Regime-dependent modelling of extremes in the extra-tropical atmospheric circulation

Frank Kwasniok

Exeter Climate Systems (XCS),
College of Engineering, Mathematics and Physical Sciences,
University of Exeter, Exeter, UK

Workshop
Non-equilibrium Statistical Mechanics
and the Theory of Extreme Events in Earth Science
Isaac Newton Institute for Mathematical Sciences, Cambridge
29 October – 1 November 2013

Outline

Outline

Extremes in the extra-tropical atmospheric circulation

- extremes of vorticity and wind speed
- hierarchy of scales in the atmosphere
- regime-dependent modelling, conditional on the large scales
- statistics, prediction and predictability of extremes
- data-driven methods from time series analysis and machine learning
- hierarchy of models of increasing complexity and realism

Outline

Cluster-weighted modelling of threshold exceedances

Exceedance of a high threshold: $y = x - u$

Mixture of generalised Pareto distributions:

$$p(y|x > u, \mathbf{z}) = \sum_{i=1}^{K} g_i(\mathbf{z}) \, \mathrm{GPD}_{\sigma_i, \xi_i}(y)$$

Clusters may be determined

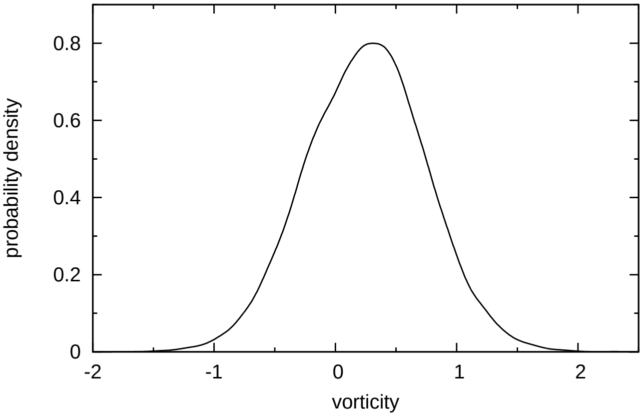
- seperately using a Gaussian mixture model or hidden Markov model (persistent or metastable regimes)
- simultaneously with the extreme model

Parameter estimation via maximum likelihood with expectation-maximization (EM) algorithm

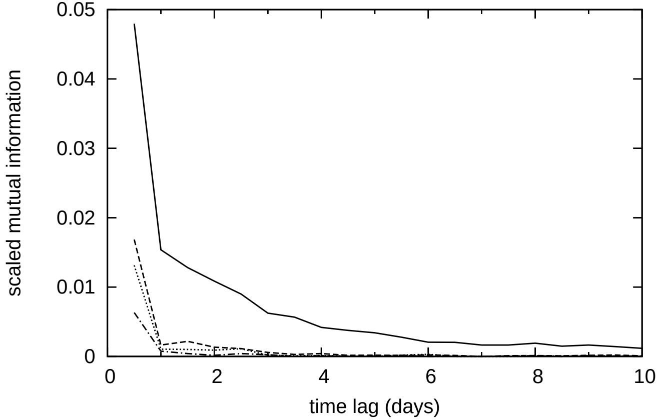
Intermediate-complexity atmospheric model

- Three-level quasi-geostrophic model on the Northern hemisphere
- Spectral truncation to T30
- Forcing determined from reanalysis data
- Realistic mean state and variability pattern
- Teleconnection patterns (PNA, NAO, . . .)

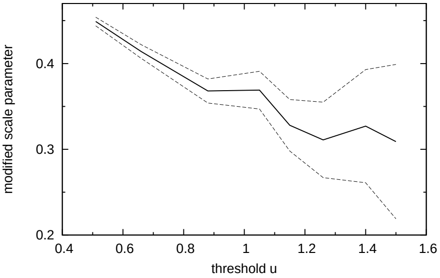
Vorticity at a grid point in the North Atlantic area

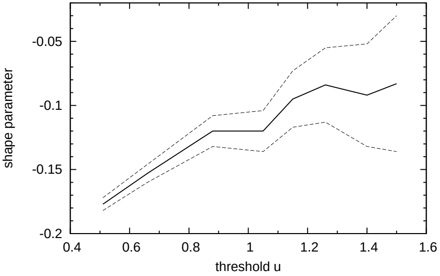


Temporal extremal dependence

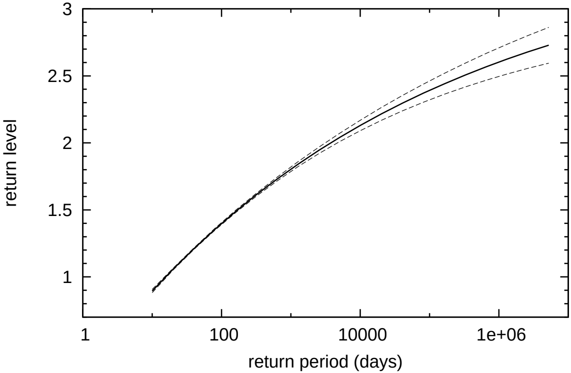


GPD parameter estimation ($\sigma = 0.26$, $\xi = -0.10$)





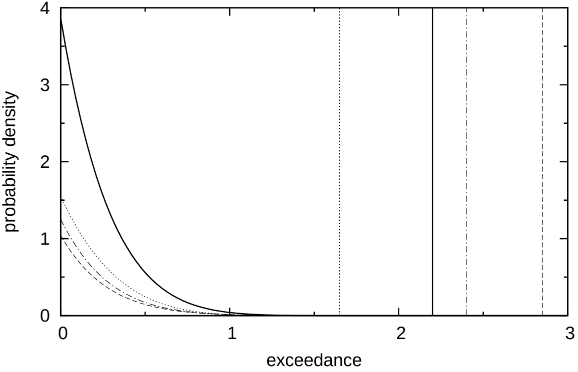
Return level plot



GPD for 3 clusters

$$\sigma_1 = 0.25, \xi_1 = -0.08 \qquad \sigma_2 = 0.25, \xi_2 = -0.11$$

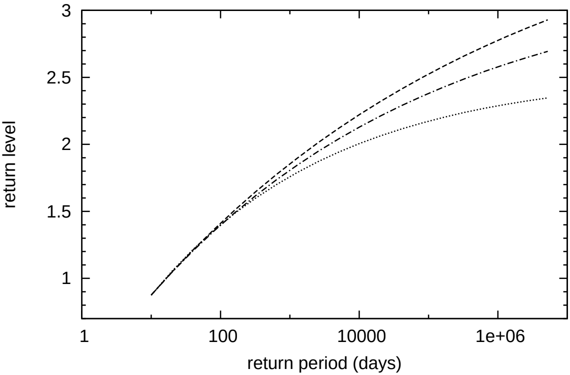
$$\sigma_3 = 0.27, \xi_3 = -0.17$$



Return level plots for 3 clusters

$$\sigma_1 = 0.25, \xi_1 = -0.08 \qquad \sigma_2 = 0.25, \xi_2 = -0.11$$

$$\sigma_3 = 0.27, \xi_3 = -0.17$$



Outline

Poisson process modelling of extremes

- extreme events often modelled as a Poisson process with rate parameter λ
- temporal clustering of extremes induced by
  - sampling variability
  - temporal variation in the rate parameter due to weather regimes
  - dependency between successive events
- inhomogeneous Poisson process whose rate parameter depends on large-scale flow variables

Cluster-weighted Bernoulli trial

Mixture of Bernoulli trial processes:

$$p(e = 1|\mathbf{z}) = \sum_{i=1}^{K} g_i(\mathbf{z})\, \pi_i$$

As before, various options for determining the clusters.

Vorticity at a grid point / 3 clusters

π	π_1	π_2	π_3
0.050	0.083	0.062	0.027
0.020	0.047	0.022	0.009
0.010	0.025	0.008	0.003

Outline

Regime-dependent prediction model for extremes

Predictive probability density:

$$p(e^\tau|\mathbf{z}^0) = \sum_{i=1}^{K} g_i(\mathbf{z}^0)\, p(e^\tau|\mathbf{z}^0, i)$$

with

$$g_i(\mathbf{z}^0) = p(i|\mathbf{z}^0) = \frac{w_i\, p(\mathbf{z}^0|i)}{\sum_{j=1}^{K} w_j\, p(\mathbf{z}^0|j)}$$

- prediction of individual extreme events at lead time τ
- relating to precursor patterns

Evaluation of forecasts

Characterisation

Utility:

$$R = \sum_{\alpha} f_{\alpha} \log(f_{\alpha}/b_{\alpha})$$

Probabilistic scores

Brier score:

$$\mathrm{Br} = \sum_{\alpha} (f_{\alpha} - e_{\alpha})^2$$

Ignorance score:

$$\mathrm{ign} = -\log f_{\alpha}$$

Extremes of energy in a barotropic atmospheric model

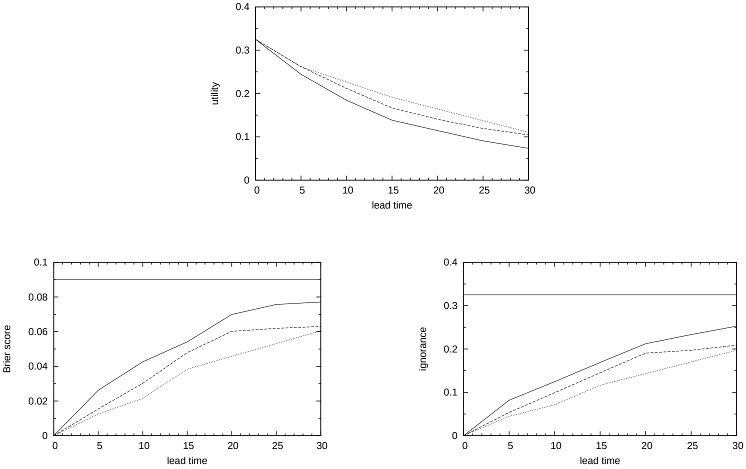
Gaussian clustering:

$$p(\mathbf{x}|i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Gamma}_i)$$

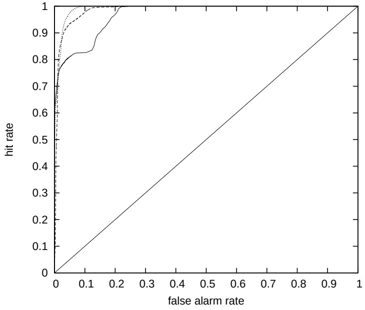
Locally constant predictive model:

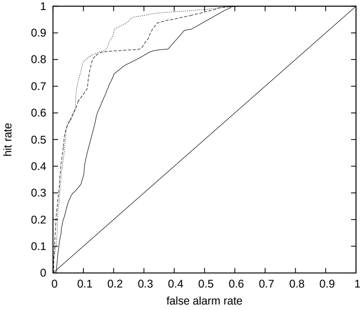
$$p(e_1^{\tau}|i) = \rho_i$$

Utility, Brier score and ignorance score for $b = 0.1$



ROC curves for $b = 0.1$ and $\tau = 10, 30$



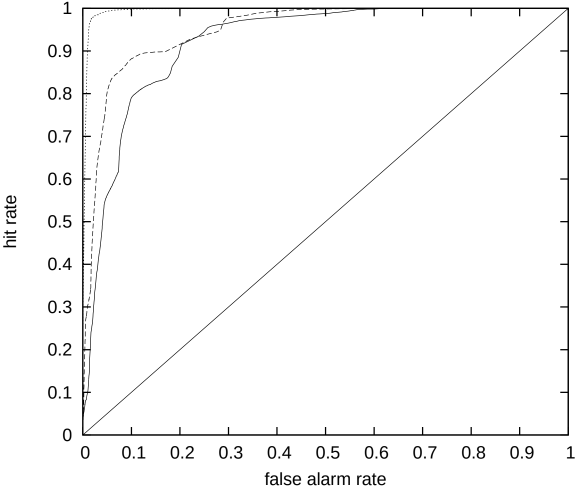


Model parameters

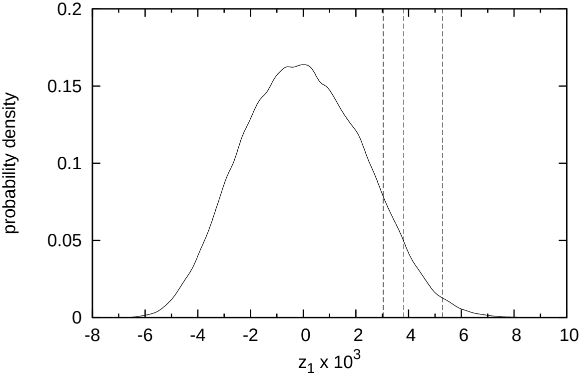
$b = 0.05$, $\tau = 25$, $K = 10$:

$w_1 = 0.124$, $\rho_1 = 0.340$
$w_2 = 0.073$, $\rho_2 = 0.038$
$w_3 = 0.082$, $\rho_3 = 0.021$
$w_4 = 0.104$, $\rho_4 = 0.018$
$w_5 = 0.078$, $\rho_5 = 0.010$
$w_6 = 0.126$, $\rho_6 = 0.003$
$w_7 = 0.158$, $\rho_7 = 0.000$
$w_8 = 0.103$, $\rho_8 = 0.000$
$w_9 = 0.081$, $\rho_9 = 0.000$
$w_{10} = 0.072$, $\rho_{10} = 0.000$

ROC curves for different event rarity; $\tau = 30$, $K = 15$



Large-scale extremes in a QG atmospheric model



Extreme model

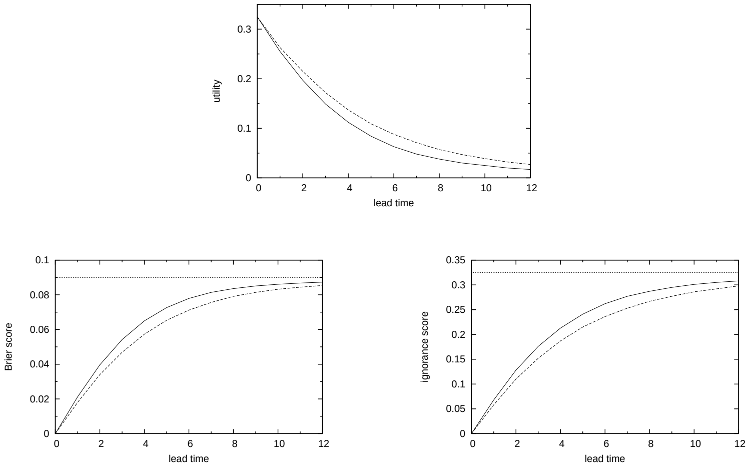
Gaussian clustering in EOF space:

$$p(\mathbf{z}|i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Gamma}_i)$$

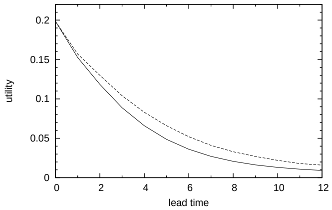
Logistic regression as predictive model:

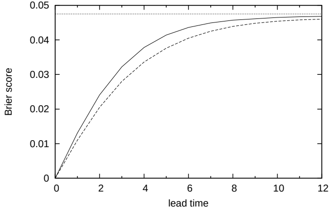
$$p(e_1^{\tau}|\mathbf{z}, i) = \frac{1}{1 + \exp(-\sum_{m=1}^{M} \beta_{i,m} z_m)}$$

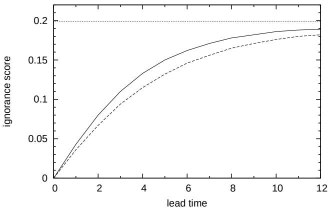
Utility, Brier score and ignorance score for $b = 0.1$



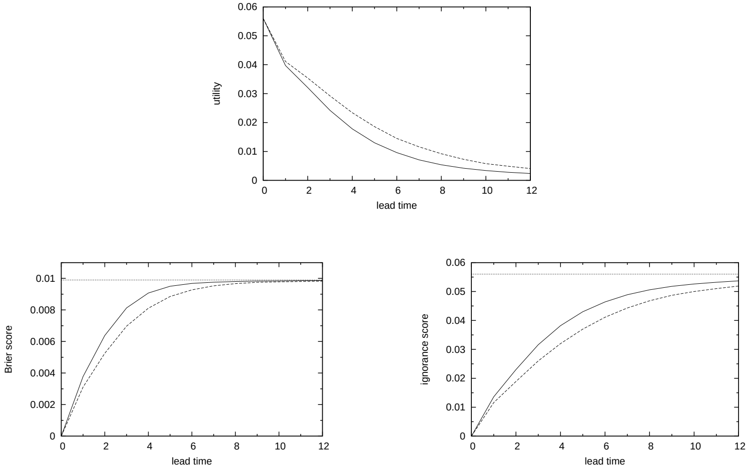
Utility, Brier score and ignorance score for $b = 0.05$



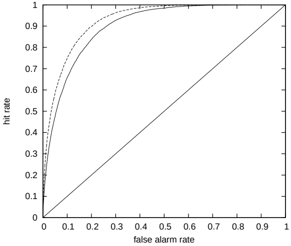


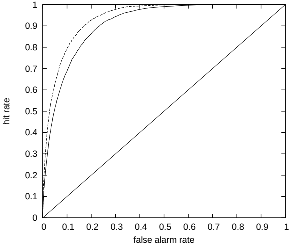


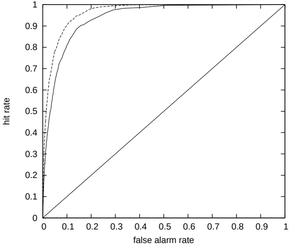
Utility, Brier score and ignorance score for $b = 0.01$



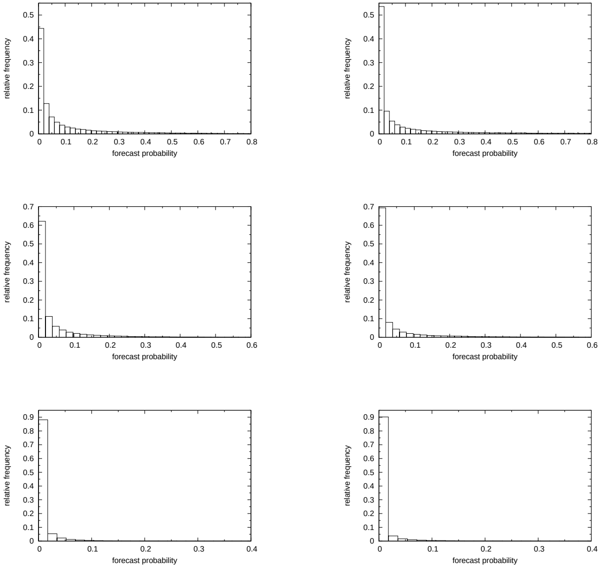
ROC curves for $b = 0.1$, $b = 0.05$ and $b = 0.01$







Distributions of forecasts



ROC curves for different event rarity; $b = 0.1$, $b = 0.05$ and $b = 0.01$