Estimation of non-stationary Processes : Part I (07.10.2008)

Illia Horenko

Research Group "***Computational Time Series Analysis*** "
Institute of Mathematics
Freie Universität Berlin (FU)

DFG Research Center ***MATHEON***
„Mathematics in key technologies"

Memo II: Probability

A probability space is a measure space such that the measure of the whole space is equal to 1.

In other words: a probability space is a triple $(\Omega, \mathcal{F}, P)$ consisting of a set Ω (called the sample space), a σ-algebra (also called σ-field) $\mathcal{F}$ of subsets of Ω (these subsets are called events), and a measure P on $(\Omega, \mathcal{F})$ such that $P(\Omega) = 1$ (called the probability measure).

Event	Probability
A	$P(A) \in [0,1]$
not A	$P(A') = 1 - P(A)$
A or B	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $= P(A) + P(B) \quad \text{if A and B are mutually exclusive}$
A and B	$P(A \cap B) = P(A\|B)P(B)$ $= P(A)P(B) \quad \text{if A and B are independent}$
A given B	$P(A\|B)$

Markov Chains: Log-Likelihood

Observed Time Series: $\{X_1, \ldots, X_T\}$, $X_t \in s_1, \ldots, s_m$

Markov-Property:

$$\mathrm{P}\left[X_t = s_j | X_1, X_2, \ldots, X_{t-1} = s_i\right] = \mathrm{P}\left[X_t = s_j | X_{t-1} = s_i\right] = P_{ij}(t)$$

Markov Chains: Log-Likelihood

Observed Time Series: $\{X_1, \ldots, X_T\}$, $X_t \in s_1, \ldots, s_m$

Markov-Property:

$$\mathrm{P}\left[X_t = s_j | X_1, X_2, \ldots, X_{t-1} = s_i\right] = \mathrm{P}\left[X_t = s_j | X_{t-1} = s_i\right] = P_{ij}(t)$$

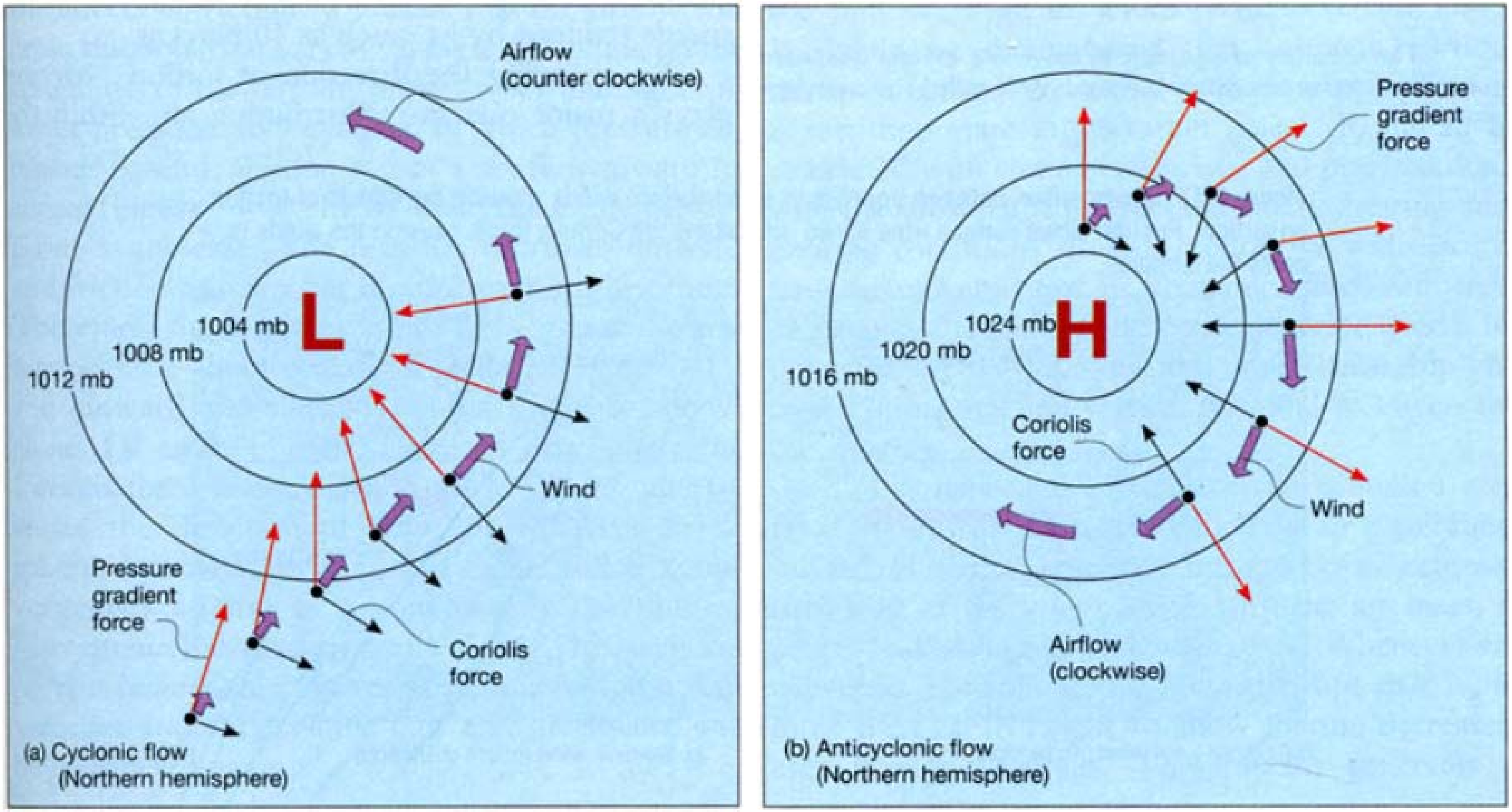
Probability of Observed Time Series (Likelihood):

$$\mathrm{P}\left[X_1, \ldots, X_T\right] = \mathrm{P}\left[X_1\right] \prod_{i,j=1}^{m} \prod_{t \in \{t_{ij}\}} P_{ij}(t)$$

Example: Circulation Patterns

Basic Circulation Patterns

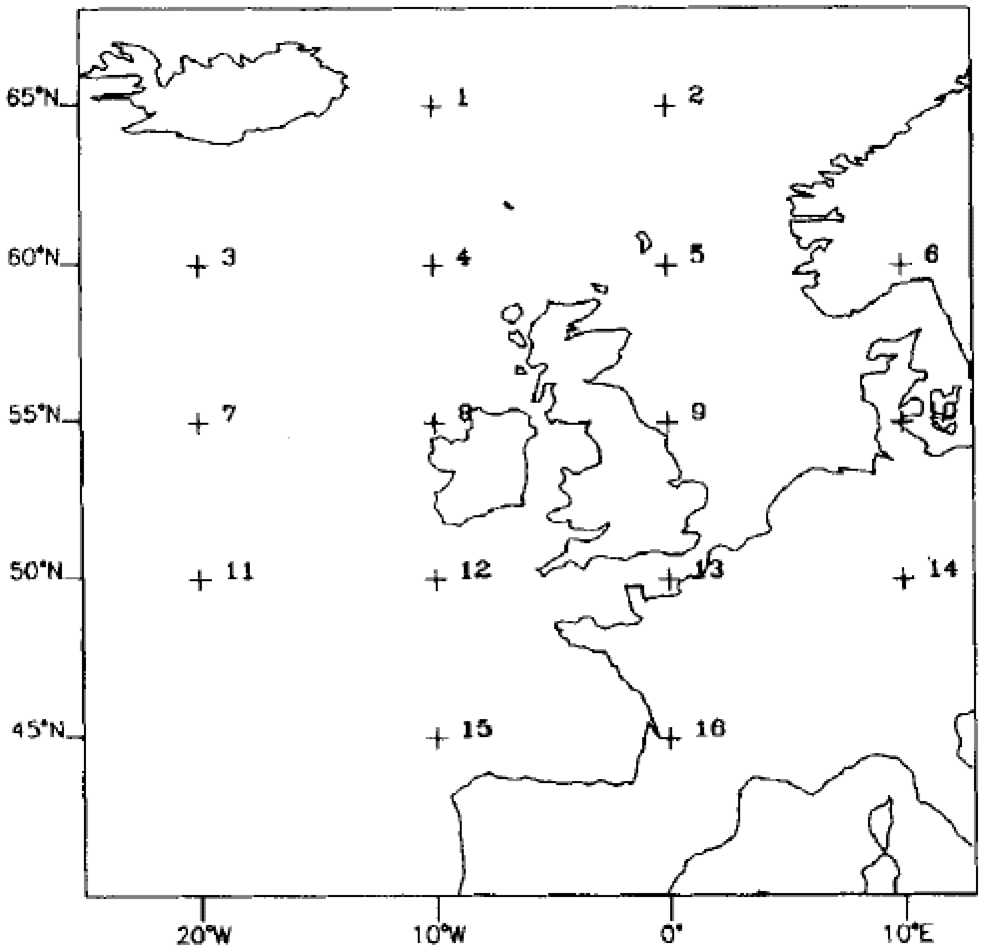
cyclonic — *anticyclonic*



www.earth.rochester.edu

Circulation Patterns over UK

Pressure Measurements on the Grid



Circulation Patterns for UK(1945-2007)

Historical Circulation Data: 27 Lamb regimes
(Data from the Univ. of East Anglia)

Lamb, Geophys. Mem. (1972),
Jones/Hulme/Briffa, Int.J.of Climat. (1977),

anticyclonic

Lamb	Description	Examples
A	anticyclonic	1888
ANE	anticyclonic north eastery	123
AE	anticyclonic easterly	186
ASE	anticyclonic south easterly	104
AS	anticyclonic southerly	82
ASW	anticyclonic south westerly	128
AW	anticyclonic westerly	369
ANW	anticyclonic north westerly	146
AN	anticyclonic northerly	161

other

Lamb	Description	Examples
NE	north eastery	100
E	easterly	362
SE	south easterly	201
S	southerly	430
SW	south westerly	400
W	westerly	1362
NW	north westerly	417
N	northerly	398
U	unclassifiable	431

cyclonic

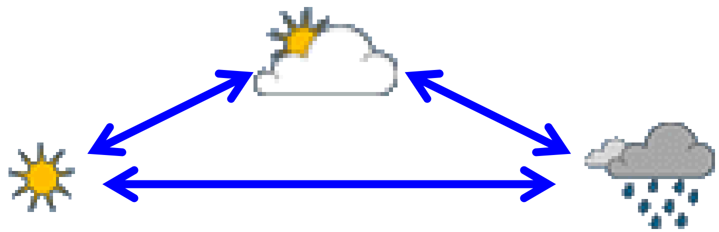
Lamb	Description	Examples
C	cyclonic	1447
CNE	cyclonic north eastery	55
CE	cyclonic easterly	113
CSE	cyclonic south easterly	39
CS	cyclonic southerly	90
CSW	cyclonic south westerly	62
CW	cyclonic westerly	212
CNW	cyclonic north westerly	81
CN	cyclonic northerly	109

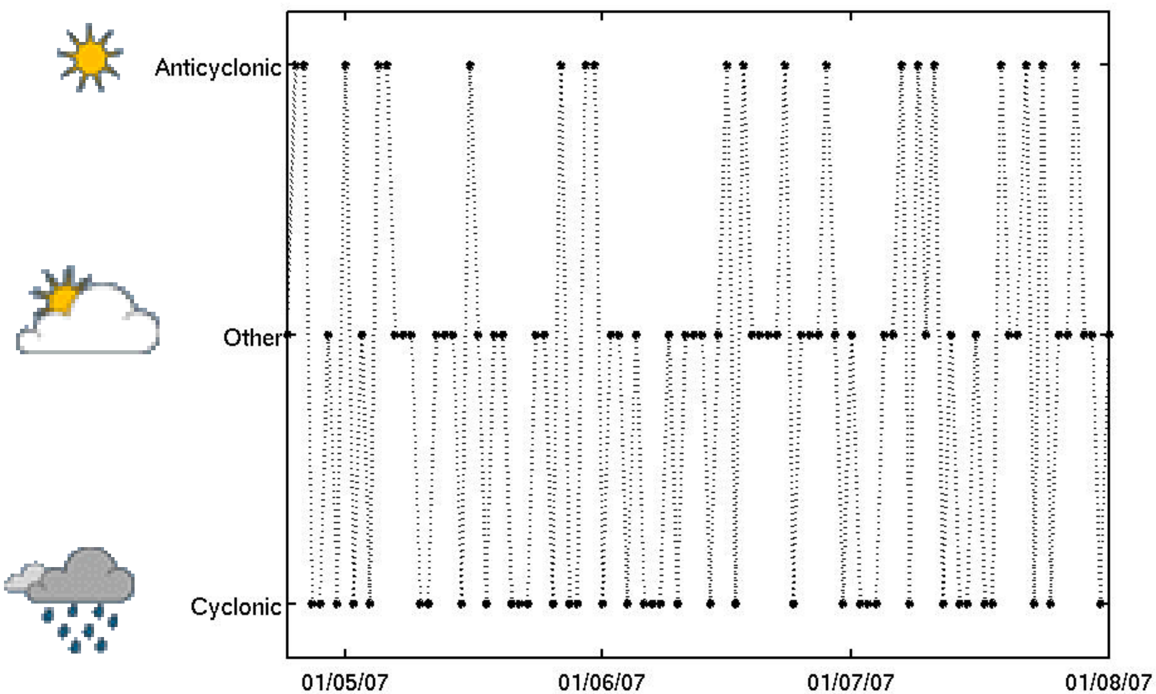
Example:Weather in UK(1945-2007)

Historical Circulation Data: weather regimes

(Data from the Univ. of East Anglia)

3 atmospherical states considered





Markov Chains: Log-Likelihood

Observed Time Series: $\{X_1, \ldots, X_T\}, \; X_t \in s_1, \ldots, s_m$

Markov-Property:

$$\mathrm{P}\left[X_t = s_j | X_1, X_2, \ldots, X_{t-1} = s_i\right] = \mathrm{P}\left[X_t = s_j | X_{t-1} = s_i\right] = P_{ij}(t)$$

Log-Likelihood:

$$\sum_{j=1}^{m} P_{ij}(t) = 1, \quad \text{for all } t, i$$
$$P_{ij}(t) \geq 0, \quad \text{for all } t, i, j$$

$$\begin{aligned} \mathbf{L}(P(t)) &= \log \mathrm{P}\left[X_1, \ldots, X_T\right] \\ &= \log \mathrm{P}\left[X_1\right] + \sum_{i}^{m} \sum_{j=1}^{m} \sum_{t \in \{t_{ij}\}} \log P_{ij}(t) \rightarrow \max_{P(t)} \end{aligned}$$

Circulation Patterns for UK(1945-2007)

Historical Circulation Data: 27 Lamb regimes
(Data from the Univ. of East Anglia)
3 atmospherical states considered

Assumption: Stationarity

	anticyclonic	other	cyclonic
anticyclonic	0.34	0.40	0.26
other	0.35	0.41	0.24
cyclonic	0.34	0.41	0.25

Locally Stationary Markov Chain

Gaussian Window: $\gamma(t, t_0) = \frac{1}{c} \exp(-\frac{(t-t_0)^2}{\sigma^2})$

Approximate Log-Likelihood:

$$\mathbf{L}\left(P(t_0)\right) \approx \log \mathrm{P}\left[X_1\right] + \sum_{i,j=1}^{m} \sum_{t \in \{t_{ij}\}} \gamma(t, t_0) \log P_{ij}(t_0)$$

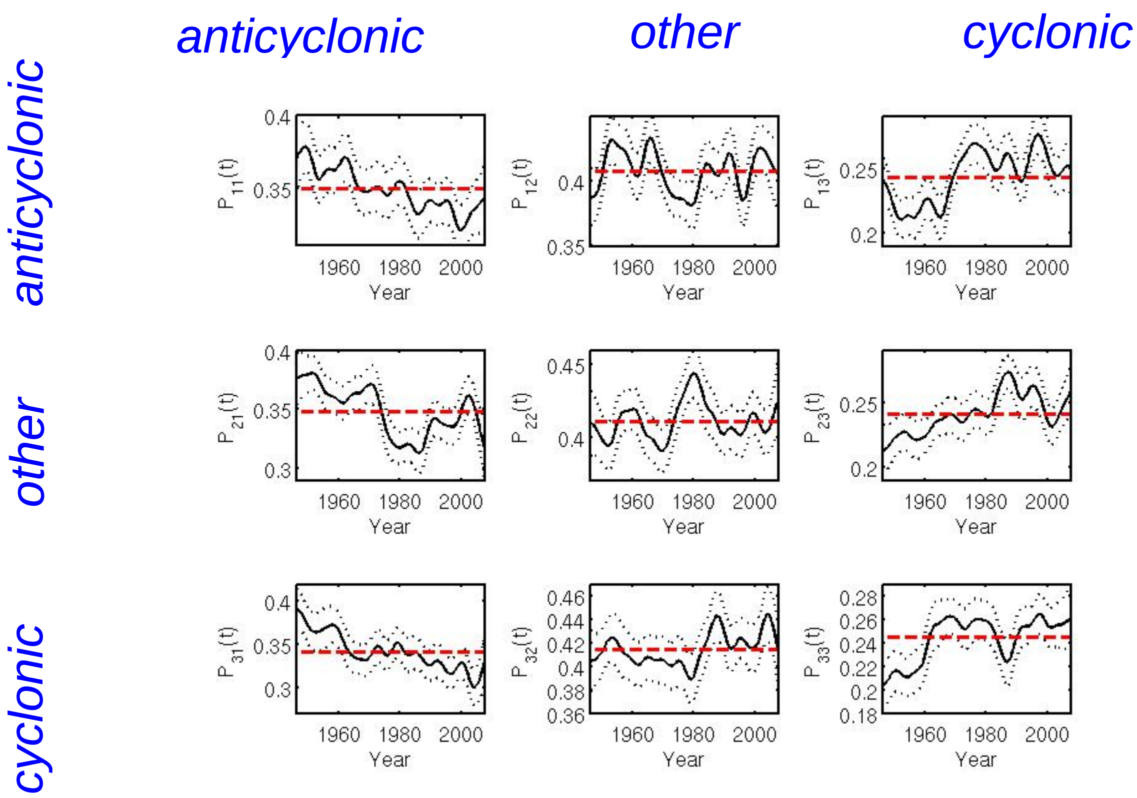
Gaussian Kernel Filtering:

$$P_{ij}(t_0) = \frac{\sum_{t \in \{t_{ij}\}} \gamma(t, t_0)}{\sum_{t \in \{t_i\}} \gamma(t, t_0)}$$

Circulation Patterns for UK(1945-2007)

Historical Circulation Data: 27 Lamb regimes
(Data from the Univ. of East Anglia)
3 atmospherical states considered

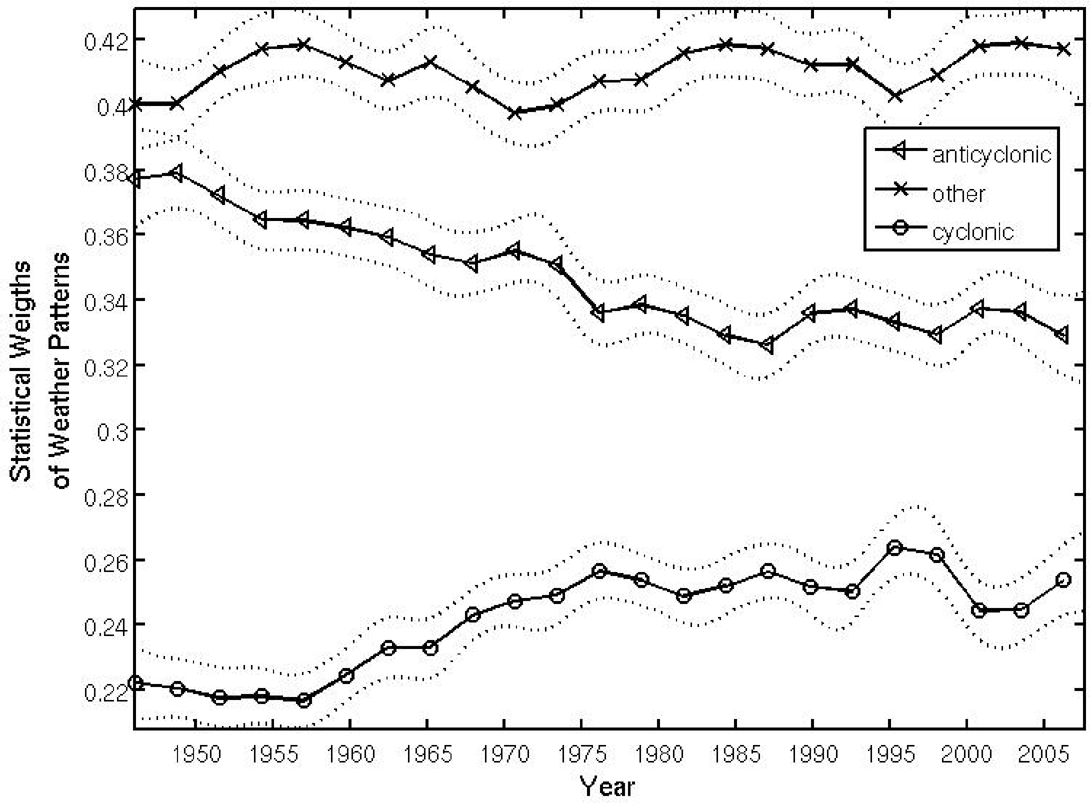
Gaussian Window vs. Homogenous Estimator



Circulation Patterns for UK(1945-2007)

Historical Circulation Data: 27 Lamb regimes
(Data from the Univ. of East Anglia)
3 atmospherical states considered

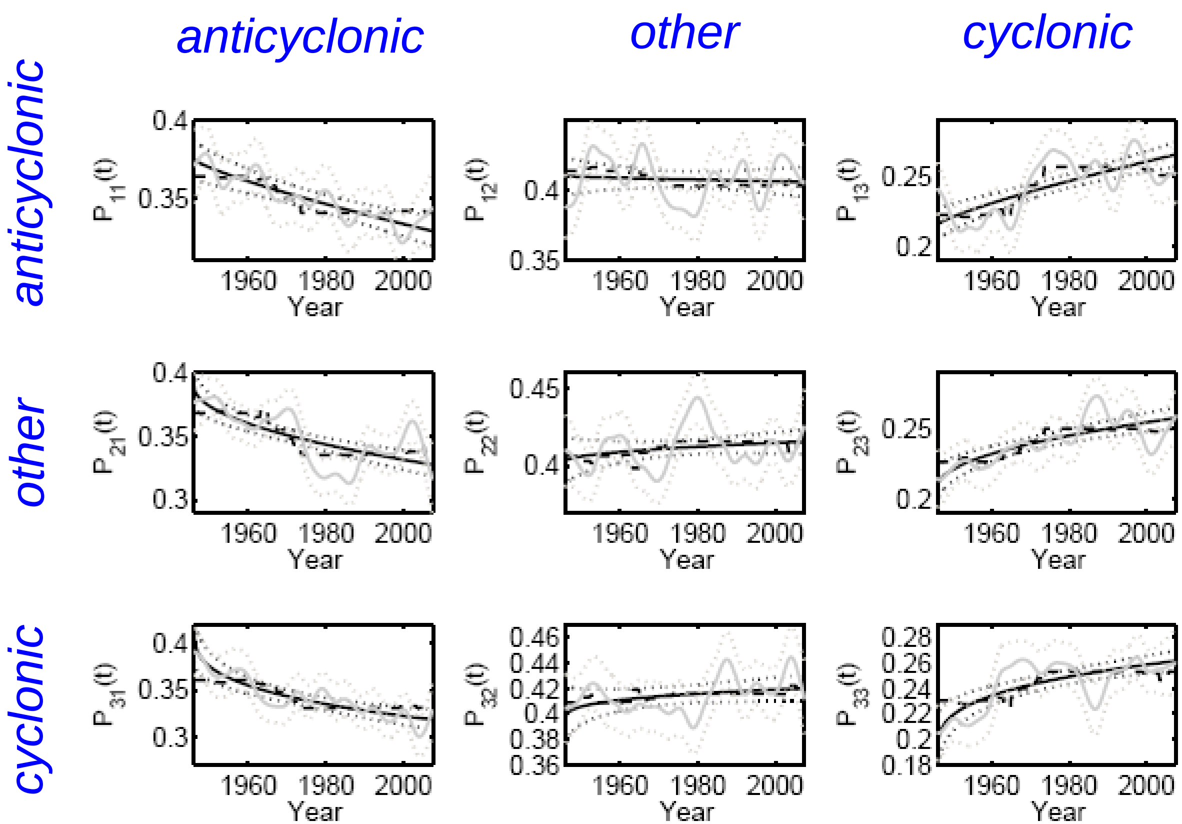
Gaussian Window vs. Homogenous Estimator



Circulation Patterns for UK(1945-2007)

Historical Circulation Data: 28 Lamb regimes
(Data from the Univ. of East Anglia)
3 atmospherical states considered

Gaussian Window vs. Homogenous Estimator



Thank you for attention!