UiO : Department of Mathematics
University of Oslo

# Stochastic Modeling of Wave Climate Using a Bayesian Hierarchical Space-Time Model with a Log-Transform

*Presented at WAVE Workshop 2011*
*Hawaii, USA, October 31. 2011*

**Erik Vanem**, Arne Bang Huseby and Bent Natvig

# Motivation and background

- Ocean wave climate important to maritime safety
  - Bad weather account for a great number of ship losses and accidents
  - Severe sea state conditions taken into account in design and operation of ships and marine structures
- Possible trends in the wave climate may need to be taken into account
  - E.g. due to climate change
- A stochastic model for significant wave height in space and time are developed
  - Including a component for long-term trends
  - Fitted to data in the North Atlantic Ocean from 1958 – 2002

# Methodology – brief summary

- Bayesian hierarchical space-time model
  - Log-transformed data to account for heteroscedasity and heterogeneous trends
  - Bayesian framework to incorporate prior knowledge
- Observation model and different levels of state models
  - Spatial model: $1^{st}$ order Markov Random Field (MRF)
  - Space-time dynamic model: Vector autoregressive model
  - Seasonal model: spatially independent Gaussian process
  - Long-term trend model: Gaussian process with quadratic trend
    - Various model alternatives with linear and no trend also tried out
- Implemented by MCMC methods
  - Gibbs sampler with Metropolis-Hastings steps; full conditionals

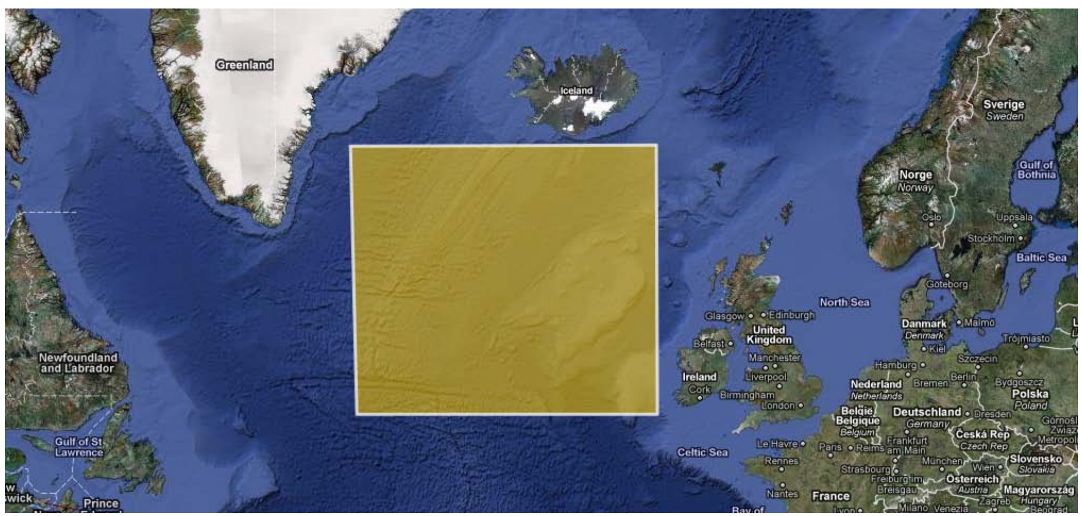
# Summary of conclusions

- Model seem to perform reasonably well overall
- Different long-term trends estimated by different model alternatives and using monthly of daily data
  - 16 – 31 cm (23-42 cm) for moderate conditions ($H_S \approx$ 3m)
  - 55 – 100 cm (76 – 140 cm) for extreme conditions ($H_S$ > 10m)
- Extrapolating the linear trends to give projections for 100 years
  - Expected increases within the range of 45-75 cm (53-90 cm) (moderate conditions) and 1.5 – 2.5 m (1.8 – 3.0 m) (extreme conditions) over 100 years
- Model selection inconclusive
- Uncertain whether the log-transform represent an improvement

# DETAILS OF THE STUDY

# Data and area description

- Corrected ERA-40 data of significant wave height(*)
  - Spatial resolution: 1.5° × 1.5° globally (some areas missing)
  - Temporal resolution: 6 hourly from Jan. 1958 to Feb. 2002 (44 years and 2 months = 64 520 points in time)
- Ocean area between 51° - 63°N and 324° - 348°E




(*) Data kindly provided by Royal Netherlands Meteorological Institute (KNMI), Dr. Andreas Sterl

# Model description – Main model

- Significant wave height at location x, time t: Z(x, t)
- Logarithmic transform: Y(x, t) = ln Z(x, t)
- Observation model:

$$Y(x, t) = H(x, t) + \varepsilon_Y(x, t)$$

with

$$H(x, t) = \mu(x) + \theta(x, t) + M(t) + T(t) \quad \text{and } \varepsilon_Y(x, t) \sim^{\text{i.i.d}} N(0, \sigma_Y^2)$$

- Alternative representation on original scale

$$Z(x, t) = e^{\mu(x) + \theta(x, t) + M(t) + T(t) + \varepsilon_Y(x, t)} = e^{\mu(x)} e^{\theta(x, t)} e^{M(t)} e^{T(t)} e^{\varepsilon_Y(x, t)}$$

- Various components represents multiplicative factors on the original scale
  - NB: Need to consider bias correction when retransforming to original (missing in the paper)
- All noise terms in the model assumed independent

# Time independent, spatial model

- $1^{st}$ order Markov Random Field

$$\mu(x) = \mu_0(x) + \alpha_\varphi \left\{\mu(x^N) - \mu_0(x^N) + \mu(x^S) - \mu_0(x^S)\right\}$$
$$+ \alpha_\lambda \left\{\mu(x^E) - \mu_0(x^E) + \mu(x^W) - \mu_0(x^W)\right\} + \varepsilon_\mu(x)$$

with the spatially specific mean,

$$\mu_0(x) = \mu_{0,1} + \mu_{0,2} m(x) + \mu_{0,3} n(x) + \mu_{0,4} m(x)^2 + \mu_{0,5} n(x)^2 + \mu_{0,6}\, m(x)n(x)$$

- $x^D$ = location of the nearest grid-point in direction D = N, S, W, E
- m(x), n(x) = longitude and latitude of location x
- $\alpha_\varphi$, $\alpha_\lambda$ : spatial dependence parameters in lateral and longitudinal directions
- $\varepsilon_\mu(x) \sim^{i.i.d} N(0, \sigma_\mu^2)$

# Short-term spatio-temporal model

- 1st order vector autoregressive model

$$\theta(x, t) = b_0\theta(x, t-1) + b_N\theta(x^N, t-1) + b_E\theta(x^E, t-1) + b_S\theta(x^S, t-1) + b_W\,\theta(x^W, t-1) + \varepsilon_\theta(x, t)$$

- Vector autoregressive parameters $b_0$, $b_N$, $b_E$, $b_S$ and $b_W$ assumed invariant in space
- $\varepsilon_\theta(x, t) \sim^{i.i.d} N(0, \sigma_\theta^2)$

# Spatially independent seasonal model

- Modeled as an annual cyclic Gaussian process

$$M(t) = c \cos(\omega t) + d \sin(\omega t) + \varepsilon_m(t)$$

- Seasonal parameters c and d assumed invariant in space
- $\omega$ related to the period of the annual sycle, e.g. $\omega = \pi/6$ for monthly data
- $\varepsilon_m(t) \sim^{i.i.d} N(0, \sigma_m^2)$
- The effect of including a semi-annual component ($2^{nd}$ harmonic) was also investigated but found to be small

# Long-term trend model

- Gaussian process with quadratic trend

$$T(t) = \gamma t + \eta t^2 + \varepsilon_T(t)$$

- $\varepsilon_T(t) \sim^{i.i.d} N(0, \sigma_T^2)$
- Model alternatives:

Model 1: $T(t) = \gamma t + \eta t^2 + \varepsilon_T(t)$ (quadratic trend model)

Model 2: $T(t) = \gamma t + \varepsilon_T(t)$ (linear trend model)

Model 3: $T(t) = 0$ (no trend model)

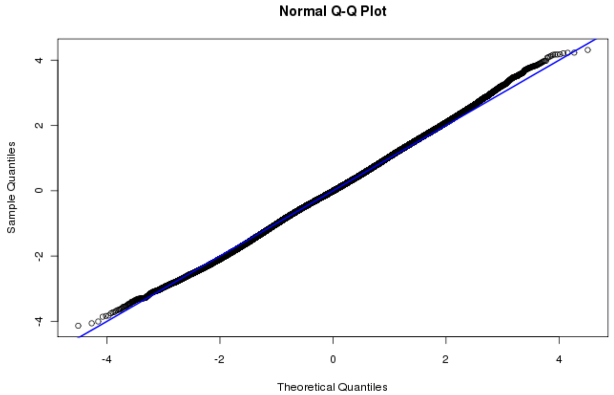
Model 4: $M(t) = c \cos(\omega t) + d \sin(\omega t) + \gamma t + \eta t^2 + \varepsilon_m(t); T(t) = 0$

Model 5: $T(t) = c \cos(\omega t) + d \sin(\omega t) + \gamma t + \varepsilon_m(t); T(t) = 0$

# MCMC simulations

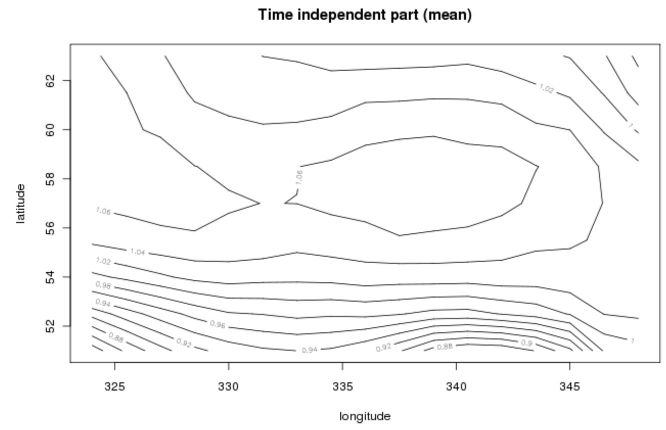
- MCMC techniques used to simulate from the model
  - Gibbs sampler with Metropolis-Hastings steps
  - 1000 samples of the parameter vector with 20,000 burn-in iterations and batch size 25 (monthly data) or 5 (daily data)
  - Convergence likely by visual inspection of trace plots, control runs with longer burn-in and different starting values
  - Plot of the residuals indicate that model assumptions are reasonable

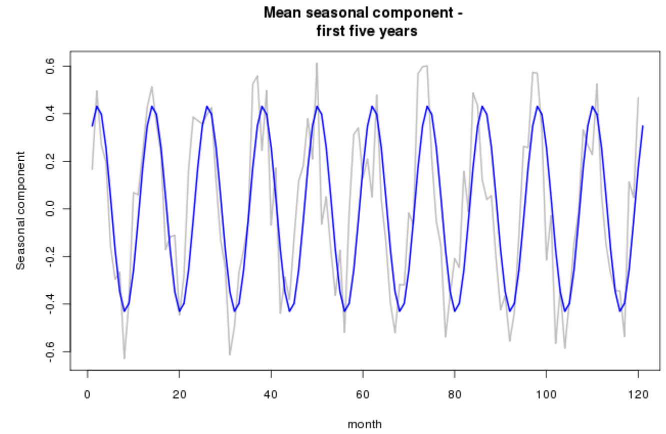
Normal probability plot of the residuals (monthly data):



# Results and predictions

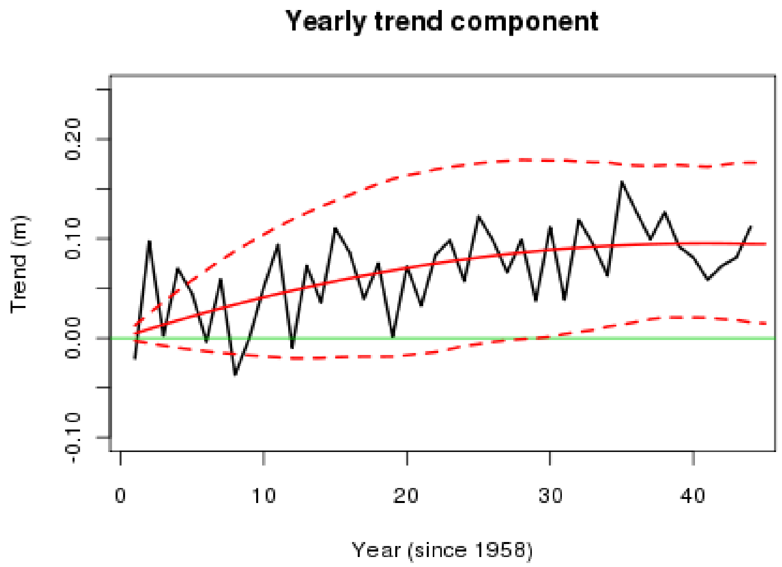
- Spatial, space-time dynamic and seasonal models perform well, with factors (monthly data)
  - $e^{\mu(x)}$ ~ 2.3 – 2.9
  - $e^{\theta(x,t)}$ ~ 0.77 – 1.5
  - $e^{M(t)}$ ~ 0.65 – 1.5 (0.67 – 1.6)
- θ(x, t) becomes more important for daily data
- Figures show spatial field and seasonal component on transformed scale (monthly data)

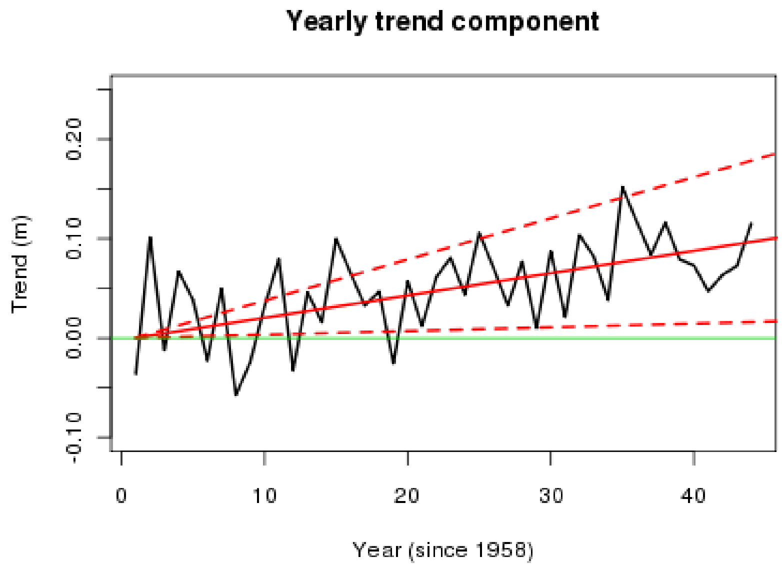


# Results – Example of estimated trends

- Quadratic and linear model, monthly data (transformed scale)



# Results – estimated expected trends

| | Normal conditions ($H_S \approx 3$ m) | | Extreme conditions ($H_S > 10$ m) | |
|---|---|---|---|---|
| | Monthly data | Daily data | Monthly data | Daily data |
| Model 1 | 30 cm (40 cm) | 22 cm (28 cm) | 1.0 m (1.3 m) | 73 cm (95 cm) |
| Model 2 | 31 cm (42 cm) | 22 cm (28 cm) | 1.0 m (1.4 m) | 72 cm (95 cm) |
| Model 4 | 19 cm (27 cm) | 16 cm (23 cm) | 63 cm (91 cm) | 55 cm (76 cm) |
| Model 5 | 26 cm (35 cm) | 19 cm (26 cm) | 88 cm (1.2 m) | 65 cm (87 cm) |

* Red values correspond to updated results with bias correction

# Future projections – 100 year trends

- Future projections made by extrapolating the linear trends (somewhat speculative)
- Critical assumption – estimated trend will continue into the future

| | Normal conditions ($H_S \approx 3$ m) | | Extreme conditions ($H_S > 10$ m) | |
|---|---|---|---|---|
| | Monthly data | Daily data | Monthly data | Daily data |
| Model 2 | 0.75 m (0.90 m) | 0.50 m (0.59 m) | 2.5 m (3.0 m) | 1.7 m (2.0 m) |
| Model 5 | 0.63 m (0.74 m) | 0.45 m (0.53 m) | 2.1 m (2.5 m) | 1.5 m (1.8 m) |

* Red values correspond to updated results with bias correction

# Model comparison and selection

- Two loss functions used for model selection, based on (short-term) predictive power
  - Standard loss function and weighted loss function
- Model selection remains inconclusive

| | Monthly data | | Daily data | |
|---|---|---|---|---|
| | $L_S$ | $L_W$ | $L_S$ | $L_W$ |
| Model 1 | 3.4119453 | 3.5604366 | 2.5615432 | 2.6654849 |
| Model 2 | 3.4247630 | 3.5459232 | 2.5729260 | 2.6842552 |
| Model 3 | **3.2667590** | 3.4683411 | 2.6002640 | 2.7280551 |
| Model 4 | 3.3168082 | 3.4679681 | **2.5569820** | **2.6550046** |
| Model 5 | 3.2979152 | **3.4545460** | 2.5692487 | 2.6816735 |

# Discussion and concluding remarks

- A Bayesian hierarchical space-time model for log-transformed significant wave height data has been presented
- Estimated expected long-term trends (1958-2002):
  - 16 – 31 cm (23-42 cm) for moderate conditions ($H_S \approx 3$m)
  - 55 – 100 cm (76 – 140 cm) for extreme conditions ($H_S > 10$m)
- Estimated expected future projections (100 years):
  - Between 45-75 cm (53-90 cm) (moderate conditions) and 1.5 – 2.5 m (1.8 – 3.0 m) (extreme conditions) over 100 years
- Trends for moderate conditions comparable to trends estimated without the log-transform
- Difficult to evaluate model alternatives – model selection inconclusive
- Possible model extensions could include regression terms with relevant meteorological parameters as covariates