



Ocean wind waves: Observing, modelling & forecasting ocean wind waves

High Impact Waves and Extreme Events

A meeting of the Challenger Society Special Interest Group on Ocean Wind Waves
19th & 20th October 2016, HR Wallingford Oxfordshire

ESTIMATING EXTREME SIGNIFICANT WAVE HEIGHTS BY INTEGRATING MODEL AND WAVEMETER DATA

F. Dentale^{1,2}, E. Pugliese Carratelli^{1,2}, P. Furcolo², F. Reale¹, R. Tomasicchio³,

(1) CUGRI –University Centre for Research on Major Hazards, Fisciano 84084, Italy - e-mail: epc@unisa.it

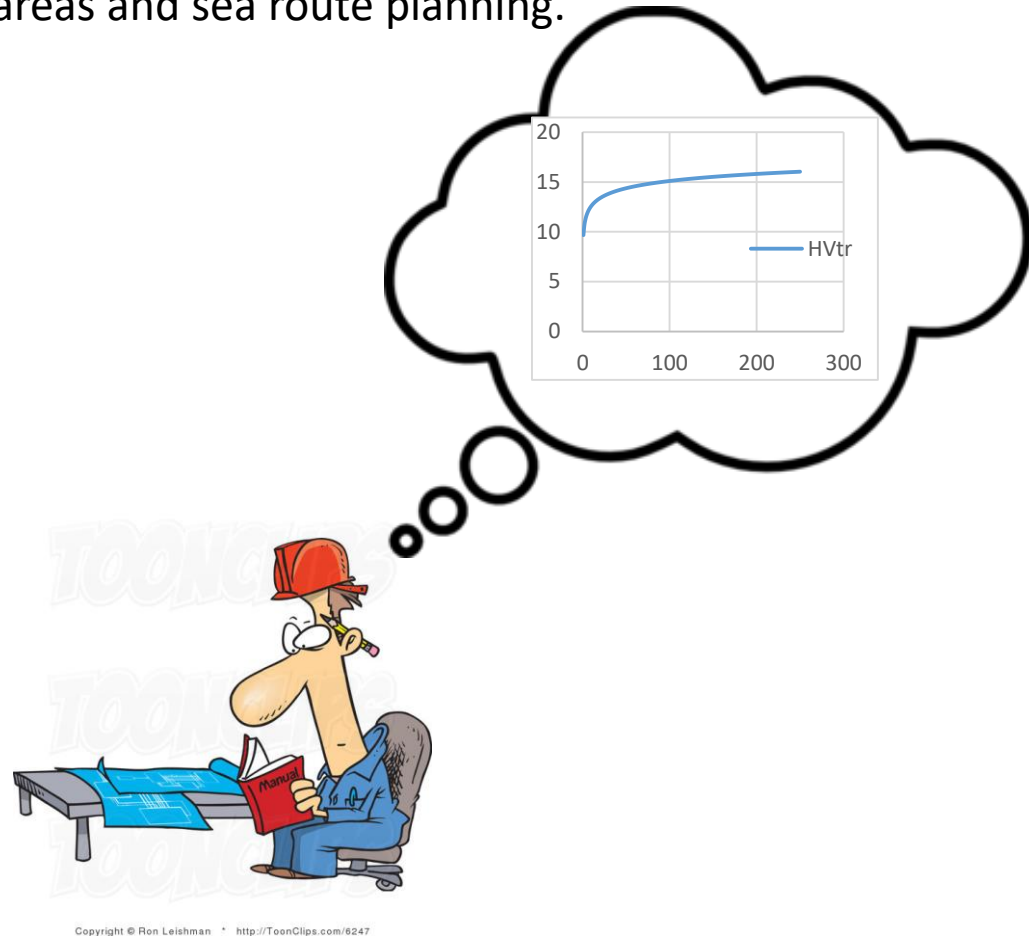
(2) Dipartimento Ingegneria Civile, University of Salerno, Fisciano 84084, Italy

(3) Dipartimento di Ingegneria dell'Innovazione dell' Università del Salento. Lecce, Italy

What's is it all about:

The evaluation of wave climate and particularly of storm extremes is one of the most important aspects of sea related activities, such as coastal and offshore constructions, civil protection of coastal areas and sea route planning.

;



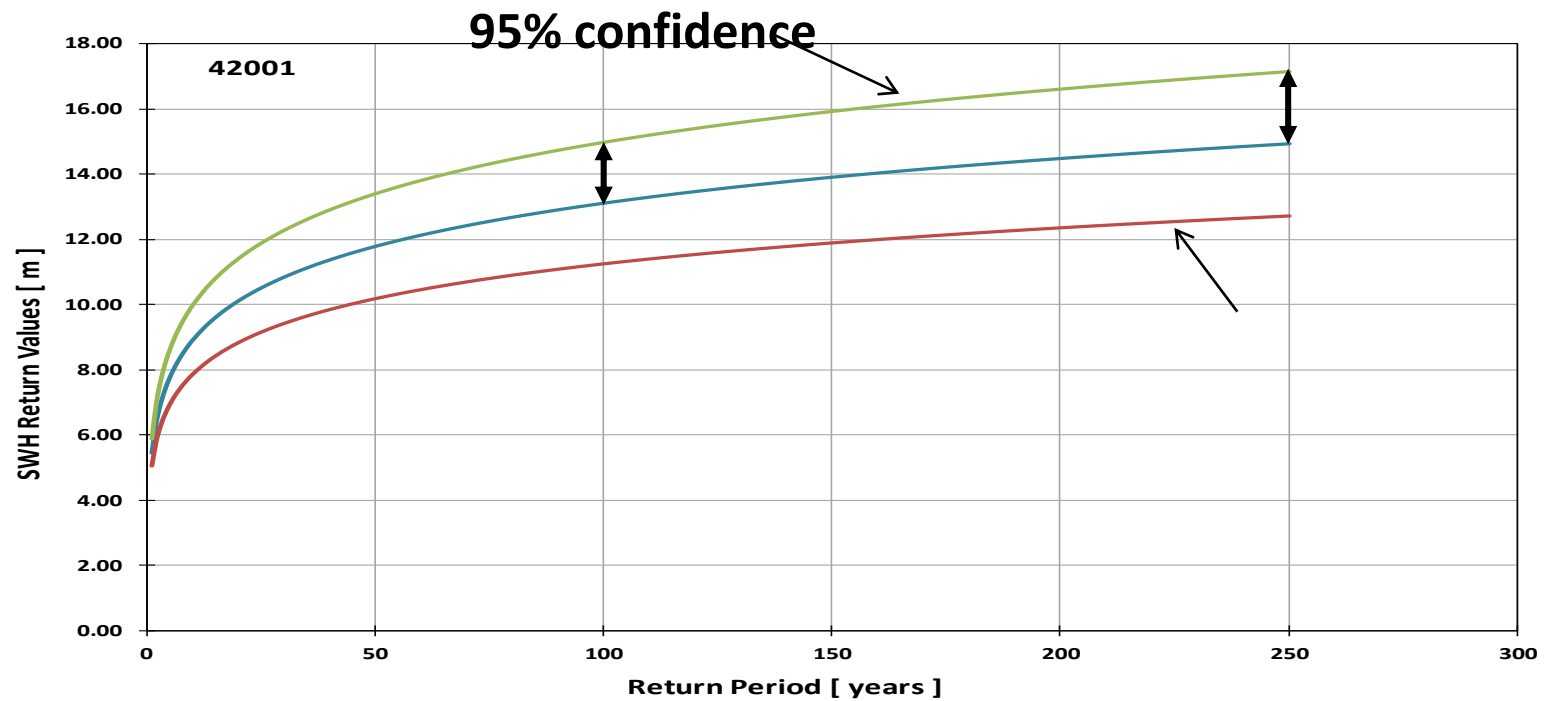
Copyright © Ron Leishman * <http://ToonClips.com/6247>

The most important tool of coastal and offshore engineering is probability distribution of Significant Wave Height (SWH), generally described by the SWH/Return Time function $SWH(Tr)$

Significant wave height SWH as a function of its return period

SWH (Tr).

The main tool of the Coastal and offshore engineer



Weather and wave models

On sites with a long historical record of wavemeter data, **the use of measured data is** the obvious choice. However, on most locations there is no adequate history of recorded data.

The availability of data generated by global and regional wind and wave model chains (in the following: *WeWaM*) have brought radical changes to the estimation procedures of extremes. Models are routinely run all over the world and SWH time series for each grid point are computed and published after assimilation (analysis) of sea truth data, generally produced by satellite altimetry

- 1 **Global Weather Model Archive Data.**
 - 2 **Local Area Weather Model(s).**
 - 3 **Wave Generation and Propagation Model.**
 - 4 **Statistical analysis of the synthetic wave data on the site.**
- } WeWaM

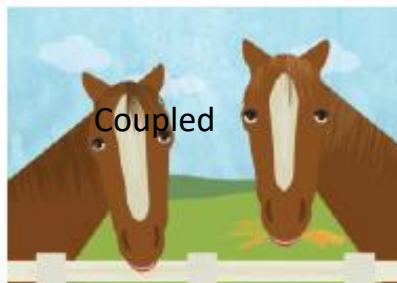
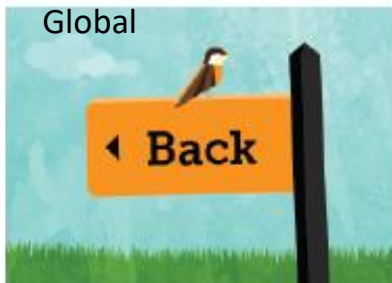
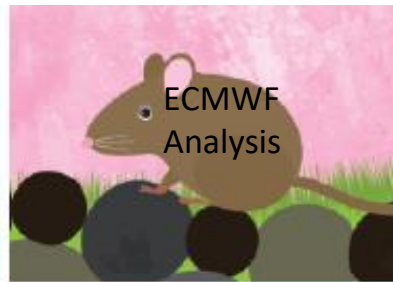
(Wave transformation on shallow water if necessary can be added to either step **3** or **4**.)

numerical weather forecast models are now widely available,
together with long records of past analyses and forecasts:
ECMWF, NECP, State Agencies...

Often at price!



There is a zoo of models, with a whole biodiversity of options



How to choose?



Or, better, how to evaluate the reliability i.e. the probability of errors in the curve I will be using?



Most models assimilate SWH data from satellite altimeters.

Altimetry data are only available at long time intervals, depending on the satellite coverage, so the probability that the assimilation is carried out during an extreme event is low: this may induce a bias and reflects badly on the reliability of the highest simulated SWHs and therefore on the quality of the extrapolated SWH(Tr).

In the applications, we are particularly concerned about high return times: 50, 100, 200 yrs.

Tests can be carried out for extreme values on model vs. buoy data; the results show relevant bias and scatter specially and generally (highest percentiles). SWH(Tr) derived from model data can thus be affected by a considerable error

Buoy data are mostly used for verification:



From the venerable pioneering work:

Challenor and Cotton, 1997; Queffeulou, 1996; Ardouhin et al 2007)....

to the recent production:

JCOMM ETWCH Expert Team on Waves and Coastal Hazards Forecasting Systems : Monthly verification

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., Bidlot, J.-R., 2011. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. *J Geophys Res* 116, 1–18

Aarnes, O. J., Breivik, Ø., Reistad, M., 2012. Wave Extremes in the Northeast Atlantic. *J Climate* 25, 1529–1543

Breivik, Ø., O. J. Aarnes, J.-R. Bidlot, A. Carrasco, and Ø. Saetra, 2013: Wave Extremes in the North East Atlantic from Ensemble Forecasts. *J Climate*, 26, 7525–7540, arXiv:1304.1354, Extreme waves ENS/Buoys

Breivik, Ø., O. Aarnes, S. Abadalla, J.-R. Bidlot, and P. Janssen, 2014a: Wind and Wave Extremes over the World Oceans From Very Large Ensembles. *Geophys Res Lett*, 41 (14), ERA Interim / ENS 100 yr rt [SWH\(TR\) from ensemble forecast](#)

Sartini, L.; Mentaschi, L. & Besio, G., (2015b). Comparing different extreme wave analysis models for wave climate assessment along the Italian coast. *Coastal Engineering*, No 100, 37–47.



Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., Bidlot, J.-R., 2011.
 A high-resolution hindcast of wind and waves for the North Sea, the Norwegian
 Sea, and the Barents Sea. J Geophys Res 116, 1–18,

H_s	N	ME	RMSE	MAE	SI	Corr	M	P95	P95	P95	P99	P99	P99	P99
	obs						stn	NME	SI	Corr	NME	SI	Corr	
WAM10	235368	0.07 m	0.47 m	0.33 m	20.8 %	0.95	40	5.0 %	8.9 %	0.96	4.7 %	9.0	0.95	
ERA-40	235368	-0.01 m	0.55 m	0.40 m	24.2 %	0.93	40	-5.0%	11.80 %	0.91	-5.7 %	13.0	0.88	

Table 5. Significant wave height observed v. WAM10 and ERA-40 at the 40 coastal and offshore stations [Saetra and Bidlot, 2004]. A total of 235,368 observations were recorded and quality controlled. The observations were averaged over 4 hours. The right part of the table compares the normalized mean error (NME) and the scatter index (SI) of the 95 and 99 percentiles for M=40 stations.

Observations over 4 hours do not provide a reliable maximum for engineering purposes. SWH may oscillate with a correlation time much lower than that

Reistad, M., Breivik, Ø., Haakenstad, H., Aarnes, O. J., Furevik, B. R., Bidlot, J.-R., 2011.
A high-resolution hindcast of wind and waves for the North Sea, the Norwegian
Sea, and the Barents Sea. J Geophys Res 116, 1–18,

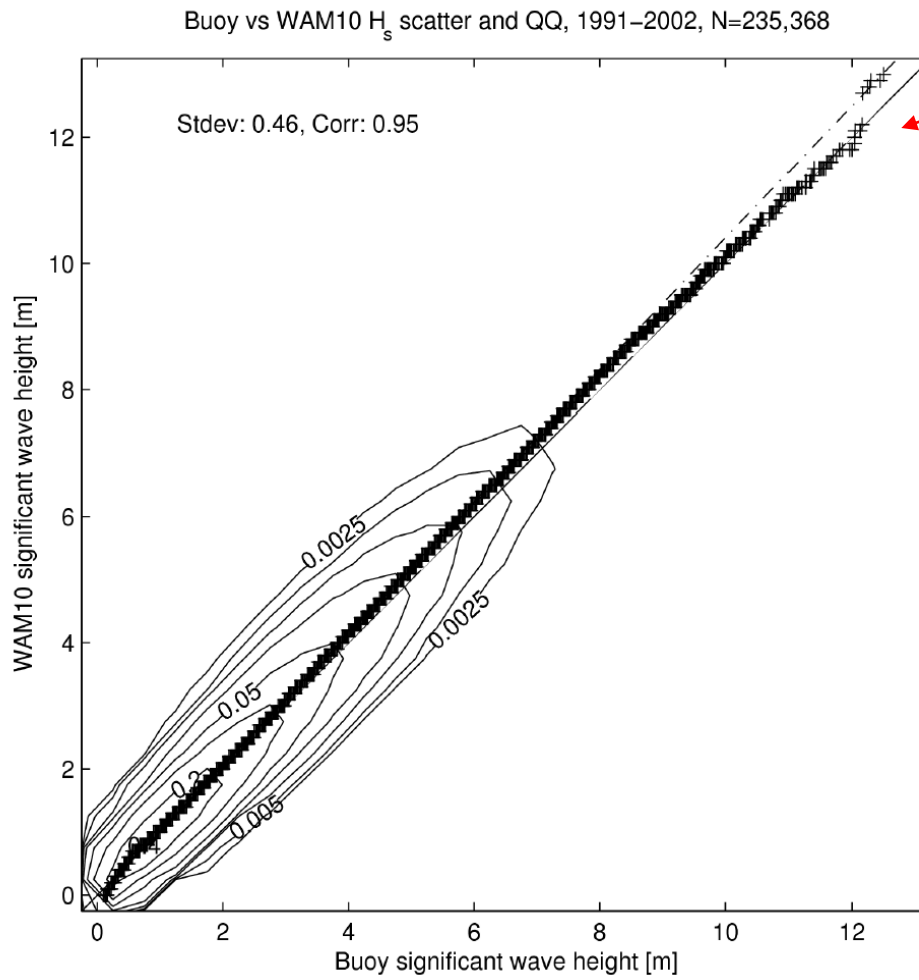


TABLE 1. Statistical comparison of modeled and observed H_s for $H_s > 0$, $H_s > H_s(p_{95})$ and $H_s > H_s(p_{99})$ at Ekofisk (2001–2009), Gullfaks C (1999–2009), and Draugen (1996–2009). The 3-hourly NORA10 data are validated against the maximum observed $H_s(20 \text{ min})$ within $\pm 1.5 \text{ h}$ of NORA10, the mean observed H_s over periods of 20 min and 1, 3, and 6 h, centered at the time of NORA10. Statistical measures shown are the scatter index (SI, %), the **NORA10-OBS bias (m)**, the correlation coefficient (R), and the regression line of $\text{NORA10} = a + b \times \text{obs}$.

	Obs period	Ekofisk				Gullfaks C				Draugen			
		SI	Bias	R	$a + bx$	SI	Bias	R	$a + bx$	SI	BIAS	R	$a + bx$
$H_s > 0$	max 20 min	16.50	-0.17	0.97	$0.07 + 0.89x$	16.29	-0.24	0.95	$-0.01 + 0.92x$	20.83	-0.18	0.94	$0.15 + 0.88x$
	20 min	17.54	0.07	0.96	$0.03 + 1.02x$	17.26	0.01	0.95	$0.08 + 0.97x$	21.17	0.05	0.94	$0.23 + 0.93x$
	1-h mean	16.03	0.06	0.97	$0.00 + 1.03x$	16.60	0.01	0.96	$0.06 + 0.98x$	20.42	0.06	0.94	$0.23 + 0.94x$
	3-h mean	15.48	0.06	0.97	$-0.01 + 1.04x$	15.86	0.01	0.96	$0.04 + 0.99x$	19.60	0.05	0.95	$0.22 + 0.94x$
	6-h mean	15.23	0.06	0.97	$-0.04 + 1.05x$	15.07	0.01	0.96	$-0.01 + 1.01x$	18.31	0.05	0.95	$0.18 + 0.95x$
$H_s > H_s(p_{95})$	max 20 min	11.33	-0.56	0.82	$1.05 + 0.73x$	11.21	-0.43	0.79	$-0.18 + 0.96x$	12.30	-0.56	0.78	$-0.24 + 0.96x$
	20 min	12.44	0.10	0.78	$0.48 + 0.93x$	12.76	0.02	0.76	$-0.06 + 1.01x$	13.92	-0.10	0.75	$-0.35 + 1.04x$
	1-h mean	10.47	0.16	0.85	$0.22 + 0.99x$	12.22	0.05	0.79	$-0.35 + 1.06x$	13.29	-0.07	0.78	$-0.56 + 1.07x$
	3-h mean	9.76	0.18	0.87	$0.19 + 1.00x$	11.70	0.09	0.82	$-0.56 + 1.10x$	12.84	-0.06	0.80	$-0.60 + 1.08x$
	6-h mean	9.88	0.25	0.86	$0.23 + 1.00x$	11.75	0.20	0.82	$-0.61 + 1.13x$	12.85	0.02	0.80	$-0.59 + 1.09x$
$H_s > H_s(p_{99})$	max 20 min	11.18	-0.80	0.73	$2.20 + 0.61x$	9.86	-0.45	0.71	$0.56 + 0.88x$	10.48	-0.53	0.73	$0.35 + 0.90x$
	20 min	11.93	0.02	0.81	$0.97 + 0.86x$	11.17	0.04	0.65	$0.65 + 0.93x$	11.82	0.02	0.69	$0.00 + 1.00x$
	1-h mean	9.43	0.13	0.78	$0.17 + 0.99x$	10.59	0.13	0.70	$0.19 + 0.99x$	11.54	0.08	0.73	$-0.08 + 1.02x$
	3-h mean	8.68	0.16	0.82	$0.37 + 0.97x$	9.91	0.22	0.73	$0.34 + 0.99x$	11.29	0.14	0.73	$0.32 + 0.98x$
	6-h mean	9.27	0.26	0.79	$0.29 + 1.00x$	10.36	0.42	0.71	$0.97 + 0.93x$	11.48	0.35	0.71	$0.74 + 0.95x$

Maximum Observed within 1.5 h



Aarnes, O. J., Breivik, Ø., Reistad, M., 2012. Wave Extremes in the Northeast Atlantic. J Climate 25, 1529–1543, doi:10.1175/JCLI-D-11-00132.1

Working on the 99° percentile

So, there might be a problem on extreme values as obtained from model derived historical data series. This would reflect on SWH(Tr) curves



Also, not all WeWaM (model chains) are equal, and not everywhere equally accurate; a simple test of their efficiency would be welcome

Why don't we try and to work on (extrapolated) return time curves rather than on SWH output data?

For engineering purposes, we would like a procedure to evaluate the reliability of SWH(Tr) curves derived from model data for any location at sea; in the following such a procedure is proposed and tested, in areas where an adequate number of in situ wave meters are available.

SWH(Tr) curves are derived by analysing WeM+WaM model time series and by integrating them with wavemeter time series in the same geographical area. This provides an estimate of SWH(Tr) for any given geographical point in the area and at the same time helps to assess the quality of the whole model system.

The basic concept of the procedure is that the parameters of any SWH(Tr) extreme value (A,B,k) curve are themselves randomly distributed, and that the distribution of such parameters can be estimated by analysing in situ data for any given area.

In principle, there is nothing new in this: similar approaches have been used by meteorologists or hydrologist for many years (regionalization of extremes)



So...

As a first step a Weibull SWH(Tr) between return time Tr and significant wave height SWH(Tr) is computed from the historical data sets of wave meters in three meteorologically and oceanographically homogenous areas, ie: Southern Italian Seas, Iberian Atlantic Waters, and Northern Gulf of Mexico (in the following respectively SIS, IAW, NGM)

The Peak Over Threshold (POT) method upon with a Weibull distribution

e.g Goda, or Sartini et al. (2015b)



$$F(SWH) = 1 - \exp \left[- \left(\frac{SWH - B}{A} \right)^k \right]$$

where A, B and k are known respectively as scale, position and shape parameters.

Once the distribution parameters are known, the H return value for a given return period Tr (in years) is evaluated by

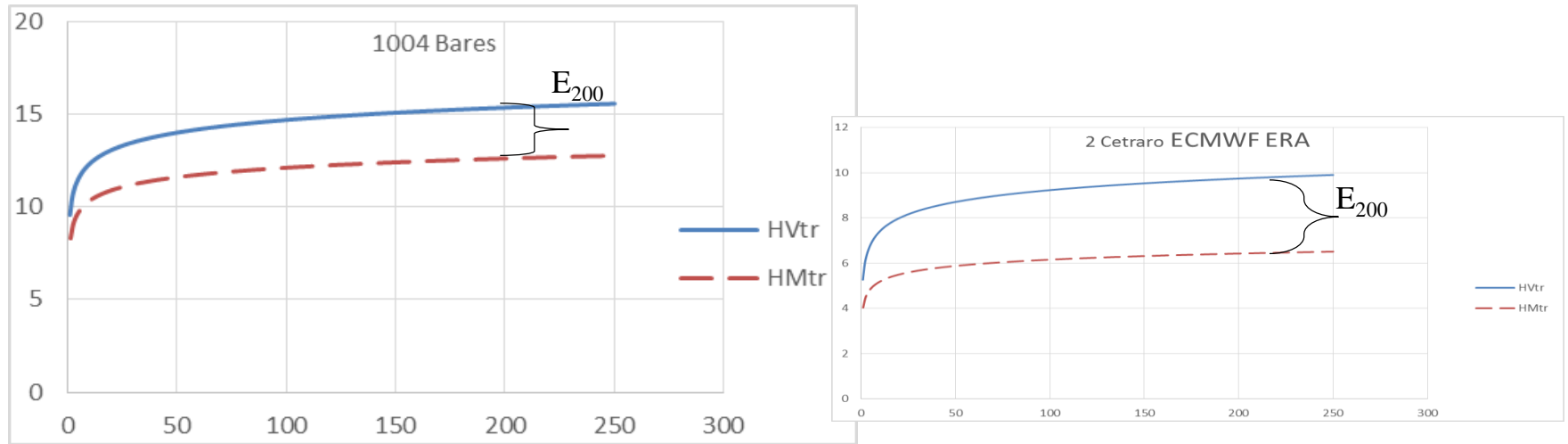
$$SWH = B + A[\ln(\lambda T_r)]^{\frac{1}{k}}$$

An important choice to be made is the threshold over which the SWH extreme values are to be chosen. The value was chosen by making sure that the number of events considered should be roughly equal for all the samples, i.e. by assuming similar values for the parameter $\lambda = NT/n$, NT being the number of events considered and n the observation length in years. (Goda)



But at this stage, we are not concerned about at evaluating, discussing or recommending one particular form of SWH(Tr), or any particular procedure to estimate its parameters: the only requirement is that such form and procedure should be uniform throughout the whole analysis.

The same is done for the model data (colocated) at the buoy locations



Typical Weibull SWH (Tr) curves from **model** (HMtr) and **buoy** (HVtr) data.

$$\text{Error } E_{Tr} = HV_{Tr} - HM_{Tr}$$

A comparison is done between the extreme value distribution obtained by making use of buoy wavemeter measurement (HV) and model data (HM) in meteorologically and oceanographically homogenous areas (i.e. Southern Italian seas, North Atlantic Iberian waters)

We then look for a correlation between the two curves

$$X_{est} = X_{model} + X_{model} * e(\mu, \sigma) :$$

Multiplicative error

Other forms are possible

$$X_{est} = X_{model} + E(\mu, \sigma)$$

$$X_{est} = X_{model} + E(\mu(X_{model}), \sigma(X_{model}))$$



A stated before, Weibull e.v. distribution connecting return time T_r and **significant wave height H** is computed for historical data sets for both model HM_{T_r} and buoy HM_{T_r} value

using the highest available sampling frequency For both model (6hrs, 3hrs, 1hr) and buoys (usually 30', 1h)

Symbols changed for convenience:
 H is Significant Wave height

So we assume an estimator $HST(T_r)$ given by

$$HST_{T_r} = HM_{T_r} + HM_{T_r} * e_{T_r}$$



The relative error e_{T_r} is assumed to be normally distributed with average μe and rms σe . The value of the parameters can estimated from the “True” values of H at the n buoy locations; so the relative error at location l is

For each T_r

$$e_{T_r}^l = [HV_{T_r}^l - HM_{T_r}^l] / HM_{T_r}^l$$

In each area there are N buoy locations

$HV_{T_r}^l$ Weibull buoy value of SWH at location l for given T_r (“true”) value

$HM_{T_r}^l$ Weibull model value of SWH at location l for given T_r (“model”) value

The average regional error is then estimated as

$$\mu e_{Tr} = [\sum e_{Tr}^1] / N$$

and its root mean square

$$\sigma e_{Tr} = \sqrt{[\sum (e_{Tr}^1 - \mu e_{Tr})^2 / (N - 1)]}$$

(sum over l)
N=number of
Stations
HV «buoy, true»
HM «model»



An estimated value HST, for any location in the area can thus be computed as

$$HST_{Tr} = HM_{Tr} + HM_{Tr} * \mu e_{Tr}$$

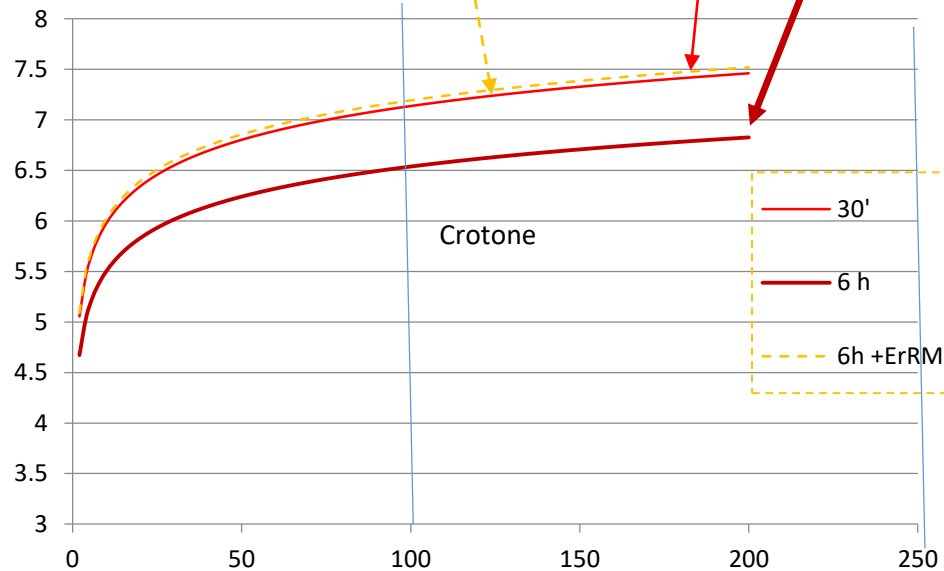
Value of HST_{Tr} with a given probability p is given by

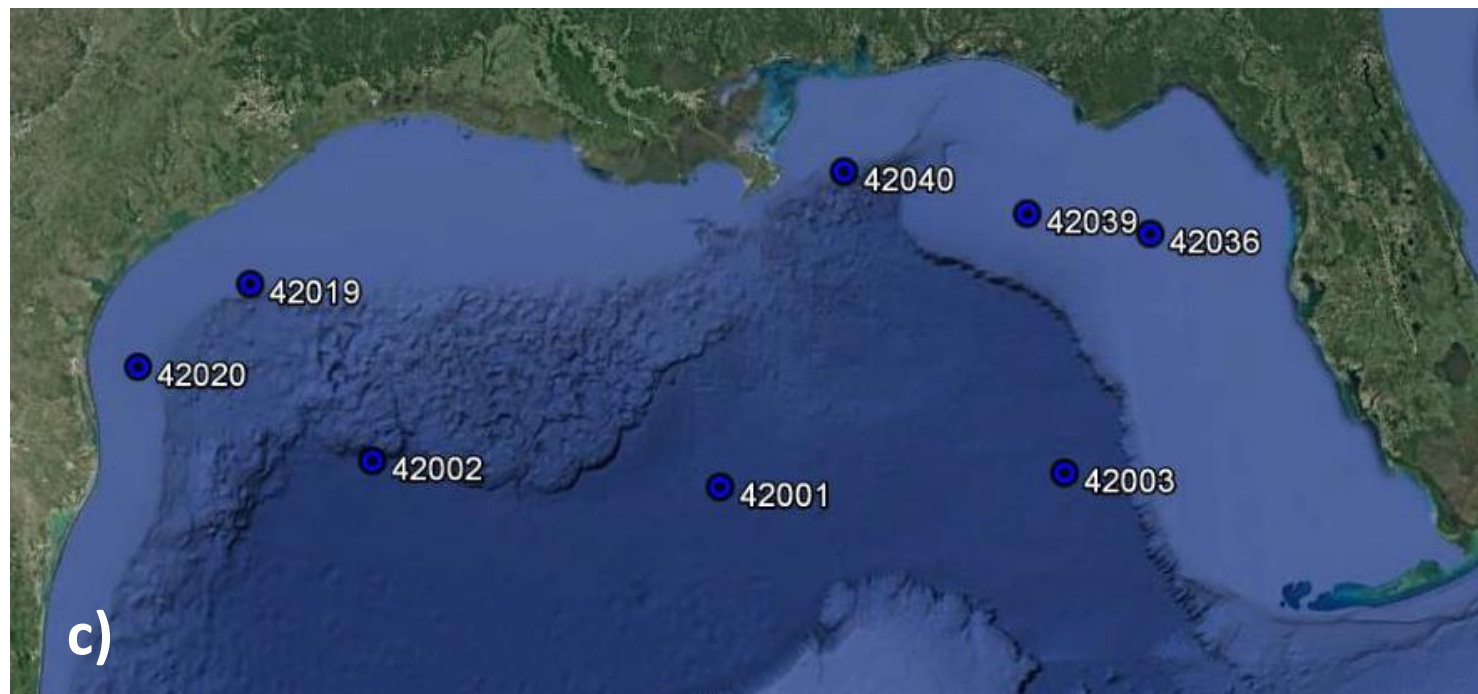
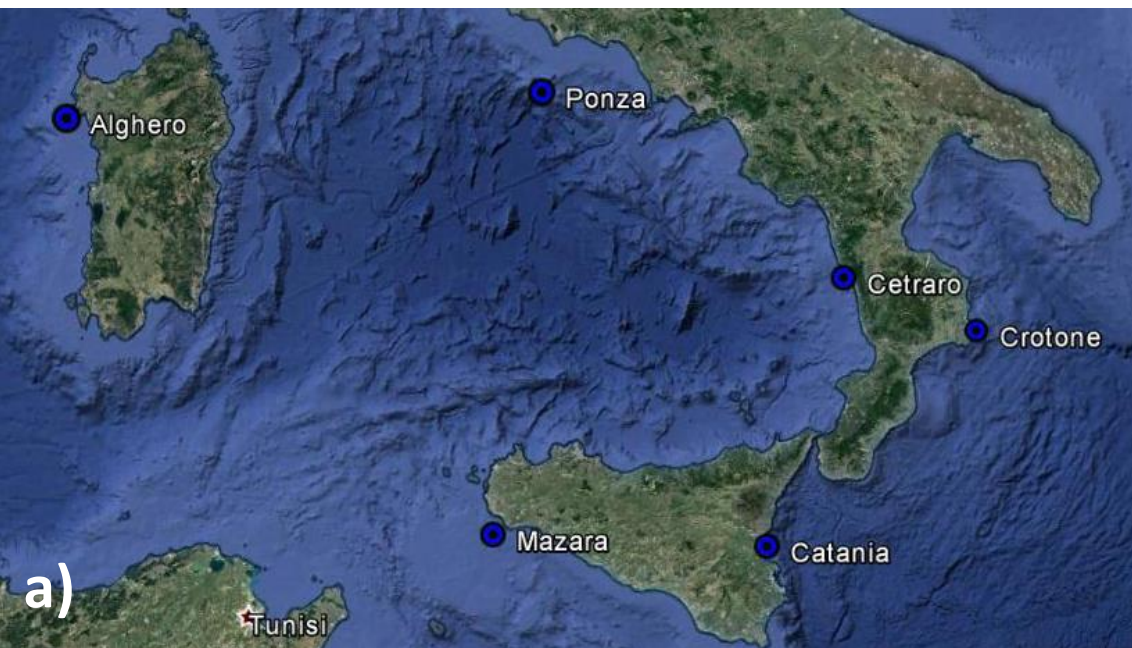
$$HST_{Tr}^p = HM_{Tr} + HM_{Tr} * \mu e_{Tr} + Up * HM_{Tr} * \sigma e_{Tr}$$

Results show the comparison between "true" and model values, as well as the estimated

$$HST_{Tr} = HM_{Tr} + HM_{Tr} * \mu e_{Tr}$$

obtained by adding the mean regional error $HM_{Tr} * \mu e_{Tr}$ to the model value HM_{Tr}





Buoy	from	to	sample	Total data	Validate data	Eff. obs. years (n)
Alghero	01 Jul 1989	05 Apr 2008	30'	125443	109817	17.13 years
Catania	01 Jul 1989	05 Oct 2006	30'	101394	91075	14.91 years
Cetraro	01 Jan 1999	05 Apr 2008	30'	95726	90327	7.79 years
Crotone	01 Jul 1989	15 Jul 2007	30'	119854	110542	16.34 years
Mazara	01 Jul 1989	04 Apr 2008	30'	121714	105638	15.80 years
Ponza	01 Jul 1989	31 Mar 2008	30'	115651	100256	16.17 years

Buoy	from	to	sample	Total data	Validate data	Eff. obs. years (n)
Bilbao	17 Apr 2002	27 Nov 2010	1 h	64289	62692	7.09
C. De Penhas	13 Jan 1998	08 Nov 2010	1 h	90952	88881	10.15
C. Silleiro	06 Jul 1998	27 Nov 2010	1 h	91945	91192	10.29
Cadiz	09 Nov 1997	19 Dec 2010	1 h	103721	101131	11.53
Estaca Bares	15 Jan 1998	01 Sep 2010	1 h	79701	76082	8.68
Villano Sisar.	12 May 1998	27 Nov 2010	1 h	86093	81764	9.34

Buoy	from	to	sample	Total data	Validate data	Eff. obs. years (n)
42001	01 Jan 1979	31 Dec 2007	1 h	231514	219719	26.28
42002	01 Jan 1979	31 Dec 2007	1 h	234259	220441	26.33
42003	01 Jan 1979	31 Dec 2007	1 h	223416	205583	24.50
42019	25 May 1990	31 Dec 2007	1 h	134778	128379	14.83
42020	24 May 1990	31 Dec 2007	1 h	135087	124928	14.49
42036	01 Jan 1994	31 Dec 2007	1 h	112621	100896	11.83
42039	12 Dec 1995	31 Dec 2007	1 h	101544	98329	11.51
42040	04 Dec 1995	31 Dec 2007	1 h	100789	98959	11.51



Southern Italian Waters

Extreme SWH(Return Time)

Model(NCEP,ECMWF) vs buoys data

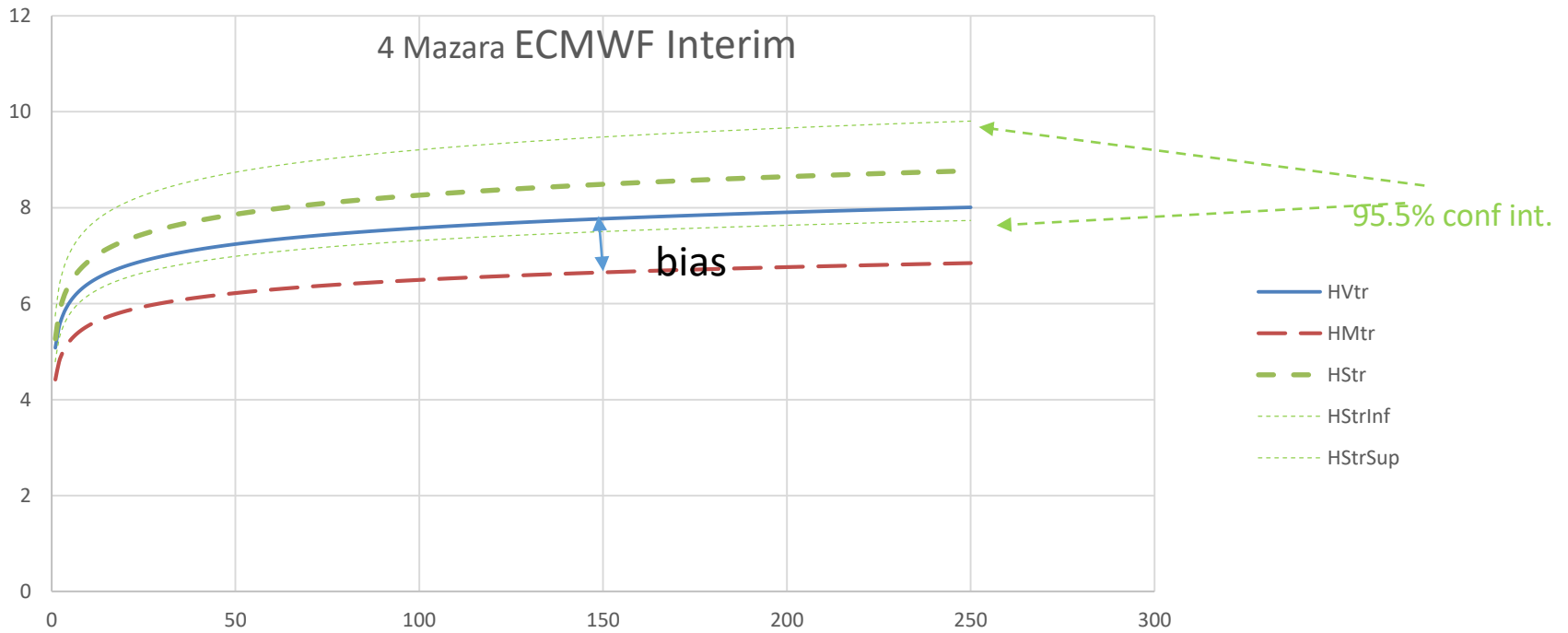
Wavewatch III Reanalysis 10' grid 3 h (Integrated model)
ECMWF Mediterranean ? Reanalysis

These wave buoys have been operating for many years, and between 1989 and 2008 (circa) have provided significant wave height (H_s) values at 30' intervals. The table shows the data availability for RON buoys considered.

There are of course long periods of missing data for various reasons. Not all the data were used

(Italian National Wavemeter Buoys (RON –ISPRA) data series)

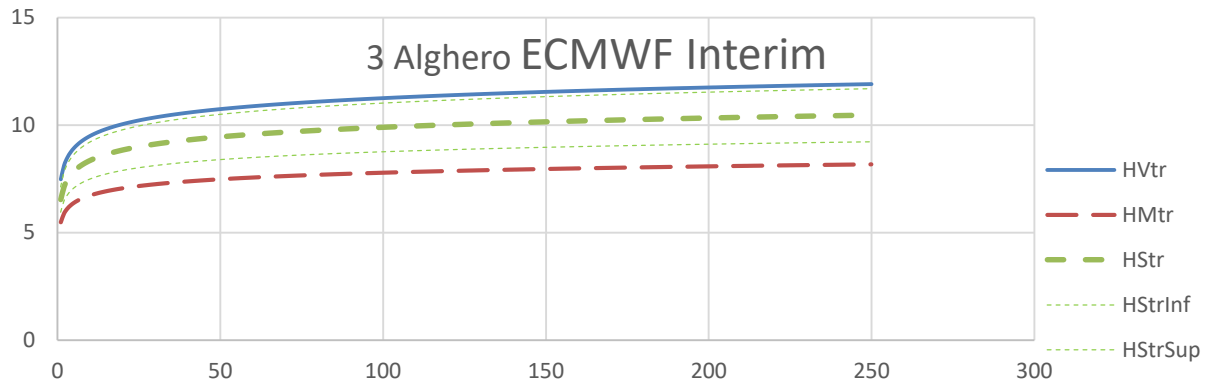
Buoy	Start	End
Alghero	01 - July - 1989	05 - April - 2008
Catania	01 - July - 1989	05 - October - 2006
Cetraro	01 - January - 1999	05 - April - 2008
Crotone	01 - July - 1989	15 - July - 2007
Mazara	01 - July - 1989	04 - April - 2008
Ponza	01 - July - 1989	31 - March - 2008



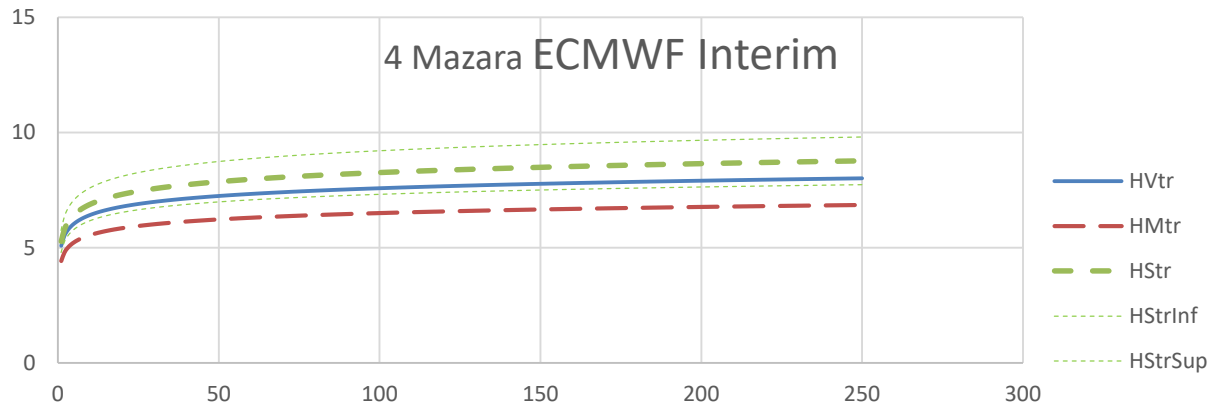
Weibull curves at buoy location: **HV(Tr): buoy ('true') value**; **HMt(Tr): model value**; **HV(Tr): estimated with model and local regional error** ; **HStrInf, HSSup 5% (Tr): confidence interval estimated with rms of difference between model and buoy dat**

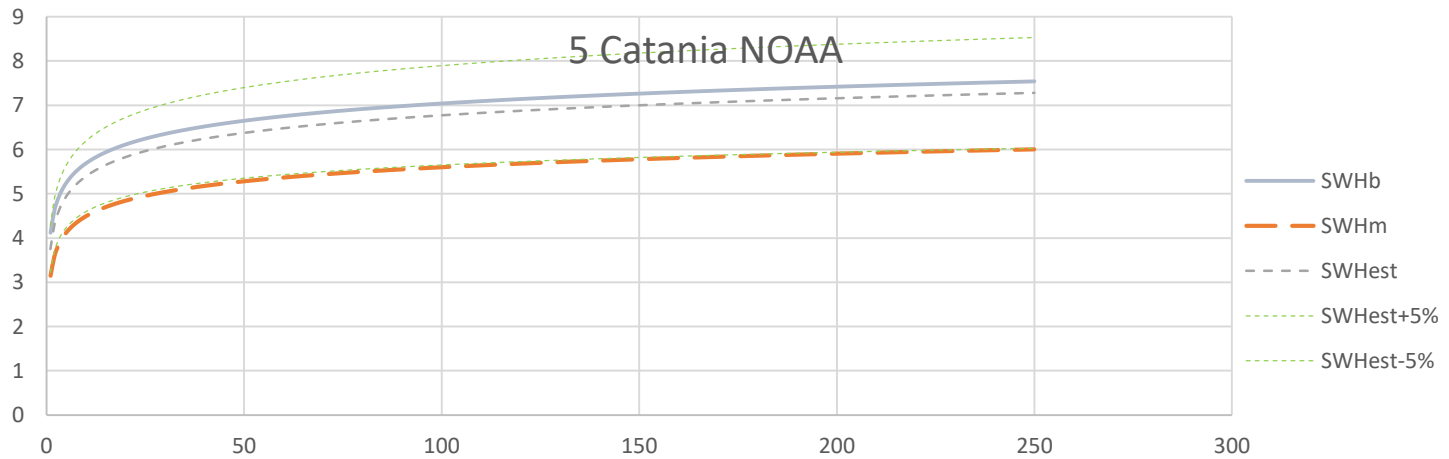
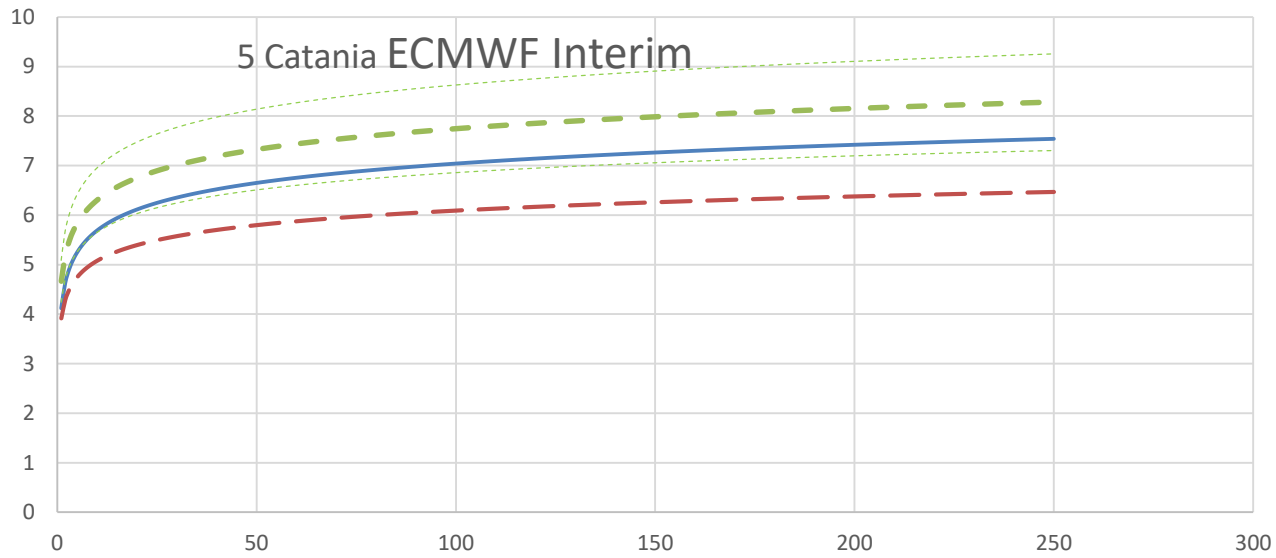
Measures uncertainty deriving from using model rather than buoy data





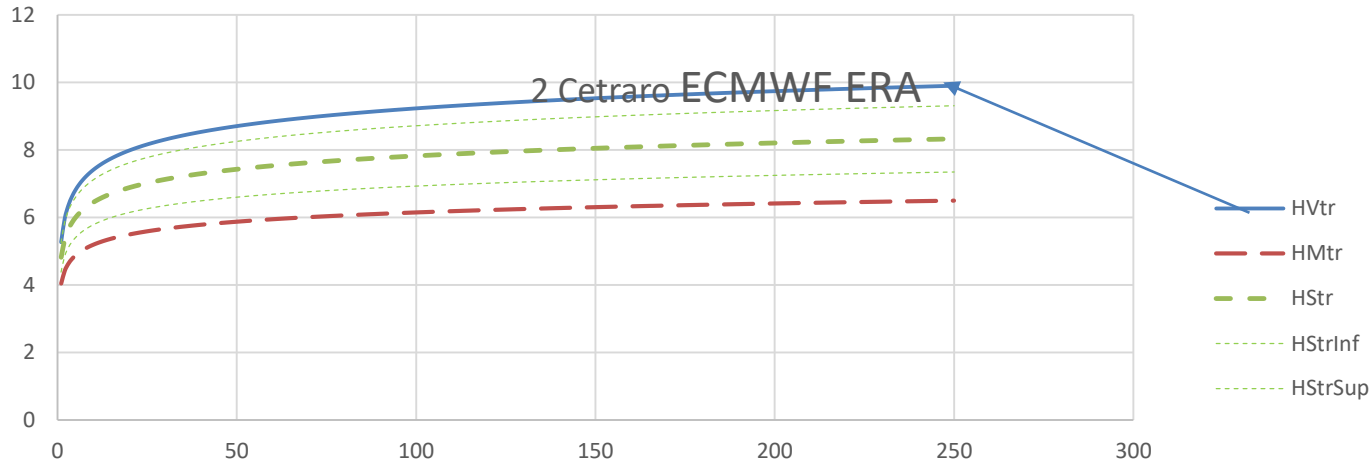
95.5% conf int.





By and large, the «true» model data within the 95% confidence curve of the «corrected» value (Model + computed bias)

With some exceptions...



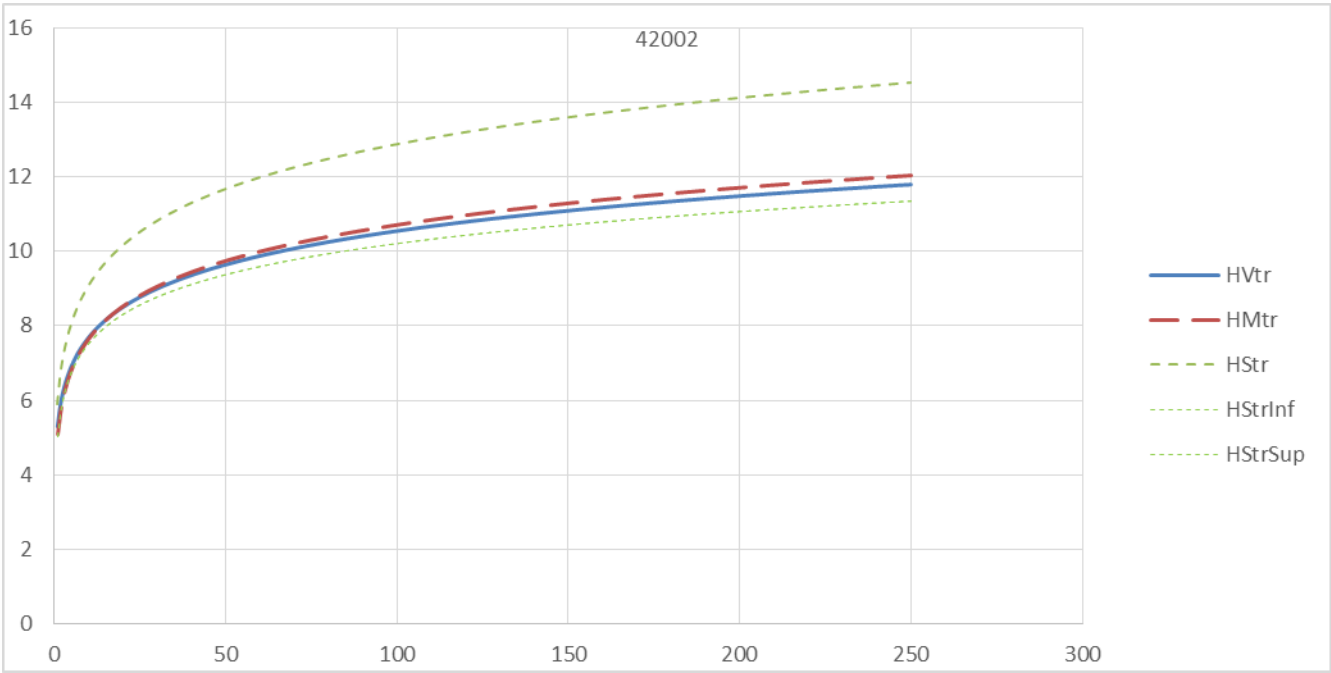
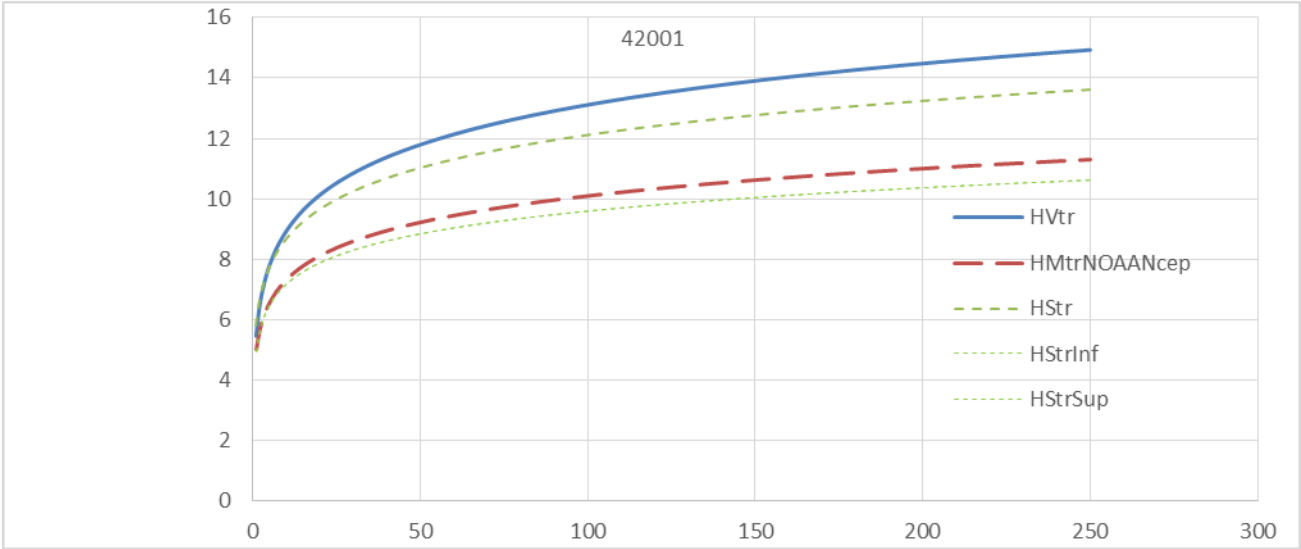
Maybe some **detail** went wrong

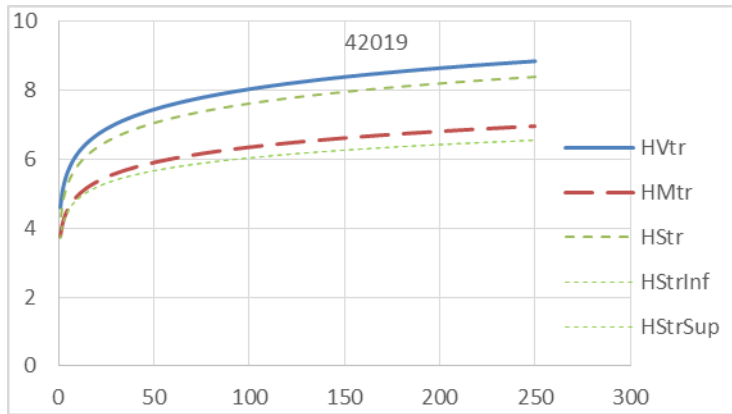
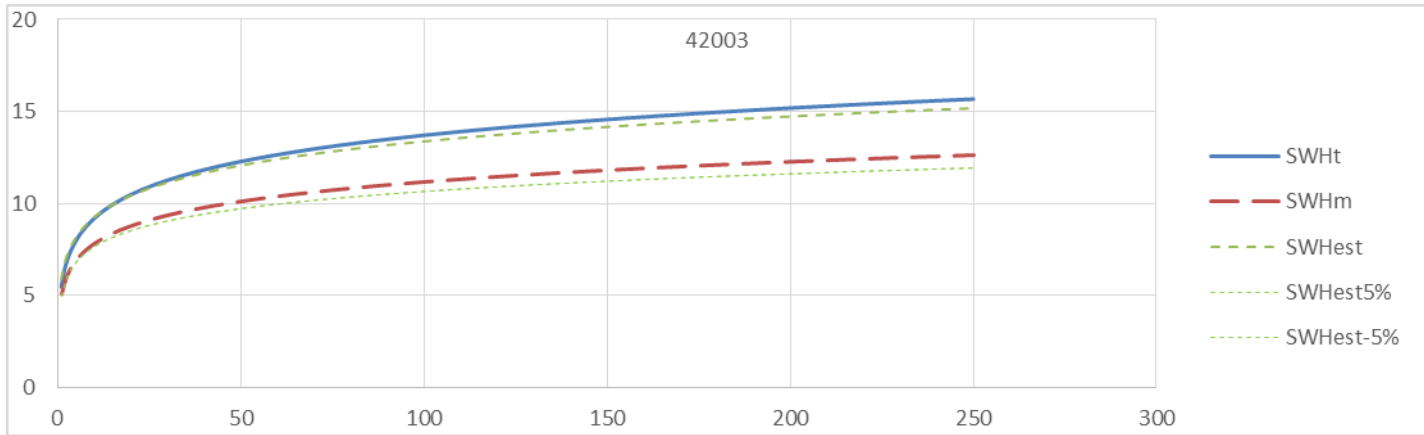


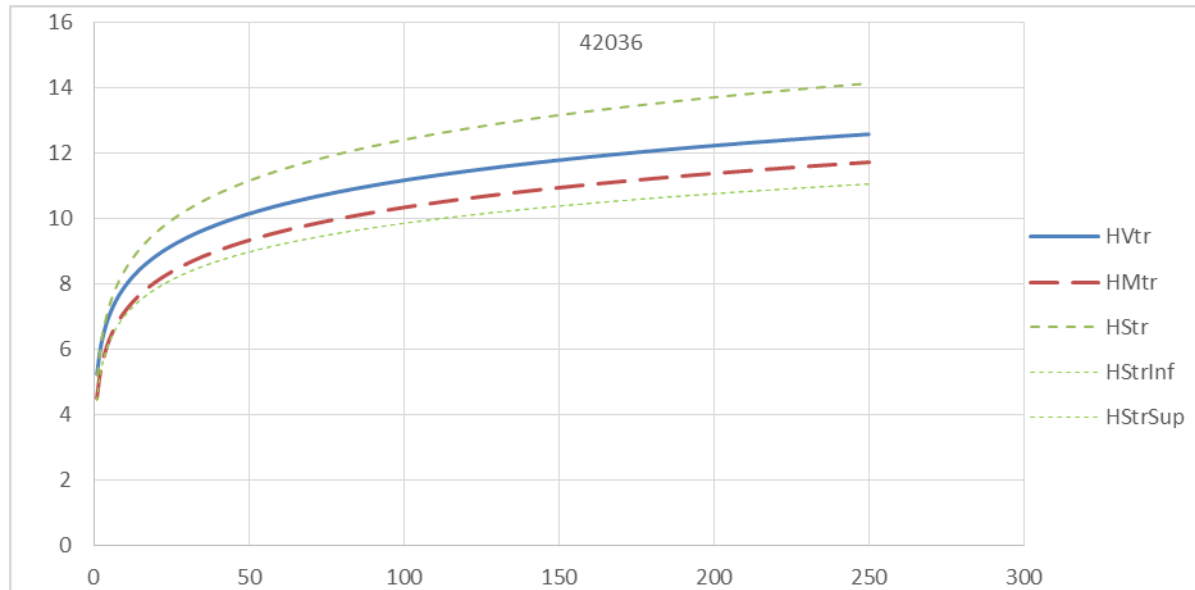
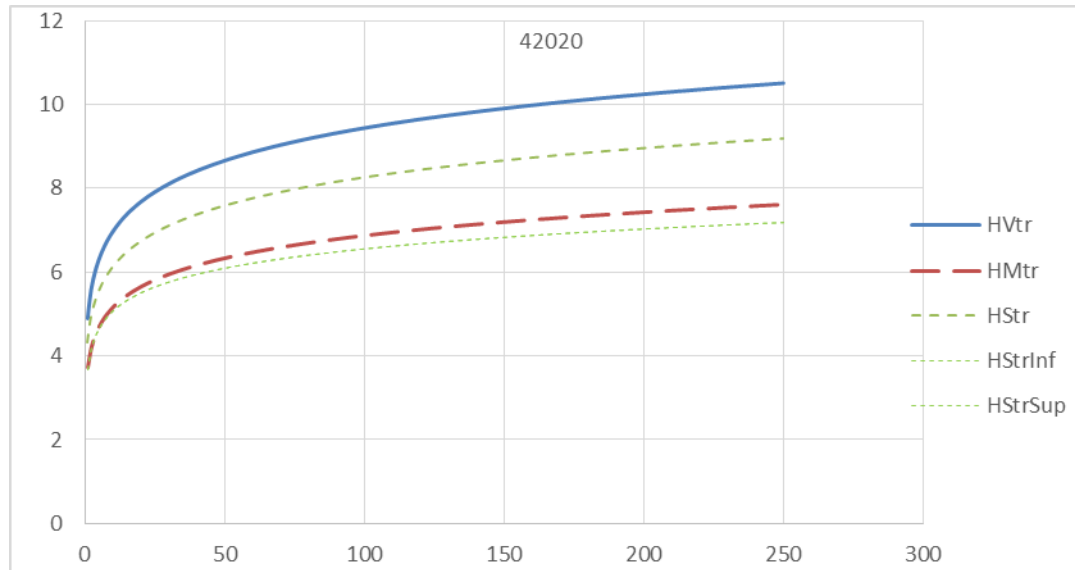


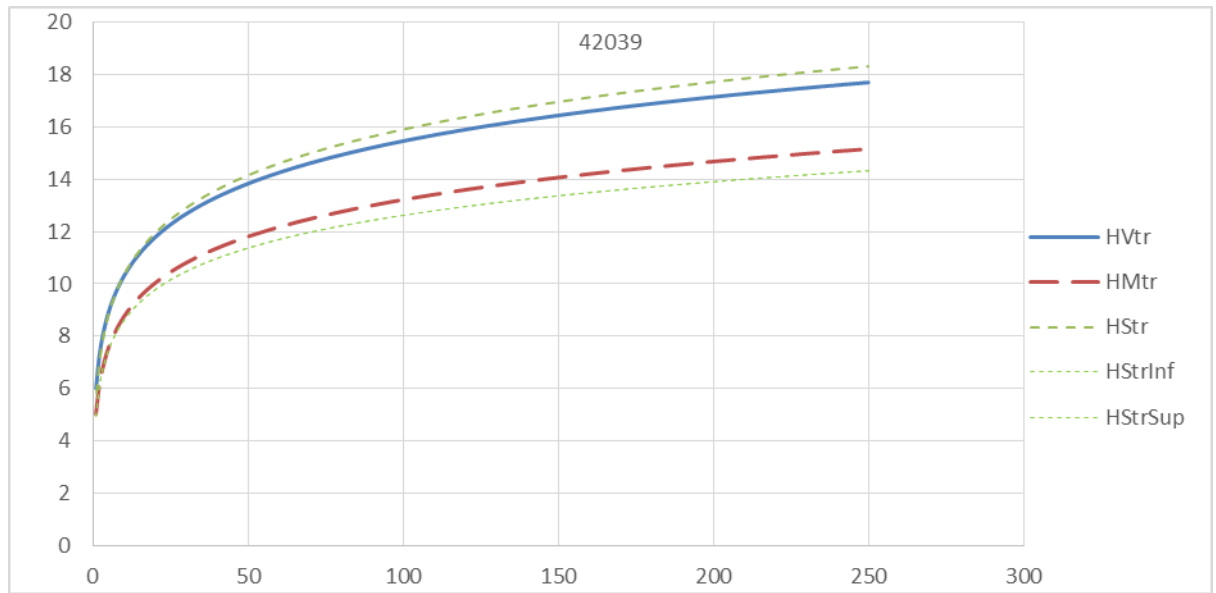
Gulf of Mexico Extreme SWH(Return Time) Model (NCEP) vs buoys data

Wavewatch III Reanalysis 10' grid 3 h (Integrated model)











Buoy	Lat.	Only hourly Data Extension
Bilbao Vizcaya	43.64° N	From 17-04-2002 to 27-11-2010
Cabo de Penhas	43.75° N	From 13-01-1998 to 08-11-2010
Cabo Silleiro	42.12° N	From 06-07-1998 to 27-11-2010
Estaca de Bares	44.12° N	From 15-01-1998 to 01-09-2010
Golfo de Cadiz	36.48° N	From 09-11-1997 to 19-12-2010
Villano Sisargas	43.50° N	From 12-05-1998 to 27-11-2010

Atlantic- Iberian waters – Puertos de l’Estado

Iberian Atlantic Waters

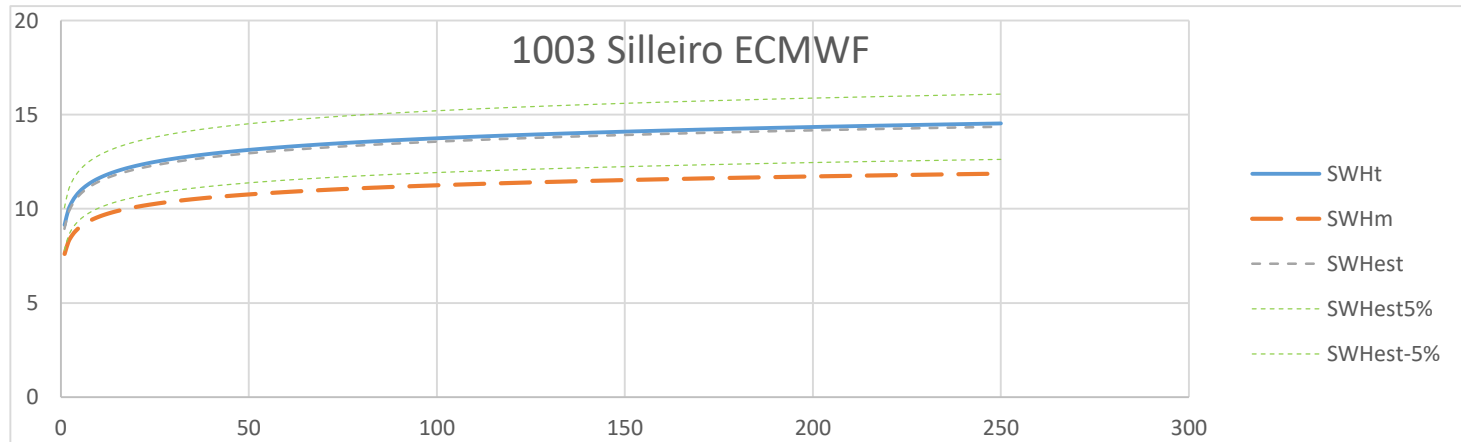
Extreme SWH(Return Time)

Model (NCEP,ECMWF) vs buoys data

Spanish buoy data are available with 1 h sampling rate

ECMWF Re-analysis ex MED 1°grid 6h (Integrated model)

95.5% conf int.



SWH(Tr) Weibull curves at buoy location: SWHt(Tr): buoy value; SWHt(Tr): model value; SWHest(Tr): estimated with model and local regional error ; SWHest5% (Tr): confidence interval estimated with model and local regional error

Iberian Atlantic Waters

Wave model NCEP, ECMWF...

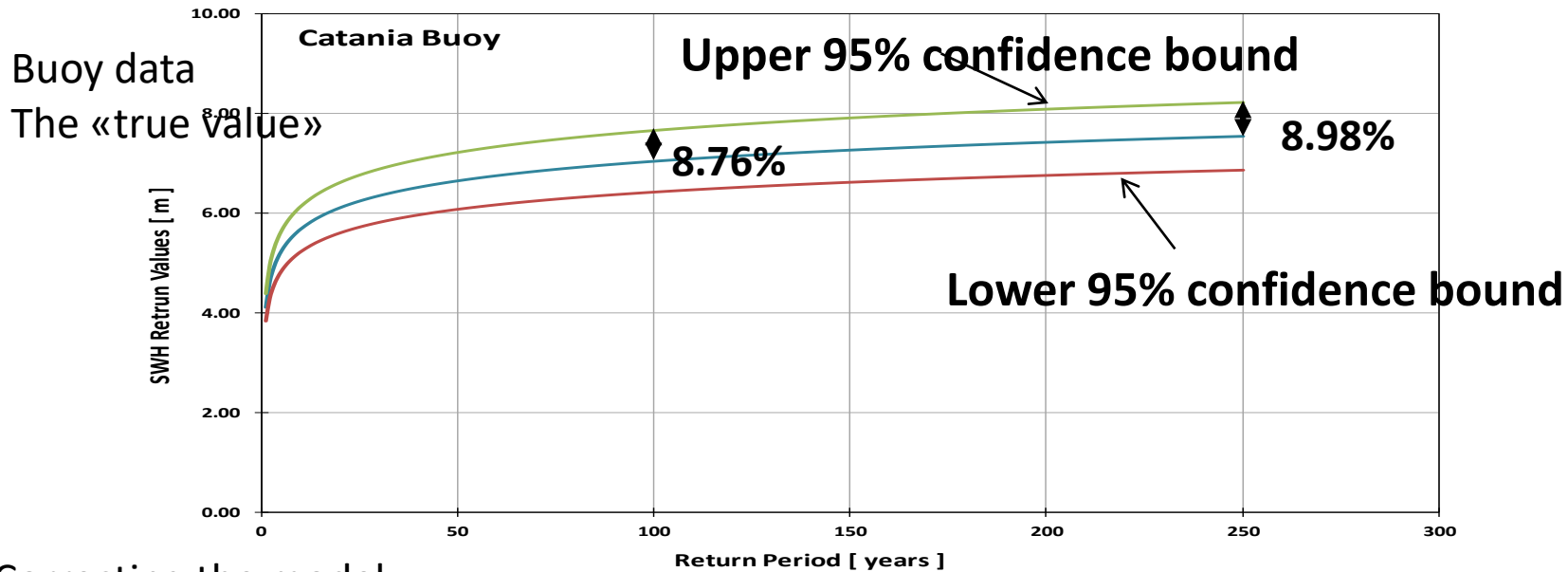
OOPS!



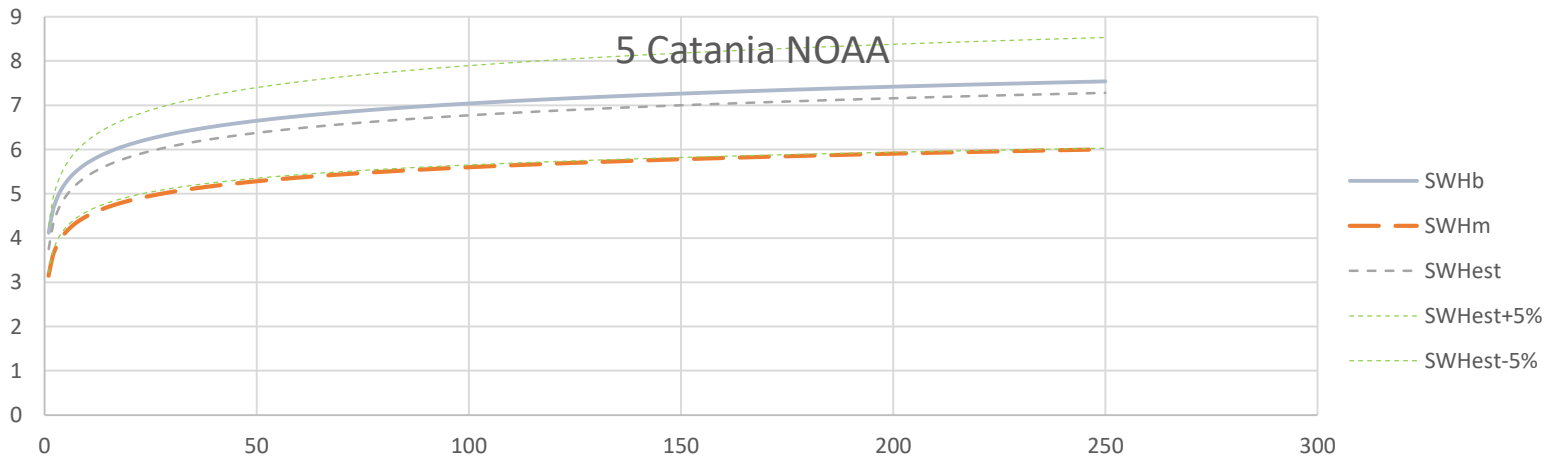
Looking at the confidence intervals – an information of the quality of our estimate of SWH (Tr) deriving from model uncertainty - : most of the «true « values fall within the 95% confidence bound. The uncertainty is not negligible

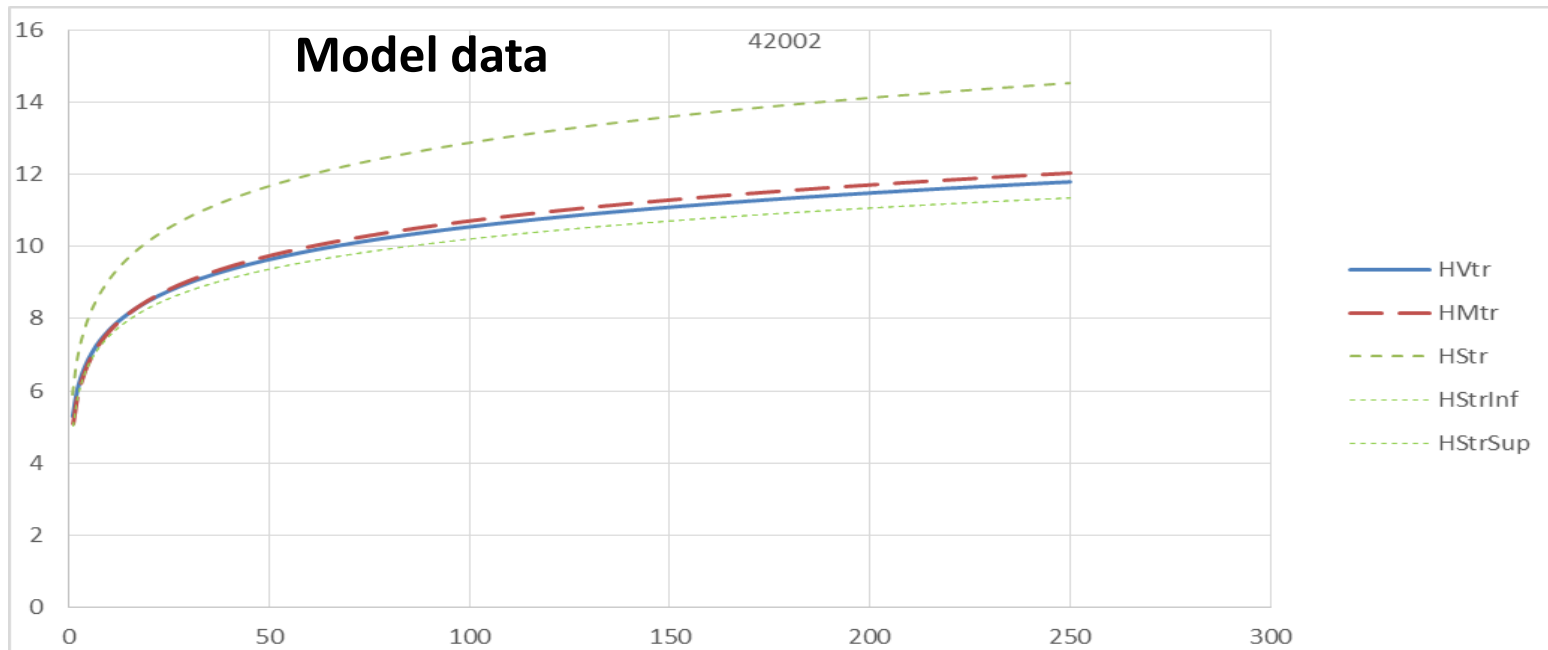
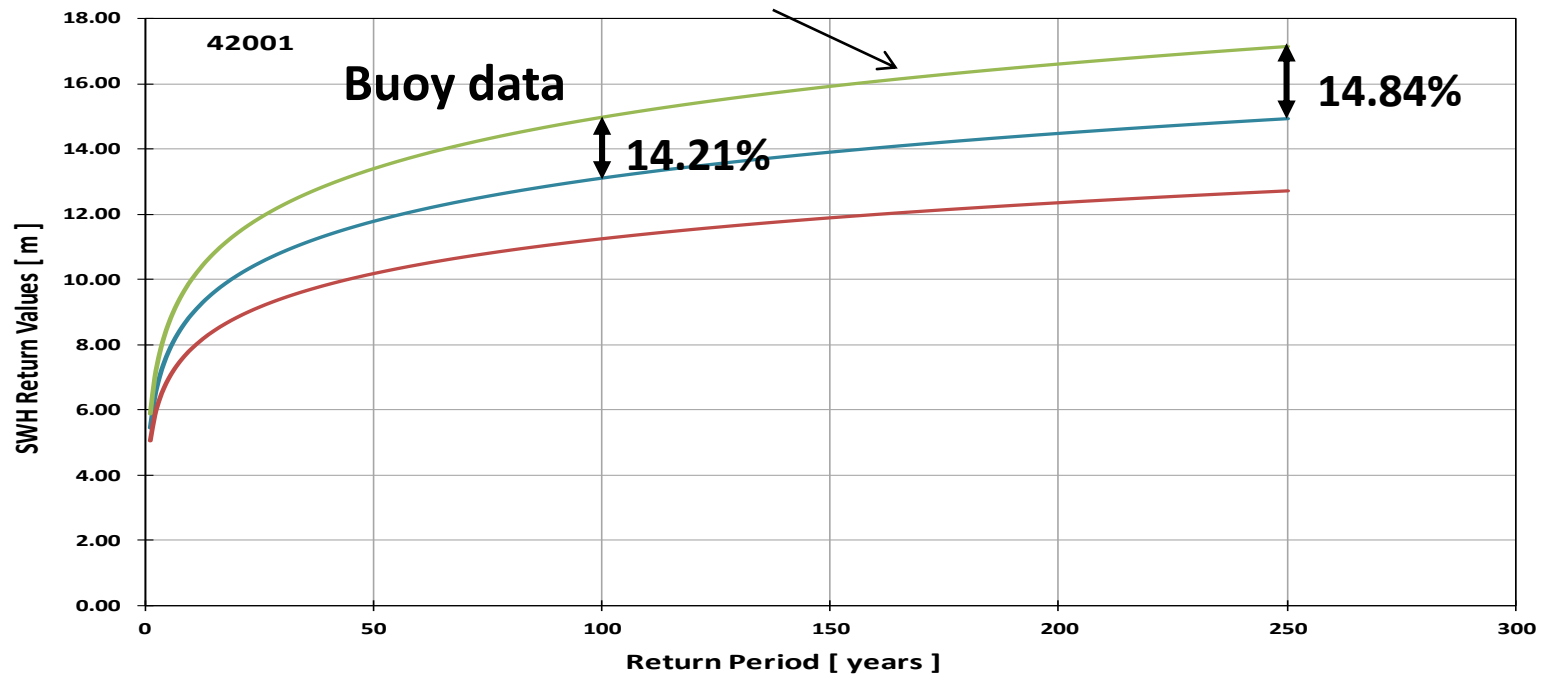
But how does it compare with the confidence intervals of a curve obtained from wavemeter data, i.e. the uncertainty deriving from fitting the curve to the data:





Correcting the model
Bias with the «regional»
error





Conclusions; and where we go from there

We need extreme SWH(Tr) curves from historical model data; and possibly, and estimate of their reliability

We propose to

- 1) Work directly SWH(Tr) curves, rather than with single SWH values:
- 2) Investigate on the statistical distribution of such curves, and to estimate the parameters of this distribution by making use of buoy data in the same area (For each return time)

We find that

SWH(Tr) curves and their confidence intervals can be found easily enough: while the uncertainty of the estimation is not negligible, it is however of the same order of magnitude of the uncertainty deriving from extreme value analysis of the buoy data

What next?

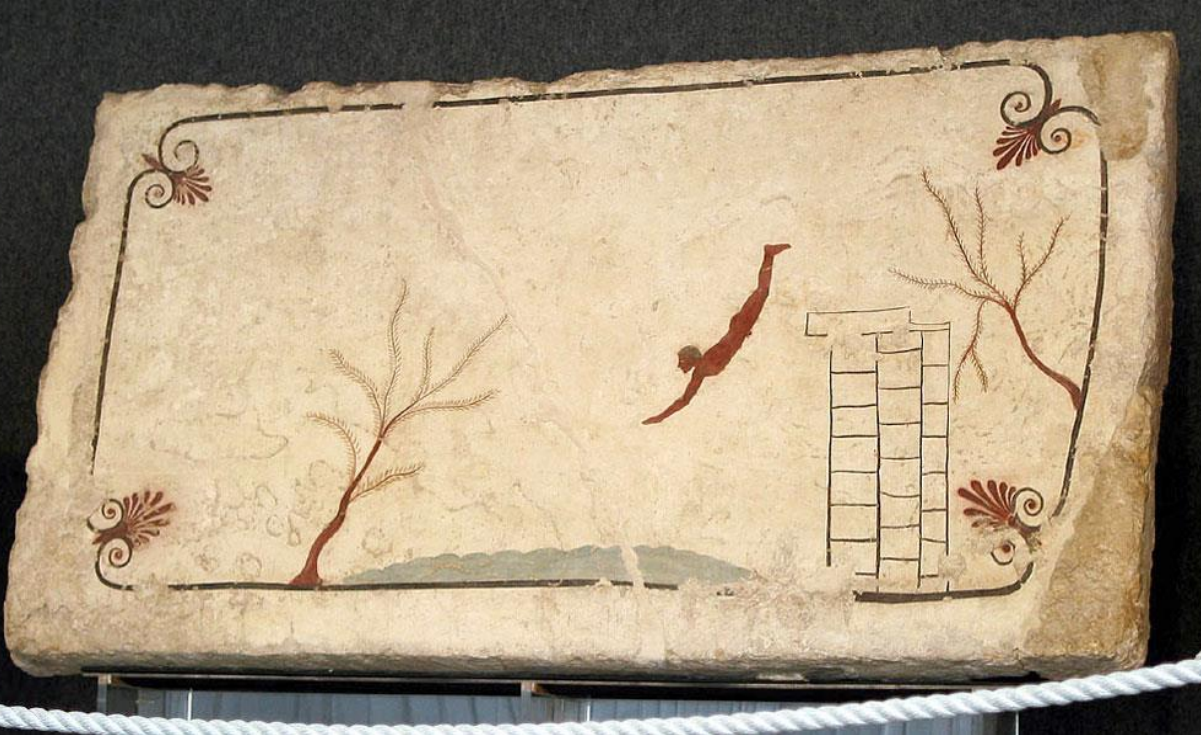
So far, we have only tested public domain model data series (ECMWF, NOAA): easily accessible but generally low resolution; tests should be carried out now with some of the many high resolution models now available.

If you calibrate,
Calibrate for the
parameter you
Need SWH(Tr)

Work funded and supported by CUGRI - University Centre for Research on Major Hazards between the Universities of Salerno and Federico II in Naples

Data provided by:

ECMWF Meteorological Archival and Retrieval System (MARS); NOAA
Altimeter: RADS (Radar Altimeter Database System Satellite) and ESA/EO Project 1172 "Remote Sensing of Wave Transformation", GlobWave
Buoy data: (Italian Environmental Agency) ISPRA; Civil Protection Service of the Campania Region; Puertos del Estado; NOAA



The diver's tumb
Paestum , Italy
(ca. 480-470 BC)

