Development of a wave ensemble system at the Met Office

Francois-Xavier Bocquet

11th Wave Workshop, Halifax, 18-23/10/2009

Table of Contents

Motivation

- Address weakness of deterministic models
  - No indication of possible alternatives
  - Can miss extreme events
  - No measure of situation-specific uncertainty (general uncertainty info available from routine validation)
  - Doesn't allow to quantify risk
- To quantify confidence in the forecast as it varies between forecast cycles (flow-dependent uncertainty)
- To obtain reliable probabilities of wave events happening
  - By sampling from a range of possible forecasts in a way which is consistent with the error structure of the observations and of the model

Method

Investigate feasibility of running a short-range wave ensemble system on the MO computer architecture

- Run test cases to determine ideal configuration of a trial system:
  - influence of initial conditions
  - Influence of boundary conditions
- Develop an ensemble suite based on operational wave model: global EPS (T+72) and nested regional EPS (T+54)
- 3 months trial period from Feb-Apr 2009
- Validation: spread-skill relationship and reliability of probabilities
- Development of trial products

Conclusions

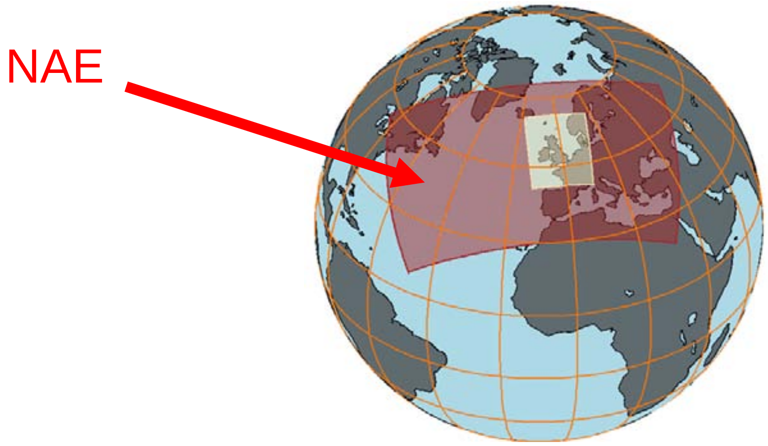
- Optimal run configuration determined consisting of global 90km and regional 24km ensemble
- Tested over a 3 month period
- Forcing winds are good quality and provide a suitable forcing to the system
- Ensemble mean performance very close to that of the control run
- Spread is a good indicator of forecast uncertainty but further work required (e.g. perturbing model physics)
- Forecast probabilities are good (esp. 1m and 2m thresholds)
- Potential products have been developed to aid interpretation of the results

Description of the models

Atmospheric ensemble (MOGREPS) and wave model

MOGREPS – The Met Office short-range ensemble

- 24-member ensemble designed for short-range forecasting
  - Global ensemble (~90km resolution, 38 levels) to T+72
    - Also runs to 15 days at ECMWF for multi-model ensemble research
  - Regional ensemble over N. Atlantic and Europe (NAE) (24km resolution, 38 levels) to T+54
  - ETKF for initial condition perts
  - Stochastic physics
  - Global run at 0Z and 12Z. Regional run at 6Z & 18Z



WAVEWATCH III

- Operational version of WAVEWATCH III with 2nd order advection scheme
- Spectral resolution: 24 directions X 25 frequencies
- Tolman-Chalikov source terms
- Snl evaluated with DIA scheme

Development of the wave ensemble system

System design

Reliable system can be obtained through a choice of the following:

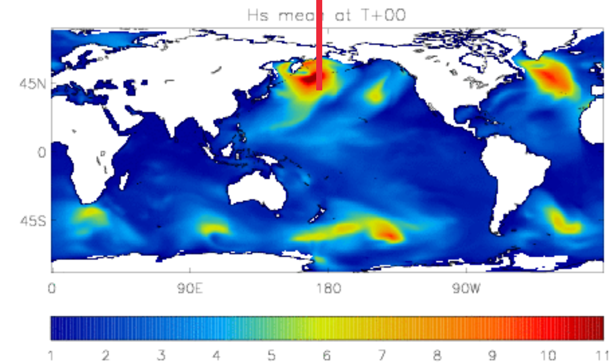
- Atmospheric Forcing (MOGREPS)
- Initial Conditions: - start from single field
  - use each previous run for IC's
  - perturb initial conditions
- Boundary Conditions : - determ. BC for NAE
  - ensemble BC for NAE
- Perturbed physics in the model
  - Outside the scope of this study

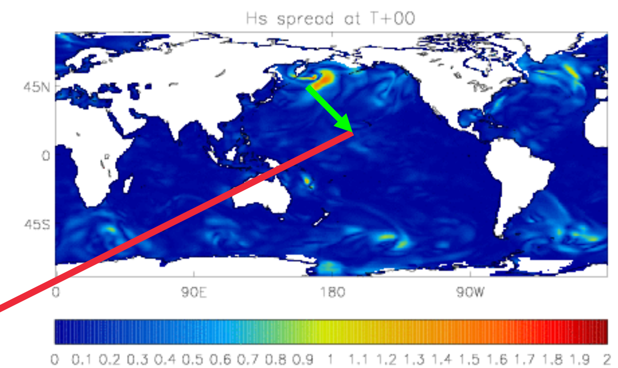
Our aim is to obtain a reliable Wave EPS for a regional domain (NAE)

Initial conditions (1)

Swell generated here at T+0

- System not expected to depend strongly on IC due to the weakly nonlinear, highly dissipative nature of the wave equations
- Tested in 2 phases:
  1. Starting from identical condition, run model with 24 different forcing winds for five days (using first 12 hours of each MOGREPS cycle)
  2. Allow system to relax under identical forcing winds
- Results: most of the spread vanishes within 12-24 hours
- Memory of initial wind-sea almost totally lost after 12-24 hours
- **Memory of swell retained up to T+60 and beyond**





Residual spread in swell field at T+60

Initial conditions (2)

Three options:

1. Perturb the initial wave field using an algorithm such as ETKF or SV – requires either DA system or linearised tangent operators
2. Use the latest T+0 “analysis” field for all ensemble members =>no spread present in the system initially, but start from a best guess field
3. Restart each ensemble member from the previous cycles’ forecast => start with an initial spread

Option 3 was selected to maintain spread at low lead times

Boundary conditions (NAE)

Two possible approaches:

1. Use the same boundary condition for all regional ensemble members – computationally less expensive but potential loss of spread in the swell part of the spectrum
2. Use BC generated by a global ensemble

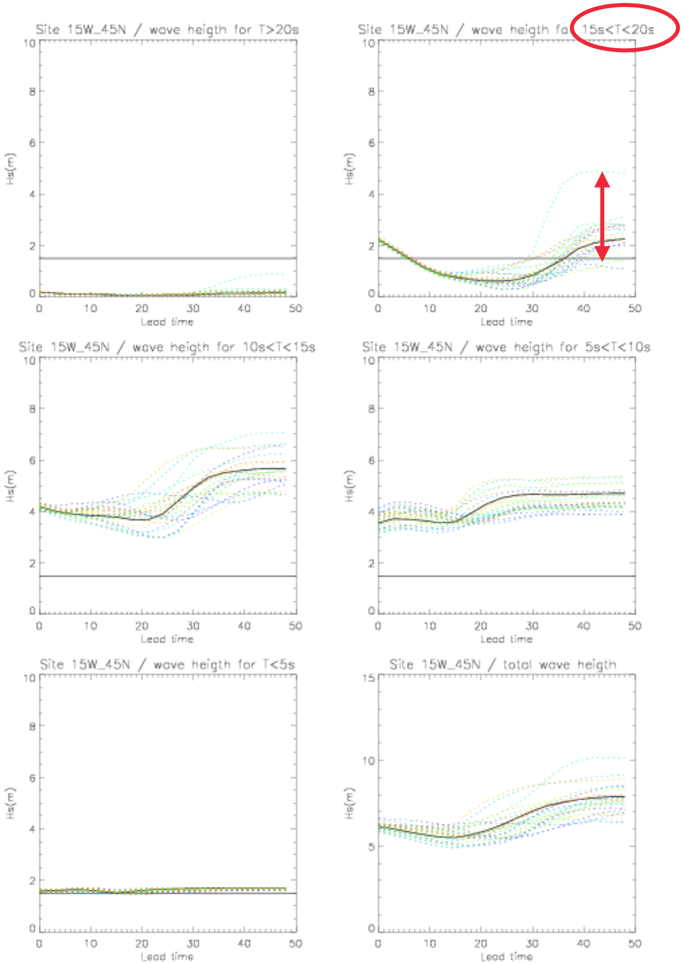
Similar overall statistics (will only be noticeably different in swell dominated seas)

Quite different when looking at certain periods (15s<Tm<20s)

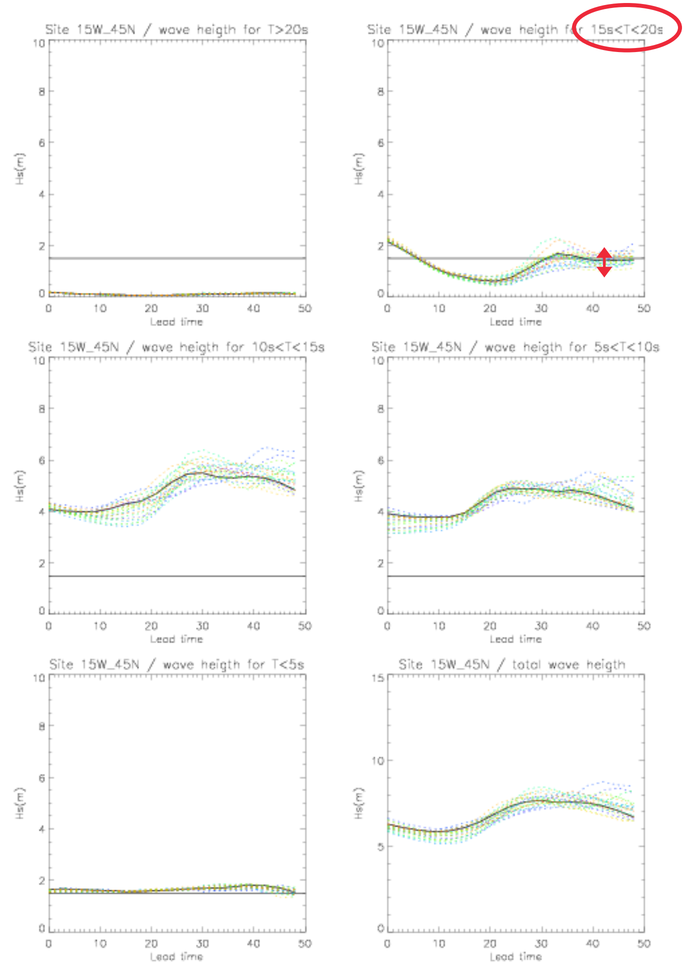
Option 2 was selected

Results of BC tests

Option 2: ensemble BC

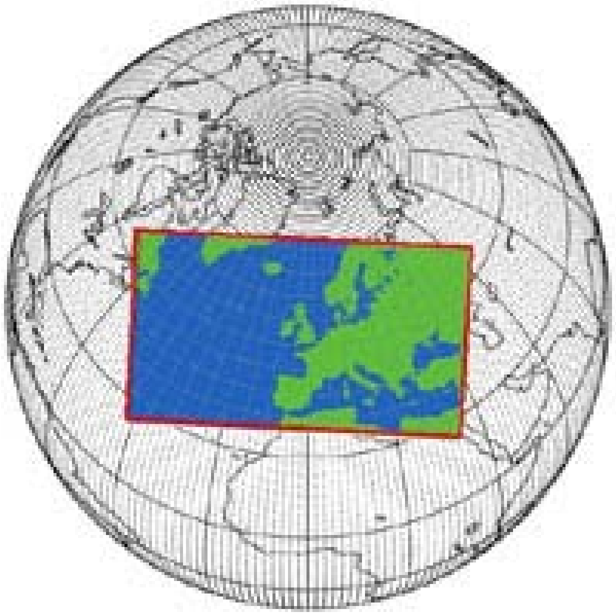


Option 3: determ. BC



Trial system design

- Global ensemble:
  - 24 members
  - each member restarts from previous cycle's T+12 field
  - 90km run
  - 24 directions X 25 freq.
  - run to T+72
  - Provides BC for NAE
- NAE ensemble:
  - 24 members
  - each member restarts from previous cycle's T+12 field
  - 24km run
  - same spectral res.
  - run to T+54
- All forcing from MOGREPS

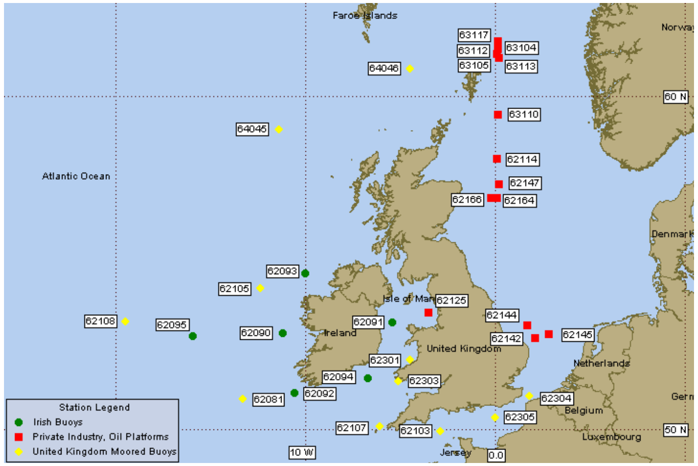


Verification results (focus on NAE)

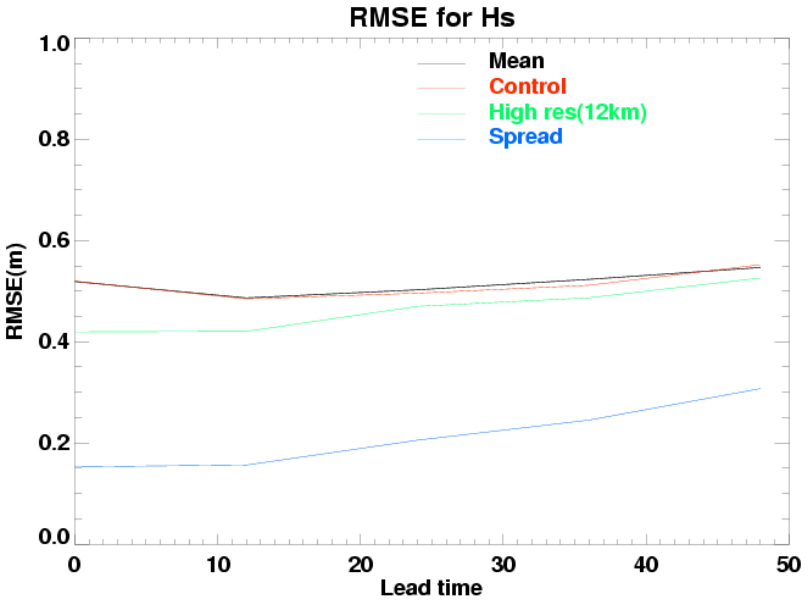
Aims of the verification

- Assess deterministic value of ensemble mean
- Check if the ensemble spread is a good indicator of forecast uncertainty
- Verify the probabilities derived from the system

- All this work has been done using wave buoy data



Spread and error statistics



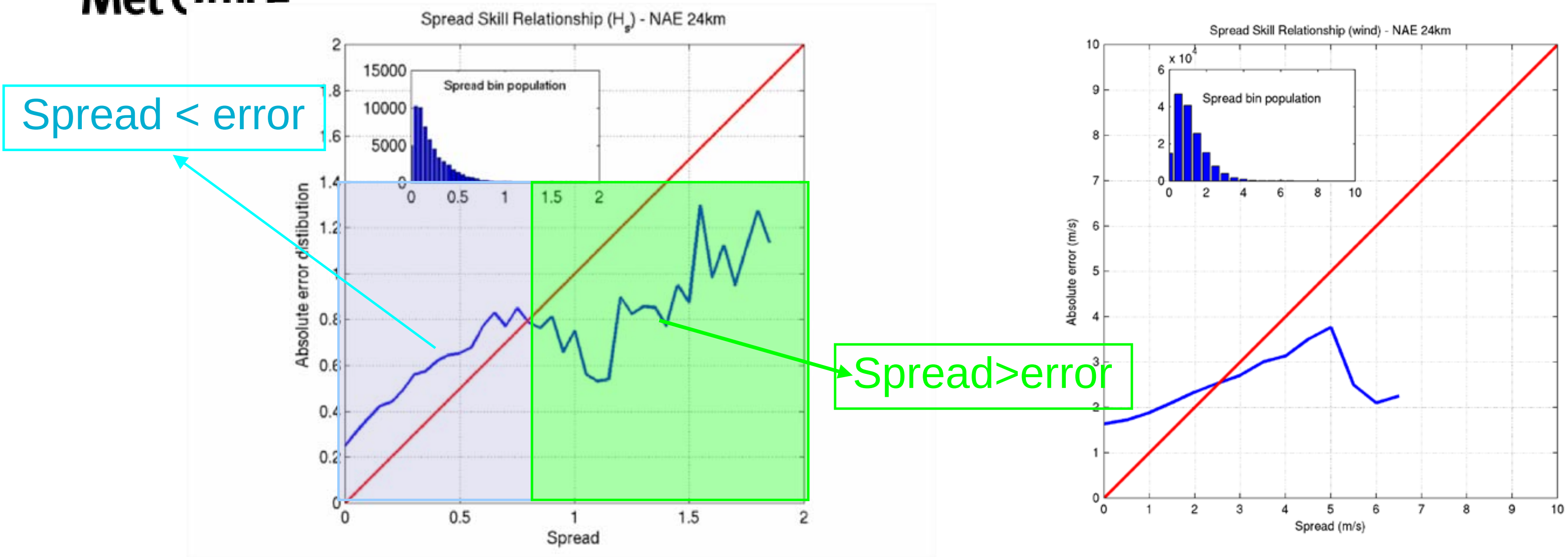
- Ensemble mean fc expected to outperform control member => not clear
- Highres initially has lowest RMS error but gap closes with incr. lead time
- Spread increases with lead time as expected and explains ~50% error
- Contributions to error:
  1. Forecast uncertainties (IC,forcing) -> spread
  2. Model uncertainties (physics,not considered here)
  3. Observational errors (~10% of observed Hs)

Spread-skill diagrams (1)

- How do we check if the spread is a good indicator of forecast uncertainty?

- Can be verified with spread-skill diagrams
  - X-axis: spread
  - Y-axis: 80th percentile of the error for each spread bin
  - Ideal spread-skill: 1-1 relationship
  - Residual errors linked to imperfect model/obs and finite size of the ensemble

Spread-skill diagrams(2)



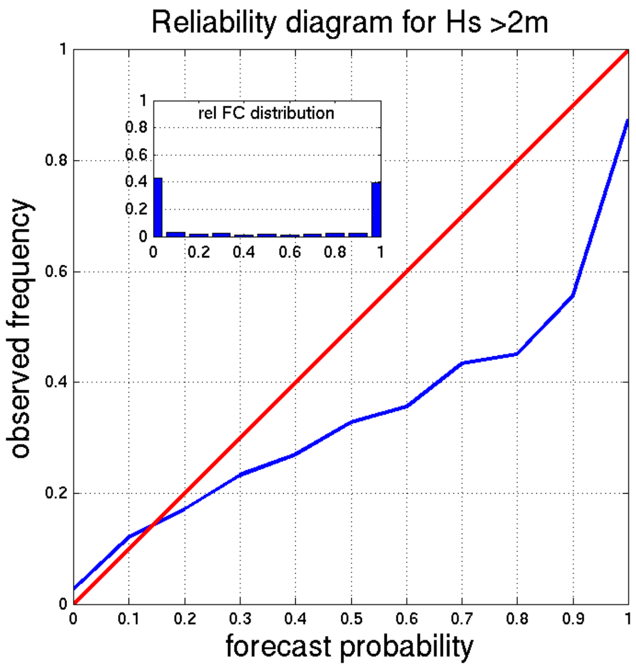
- Spread is good predictor of forecast error (80th percentile of absolute error on the ensemble mean)
- Main population in lower spread bins
- When the spread in the FC is low, the spread underestimates FC error (more confident in this)
- When the spread in the FC is high, the spread overestimates FC error
  - ⇒ Prudent approach

Verifying probabilities

- Deterministic models: predict event and verify against the observational outcome
- Probabilistic models: no true or false forecast
- “If we predict that an event will happen with a probability of 60% and the event happens, where we right or wrong?”
- Right if in 60% of those cases, the event actually happens => reliability diagrams
- Requires long period to verify

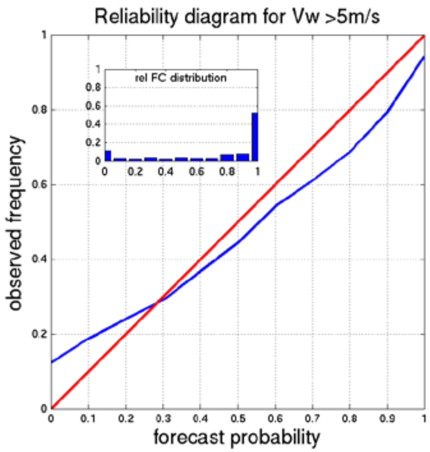
Reliability Diagrams

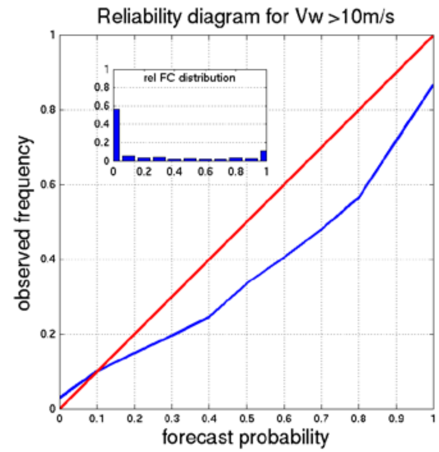
- Graph of observed frequency of an event vs. forecast probability of the event
- If event probability is 60%, it should occur 60% of the time in the long run
- Graph should be a diagonal for a perfect system
- Ref: Wilks 1995

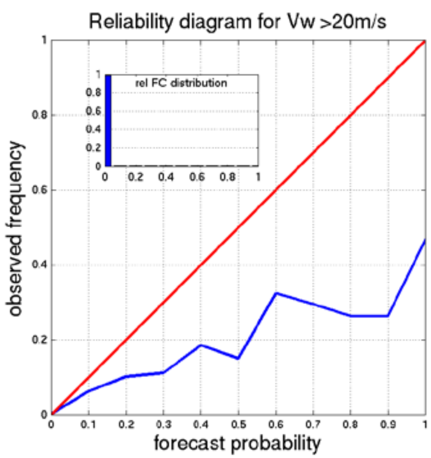


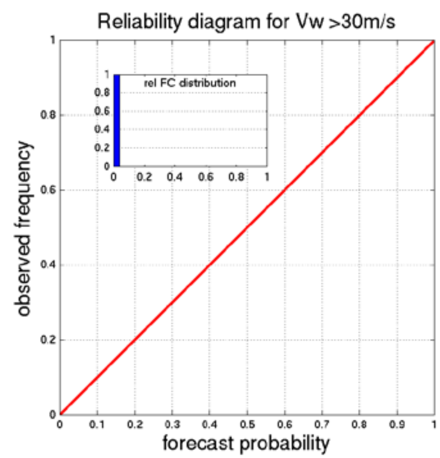
Reliability diagrams (Vw)

- Tendency to underestimate probabilities
- Pretty good performance for lower wind speeds
- Lack of extreme events in trial period

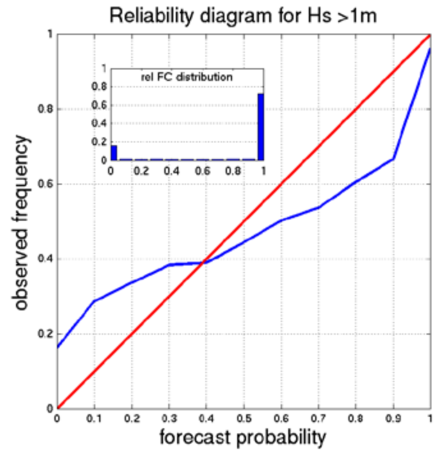


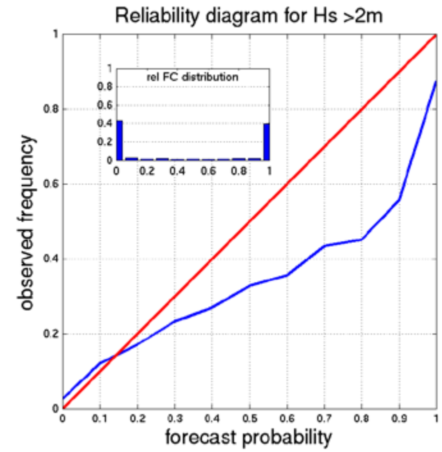


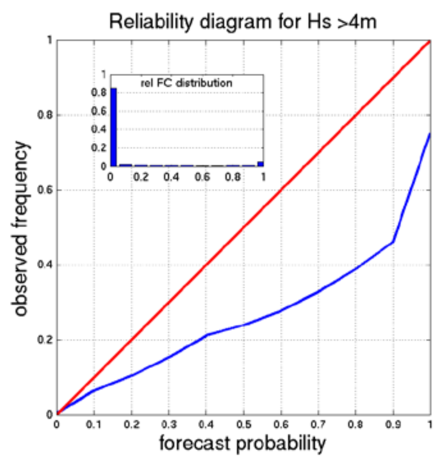


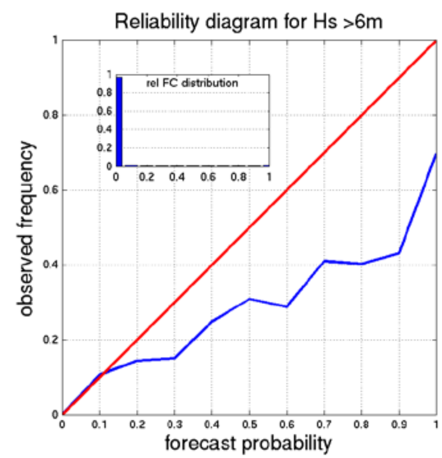


Reliability diagrams (Hs)









Potential products

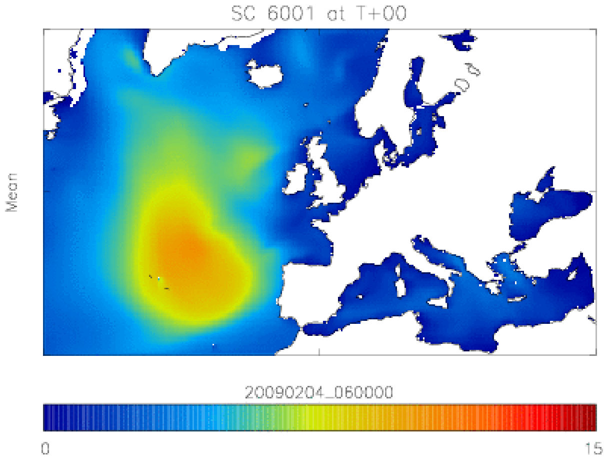
Challenges in conveying probabilistic information

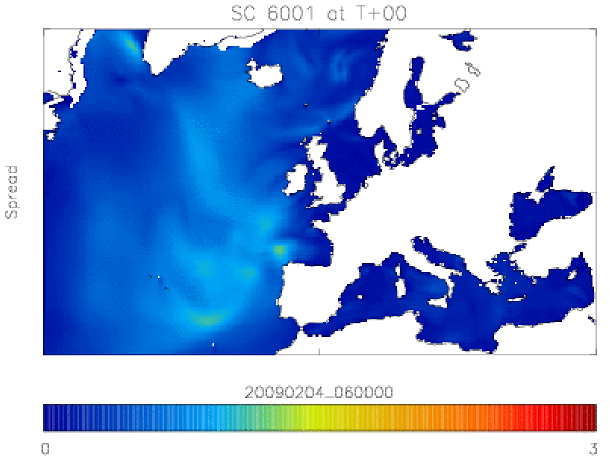
- People (and some journalists) don’t understand probabilities (e.g. BBQ summer coverage)
- Customers want to make a yes/no decision based on the forecast – more complex in the

> The office now claims that it is "66% certain" that next winter will be warmer and wetter than last. The figure is an ominously precise advance on the 65% certainty of a warm summer. The information is useless without knowing the likelihood of the "66%" being correct.

Products – Mean/Spread plots

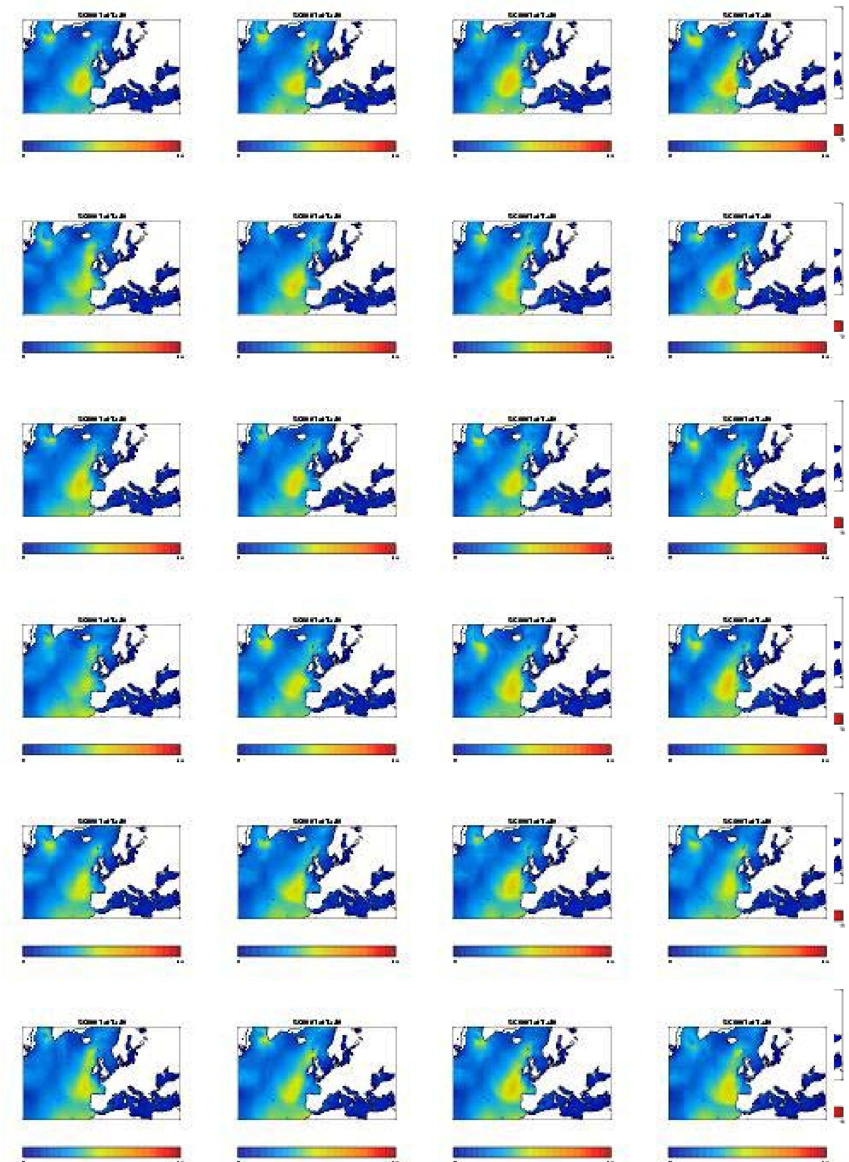
- Shows mean forecast (top)
- Ensemble spread (bottom) shows where the forecast uncertainties are located geogr.
- Two options:
  1. Location of interest is in low spread zone -> forecast reliable -> use as det. Forecast but with higher confidence
  2. Location of interest is in high spread zone -> more detailed look required ->see next slides





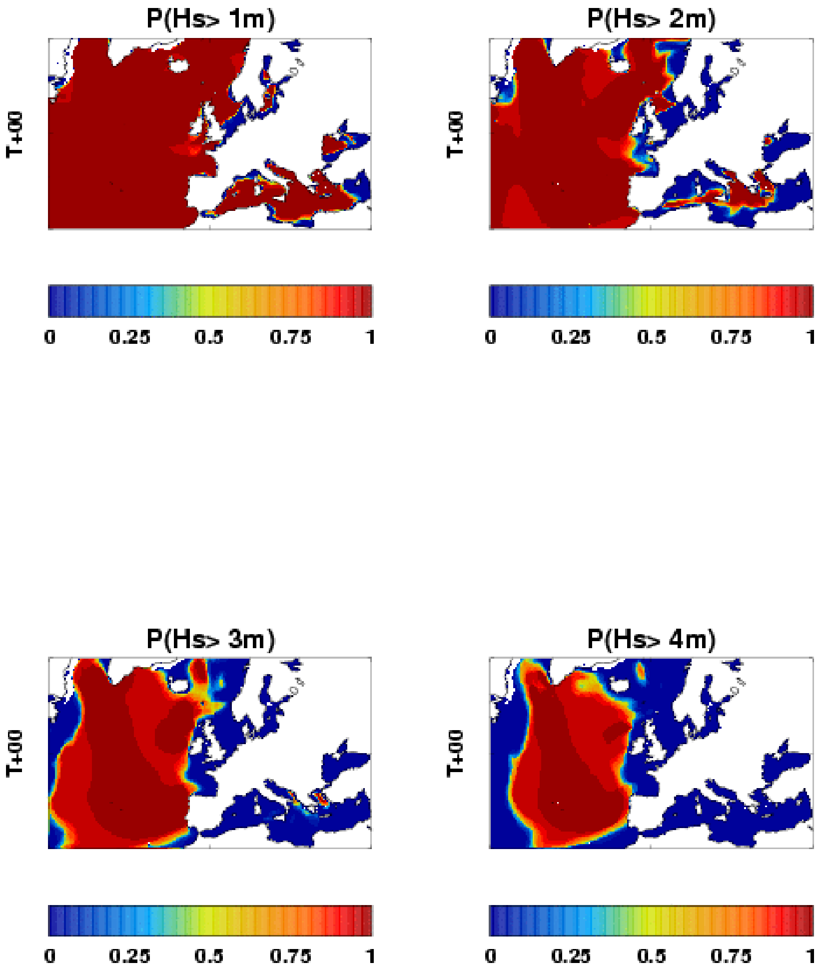
Products – Postage stamps

- Shows all forecasts
- Overview of each ensemble member
- Allows visual clustering
- Differences for instance at T+40 south of Greenland

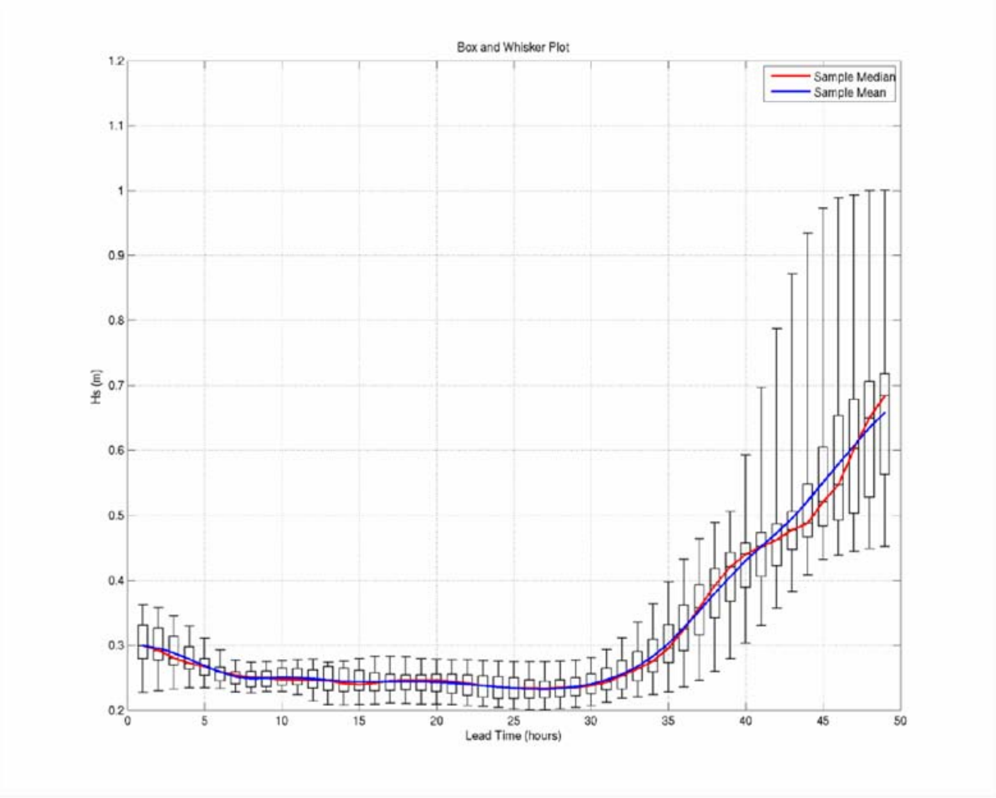


Products – Probability maps

- Shows probability of exceeding different thresholds [1m,2m,3m,4m]
- Help for operational decisions (e.g. operation with 4m thresh for RN)

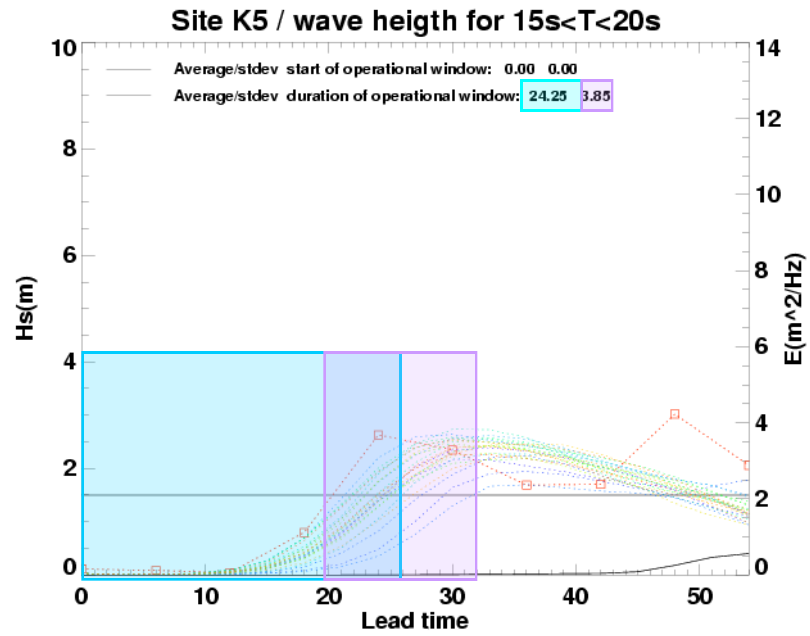


Products - meteograms



- Site-specific product
- Shows probabilistic forecast for all lead times
- Shows mean, median and quartiles

Products – Operational windows



- Some offshore operations are mainly sensitive to energy in certain frequency bands (e.g. heavy-lift and low period swells) => select band and operational threshold
- Integrate E spectrum over that band to obtain an equivalent wave-height
- Check for operational window where threshold is not exceeded

Conclusions / Future work

Conclusions

- Optimal run configuration determined consisting of glbal and regional ensemble
- Tested over a 3 month period
- Forcing winds are good quality and provide a suitable forcing to the system
- Ensemble mean performance very close to that of the control run
- Spread is a good indicator of forecast uncertainty but further work required (e.g. perturbing model physics)
- Forecast probabilities are good (esp. 1m and 2m thresholds)
- Potential products have been developed to aid interpretation of the results

Future work

- Investigate ways of representing model uncertainty in the forecast (e.g. perturbed physics schemes)
- Extend validation to include satellite data to improve our understanding of the performance of the system away from coastal waters
- Implement Relative Economic Value score to understand how valuable the system would be
- Target a period with more severe conditions to improve stats on severe sea state events