Bayesian hierarchical space time model applied to high-resolution hindcast data of significant wave height

Presented at Wave workshop 2013, Banff, Canada

Erik Vanem
31.10.2013

MANAGING RISK DNV

Introduction

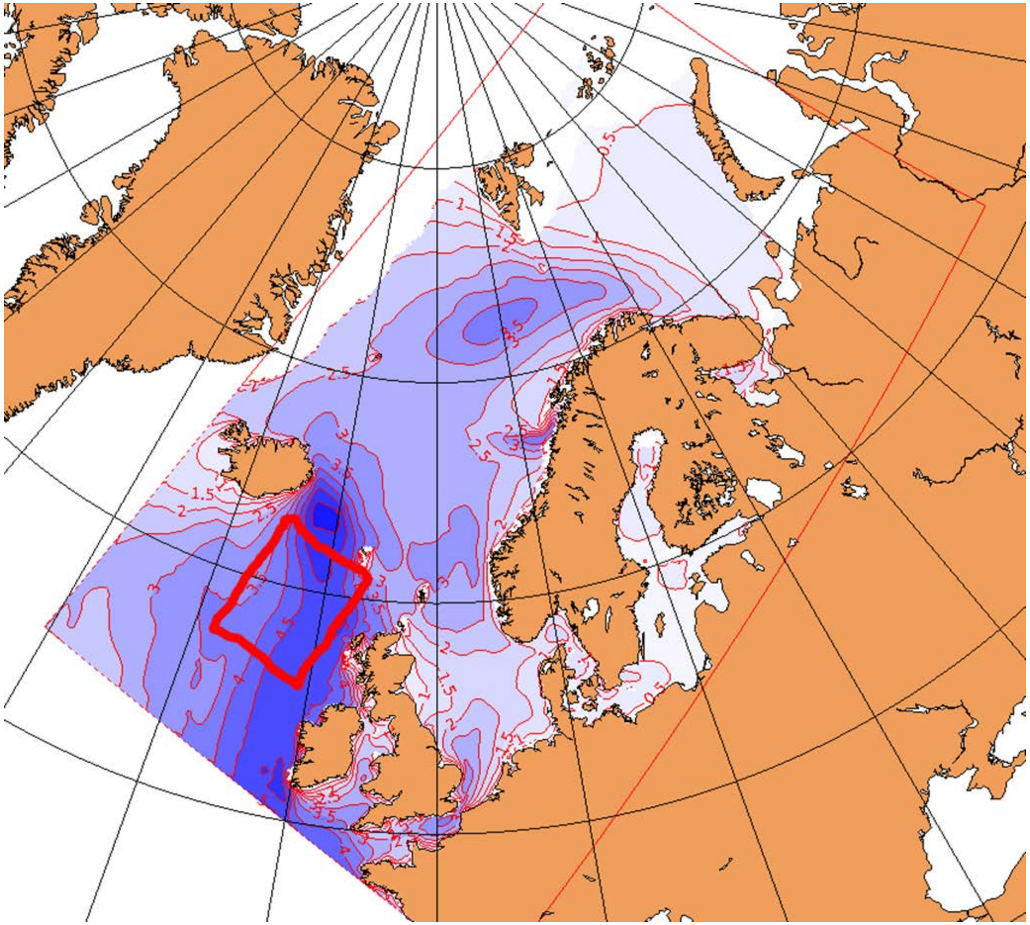
- The Bayesian hierarchical space-time models were developed within a recent PhD-project at UiO and successfully applied to C-ERA-40 data of significant wave height
  - Increasing long-term trends were identified
- Since then, improved data on SWH have become available – NORA10
  - Improved spatial resolution
  - Reduced bias and lower RMSE
- The same Bayesian hierarchical model has been applied to NORA10 data of SWH and results will be compared
  - Uses a regression component on CO_2 levels for long-term trends

Summary of conclusions

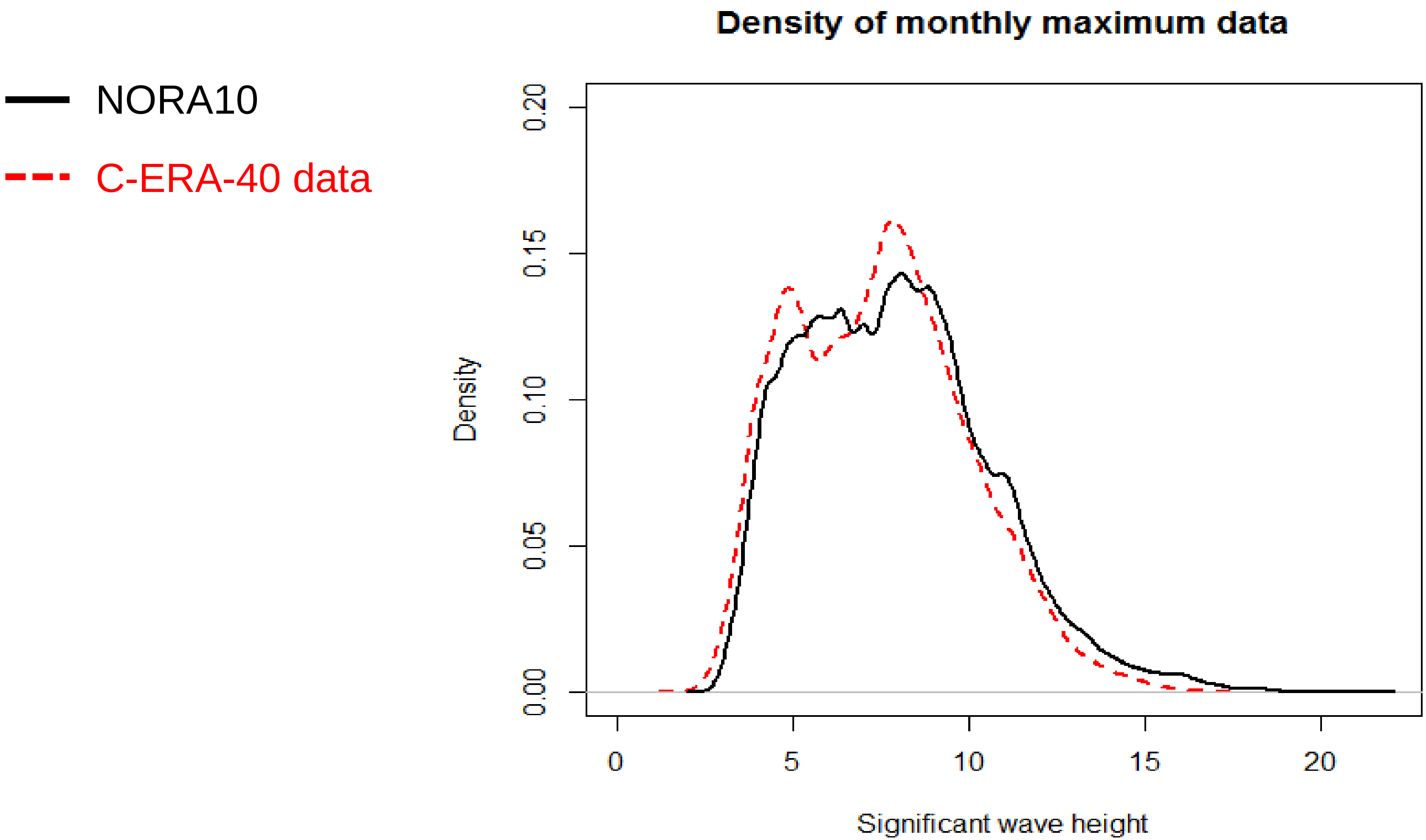
- Results seem to contradict previous results from C-ERA-40 data in that there are no statistically significant trends in the data

NORA10 - Data description

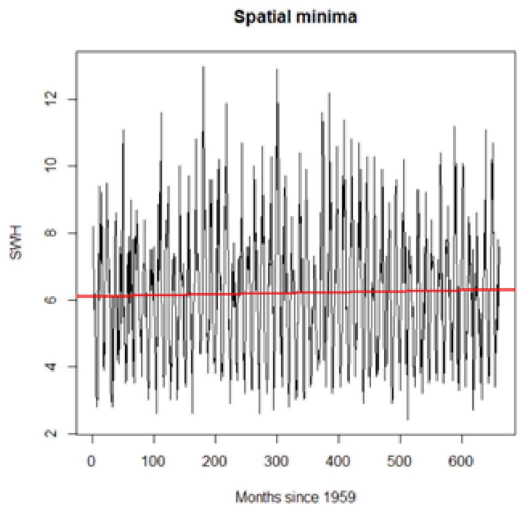
- Combined high-resolution atmospheric downscaling and wave hindcast based on ERA-40
  - 3 hourly wave fields
  - 10 km spatial resolution
  - Area in the northeast Atlantic
  - Covers the period 1958 – 2012
- Monthly maximum data
  - Overall mean: 7.9 m **(C-ERA-40: 7.5m)**
  - Min value: 2.4 m **(C-ERA-40: 1.9 m)**
  - Max value: 21.7 m **(C-ERA-40: 17.1 m)**
  - Higher than C-ERA-40 for all months
    - Difference up to 1.91 m for March

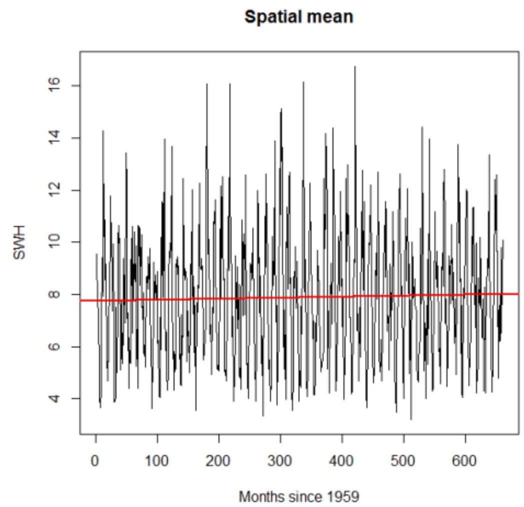


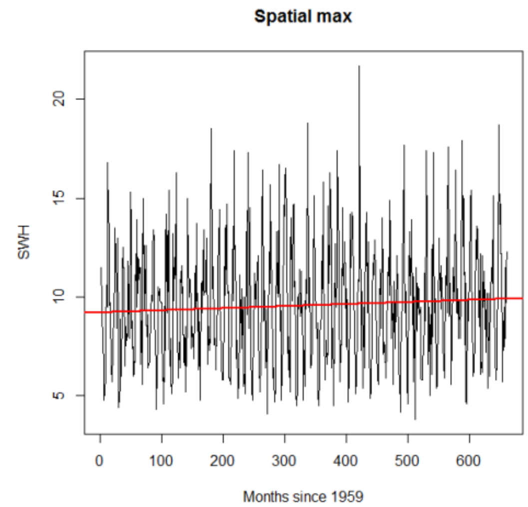
Density of monthly maximum data



Initial trend analysis – fitting a straight line by least squares



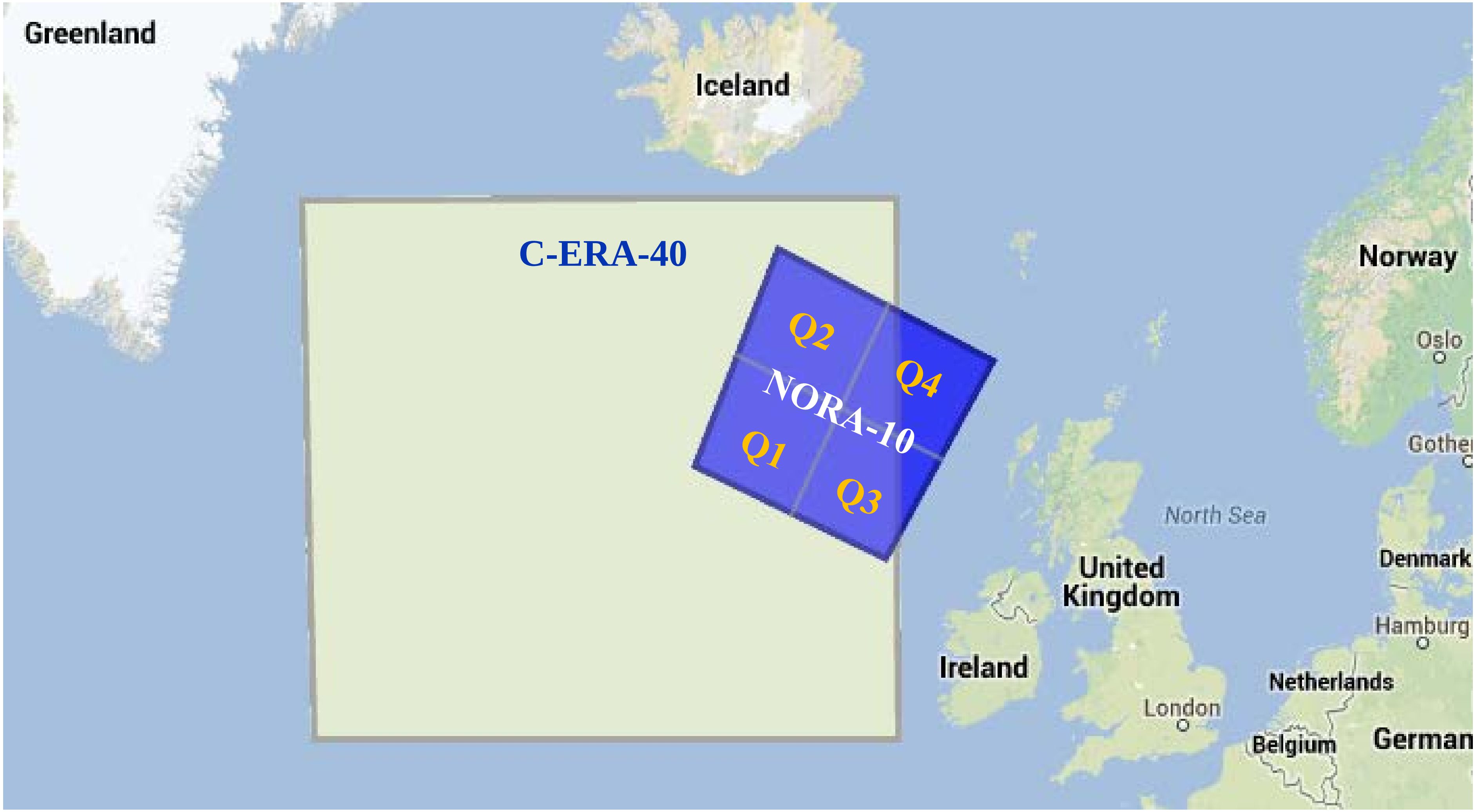




	Spatial min	*Spatial mean*	*Spatial max*
Intercept	6.1058	7.7474	9.2130
Slope	0.0003138	0.0004393	0.001056
p-value of the slope	0.398	0.463	0.0854
Annual trend (cm)	0.3765	0.5272	1.268
Accumulated trend (1959 – 2012) (cm)	20.71	28.99	69.72

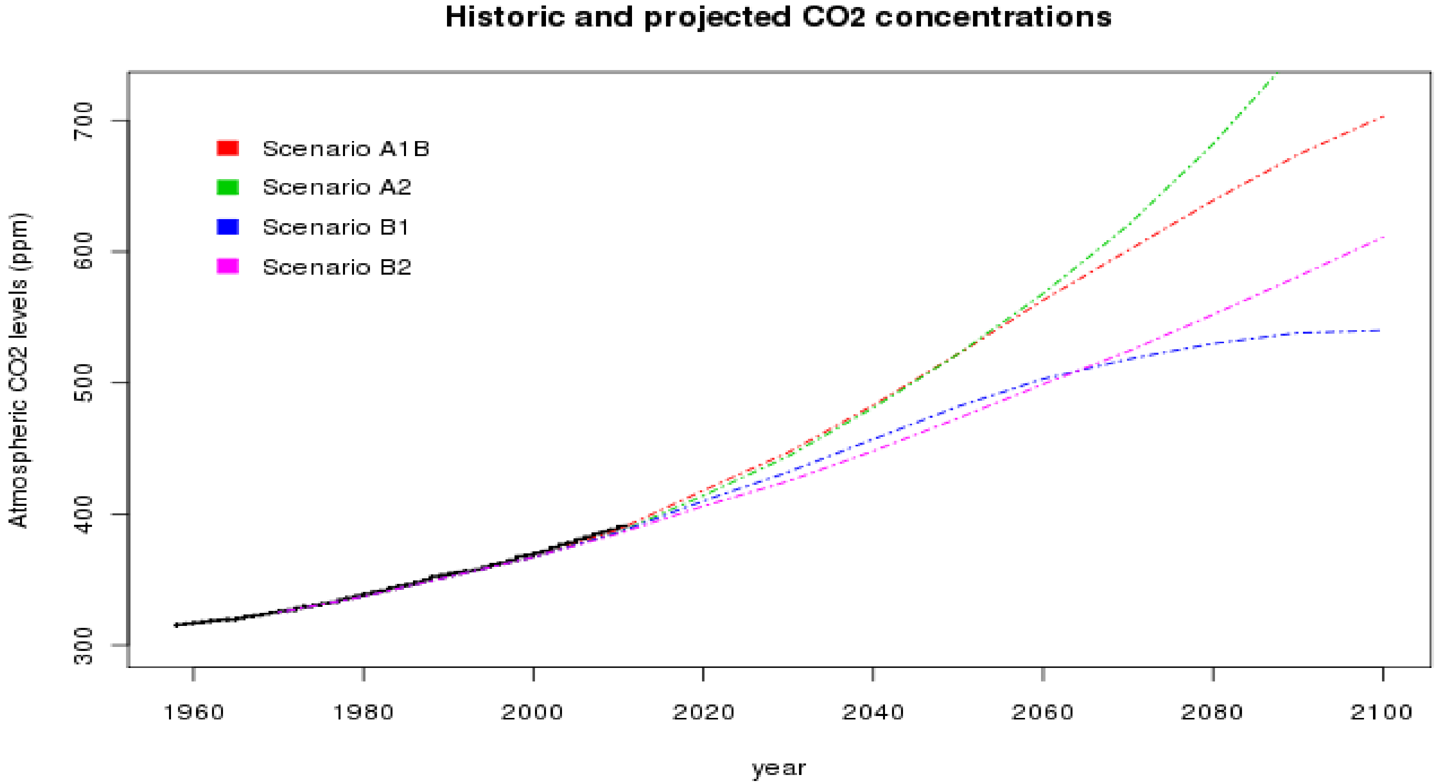
MANAGING RISK

Area description



CO_2 data as covariates

- Historical data for model fitting
- Future projections for predictions
  - A2 and B1 IPCC emission scenarios



The Bayesian hierarchical space-time model

$$Z(x, t) = H(x, t) + \varepsilon_Z(x, t)$$

With

$$H(x, t) = \mu(x) + \theta(x, t) + M(t) + T(t)$$

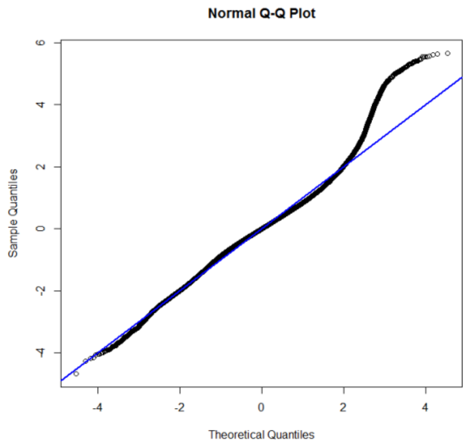
and

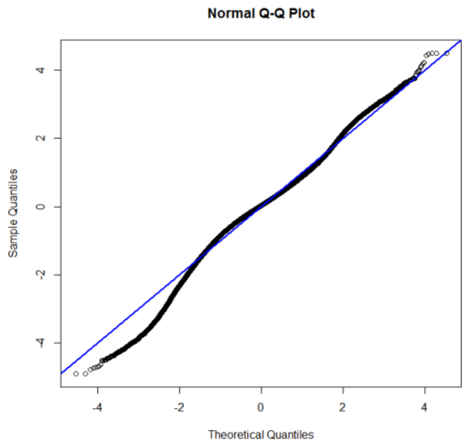
$$\varepsilon_Z(x, t) \sim^{i.i.d} N(0, \sigma_Z^2)$$

- $\mu(x) = \mu_0(x) + a_\varphi \{\mu(x^N) - \mu_0(x^N) + \mu(x^S) - \mu_0(x^S)\} + a_\lambda \{\mu(x^E) - \mu_0(x^E) + \mu(x^W) - \mu_0(x^W)\} + \varepsilon_\mu(x)$
  - $\mu_0(x) = \mu_{0,1} + \mu_{0,2} m(x) + \mu_{0,3} n(x) + \mu_{0,4} m(x)^2 + \mu_{0,5} n(x)^2 + \mu_{0,6}\, m(x) n(x)$
- $\theta(x, t) = b_0 \theta(x, t-1) + b_N \theta(x^N, t-1) + b_E \theta(x^E, t-1) + b_S \theta(x^S, t-1) + b_W\, \theta(x^W, t-1) + \varepsilon_\theta(x, t)$
- $M(t) = c \cos(\omega t) + d \sin(\omega t) + f \cos(2\omega t) + g \sin(2\omega t) + \varepsilon_m(t)$
- $T(t) = \gamma G(t) + \eta \ln G(t) + \varepsilon_T(t)$

Simulation setup

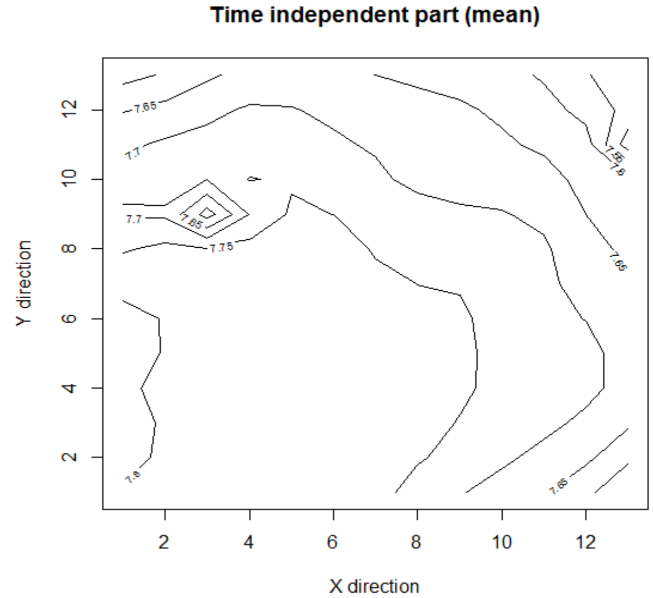
- Five simulations – complete area and four sub-areas
  - 13 x 13 = 169 locations and 648 monthly maxima: 109 512 data points in space and time
- 60 000 iterations to obtain 1000 samples of the posterior
  - Burn-in period of 40 000 samples
  - Batch size of 20
- Gibbs sampler and Metropolis-Hastings steps
  - MH steps repeated 6 times every iteration – 73-78% acceptance rate
- QQ-plots OK, but not perfect

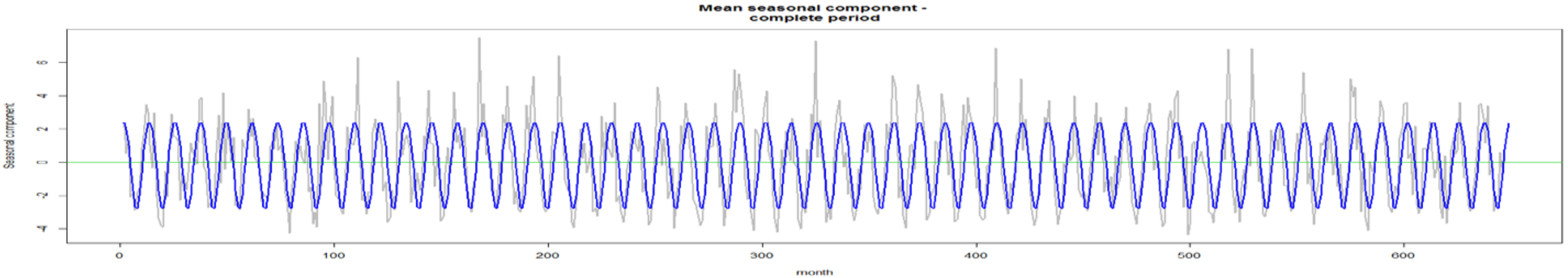




Results – Complete area

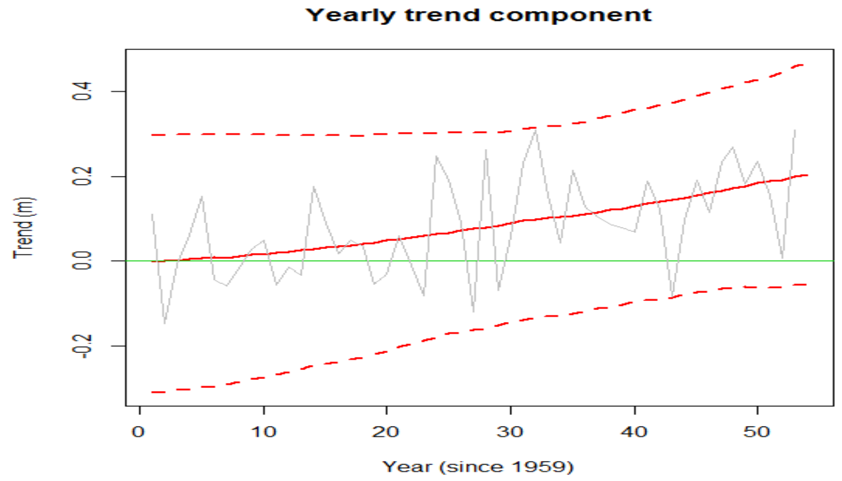
- Mean time independent part: spatial variation between 7.5 – 7.8 m (average 7.7 m)
  - Less spatial variability than C-ERA-40
  - Smaller geographical area, so this is reasonable
- Mean space-time interaction part: Contributions from -2.1 to 1.4 m
  - Slightly larger than for C-ERA-40
  - Higher spatial resolution, so this is reasonable
- Mean seasonal part between -3.0 and 2.5 m
  - Annual variation of 5.5 m
  - Slightly larger than for C-ERA-40

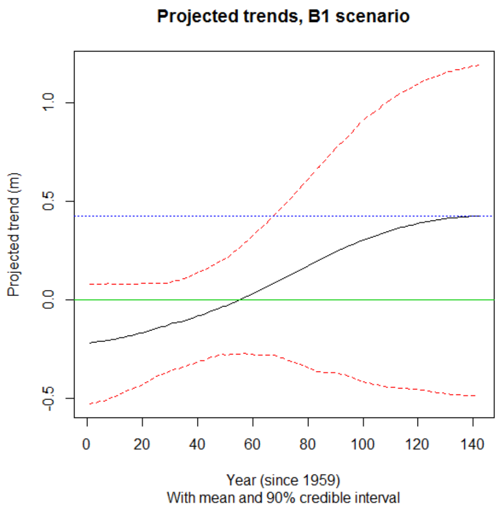


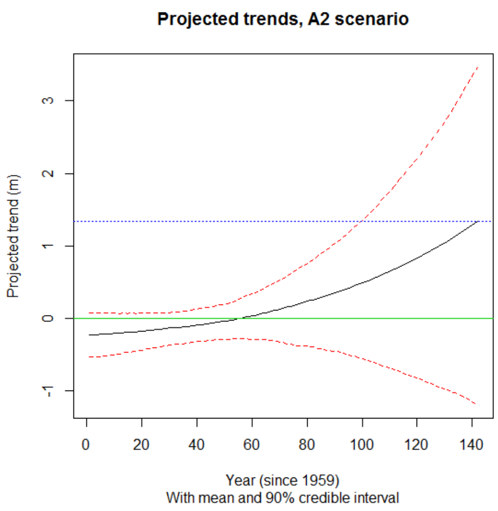


Simulation results – trends and future projections

- Mean increasing trend: 21cm
- Not statistically significant at 90% level
  - 90% credible interval from -5.6 – 48 cm
  - C-ERA-40 trends were statistically significant and stronger
- Future projections towards 2100
  - Increasing but not statistically significant
  - B1 scenario: mean 0.43 m
    - 90% credible interval: -0.48 – 1.2 m
  - A2 scenario: mean 1.4 m
    - 90% credible interval: -1.2 – 3.5 m
  - Less than for C-ERA-40 data

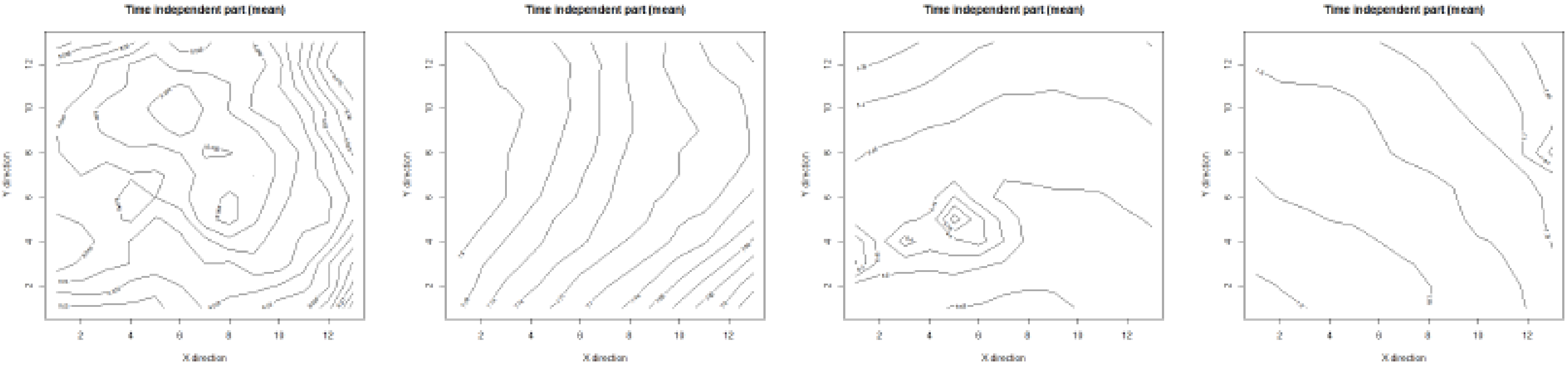






Results – sub-areas Q1 – Q4

- Reduced spatial coverage but increased spatial resolution
- Results generally consistent with the results for the complete area except perhaps for the average level of the spatial fields
  - Spatial variability in the fields seem OK, but the average level is perhaps not estimated correctly
- Other components are consistent across the simulations



Results for long-term trends – sub-areas Q1 – Q4

- Long-term trends increasing but not statistically significant for Q1 – Q3
  - Statistically significant increase for Q4: 35 cm with 90% credible interval 2.5 – 62 cm
- Expected future projections increasing, but not statistically significant
- Trends and projections for the complete area falls within the estimates for the four sub-areas and are hence consistent

TABLE 5. Range of estimated values for the different model components (m)

	Complete area	Q1	Q2	Q3	Q4
Spatial field	7.5 - 7.8	8.0 - 8.1	7.6 - 7.8	6.3 - 6.6	7.6 - 7.9
$\theta(x, t > 0)$	-2.1 - 1.4	-1.1 - 0.70	-0.72 - 0.81	-1.6 - 1.2	-0.95 - 0.97
Seasonal component	-3.0 - 2.5	-3.0 - 2.5	-3.0 - 2.5	-2.9 - 2.4	-3.0 - 2.5
Long-term trend	0.21	0.04	0.16	0.11	0.35
A2 projections	1.4	0.27	1.0	0.61	2.3
B1 projections	0.43	0.09	0.33	0.20	0.72
Trend: 90% c.i.	-0.056 - 0.48	-0.23 - 0.33	-0.14 - 0.54	-0.14 - 0.32	0.025 - 0.62
A2: 90% c.i.	-1.2 - 3.5	-2.5 - 3.2	-1.4 - 4.1	-2.5 - 3.2	-0.97 - 4.7
B1: 90% c.i.	-0.48 - 1.2	-0.95 - 1.1	-0.6 - 1.6	-0.97 - 1.2	-0.51 - 1.7

Some control runs

- To check convergence, some control runs with longer burn-in period were performed
  - Complete area, Q3 and Q4
- Results are consistent for all components except the spatial fields
  - Spatial variability is similar, but the average value is very different
  - Trace plots of the spatial fields indicate that this component has not converged sufficiently
  - Slightly different expected trends, but with higher uncertainty bands.
  - Trends are still statistically significant for the sub-area Q4. Otherwise, increasing but not statistically significant trends

Discussion and comparison of results

- Overall, results are found to be internally consistent
  - Except the average level of the mean spatial fields
    - Could this be explained by lack of convergence?
    - Is the model over-parameterized so that many solutions give reasonable fit?
    - Is the Gaussian assumption unrealistic?
- Compared to C-ERA-40 data, NORA10 data display less long-term trends
  - C-ERA-40 trends are statistically significant, whereas NORA10 trends are not
  - Could it be explained by the different geographical areas?
  - Could it be explained by the different temporal span?
    - NORA10 data contains 10 more years of data
    - Crude investigation of the data indicate that for the spatial mean and max time series, there is an increasing trend from 1958 – 2002, followed by a VERY slight decrease from 2002 – 2012 (none statistically significant)
      - However, NORA10 increases are not statistically significant for the period 1958 – 2002 either
- A similar analysis of ERA-40 wind speeds indicated that there are no trends in the monthly maximum wind speed over the area
  - Could trends have been overlooked in the NORA10 data due to this?

Summary and conclusions

- Bayesian hierarchical space-time models have been applied to NORA10 data
- Results seem to contradict previous results from C-ERA-40 data in that there are no statistically significant trends in the data
- How can this be explained?
  - In particular since NORA10 data is derived from ERA-40 data
  - ERA-40 data of SWH indicate statistically significant increasing trends over the area
  - ERA-40 data of wind speed indicate no trend over the area
- Future projections are highly uncertain
  - Expected future trends are increasing, but are not statistically significant
  - 90% credible intervals range from negative to positive trends and future projections

Safeguarding life, property and the environment

www.dnv.com

MANAGING RISK

DNV