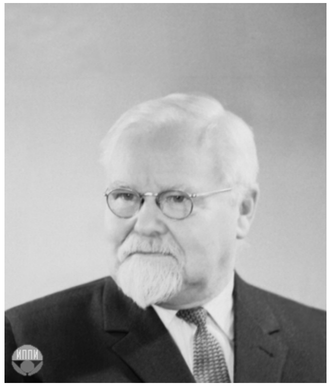
Incorporation of Temporal Information

Regularization : $\Theta(t)$ *should be much "*slower*" than* x_t

$$\mathbf{L}(\Theta,\Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \rightarrow \min_{\Gamma(t),\Theta},$$

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) = 1, \quad \forall t \in [0,T]$$
$$\gamma_i(t) \geq 0, \quad \forall t \in [0,T], i = 1,\dots,\mathbf{K}$$

$$|\gamma_i|_{\mathcal{H}^1(0,T)} = \| \partial_t \gamma_i(\cdot) \|_{\mathcal{L}_2(0,T)} = \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \leq C_\epsilon^i < +\infty,$$



A. N. Tikhonov
(http://en.wikipedia.org)

Incorporation of Temporal Information

Regularization : $\Theta(t)$ *should be much "slower" than* x_t

$$\mathbf{L}(\Theta,\Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \longrightarrow \min_{\Gamma(t),\Theta},$$

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) = 1, \quad \forall t \in [0,T]$$
$$\gamma_i(t) \geq 0, \quad \forall t \in [0,T], i = 1,\dots,\mathbf{K}$$

$$|\gamma_i|_{\mathcal{H}^1(0,T)} = \| \partial_t \gamma_i(\cdot) \|_{\mathcal{L}_2(0,T)} = \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \leq C_\epsilon^i < +\infty,$$

Regularized clustering functional:

$$\mathbf{L}^\epsilon(\Theta,\Gamma(t),\epsilon^2) = \mathbf{L}(\Theta,\Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \longrightarrow \min_{\Gamma(t),\Theta}$$

(H. 09, SIAM J. of Sci. Comp.)

FEM: Regularized Clustering Functional

Regularized clustering functional:

$$\mathbf{L}^{\epsilon}(\Theta, \Gamma(t), \epsilon^2) \quad = \quad \mathbf{L}(\Theta, \Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T \left(\partial_t \gamma_i\,(t)\right)^2 dt \longrightarrow \min_{\Gamma(t),\Theta}$$

(H. 09, SIAM J. of Sci. Comp.)

$$\begin{aligned} \sum_{i=1}^{\mathbf{K}} \gamma_i(t) &= 1, \quad \forall t \in [0,T] \\ \gamma_i(t) &\geq 0, \quad \forall t \in [0,T], i = 1,\ldots,\mathbf{K} \end{aligned}$$

FEM: Regularized Clustering Functional

Regularized clustering functional:

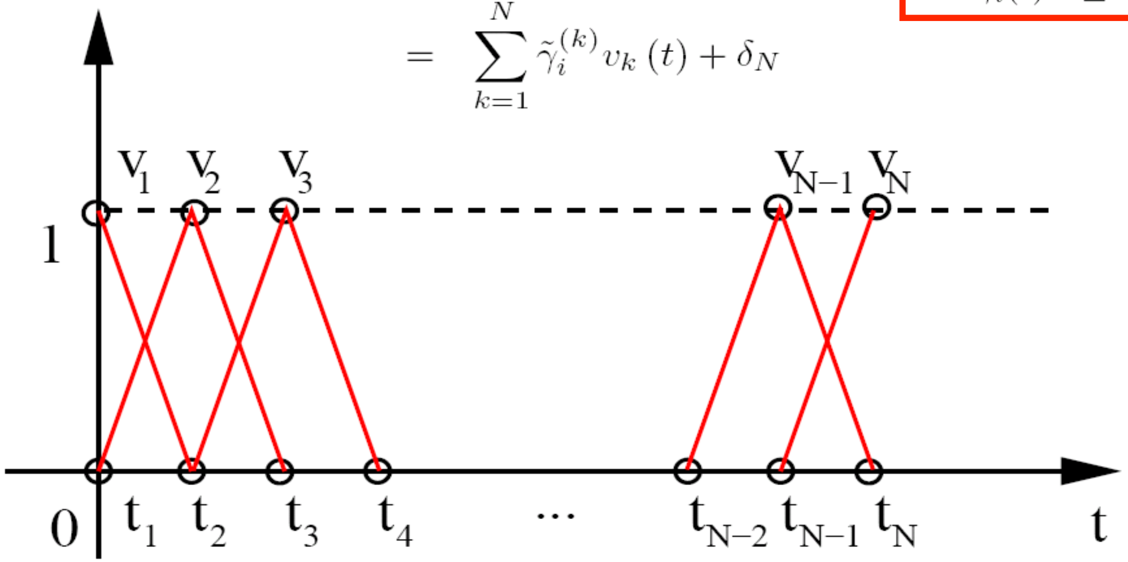
$$\mathbf{L}^{\epsilon}(\Theta, \Gamma(t), \epsilon^2) \quad = \quad \mathbf{L}(\Theta, \Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \rightarrow \min_{\Gamma(t),\Theta}$$

(H. 09, SIAM J. of Sci. Comp.)

$$\begin{aligned} \sum_{i=1}^{\mathbf{K}} \gamma_i(t) &= 1, \quad \forall t \in [0,T] \\ \gamma_i(t) &\geq 0, \quad \forall t \in [0,T], i = 1,\ldots,\mathbf{K} \end{aligned}$$

Galerkin-Ansatz :

$$\begin{aligned} \gamma_i(t) &= \tilde{\gamma}_i(t) + \delta_N \\ &= \sum_{k=1}^{N} \tilde{\gamma}_i^{(k)} v_k(t) + \delta_N \end{aligned}$$



Linear Finite Elements

FEM-Discretized Clustering Functional

$$\tilde{\mathbf{L}}^{\epsilon} = \sum_{i=1}^{\mathbf{K}} \left[a^{\mathbf{T}}(\theta_i)\bar{\gamma}_i + \epsilon^2 \bar{\gamma}_i^{\mathbf{T}} \mathbf{H} \bar{\gamma}_i \right] \to \min_{\bar{\gamma}_i, \Theta}$$

subjected to

$$\sum_{i=1}^{\mathbf{K}} \tilde{\gamma}_i^{(k+1)} = 1, \quad \forall k = 1, \ldots, N,$$

$$\tilde{\gamma}_i^{(k+1)} \geq 0, \quad \forall k = 1, \ldots, N; i = 1, \ldots, \mathbf{K}.$$

Iterative Subspace Minimization:
sparse QP can be used

where $a(\theta_i) = \left(\int_{t_1}^{t_2} v_1(t) g(x_t, \theta_i) dt, \ldots, \int_{t_{N-1}}^{t_N} v_N(t) g(x_t, \theta_i) dt \right)$

- quadratic programming problem wrt. $\Gamma(t)$ for a fixed Θ
- convex minim. problem wrt. Θ for a fixed $\Gamma(t)$

•Horenko, *SIAM Journal of Sci. Comp., 2008*
•Horenko, *JAS, 2009*
•Horenko, *DYNAT, 2009*
•Franzke/Horenko/Majda/Klein, *JAS, 2009*

BV-case: making problem diff.-able

BV-constraint in the clustering functional minimization

$$\mathbf{L}(\Theta, \Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g(x_t, \theta_i) \to \min_{\Gamma(t), \Theta},$$

$$\sum_{i=1} \gamma_i(t) = 1, \quad \forall t \in [0, T]$$

$$\gamma_i(t) \geq 0, \quad \forall t \in [0, T], i = 1, \ldots, \mathbf{K}$$

$$\|D\gamma_i^T\|_1 \leq C \quad \forall i \qquad v_i := D\gamma_i^T$$

Theorem about the unique repr. of BV-func. (Moreau'88):

$$v_i = v_i^+ - v_i^-, \quad v_i^+ = \max(v_i, 0), v_i^- = \max(-v_i, 0)$$

Adaptive FEM-discretization:

$$\min_{x,\Theta} c^T(\Theta) x, \quad \text{subject to } A_{\text{eq}} x = b_{\text{eq}}, A_{\text{neq}} x \geq b_{\text{neq}}.$$

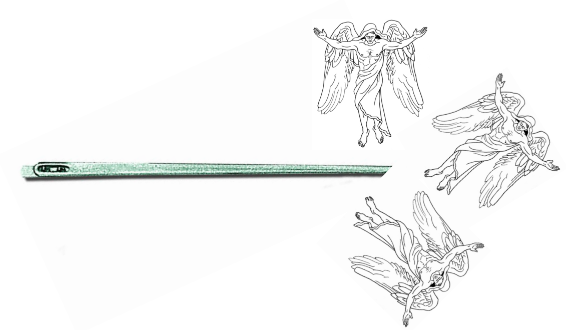
- linear programming problem wrt. x for a fixed Θ
- convex minim. problem wrt. Θ for a fixed x

H. 09, accepted for publ. In J. of Atmos. Sci.

Non-stationarity: Yes or No?

Information theory: golden ratio between numb. of parameters and "model quality"

(Ockham'XIVct.;Hartley'28;Shannon'48;Kolmogorov'68;Akaike'74;Tsai'98)



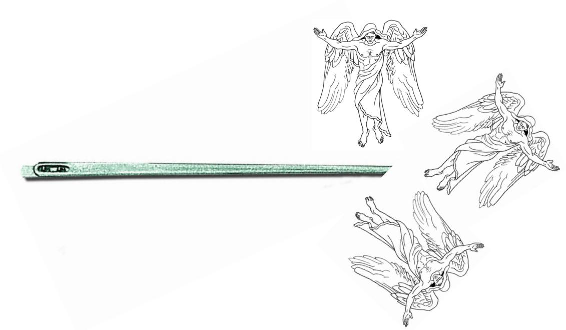
How many angels can dance on the head of a pin?

Non-stationarity: Yes or No?

Information theory: golden ratio between numb. of parameters and "model quality"

(Ockham'XIVct.;Hartley'28;Shannon'48;Kolmogorov'68;Akaike'74;Tsai'98)

Entia non sunt multiplicanda praeter necessitatem!



How many angels can dance on the head of a pin?

William of Ockham
(http://en.wikipedia.org)

Non-stationarity: Yes or No?

Information theory: golden ratio between parameters and “model quality“

(Hartley’28;Shannon’48;Kolmogorov‘68;Akaike‘74;Tsai‘98)

Let

$$G_{\mathbf{K}}(t) = \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(z_t, \theta_i\right) \qquad M(\mathbf{K}) \text{ is the number of parameters}$$

Theorem 2.1 *If for given* $\mathbf{K}$ *and optimal* Γ^* *the* $G_{\mathbf{K}}(t)$ *are independently* $\sigma_{\mathbf{K}}\chi_1^2$ *distributed for some constant* $\sigma_{\mathbf{K}}$, *then the* $\mathbf{K}$*-dependent part of the Akaike Information Criterium is given by*

$$AIC = N \log \mathbf{L}_{\mathbf{K}} + \frac{1}{2} \sum_{i=1}^{N} \log \frac{G_{\mathbf{K}}(t_i)}{\mathbf{L}_{\mathbf{K}}} + M(\mathbf{K})$$

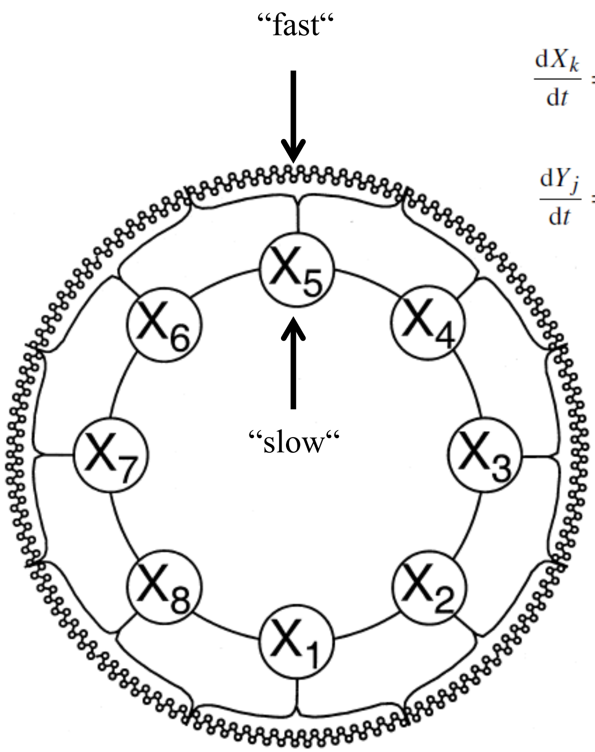
What does it mean:

Model Diskrimination and identification of optimal **K**

H. 09, accepted for publ. In Journal of Atm. Sci.
Putzig, H., Metzner 10, in preparation

Example I: Lorenz96+forcing

Wilks, *Quart.J.Royal Met.Soc.*, 2005



$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K$$

$$\frac{dY_j}{dt} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b}X_{\mathrm{int}[(j-1)/J]+1};$$ **+ *F(t)***

+  **=** **non-stationary model (time-dependent parameters)**

???

$$x_t = \mu(t) + \mathbf{A}(t)\,\phi_1(x_{t-\tau}, \ldots, x_{t-m\tau}) + \mathbf{B}(t)\,\phi_2(u_t) + \mathbf{C}(t)\,\epsilon_t$$

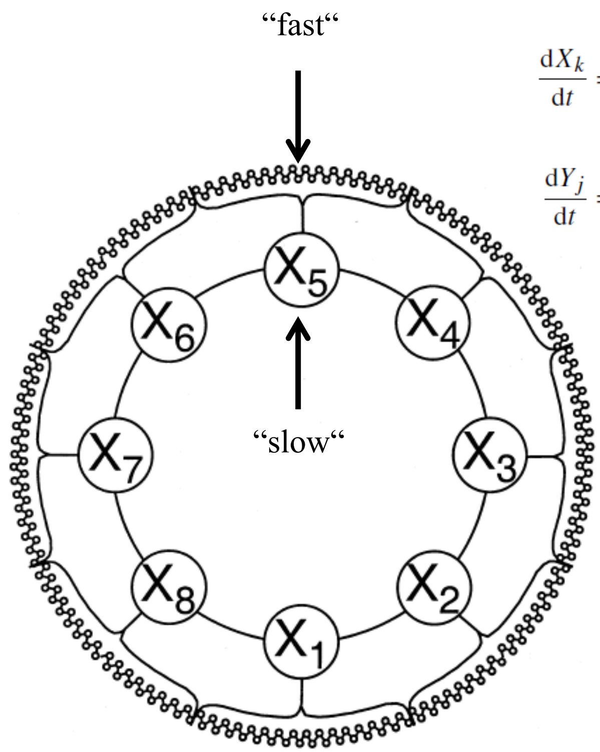
$$\mathbf{C}(t) = \mathbf{P}(t)\,\mathbf{\Lambda}(t)$$

$$g(x_t, \theta(t)) = \|x_t - \mu(t) - \mathbf{A}(t)\,\phi_1(x_{t-\tau}, \ldots, x_{t-m\tau}) - \mathbf{B}(t)\,\phi_2(u(t))\|_{\mathbf{P}(t)}$$

(H. 10, JAS, in press)

Example I: Lorenz96+forcing

Wilks, *Quart.J.Royal Met.Soc.*, 2005



$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K$$

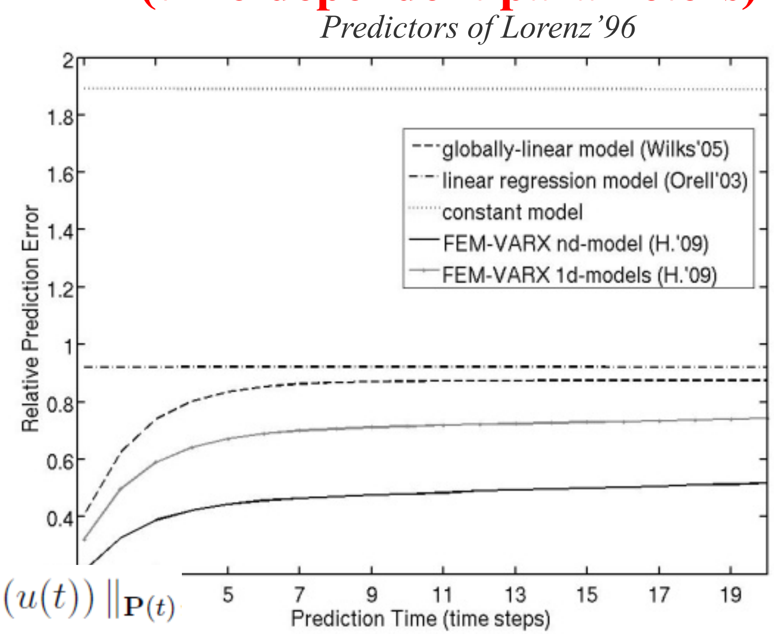
$$\frac{dY_j}{dt} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b}X_{\mathrm{int}[(j-1)/J]+1}; \quad \mathbf{+\ F(t)}$$

+

???

=

non-stationary model
(time-dependent parameters)



$$x_t = \mu(t) + \mathbf{A}(t)\,\phi_1(x_{t-\tau}, \ldots, x_{t-m\tau}) + \mathbf{B}(t)\,\phi_2(u_t) + \mathbf{C}(t)\,\epsilon_t$$

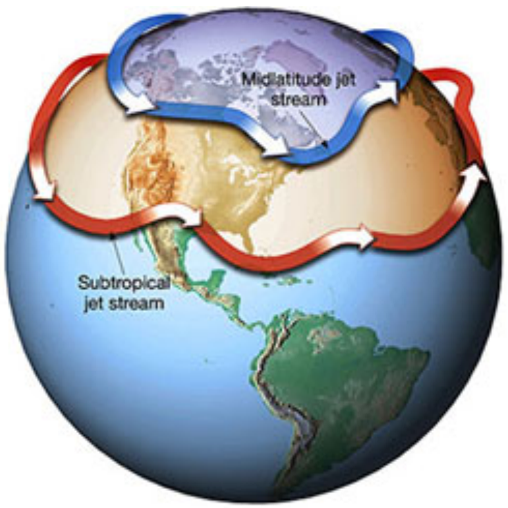
$$\mathbf{C}(t) = \mathbf{P}(t)\,\mathbf{\Lambda}(t)$$

$$g(x_t, \theta(t)) = \|x_t - \mu(t) - \mathbf{A}(t)\,\phi_1(x_{t-\tau}, \ldots, x_{t-m\tau}) - \mathbf{B}(t)\,\phi_2(u(t))\|_{\mathbf{P}(t)}$$

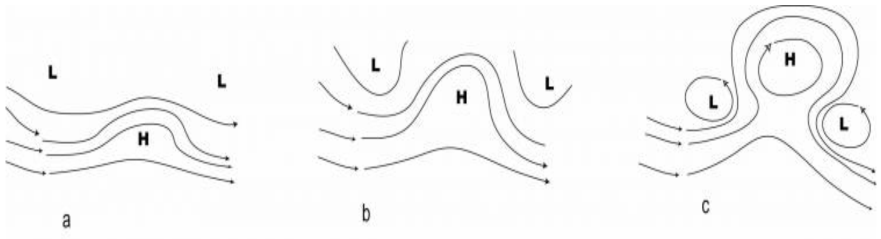
(H. 09, JAS, in press)

Example II: analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

Wind Jets transport moisture from US to Europe

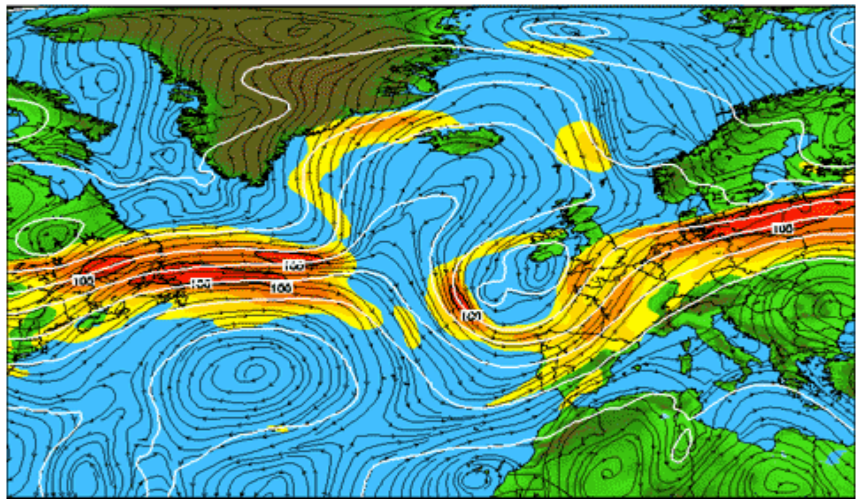


Jet-Blocking



VARX (standard stationary stochastic Model)

$$y_t \quad = \quad \mu + \sum_{q=1}^{m} \mathbf{A}_q y_{t-q\tau} + \mathbf{B}\phi\left(x_t\right) + \mathbf{C}\epsilon_t$$

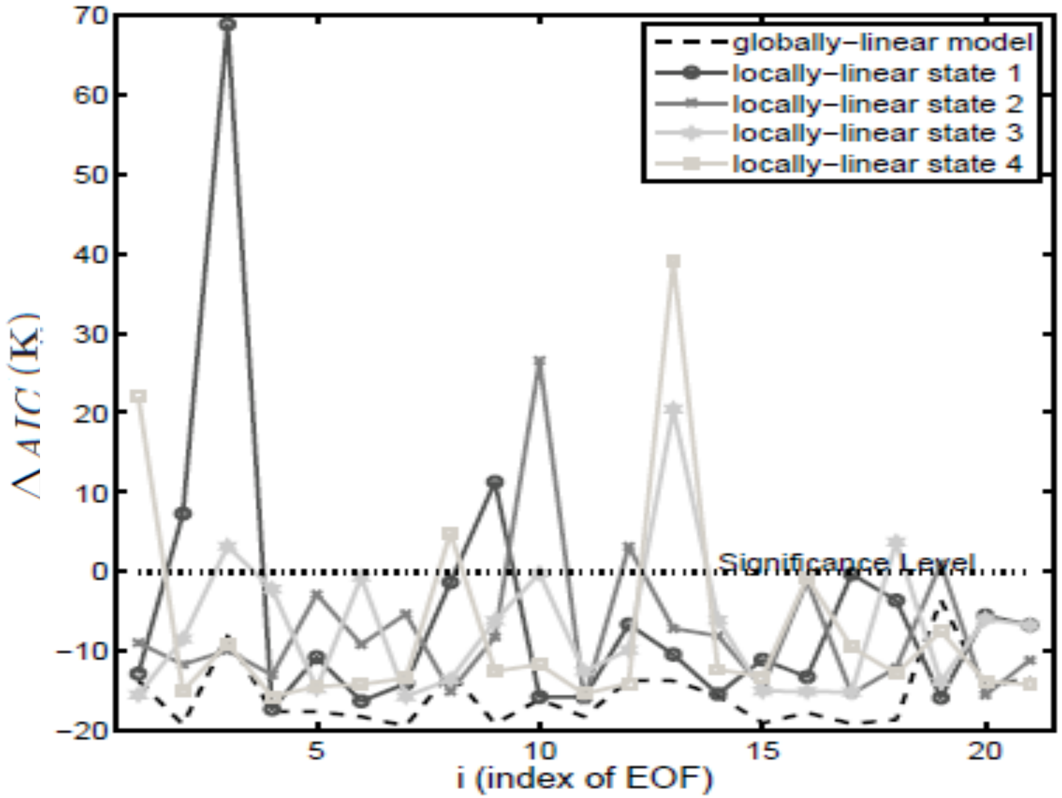


Jet-Blocking results in “indian summer“ and “blackberry cold“

Analysis of ERA40 geopotential time series (PIK, Potsdam)

Example II: analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

$$AIC(\mathbf{K}) = N \log \mathbf{L_K} + \frac{1}{2}\sum_{i=1}^{N} \log \frac{G_{\mathbf{K}}(t_i)}{\mathbf{L_K}} + M(\mathbf{K})$$



*non-stationary (**K**=4), stat. significant influence of seasonal factor and* CO_2, *influence of solar activity insignificant*

Stationarity: Yes or Not?

Answer: Not!

Example II: analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

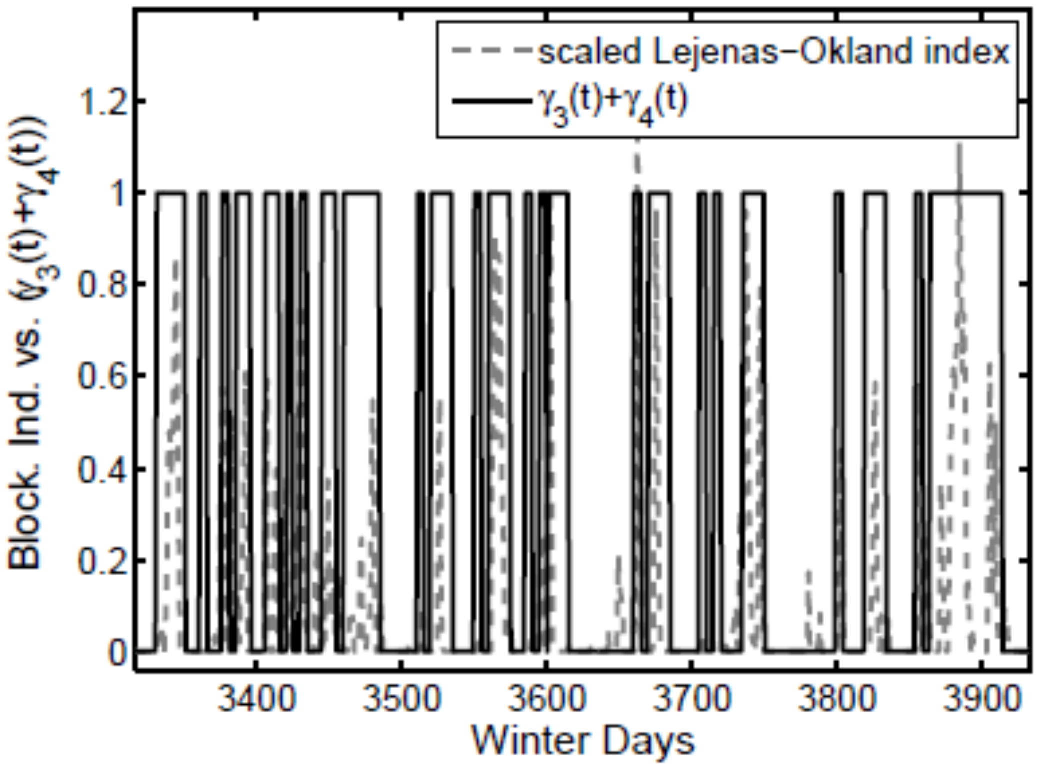


Figure 6: Comparison of the negative Lejenas-Oakland blocking index (dashed line) and the sum of cluster affiliations of locally-linear states 3 and 4 (solid line, calculated with FEM-VARX for $\mathbf{K} = 4, N = 4000, C = 3000, q = 2$).

H. 10, in press, Journal of Atm. Sci.

Example II: analysis of geopotential data (Europe, 1958-2003, analysis of climate impact factors)

$$\Delta \mathrm{H}^{(i)}(u(t)) = \mathbb{E}^{i}[x_t]$$

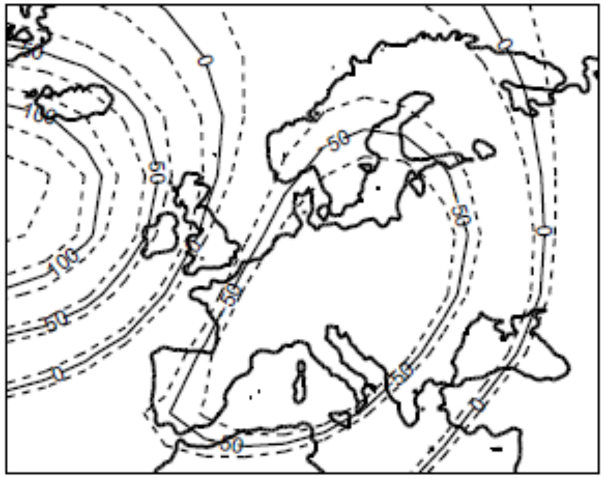
$$x_t = \mu^i + \sum_{q=1}^{m} A_q^i x_{t-q\tau} + B^i \phi(u(t)) + C^i \epsilon_t$$

=>

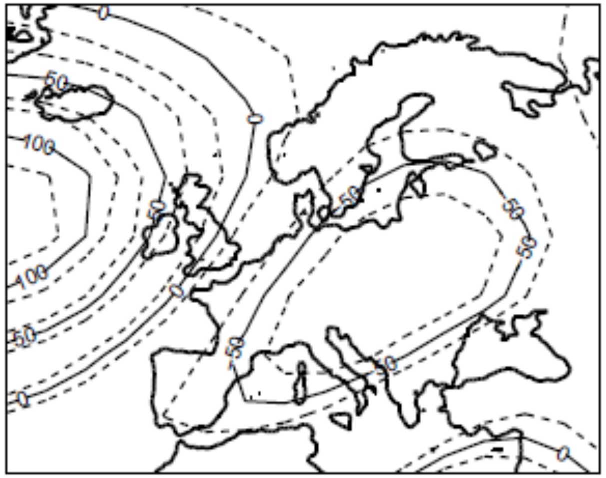
$$\Delta \mathrm{H}^{(i)}(u(t)) = \mu^i + \sum_{q=1}^{m} A_q^i \Delta \mathrm{H}^{(i)}(u(t)) + B^i \phi(u(t))$$

CO_2 impact on “indian summer“ local model (i=3)

Mean Δ H$^{(3)}$(winter 1958/1959)



Mean Δ H$^{(3)}$(winter 2002/2003)



Result: essential pressure impact factors
CO_2 and *seasonal factor* (but not *solar activity*)

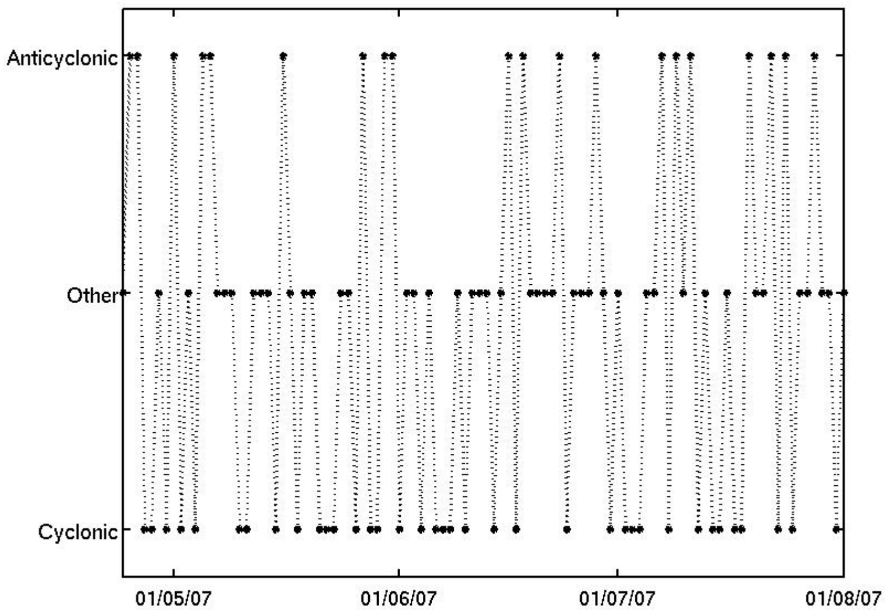
H. 10, in press, Journal of Atm. Sci.

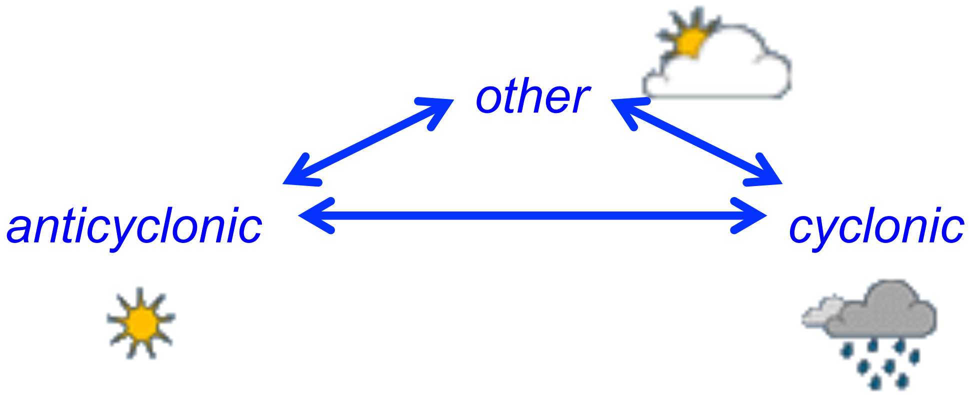
Lamb Circulation Patterns for UK (1945-2007)

Historical Circulation Data: weather regimes

(Data from the Univ. of East Anglia)

3 atmospherical states considered





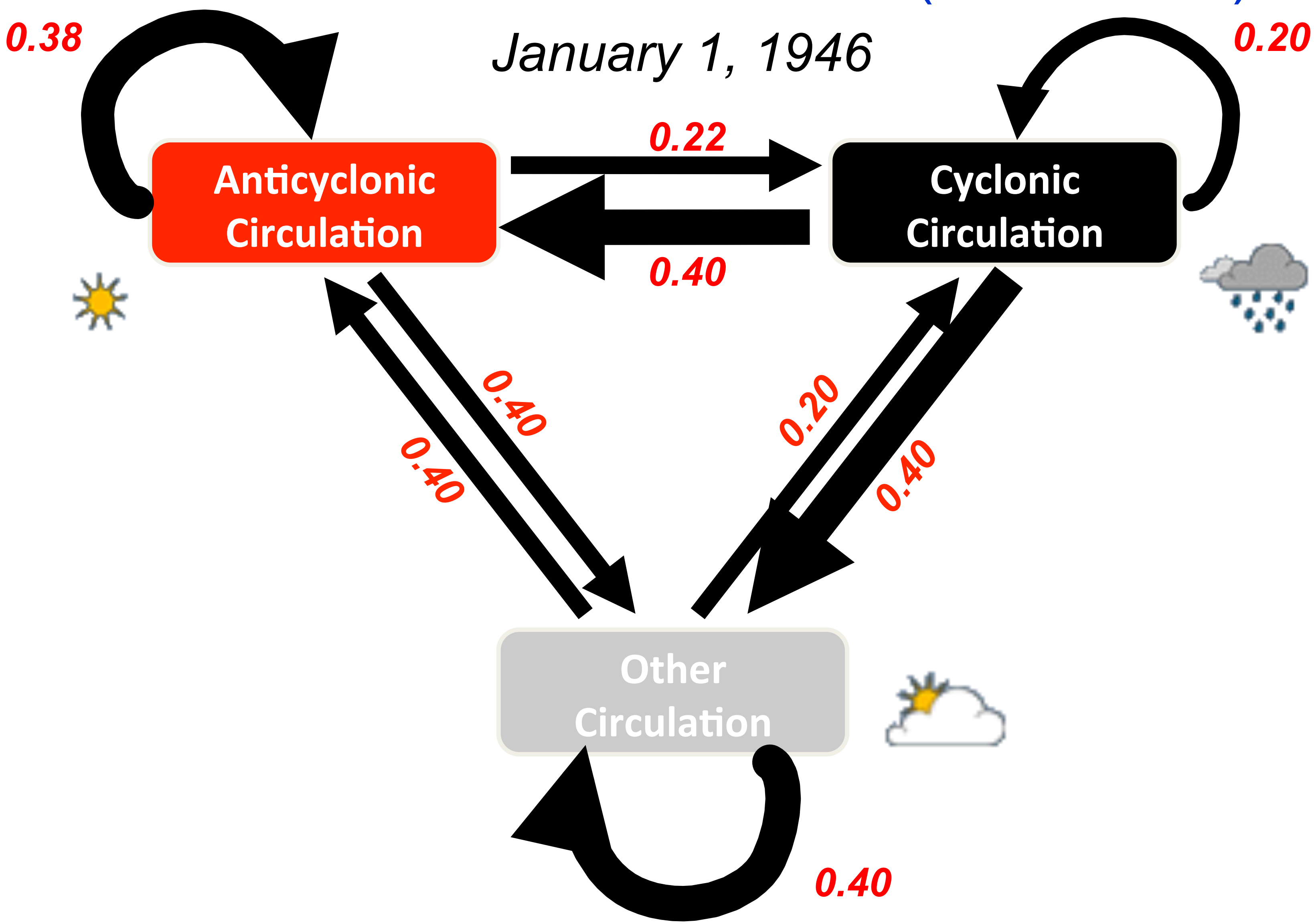
$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$$

$$x_t : X_t \rightarrow X_{t+1}$$

$$g(x_t, P^i) \quad = \quad -\log P^i_{X_t X_{t+1}}$$

Circulation Patterns for UK(1945-2007)

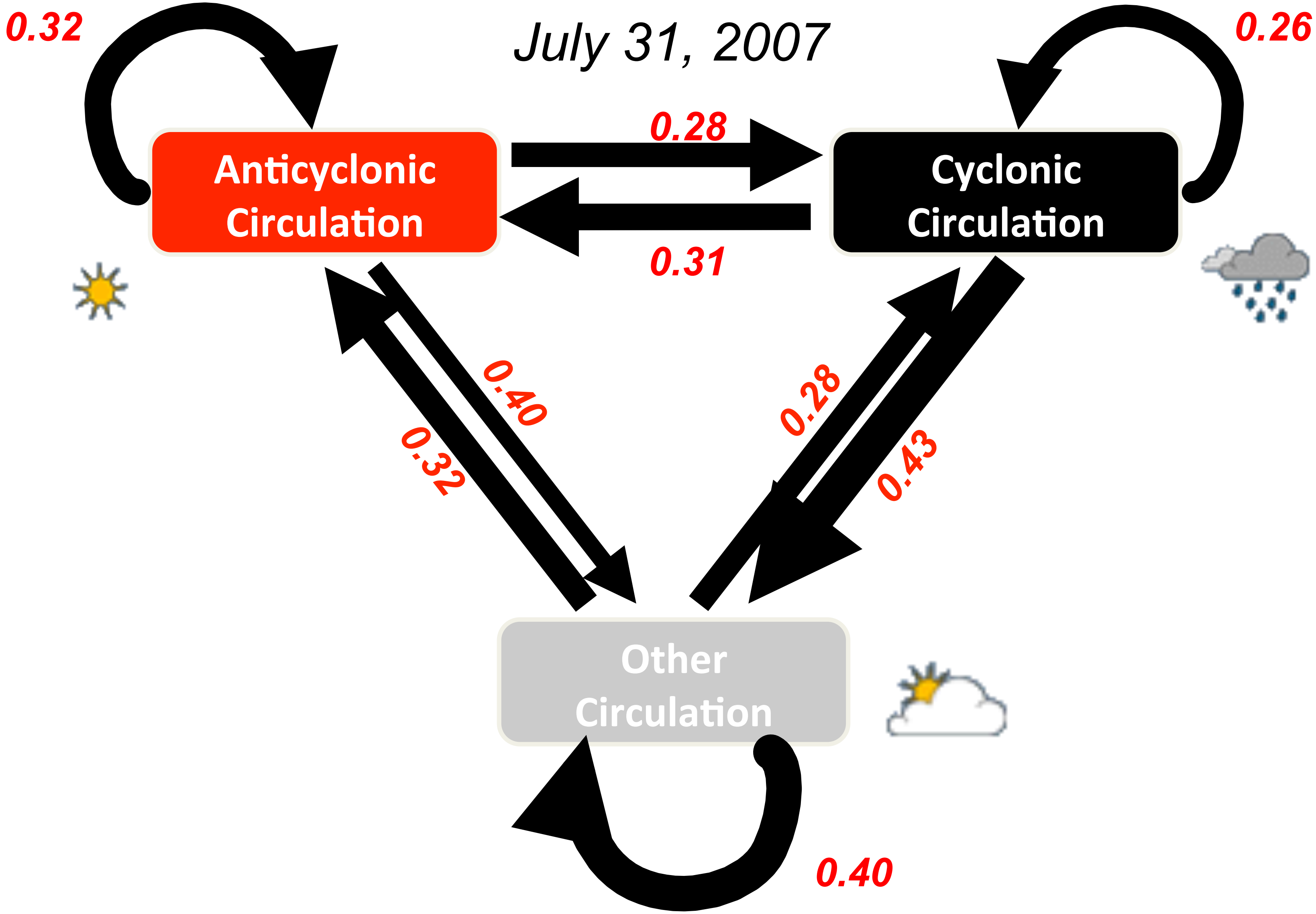
January 1, 1946



H., Journal of Atmos. Sci. (2009)

Circulation Patterns for UK(1945-2007)

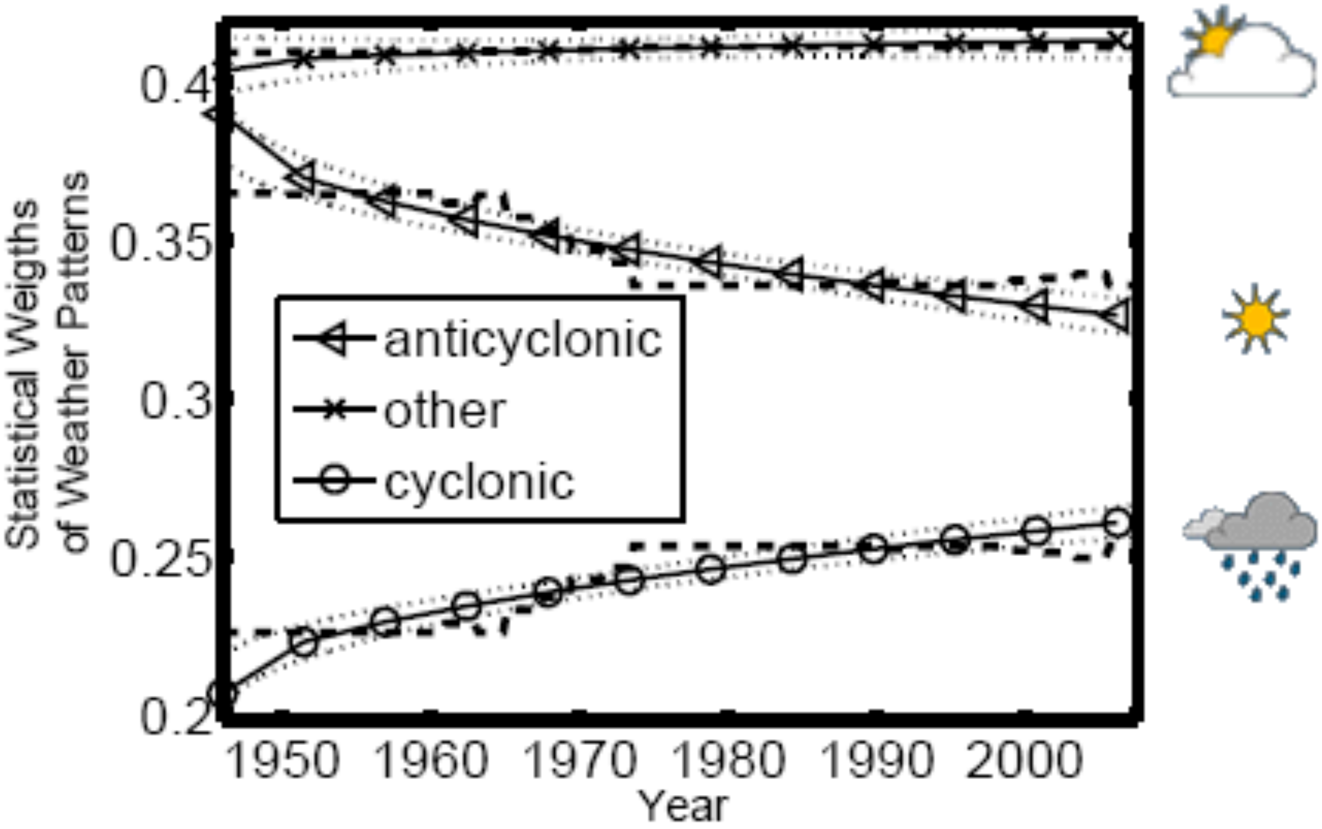
July 31, 2007



H., Journal of Atmos. Sci. (2009)

Circulation Patterns for UK(1945-2007)

Comparison of resulting a posteriori stat. weights
(obtained with two different sets of math. assumptions)



H., Journal of Atmos. Sci. (2009)

And now…

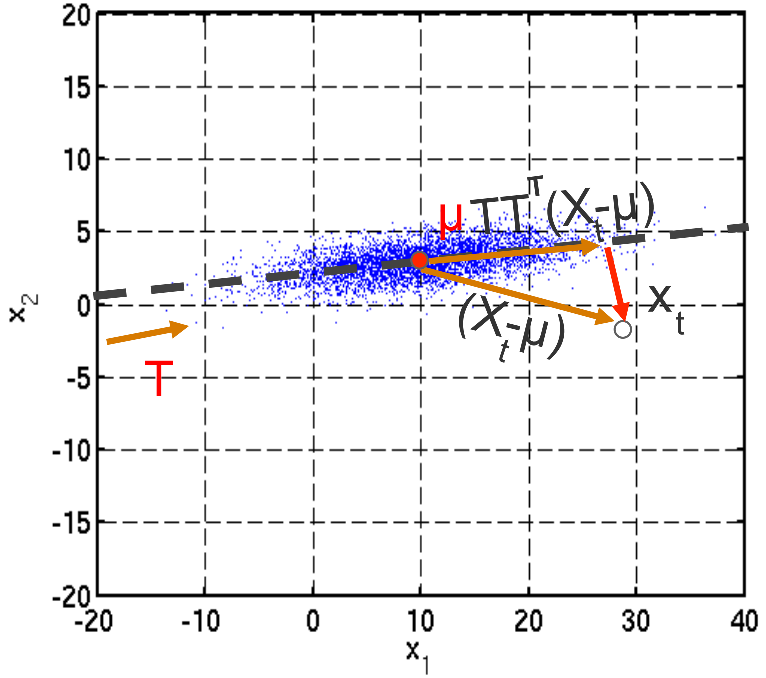

GRAHAM CHAPMAN JOHN CLEESE TERRY GILLIAM
ERIC IDLE TERRY JONES MICHAEL PALIN
MONTY PYTHON'S
And Now For Something
COMPLETELY DIFFERENT
DVD
The BEST of Monty Python's Flying Circus

Dimension Reduction+Clustering

Problem 1: $\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$

$$g\left(x, \theta_i\right) = \| x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \|^2$$

$\theta_i = \mathcal{T}_i \in \mathbf{R}^{nxm}$ ***m << n***

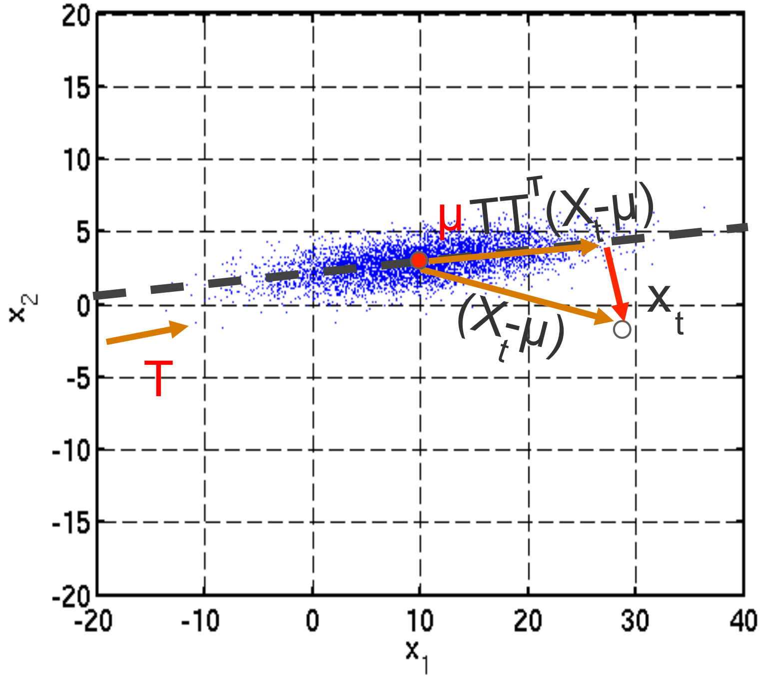


Dimension Reduction+Clustering

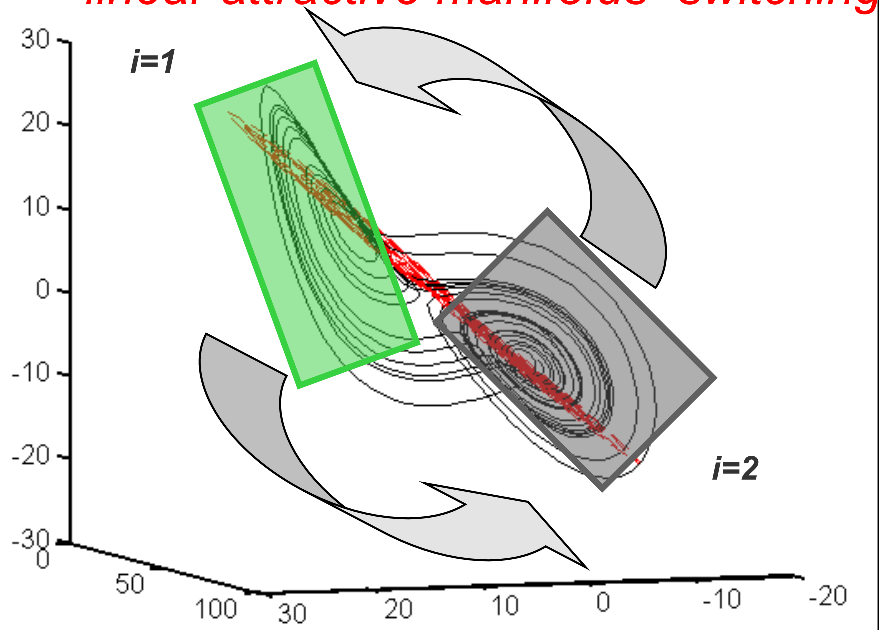
Problem 1: $$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \longrightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \, \|^2$$

$\theta_i = \mathcal{T}_i \in \mathbf{R}^{nxm}$ ***m << n***



Idea *Essential manifold* can be appproximated by *linear attractive manifolds+switching*



H./Schmidt-Ehrenberg/Schütte 07
H.,JAS, 08

Wrap-Up: Methods

Given data x_t, look for parameters $\theta(t)$

$$\mathbf{F}\left(x_t,\ldots,x_{t-m\tau},\theta(t),t\right)=0$$

(direct problem)

$$\int_0^T g\left(x_t,\theta(t)\right)dt \;\rightarrow\; \min_{\theta(t)}$$

(inverse problem)

$$g\left(x_t,\theta(t)\right) = \sum_{i=1}^{K}\gamma_i(t)g\left(x_t,\theta_i\right)$$

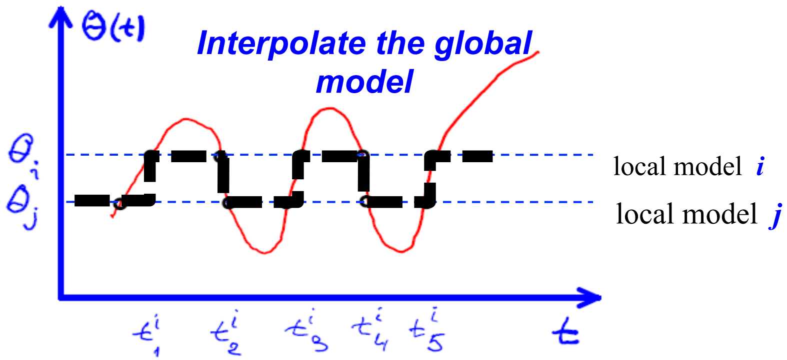
(ansatz)

- Extension from $W_{1,2}$ to BV-function spaces (allows parameter discontinuities)
- Data-based parameterization of non-stationary autoregressive factor models (stochastic modeling of sub-grid scale, **FEM-VARX**)

$$x_t = \mu(t)+\mathbf{A}(t)\phi_1\left(x_{t-\tau},\ldots,x_{t-m\tau}\right)+\mathbf{B}(t)\phi_2\left(u_t\right)+\mathbf{C}(t)\epsilon_t$$

- Deployment of *information theory* (optimal *model discrimination*)

$$\begin{aligned}
\mathbf{L}(\Theta,\Gamma(t)) &= \int_0^T\sum_{i=1}^{\mathbf{K}}\gamma_i(t)g\left(x_t,\theta_i\right)\rightarrow \min_{\Gamma(t),\Theta},\\
\sum_{i=1}\gamma_i(t) &= 1,\quad \forall t\in[0,T]\\
\gamma_i(t) &\geq 0,\quad \forall t\in[0,T],\quad i=1,\ldots,K\\
\|\gamma_i(t)\| &\leq C_i
\end{aligned}$$



Adaptive FEM-Clustering Methods: efficient numerical solvers

•Horenko, *SIAM Journal of Sci. Comp., 2009*
•Horenko, *JAS, 2009*
•Horenko, *DYNAT, 2010*