Regional Frequency Analysis for spatial extreme value analysis of ocean waves

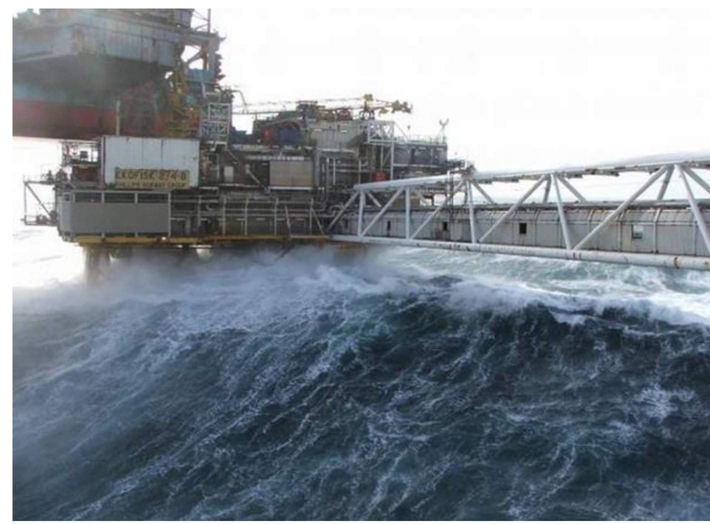
Wave Workshop 2019

Erik Vanem

14 November 2019

Motivation and background

- Reliable statistical description of extreme ocean environments are important for the design of ships and other marine structures
  - Typically interested in 20-, 25- or 100-year return values of relevant parameters, e.g. significant wave height, H_S
  - Can be achieved by establishing a probability distribution based on historical data
- Uncertainties become large for large return periods compared to available data records
  - May utilize information in data collected at different, but similar sites
- Return value estimates for specific locations where there are no data
- Spatial changes due to climate change



Introduction

- Regional Frequency Analysis is a well established methodology in various geosciences, in particular hydrology
  - Not much used for ocean waves
- The main idea is to identify homogeneous regions and use all data in those regions to more accurately estimate the quantile function of a random variable
- Main assumption in the index flood (or index wave) approach:
  - The frequency distributions at locations within homogeneous regions are identical apart from a location-specific scaling factor – the index wave

$$Q_i(F) = \mu_i q(F)$$

$Q_i(F)$ - quantile function at site i

μ_i - index wave at location i

$q(F)$ - regional growth curve for the homogeneous region

F - cumulative probability distribution

Wave data

- RFA applied to significant wave height over the North Atlantic
  - Spatial grid of 142 x 113 grid points
  - Use annual maxima
- Wave climate data obtained from ExWaCli project, based on EC-EARTH model output and WAM wave model
  - One historical run (1970 – 1999)
  - Future projections for RCP 4.5 and RCP 8.5 scenarios (2071 – 2100)

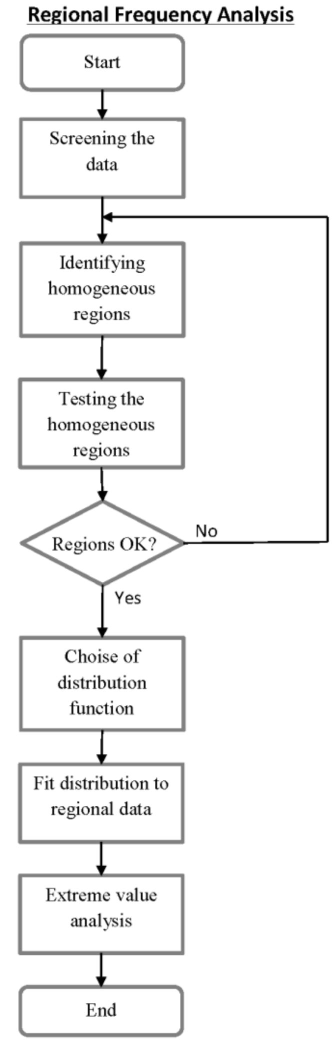
Main steps of a RFA

1. Screening the data
   a. Remove obvious errors, look for grossly discordant sites
2. Identification of homogeneous regions
   a) Clustering based on attributes, watershed filter
   b) Testing regions for homogeneity/heterogeneity, *H*-statistic
   c) Possibly reassign locations
3. Choice of frequency distribution
   a) Goodness-of-fit tests, intuition
4. Estimation of frequency distribution
   a) Method of L-moments

-> Extreme value analysis

<u>Modelling assumptions</u>

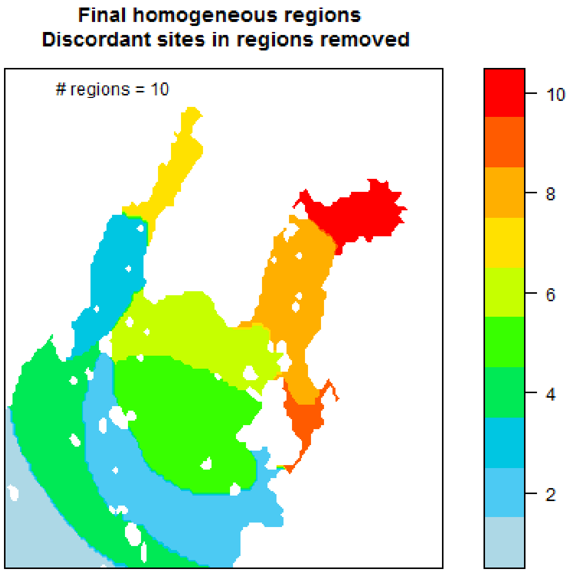
– Observations at a particular location are iid
– Observations at different sites are independent



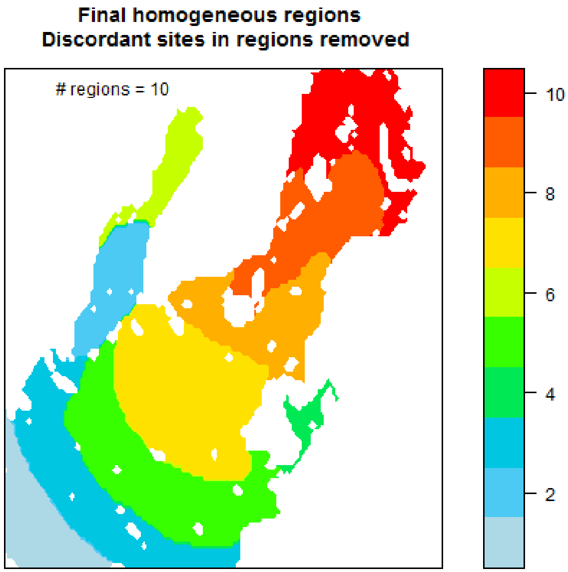
Analysis and results

Identifying Homogeneous regions

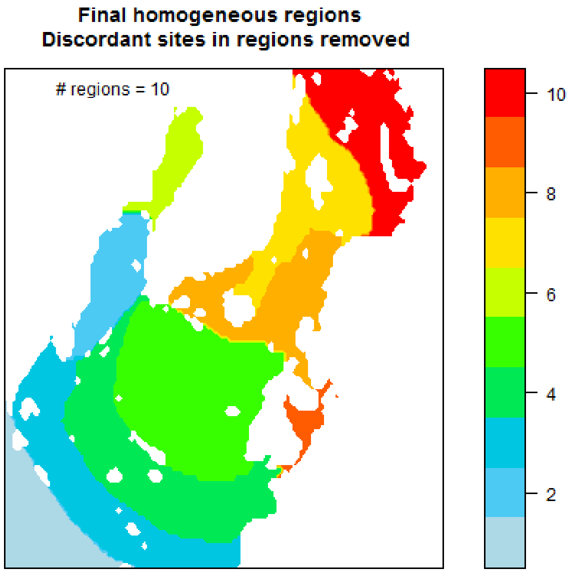
- Selecting number of regions (partly subjective) -> 10, 20 and 30 regions tried
- Assigning sites to regions and testing
  - Clustering based on attributes + watershed filter
  - Test statistics based on L-moments



Historical data



RCP 4.5 data



RCP 8.5 data

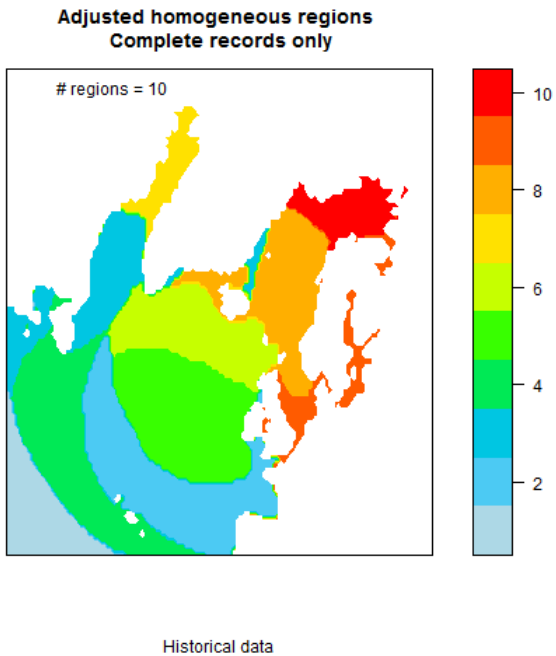
Effect of removing regionally clearly discordant sites

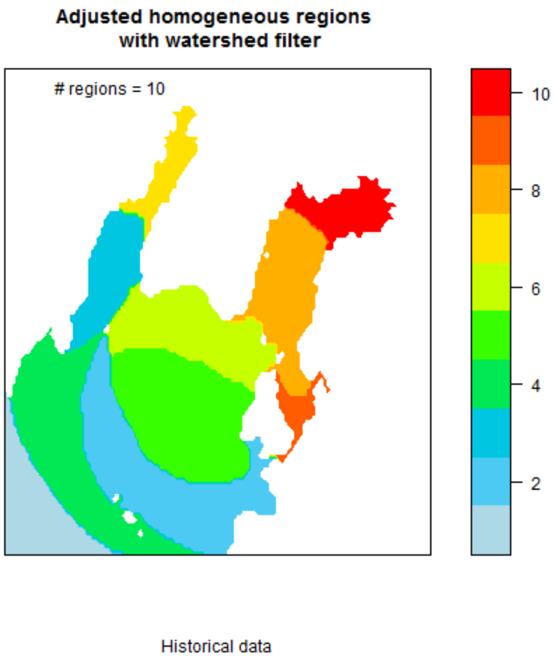
10 regions for Historical data

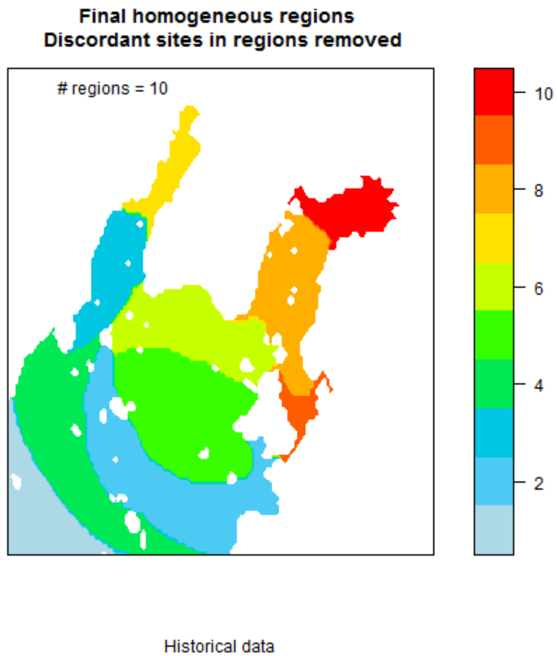
A. Regions identified by hierarchical clustering and K-means clustering on attributes

B. Applying watershed filter

C. Removing clearly discordant sites

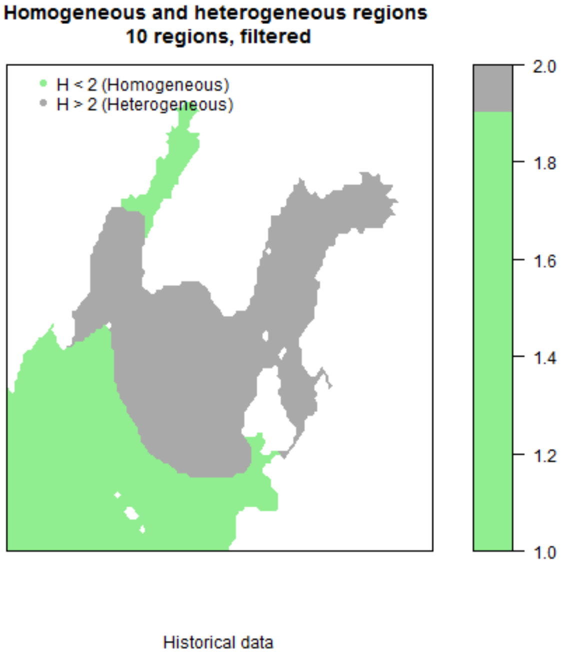


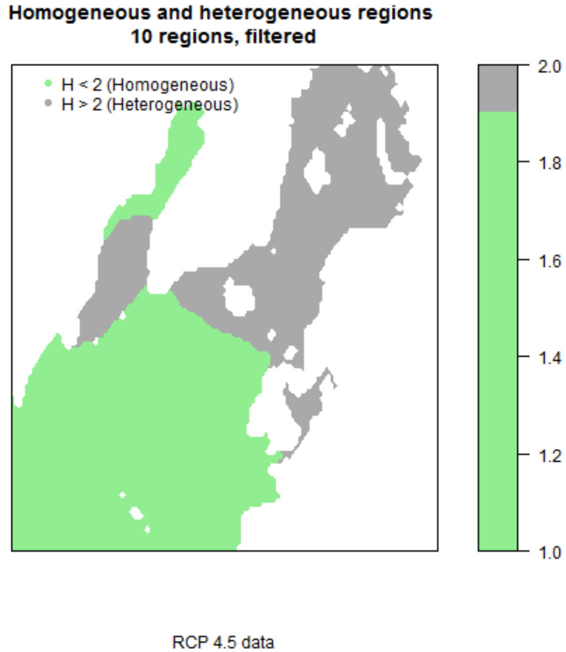


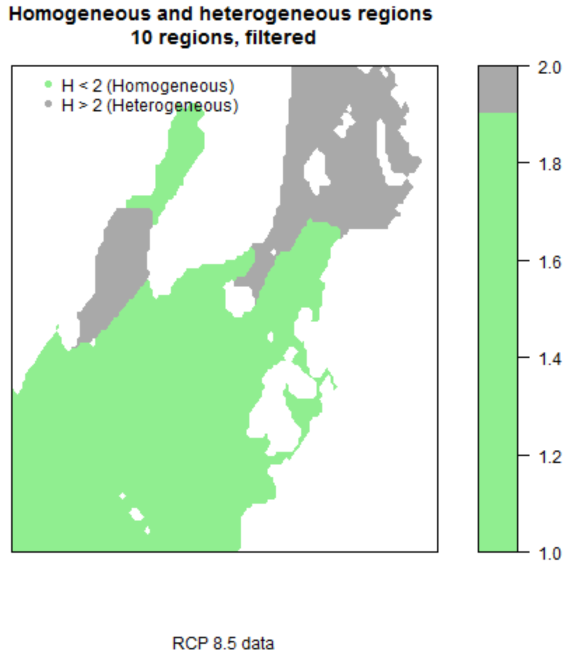


Testing for homogeneity

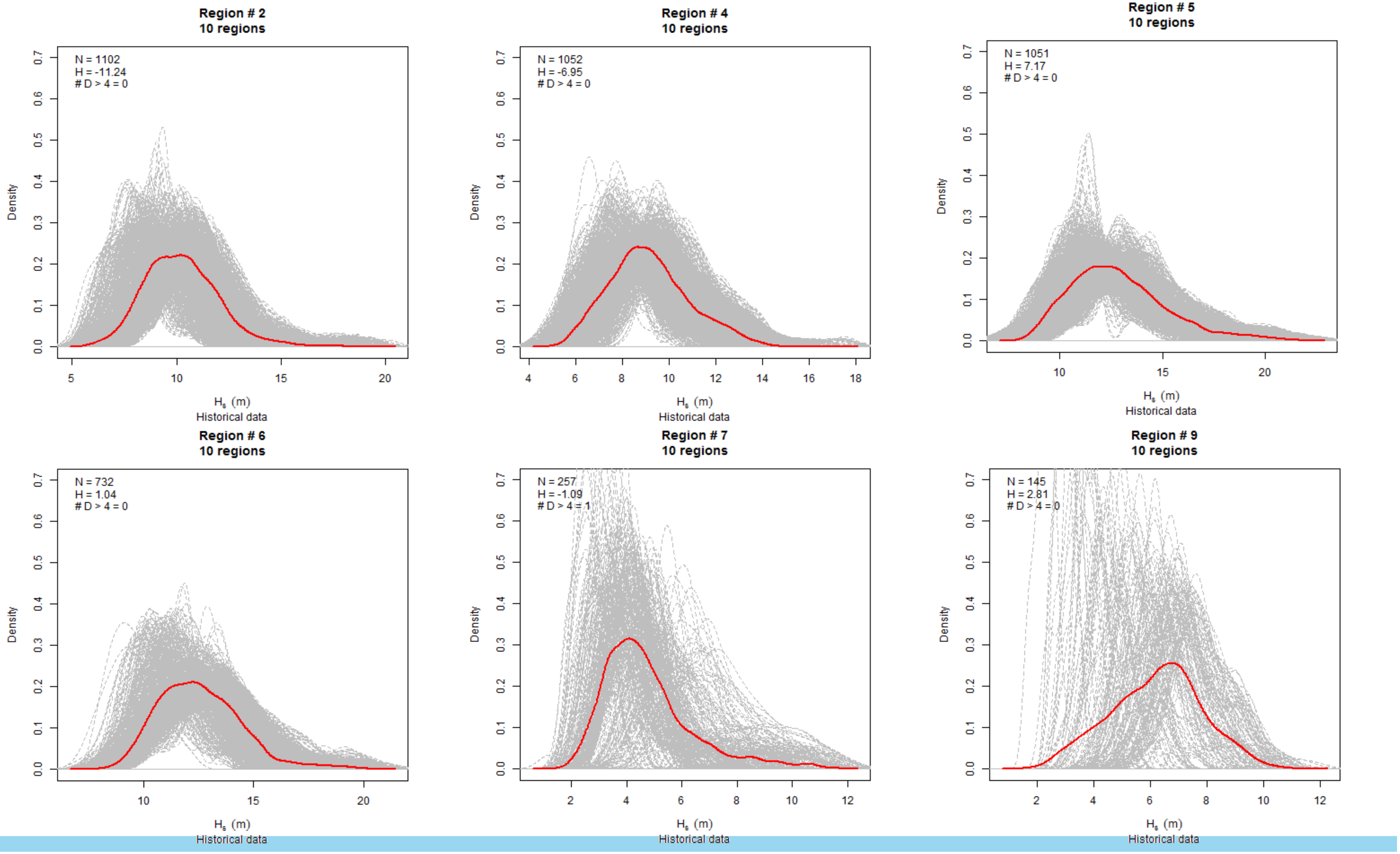
- Test for homogeneity based on the *H*-statistic
  – H ≥ 2: Possibly heterogeneous
  – H < 2: Reasonably homogeneous
  – Except for historical data, the regions appear to be fairly homogeneous in most areas South of Greenland







Comparing densities for locations within regions

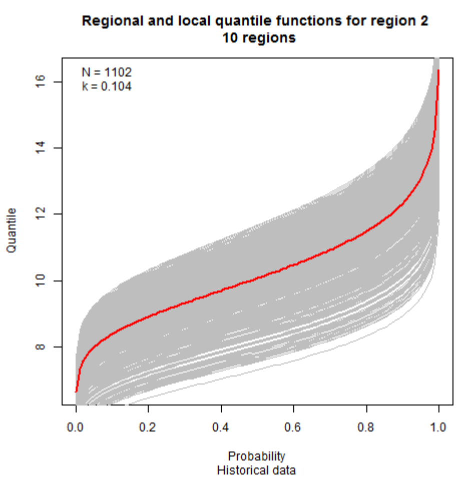


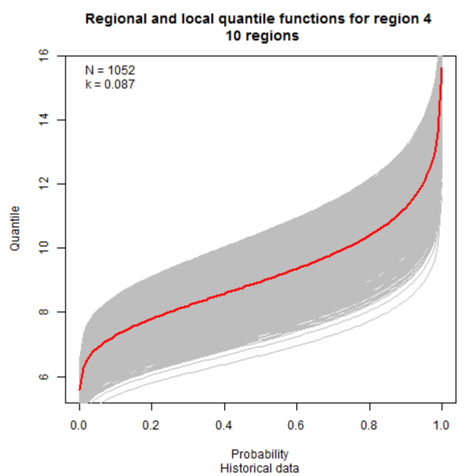
Choice of distribution function and estimation

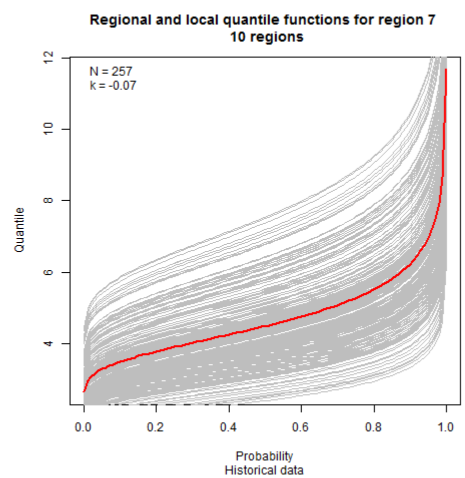
- Use annual maxima so GEV distribution is a reasonable choice
- This also compares well in terms of the goodness-of-fit statistics

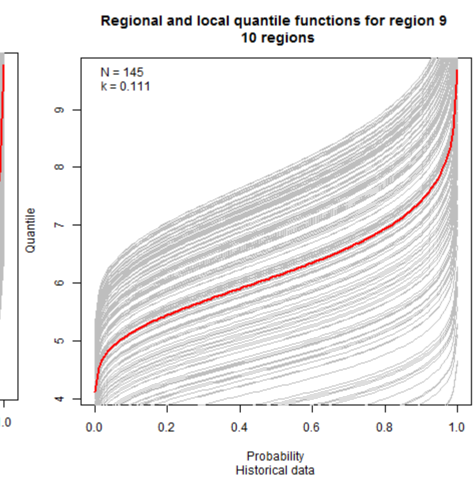
- Model parameters estimated by method of L-moments

<u>Regional growth curve together with quantile functions for individual sites:</u>



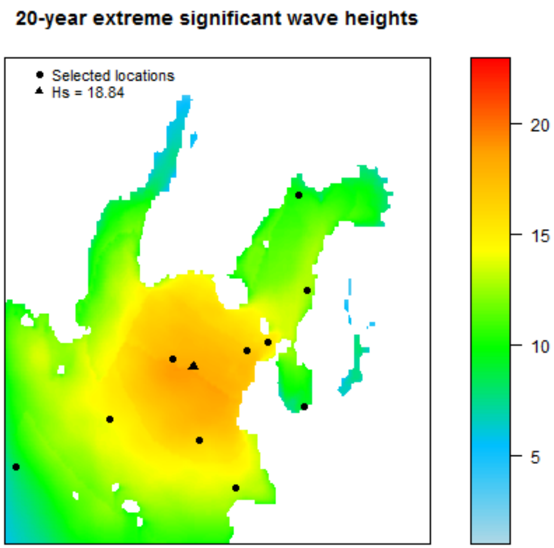




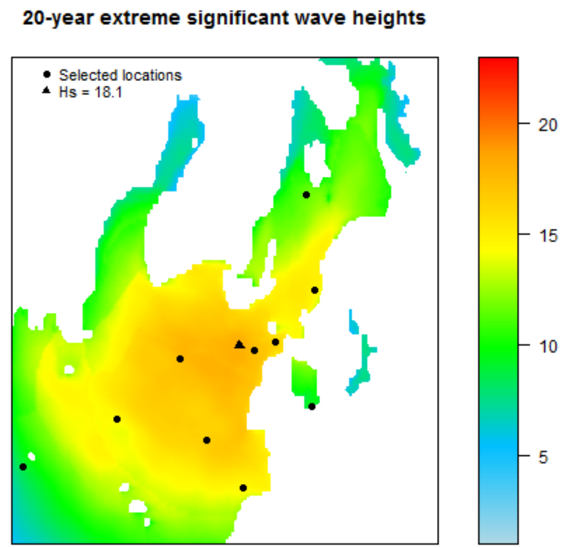


Extreme value analysis based on RFA

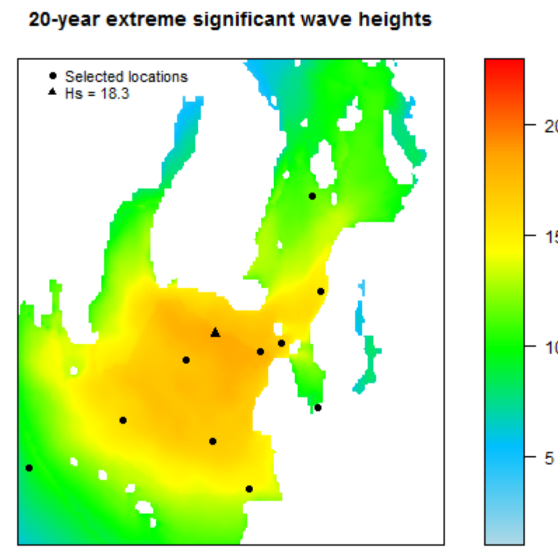
20-year return values over the whole area in the different scenarios



Historical data



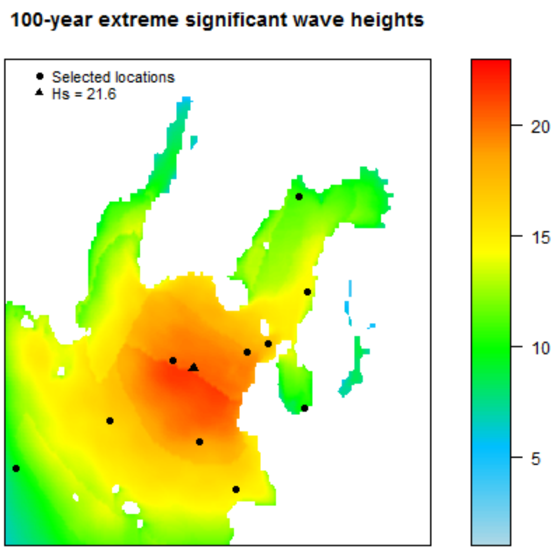
RCP 4.5 data



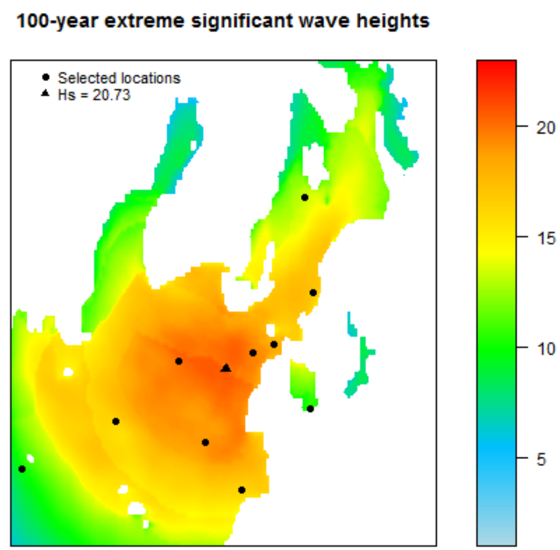
RCP 8.5 data

Extreme value analysis based on RFA

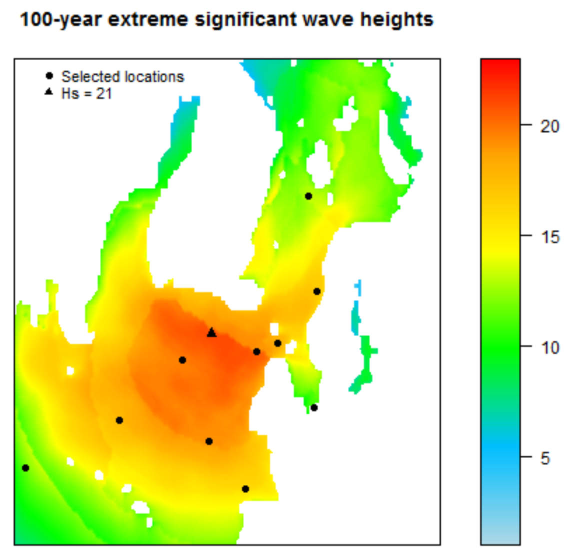
100-year return values



Historical data

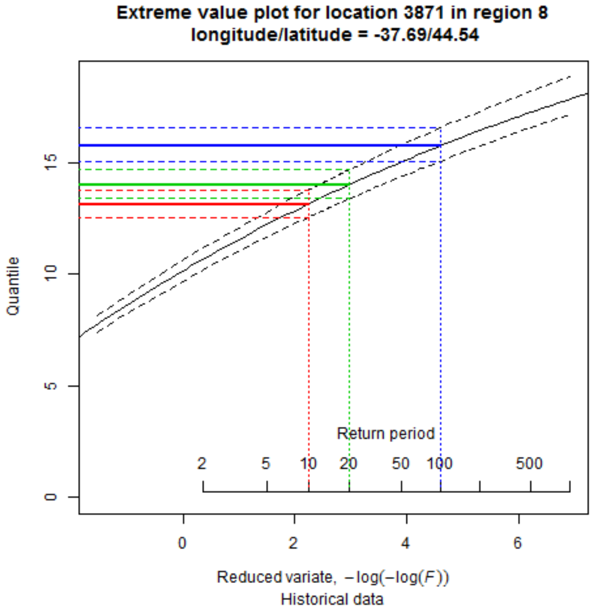


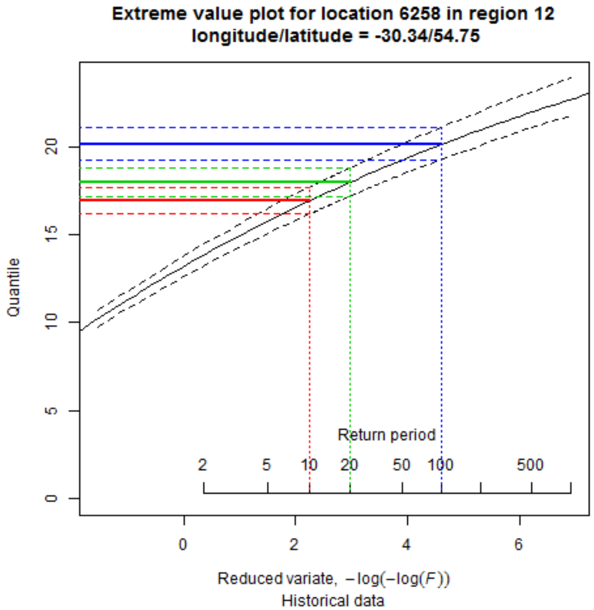
RCP 4.5 data

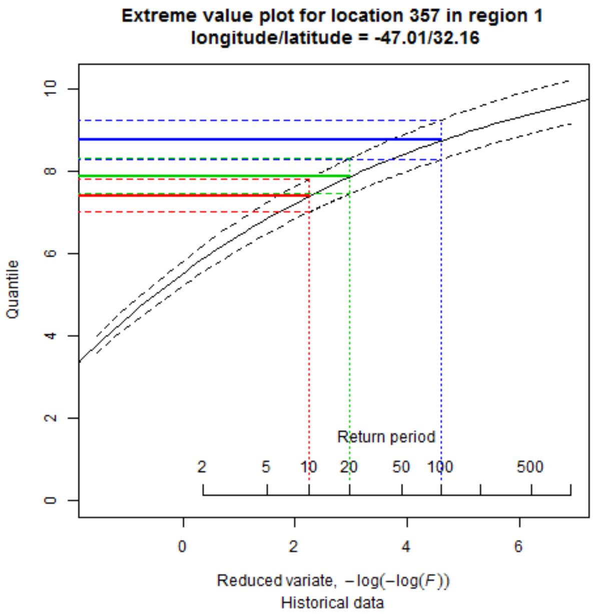


RCP 8.5 data

Extreme value analysis based on RFA – selected locations

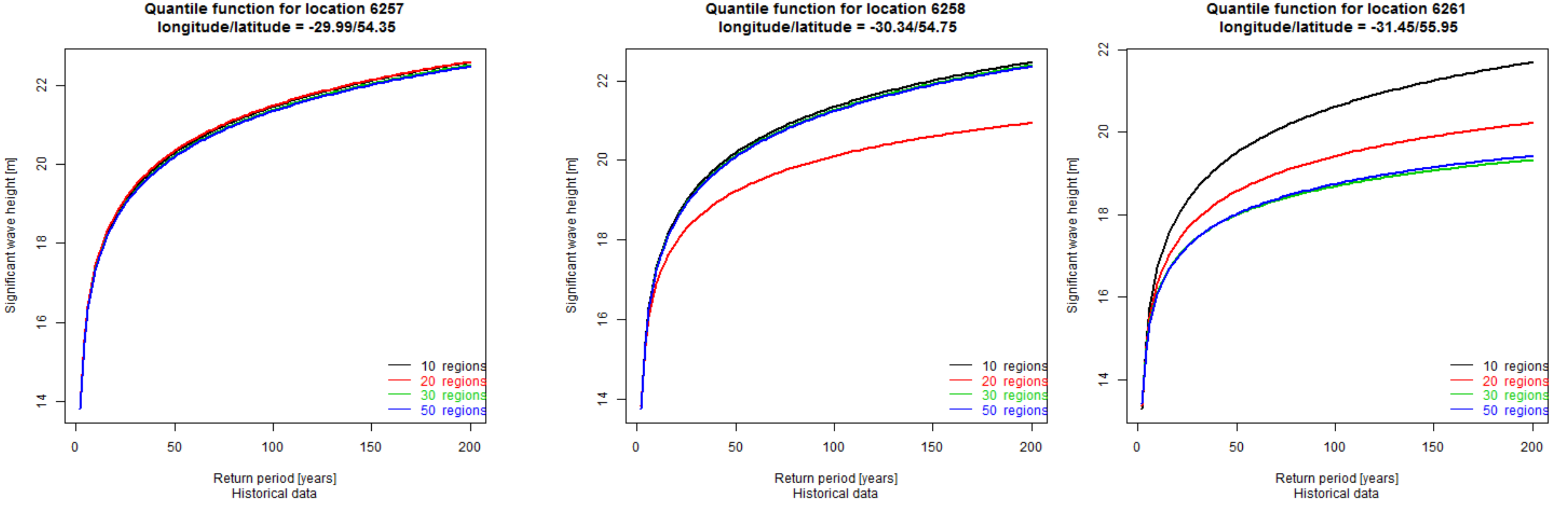






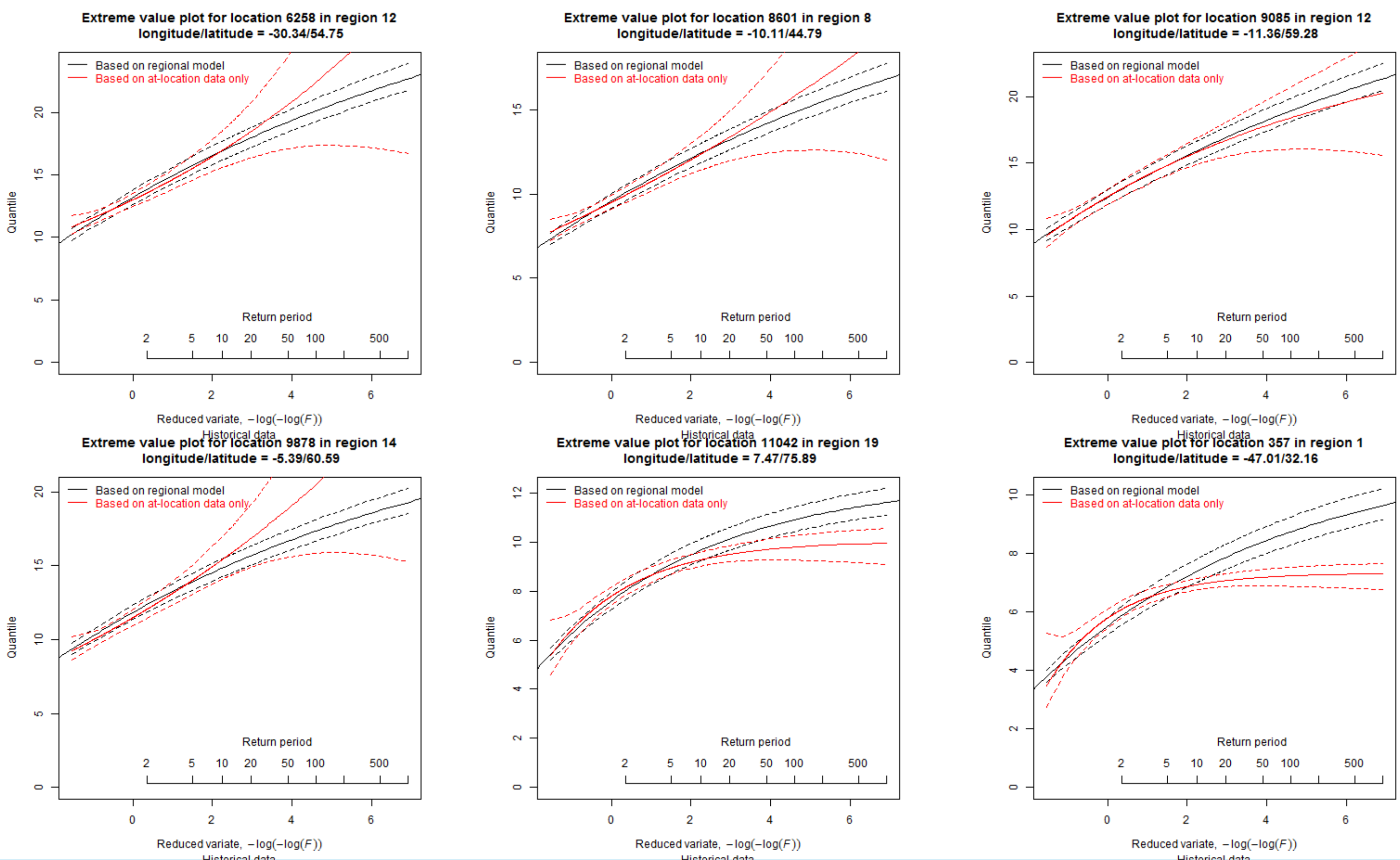
Discussion – sensitivity to the number of regions

- At some locations, the results are not sensitive to the number of regions. However, at some locations – in particular locations at the border between regions – the number of regions assumed influence the results.



Regional vs. grid-point modeling

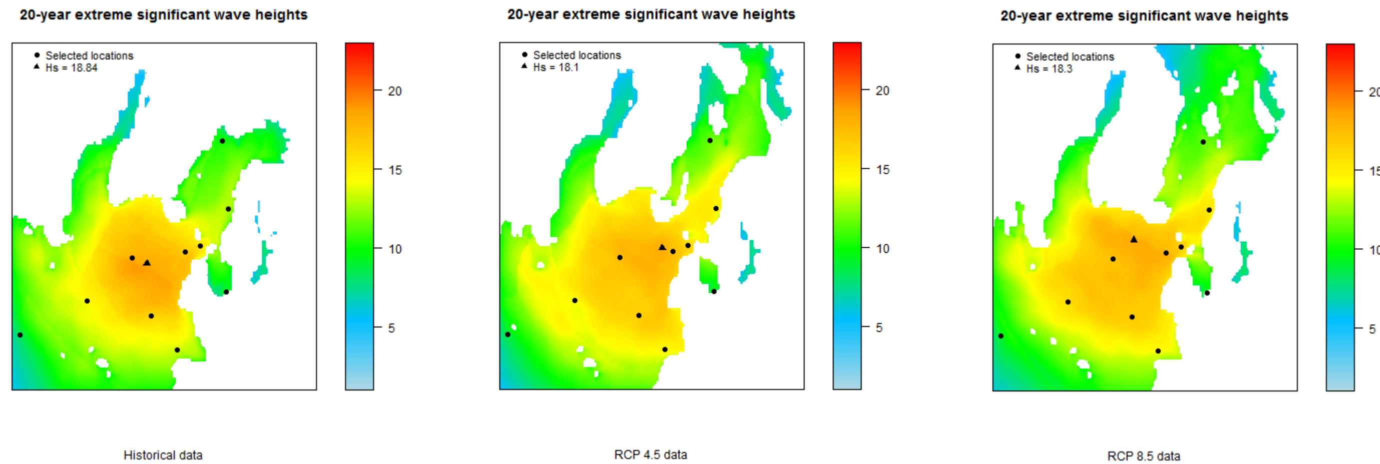
- Applying regional frequency analysis generally gives much more accurate return value estimates – especially for high return values



DNV·GL

The effect of climate change

- NB: General conclusions about climate change cannot be made based on one climate run from one climate model - but RFA is able to detect long term trends due to climate change
  – E.g. from 20-year return value maps:



- The results of this analysis suggest:
  1) Highest return values decline in a future climate
  2) Area of high return values above a certain threshold increase

Summary and conclusions

- RFA may be applied to ocean wave data
- Useful for Extreme value analysis
  - Obtains more robust estimates of high return values; Narrower confidence bands
  - Estimated return values are within a reasonable range
- Useful in spatial analysis
  - May interpolate to ungauged sites
- Main assumptions:
  - data at a particular location are iid -> probably OK for annual maxima
  - Data at different sites are independent -> may not be entirely satisfied if annual maxima from neighbouring points are from the same storm event
- Identification and assigning regions perhaps most critical step
- May be used to assess high-level changes in spatial patterns due to climate change

Paper with details have been published in ***Ocean Engineering***

Thanks for your attention!

Erik Vanem
Erik.Vanem@dnvgl.com
+47 6757 9900

www.dnvgl.com

SAFER, SMARTER, GREENER

DNV·GL