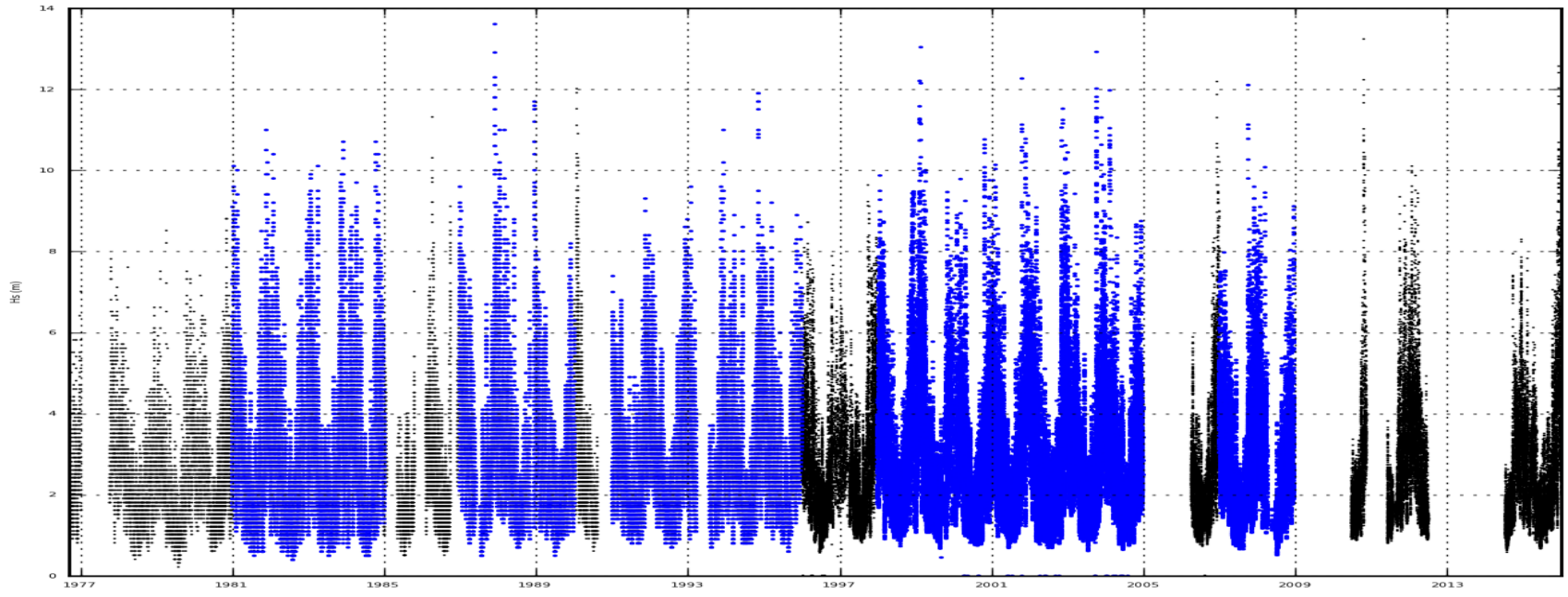
HR Wallingford
Working with water



Source file: \\hrw-uk.local\projects\live\dek2164$\data_received\NDBC\46005*gz Filter: Caires_years

Estimating extreme sea states from numerical models - some adventures in statistics

Doug Cresswell with help from Jean Bidlot, David Wyncoll, Lluis Via Estrem and others at HRWallingford, and elsewhere

20th October 2016

HR Wallingford
Working with water

Map

- Motivating questions
- A method of validation tailored to extremes analysis
- A bias correction scheme
- Uncertainty analysis on extreme estimates
- Plotting position
- Where next

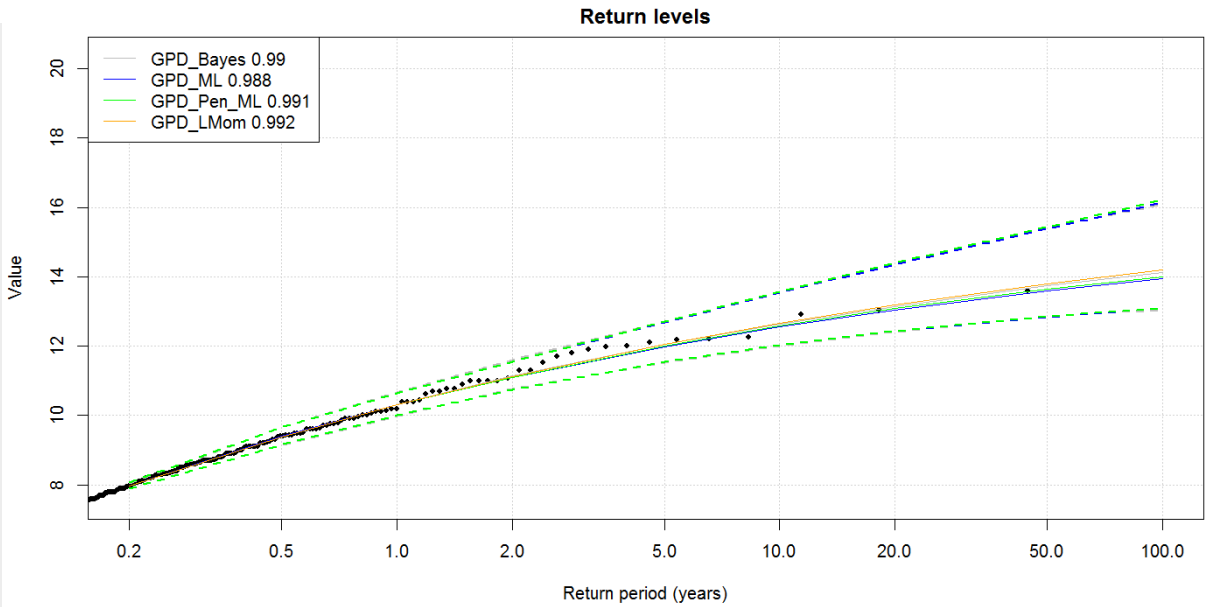
Motivation

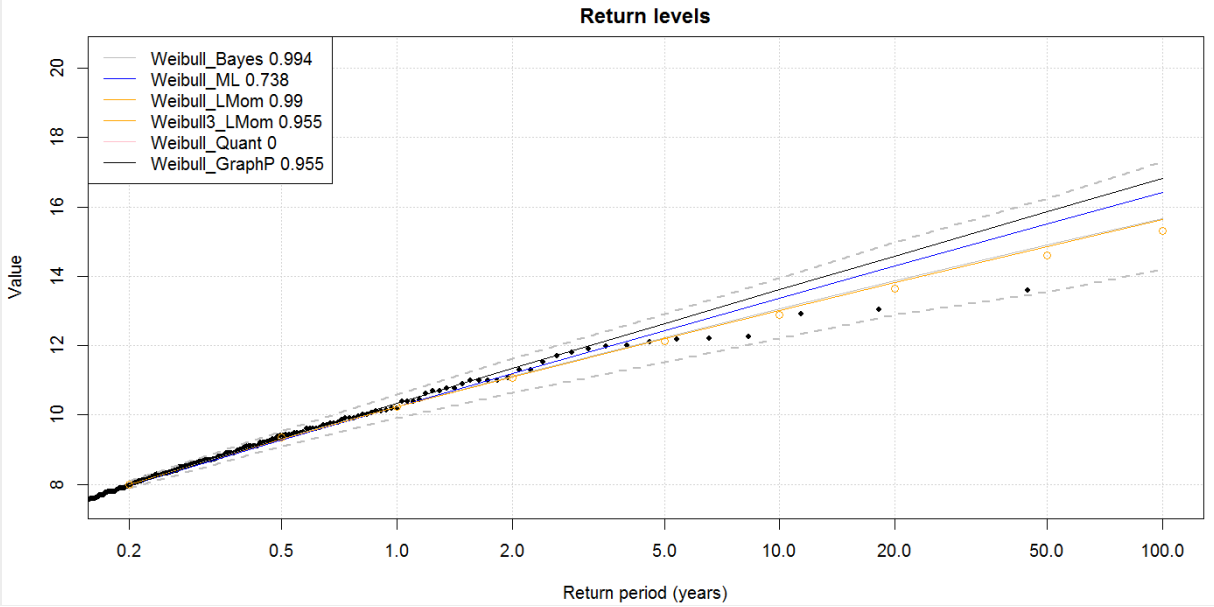
How best to estimate extreme sea states based on numerical models?

What can we learn from comparison with buoys?

Are the confidence intervals attached to estimates of extreme conditions too narrow or too broad?

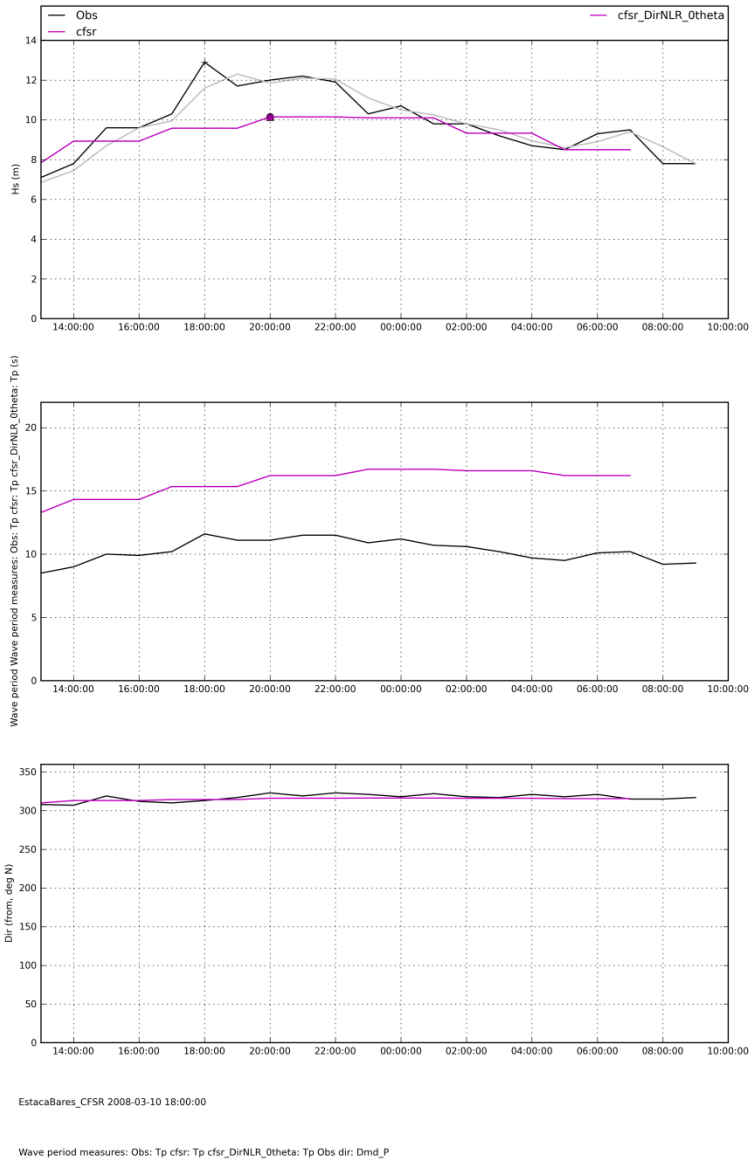
Can we identify where the uncertainty in estimates of extreme conditions comes from?





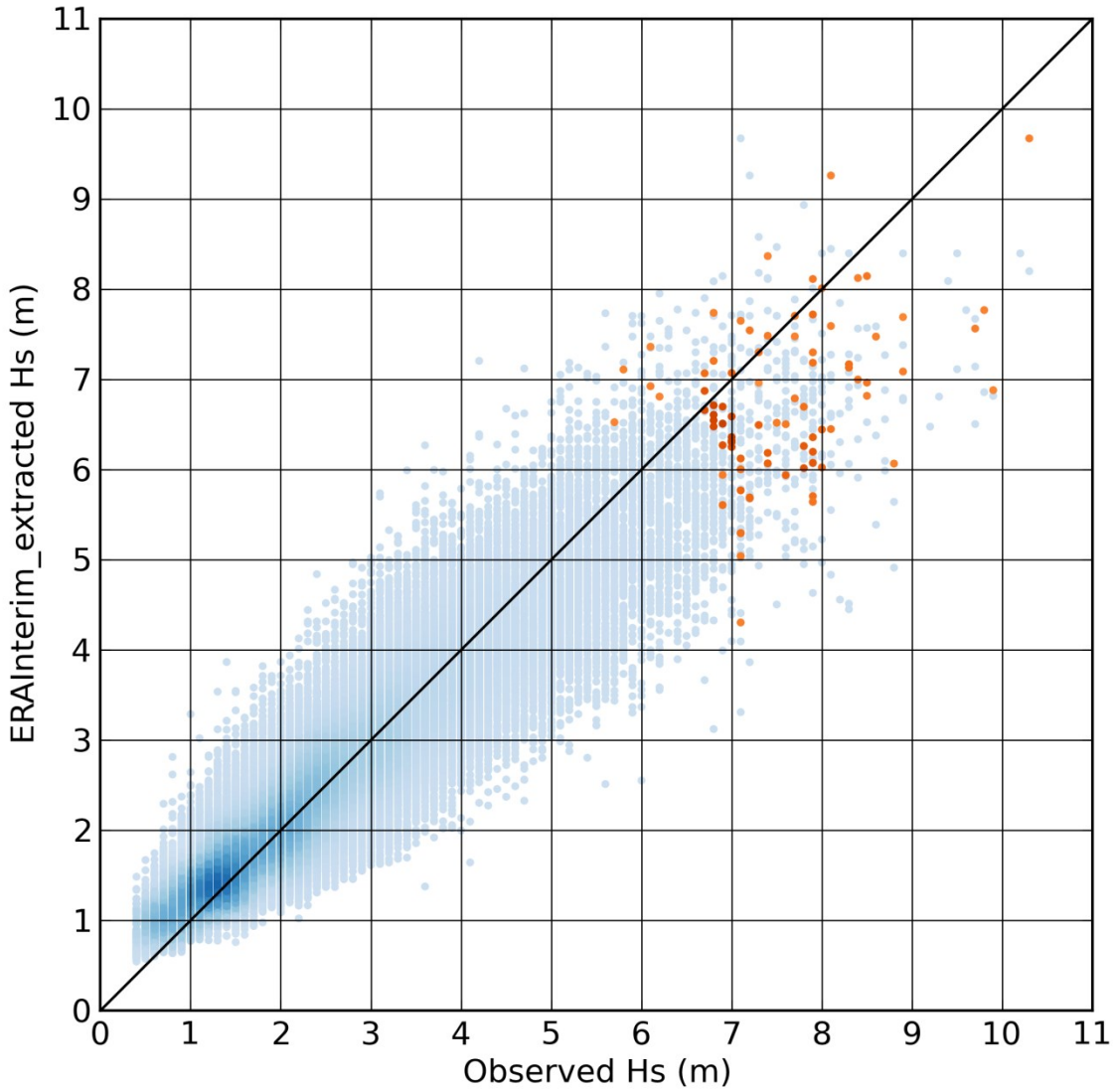
Why model peaks ≠ observed?

- Phasing of storm peaks
- -> match peaks over thresholds
- Sampling of sea states
- -> add buoy-like noise
- Model structure: parameterisations/resolution/archive frequency/data assimilated
- -> perhaps we can derive a correction?



Validation tailored for extremes

- Density of matched conditions
- By time in blue
- By peak in orange
- Distributions/biases differ



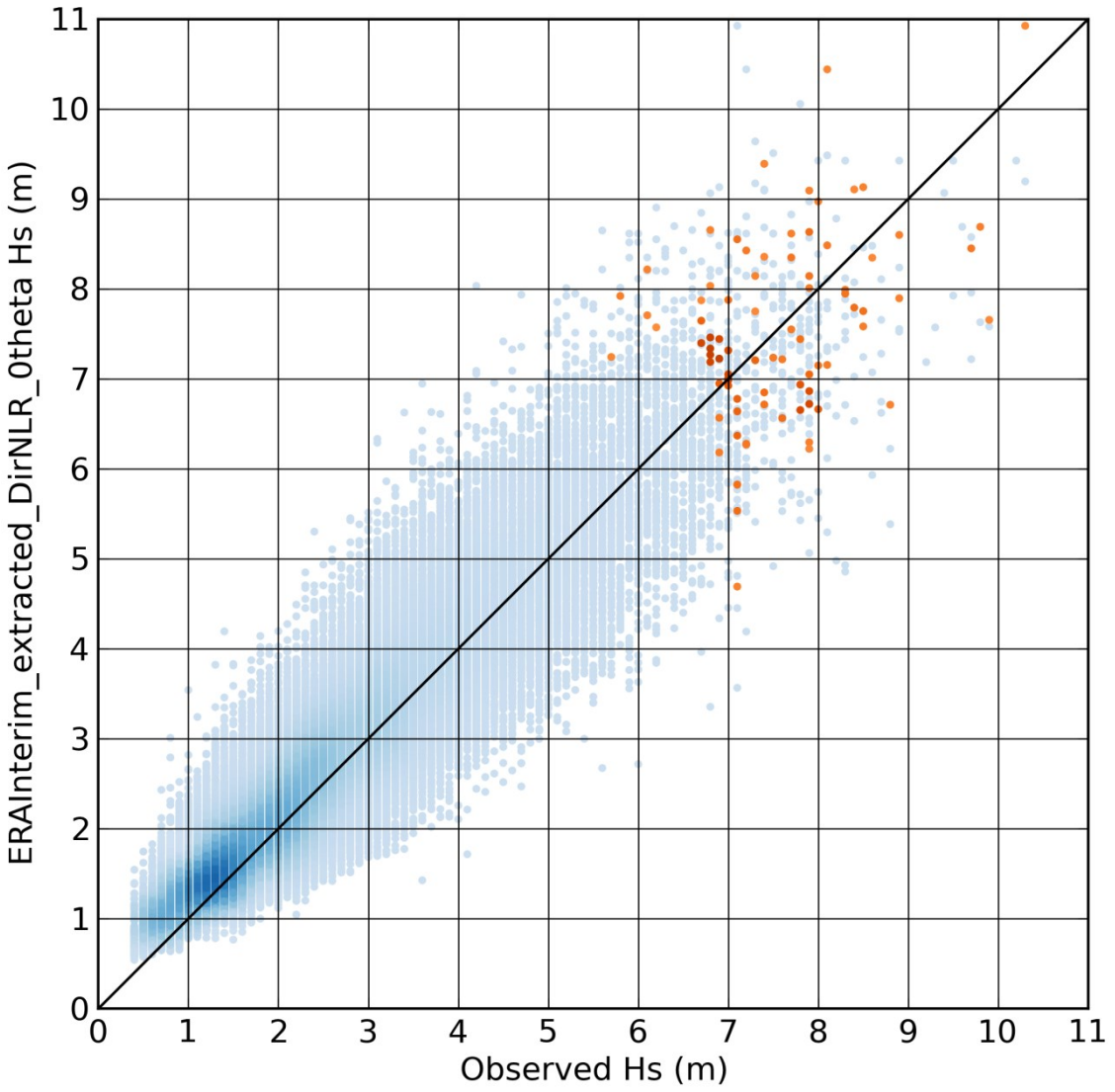
Analysis at 46206_ECMWF. Coloured: peaks, grey scale: all Filters and code spec:
Analysis: .\ERA-Int_Validation_v14.03_dcr_n4\ based on MetOcean Toolbox trunk image file : .\ERA-Int_Validation_v14.03_dcr_n4\\46206_ECMWF\summary_plots\4620

Is there a simple/mappable correction?

- Looking for a mappable bias correction
- Adopting non-directional simplification of Minguez:

$$H_{s_{cal}} = a H_{s_{orig}}{}^{b}$$

- Find [a,b] -> no bias in matched peaks
- Is there geographical consistency in [a,b]?



Analysis at 46206_ECMWF. Coloured: peaks, grey scale: all Filters and code spec:
Analysis: .\ERA-Int_Validation_v14.03_dcr_n4\ based on MetOcean Toolbox trunk image file : .\ERA-Int_Validation_v14.03_dcr_n4\\46206_ECMWF\summary_plots\46

ERA-Interim vs buoys

- Some clear consistency, e.g. shelf edge
- Coastal buoys tend to be grouped
- Resolution/fetch effects clear
- Some neighbour mismatches
- More work needed to classify and test predictive rules.

What about the rest?

Can we use the residuals to inform uncertainty analysis in estimates of extremes?

Can summarise residuals as

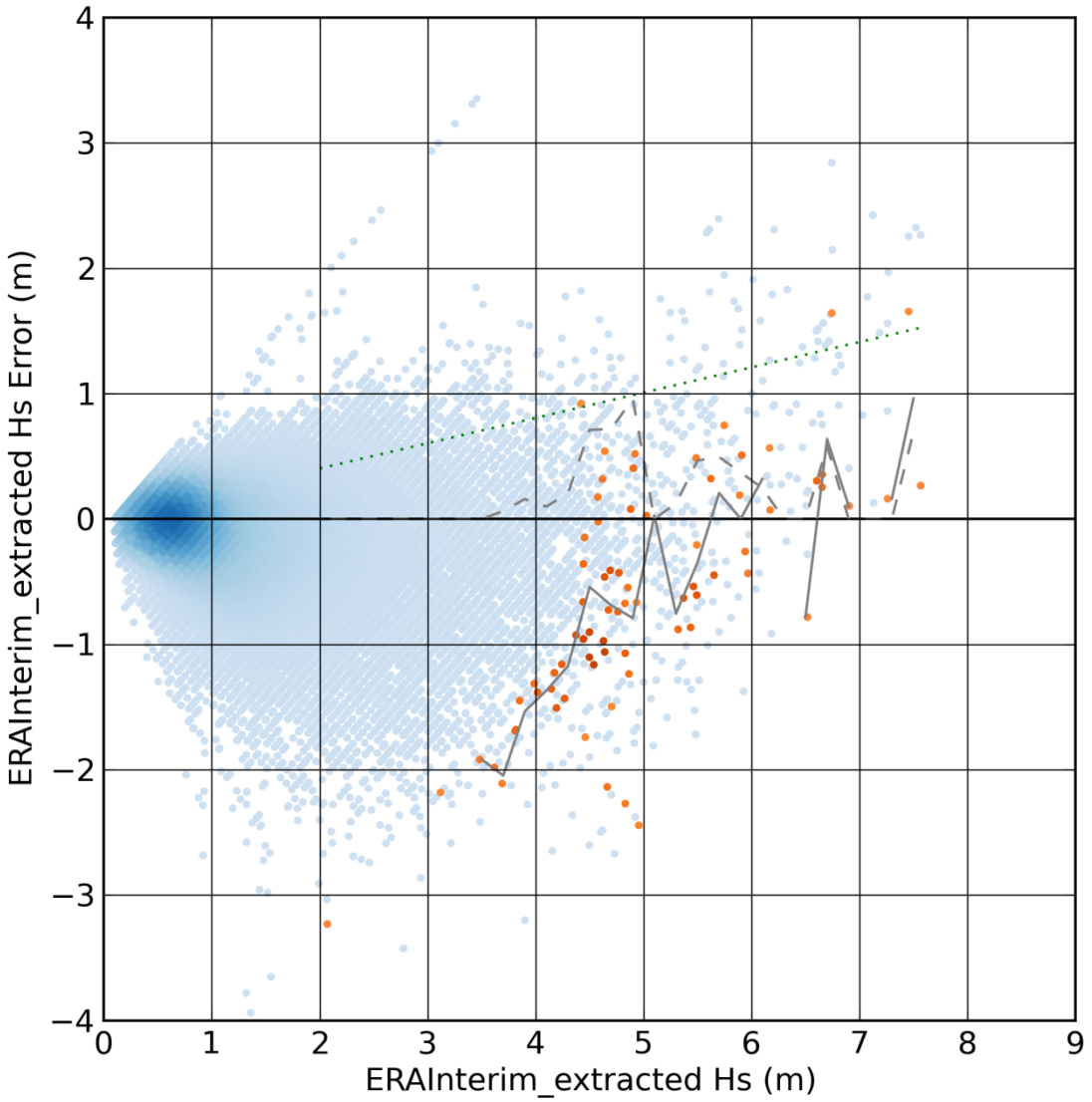
$$H_{s_{ob}} \sim H_{s_{cal}} + N(0, \varepsilon)$$

Or

$$H_{s_{ob}} \sim H_{s_{cal}} * N(1, \varepsilon)$$

Or

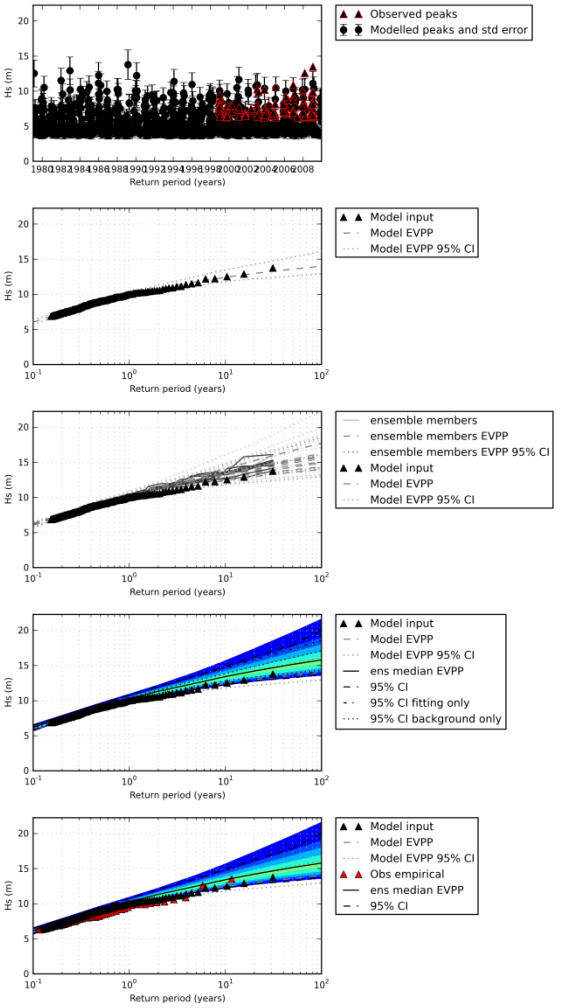
…?



Analysis at Mahon_ECMWF. Coloured: peaks, grey scale: all Filters and code spec:
Analysis: .\ERA-Int_Validation_v14.03_dcr_n4\ based on MetOcean Toolbox trunk image file : .\ERA-Int_Validation_v14.03_dcr_n4\\Mahon_ECMWF\summary_plots\Ma

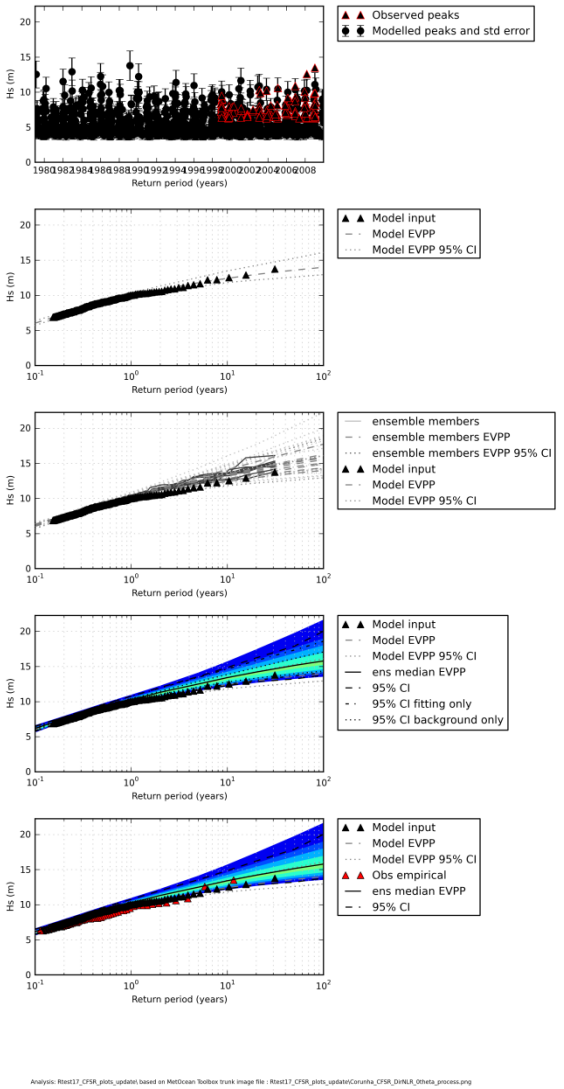
Including model skill in uncertainty estimates

- Associate uncertainty with modelled peak
- Sample -> Monte Carlo ensemble of series
- Estimate extremes on each series (using MCMC Bayesian methods)
- Combine Monte Carlo ensembles for best estimate and total uncertainty
- Notice positive increment to best estimate due to introduction of uncertainty on inputs.
- Interrogate 2D ensemble to estimate contribution to uncertainty from sampling and uncertainty on inputs

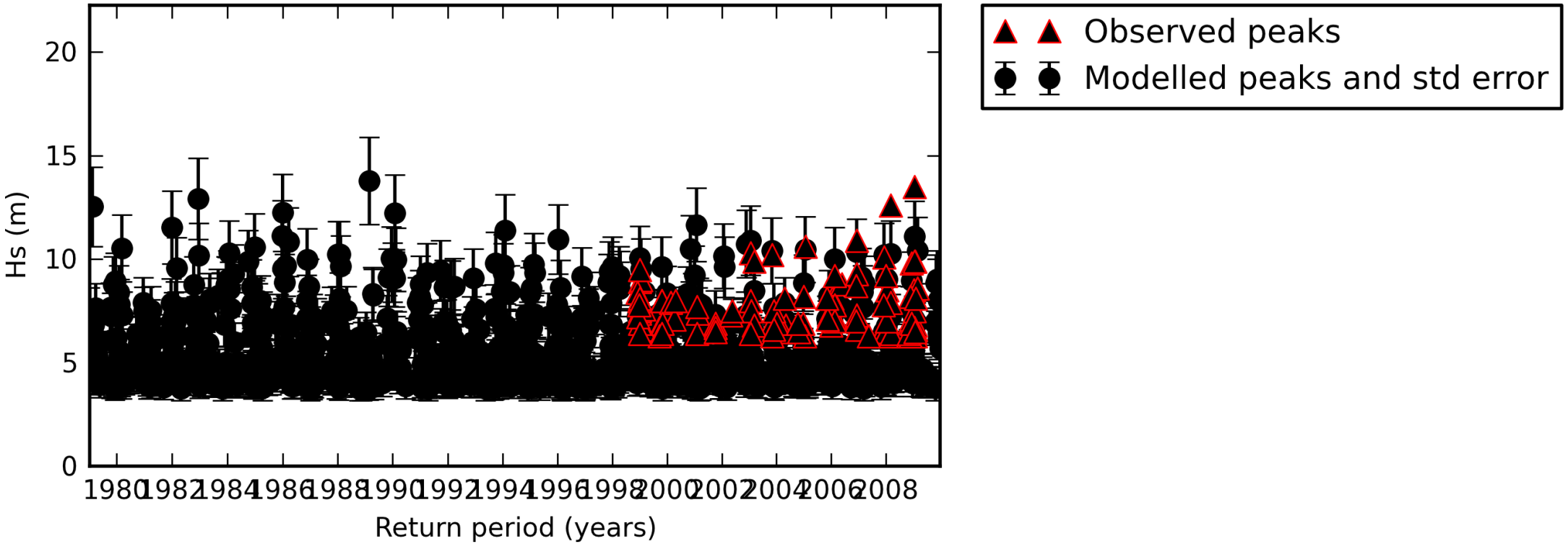


Including model skill in uncertainty estimates

- **Associate uncertainty with modelled peak**
- Sample -> Monte Carlo ensemble of series
- Estimate extremes on each series (using MCMC Bayesian methods)
- Combine Monte Carlo ensembles for best estimate and total uncertainty
- Notice positive increment to best estimate due to introduction of uncertainty on inputs.
- Interrogate 2D ensemble to estimate contribution to uncertainty from sampling and uncertainty on inputs

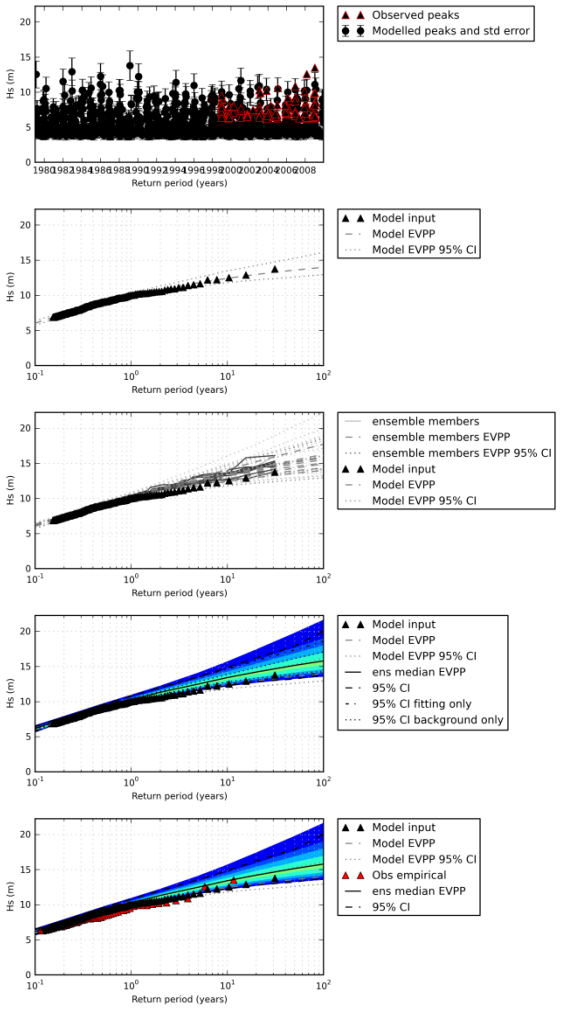


Including model skill in uncertainty estimates

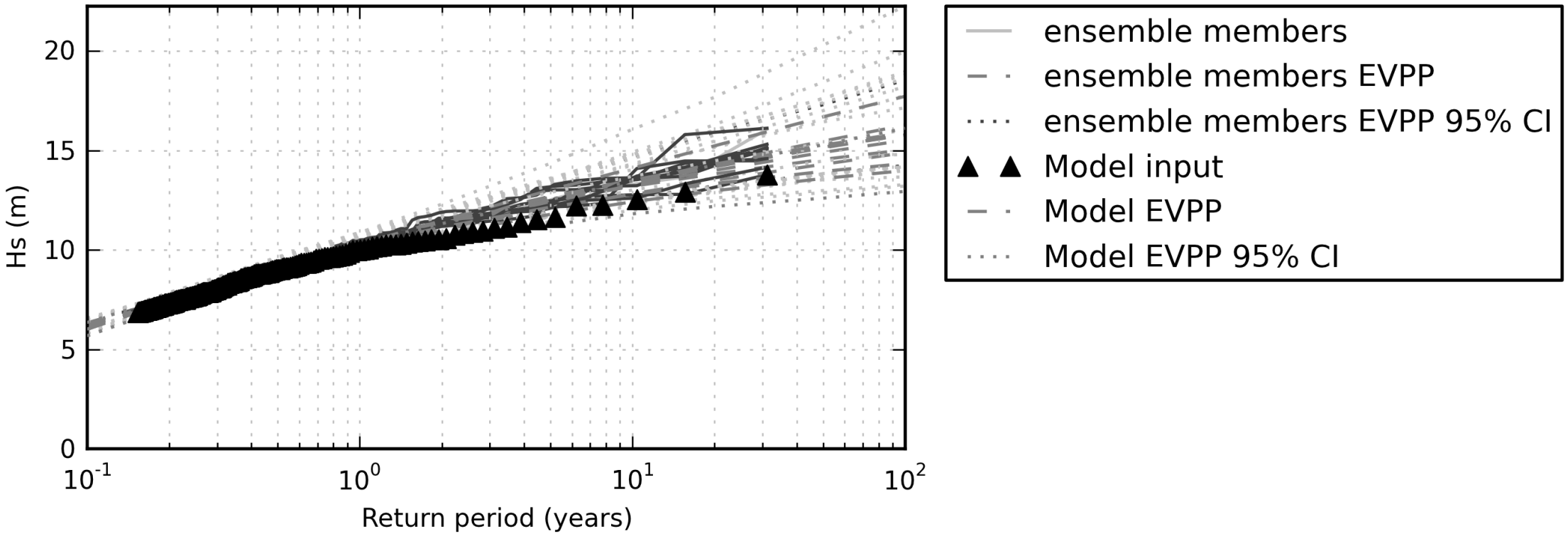


Including model skill in uncertainty estimates

- Associate uncertainty with modelled peak
- **Sample -> Monte Carlo ensemble of series**
- **Estimate extremes on each series (using MCMC Bayesian methods)**
- Combine Monte Carlo ensembles for best estimate and total uncertainty
- Notice positive increment to best estimate due to introduction of uncertainty on inputs.
- Interrogate 2D ensemble to estimate contribution to uncertainty from sampling and uncertainty on inputs

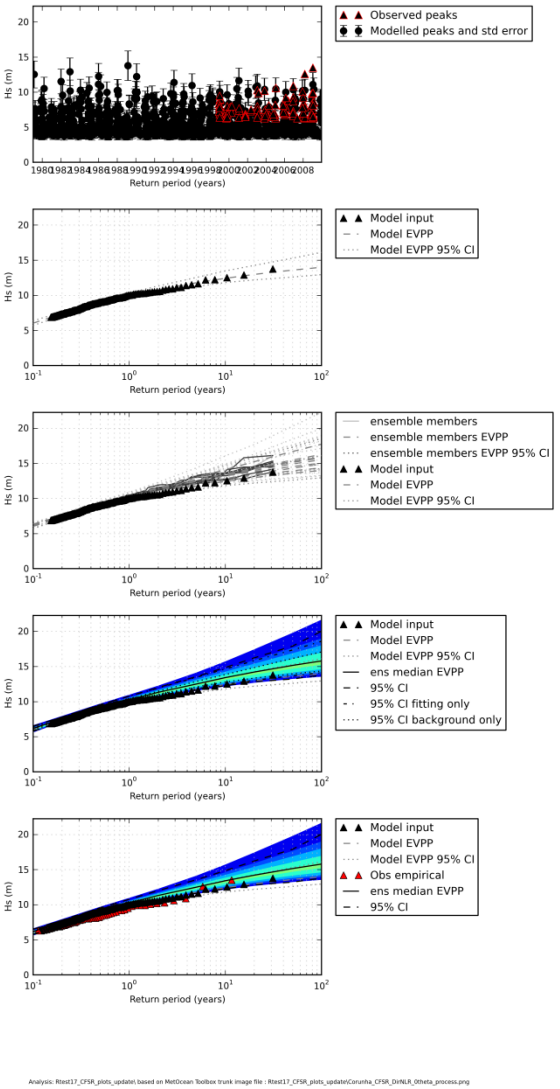


Including model skill in uncertainty estimates

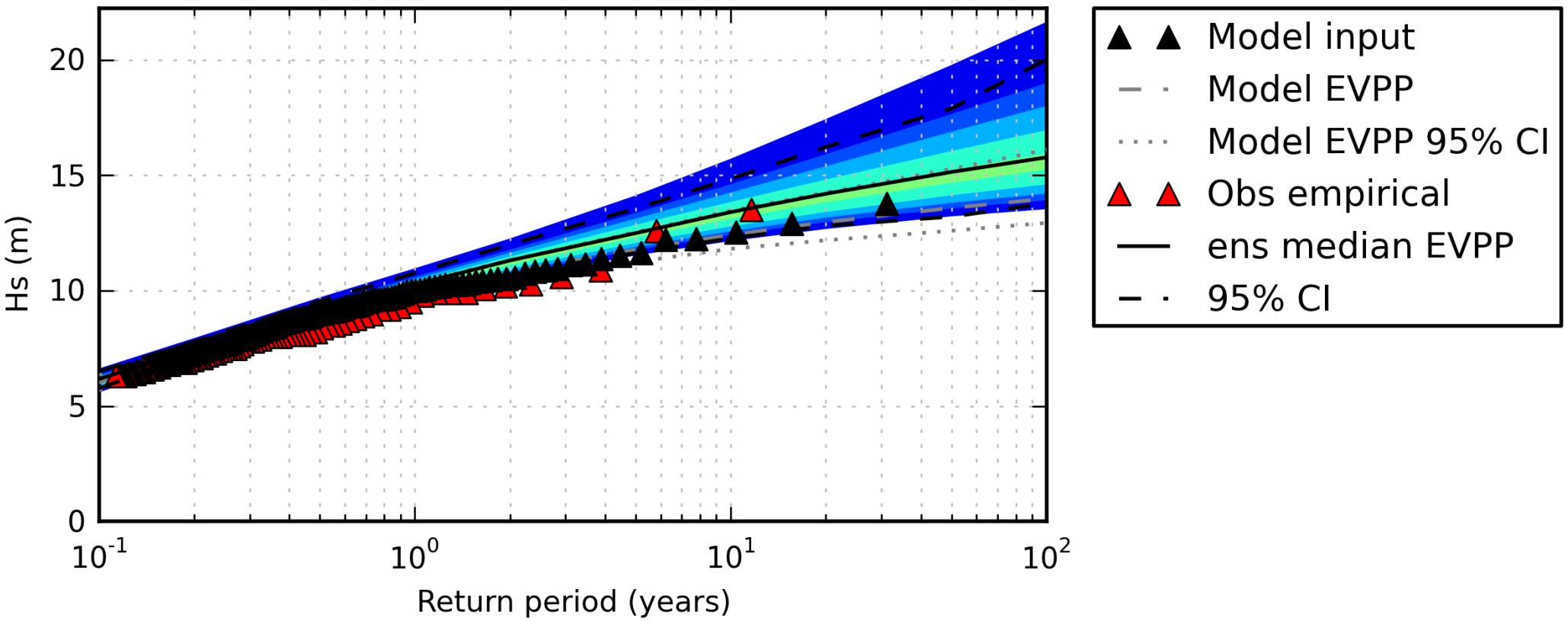


Including model skill in uncertainty estimates

- Associate uncertainty with modelled peak
- Sample -> Monte Carlo ensemble of series
- Estimate extremes on each series (using MCMC Bayesian methods)
- **Combine Monte Carlo ensembles for best estimate and total uncertainty**
- **Notice positive increment to best estimate due to introduction of uncertainty on inputs**.
- Interrogate 2D ensemble to estimate contribution to uncertainty from sampling and uncertainty on inputs

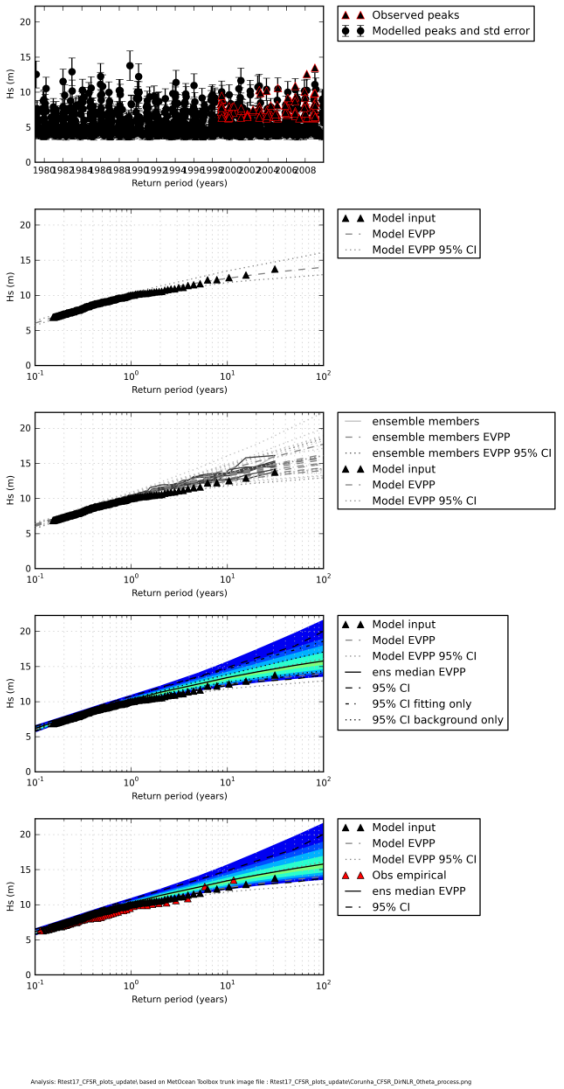


Including model skill in uncertainty estimates

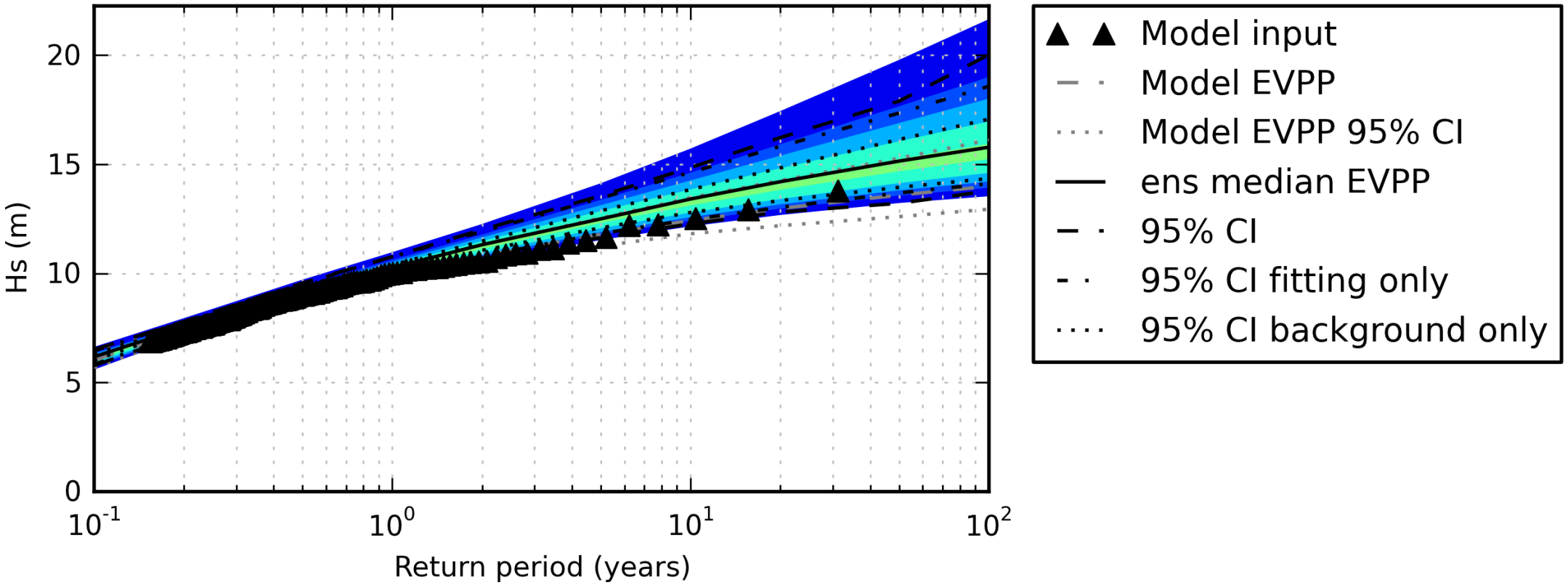


Including model skill in uncertainty estimates

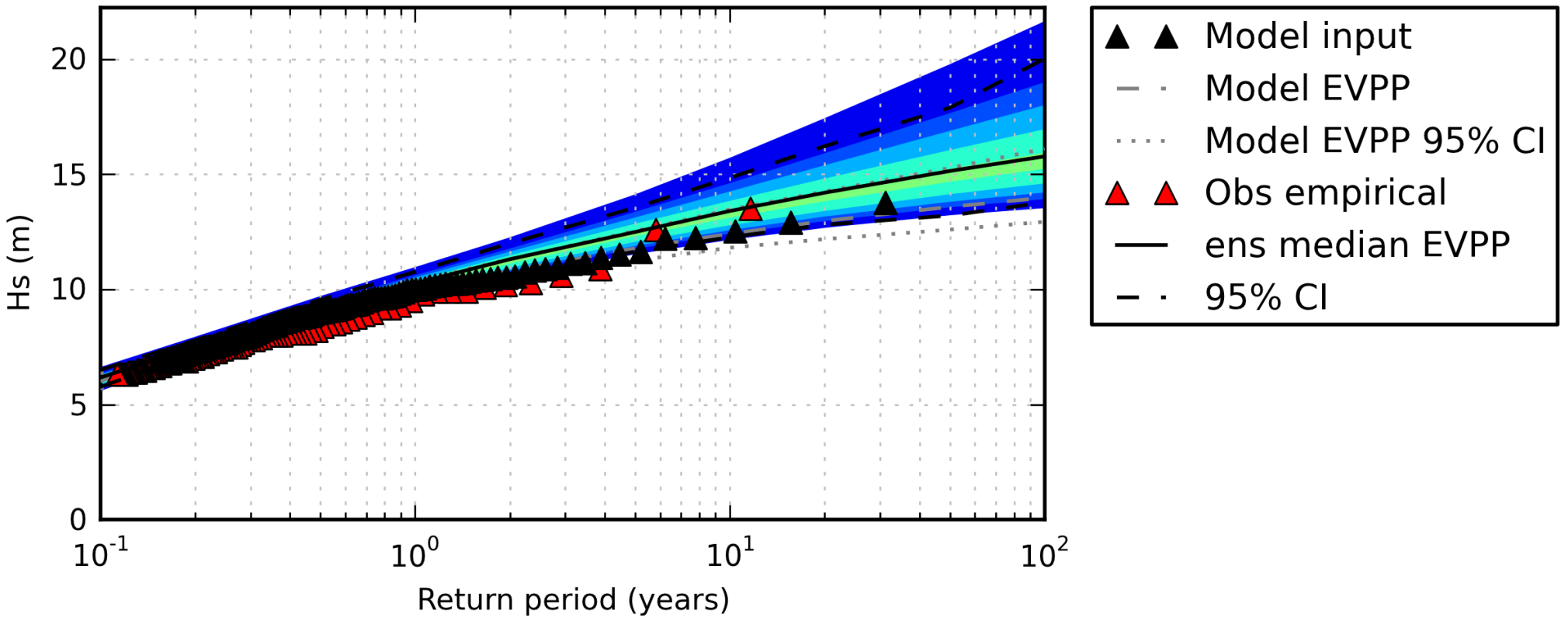
- Associate uncertainty with modelled peak
- Sample -> Monte Carlo ensemble of series
- Estimate extremes on each series (using MCMC Bayesian methods)
- Combine Monte Carlo ensembles for best estimate and total uncertainty
- Notice positive increment to best estimate due to introduction of uncertainty on inputs.
- **Interrogate 2D ensemble to estimate contribution to uncertainty from sampling and uncertainty on inputs**



Including model skill in uncertainty estimates



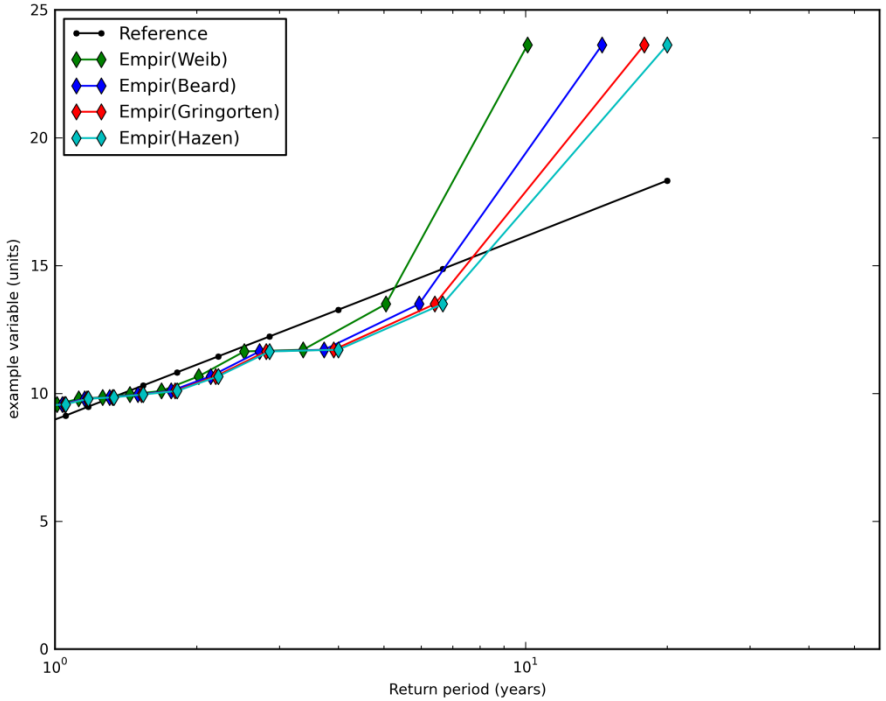
How to validate that?



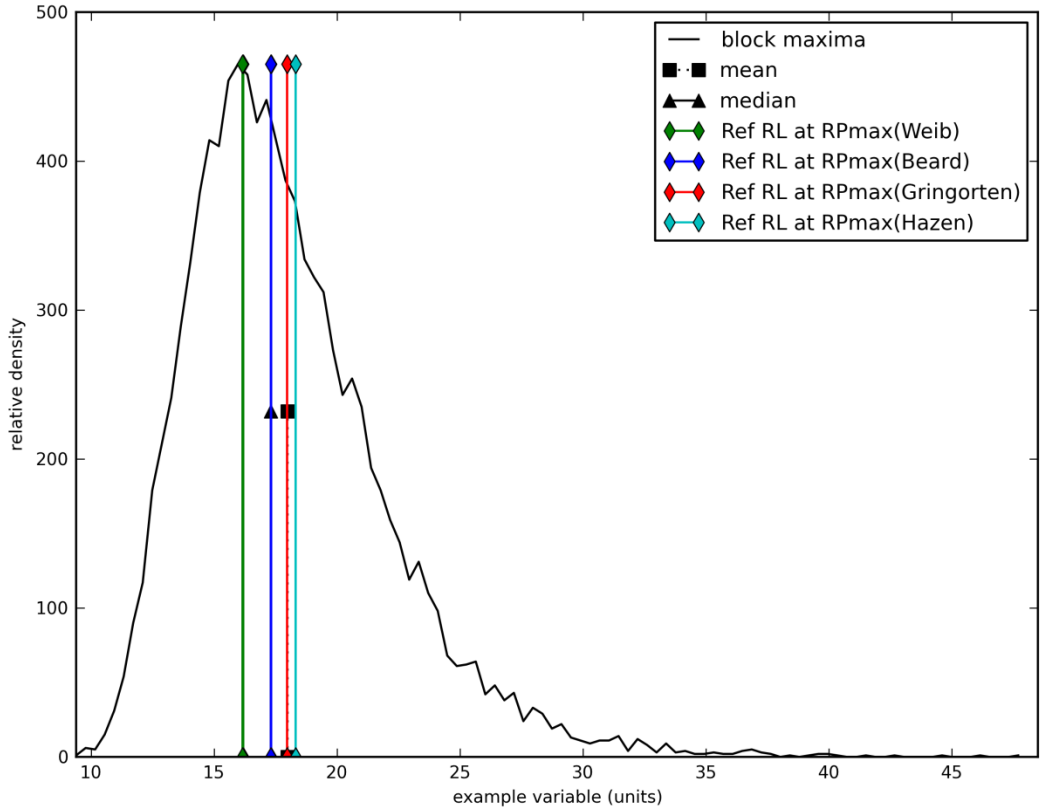
- Use large ensemble of buoys e.g. Bidlot et al?
- Verify that probability range of max (few) storms is well populated?
- -> what probability / return period to associate with observed storms?

Clarifying plotting position choice

What is the most likely return period of the maximum value in a 10 year time series?
Debated, e.g. Makkonen (2006)



GPD_0.01_2_3: shape: 0.01 loc: 2 scale: 3 nblocks: 10000, 10 years per block, 10 events per year.



GPD_0.01_2_3: shape: 0.01 loc: 2 scale: 3 nblocks: 10000, 10 years per block, mean 10 events per year.

$$P = \frac{m - a}{N + 1 - 2a}$$

Beard (a=0.31) represents median well
Gringorten (a=0.44) the mean

HR Wallingford
Working with water

Summary

- A validation technique for extremes
- Some geographical consistency in bias corrections to matched peaks
- A method for propagation of measurement/modelling errors through extreme estimation
- An investigation of return period of maxima
- More to do

HR Wallingford
Working with water

What next?

- Revisit error(modelled Hs)
- Include uncertainty in observations
- Collate uncertainty estimates across many buoys
- Validate uncertainty propagation through extremes estimation

- Use wide buoy network to establish prior distribution(s) – to inform statistical fits

- Compare short hi-res and long lo-res datasets – and look at methods of combining them

- Multivariate Extremes / Joint Probability / (Potential) Response based

That's all

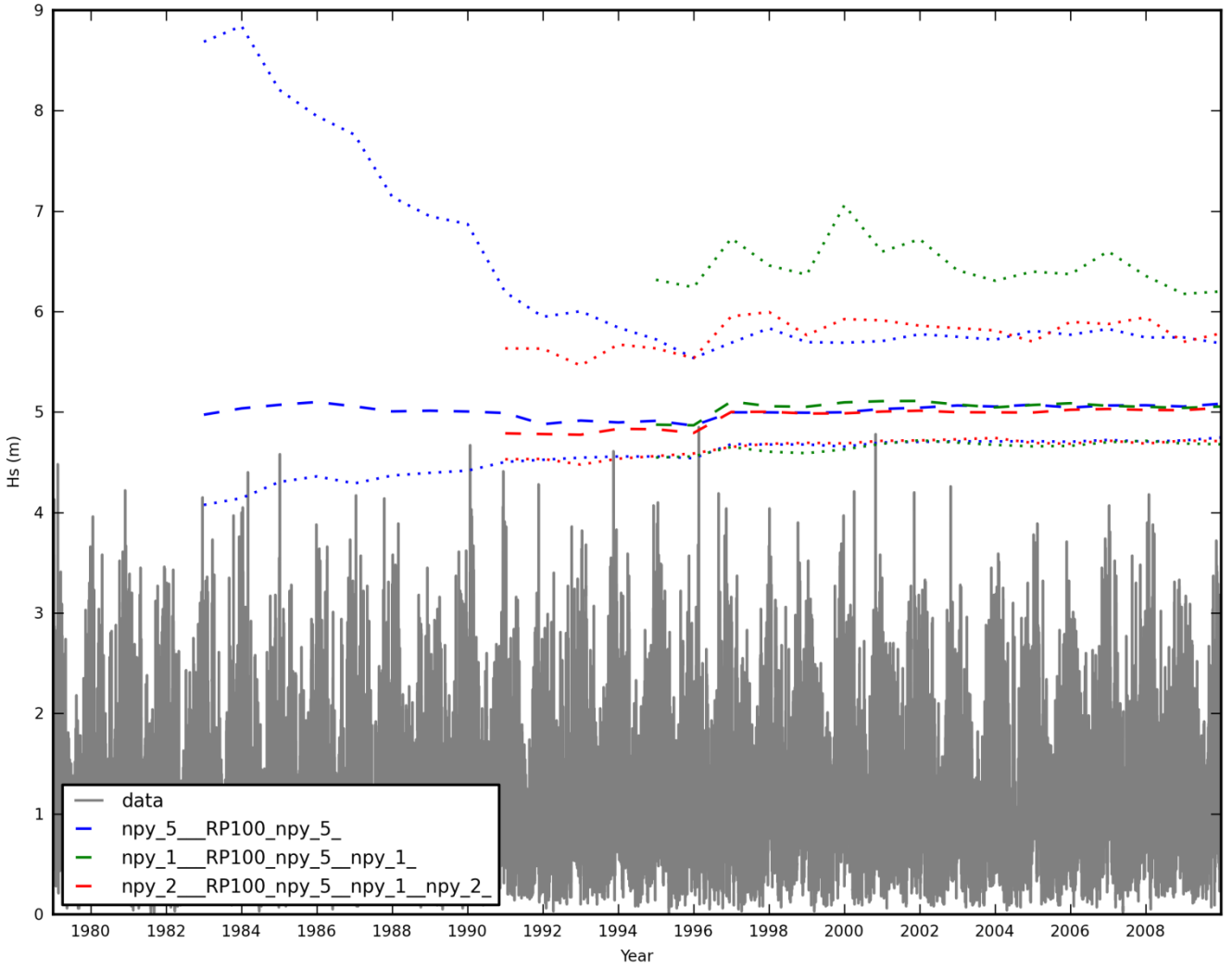
Thanks for listening

Any questions?

How often should we expect to be surprised?

Or:
We expect the unexpected, but how unexpected should we expect the unexpected to be?

Or:
how does the best estimate of Hs100 evolve with time?



HR Wallingford
Working with water

References

- Minquez, R., Espejo, A., Tomas, A., Mendez, F.J., Losada, I.J. Directional Calibration of Wave Reanalysis Databases Using Instrumental Data. J. Atmospheric and Oceanic Technology, v28, 2011
- Bidlot, J-R, Holmes, D, Wittmann, P.A., Lalbeharry, R., Chen, H.S., Intercomparison of the Performance of Operational Ocean Wave Forecasting Systems with Buoy Data, Weather and Forecasting, v17, 2002
- Bitner-Gregersen, E.M. & Magnusson , A.K. Effect of intrinsic and sampling variability on wave parameters and wave statistics, Ocean Dynamics (2014) 64:1643–1655
- Dee, D.P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi,
- U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Källberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz , J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut and F. Vitart. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137: 553–597, April 2011 A.
- Makkonen, L. Plotting Positions in Extreme Value Analysis, J. Applied Meteorology and Climatology, v45, 2006