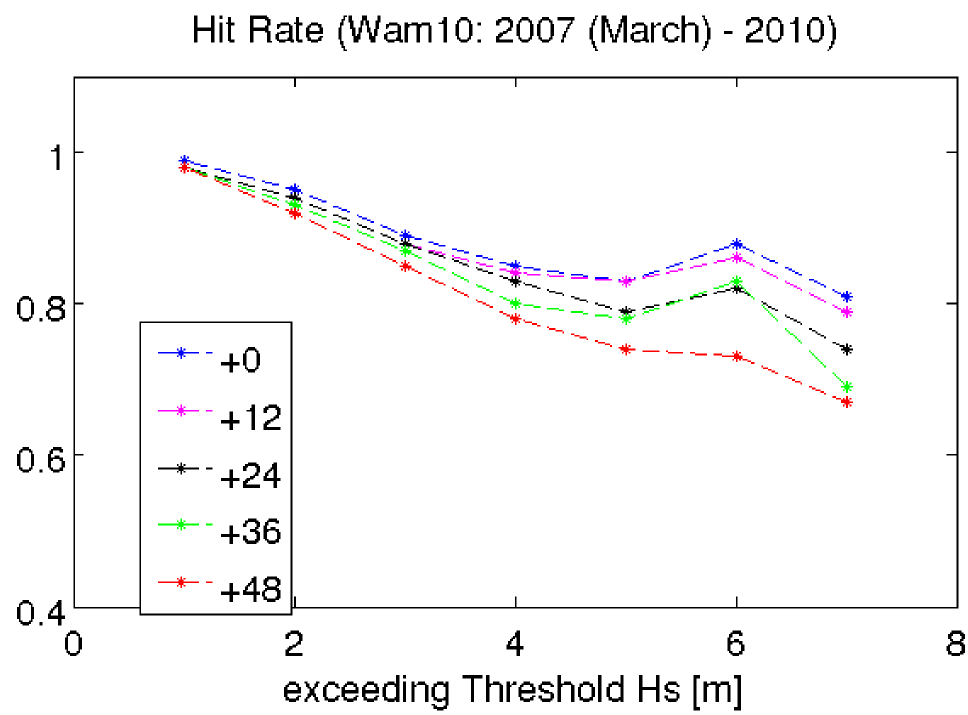
Norwegian Meteorological Institute
met.no

met.no report

Report no. 14/2011
Oceanography
ISSN: 1503-8025
Oslo, December 8, 2011

Validation of the Operational Wave Model WAM at met.no - Report 2010

Yvonne Gusdal, Ana Carrasco, Birgitte R. Furevik, Øyvind Sætra



Norwegian
Meteorological Institute
met.no

report

Number	Subject	Date	Classification	ISSN
14/2011	Oceanography	December 8, 2011	☒ Open ☐ Restricted ☐ Confidential	1503-8025

Title

Validation of the Operational Wave Model WAM at met.no - Report 2010

Authors

Yvonne Gusdal, Ana Carrasco, Birgitte R. Furevik, Øyvind Sætra

Client(s)	Client reference

Abstract

The significant wave height (Hs) from the operational wave model WAM at met.no, is validated against EnviSat Radar Altimeter (RA-2) and in-situ observations. WAM is run at 50km, 10km and 4km resolution (WAM50, WAM10 and WAM4) and is forced with 10m surface winds from the numerical weather prediction model HIRLAM to produce a 66 hour forecast. When comparing WAM10 and WAM50 for 2010, the behavior of the two models are quite similar, but WAM10 performs better than WAM50. When comparing WAM10 and WAM4, just small improvements are shown in the higher resolution model. This may be due to the fact that the available buoys are located offshore where the advantage of WAM4 can't be seen. We find that the introduction of a higher resolution model together with changes implemented in the 10m forcing over the decade 1999 - 2010, has a positive impact on the forecast of Hs. However, due to the continually upgrade of the mesh size in HIRLAM, WAM is systematically overestimating the wave height since 2003. From the Categorical Statistics we find that the forecasted Hs for the period 2007 - 2010 (WAM10) has a higher hit rate of all exceeding Hs than for the period 1999-2007 (WAM50). The false alarm ratio has also become lower for the long forecast (+36 and +48), but higher for the analysis and the +12 forecast, especially for the highest waves. We find a much higher frequency bias in the 2007 - 2010 period, meaning that the wave model is forecasting more high wave events than observed.
An artificial enhancement of the wind has been used in WAM at met.no since 1998, where the enhancement is 4% for winds between 15m/s and 25m/s. This artificial intensification of the wind were removed from the model per November 1, 2011.

Keywords

WAM, Significant wave height,Validation

Disiplinary signature	Responsible signature
ØH	Lars-Anders Breivik

Postal address
PO Box 43 Blindern
N-0313 Oslo
Norway

Office
Niels Henrik Abels vei 40

Telephone
+47 2296 3000

Telefax
+47 2296 3050

e-mail: met.inst@met.no
Web: met.no

Bank account
7695 05 00601

Swift code
DNBANOKK

1 Introduction

The aim of the validation is to estimate the forecast skill of the operational wave models at the Norwegian Meteorological Institute (met.no). The wave models run operationally at met.no are the regional wave model WAM at 50km, 10km and 4km resolution, and the nearshore wave model SWAN at 500m resolution. All models are run with winds from HIRLAM, except SWAN which is forced with UM wind. Both in-situ and EnviSat Radar Altimeter (RA-2) data are applied to validate the wave models. The model domains for WAM are shown in Fig.(1). The only wave parameter validated in this study is the significant wave height (Hs). To give a better estimate of the model skill, wave period, wave direction and the 10m wind should be studied in future work. SWAN has not been validated in this report, due to lack of observations in the SWAN model domain.
Long term statistics of WAM against Norwegian buoys is presented in section 5.1 and for the last year against ECMWF buoys in section 5.3. In section 5.2, a comparison with satellite altimeter data is included. The data and methods are presented in sections 2-4.

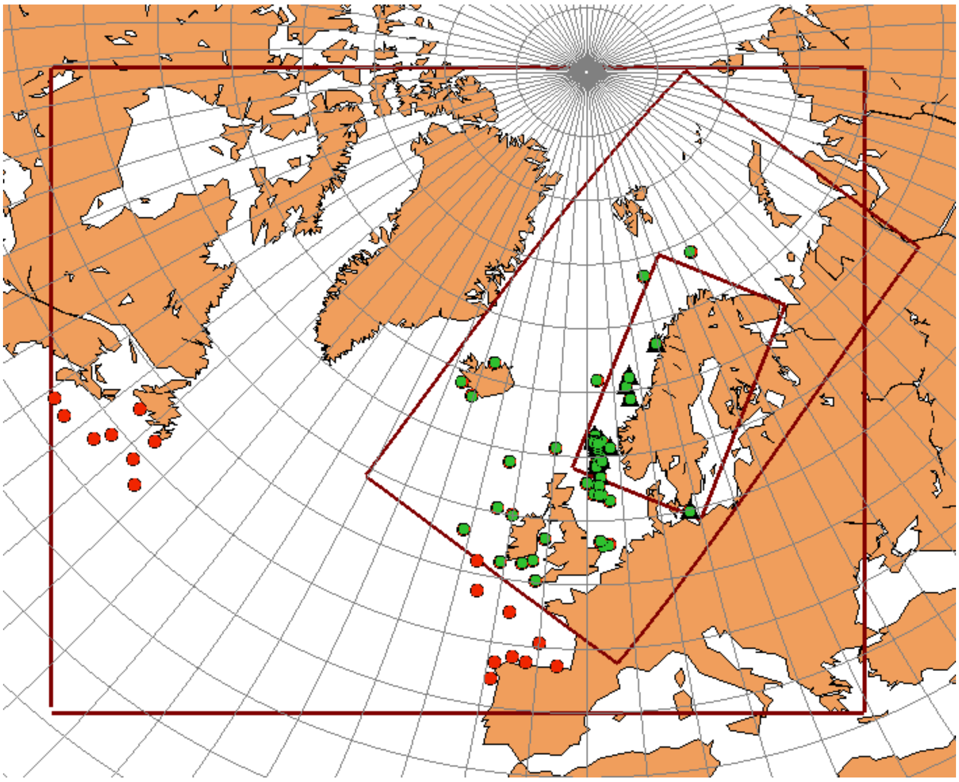


Figure 1: Buoys and domains of WAM50, WAM10 and WAM4. The large domain corresponds to WAM50 the middle to WAM10 and the smallest to WAM4.

2 Model

2.1 WAM

The operational wave prediction model at met.no is the third generation spectral wave model, WAM, initially developed by an international group of scientists [WAMDI Group (1988); Sætra et al. (2004); Komen et al. (1994)]. At met.no, WAM50 is run four times a day at 50km resolution, with wind from HIRLAM12[1] as input data. Additionally, a WAM model with 10 km and 4 km resolution (WAM10 and WAM4) is run twice a day, forced with wind data from HIRLAM8[2] and HIRLAM4[3] respectively. WAM10 is nested into the 50km model while WAM4 is nested into WAM10. The higher resolution model WAM4 primarily covers the Norwegian coastal waters as shown in Fig.(1). The forecast period for each model is 66 hours. Wave measurements from ERS-2[4] and EnviSat[4] satellites are used to correct the initial state of the WAM model. The WAM model computes two-dimensional wave spectra. From the two-dimensional spectra, several parameters are computed, e.g. significant wave height, peak wave period, mean wave period, peak wave direction and mean wave direction. The wave parameters are computed for total sea, and for wind sea and swell.

3 Data

3.1 EnviSat RA-2

The EnviSat RA-2 instrument operates on both Ku- and S-band. Former work [Abdalla (2005)] shows that the Ku-band Hs is of higher quality than the S-band Hs. Therefore, in this study we only apply the Ku-band Hs. Before collocating the observations and the model results, the altimeter data is quality controlled. Close to the coast and the ice edge some bad quality data occurs. These observations are removed from the data set. It is important to note that from experience with EnviSat and buoy observations, it is determined that the EnviSat wave height is slightly overestimated by 3-4 %, [Abdalla (2005)]. Further, to perform a proper validation, the scale of the observations must match the scale of the model. For our purpose, the resolution of the EnviSat RA-2 measurements (8km) are much higher than the model resolution of WAM50. Also the model resolution of WAM10 has a slightly higher mesh size than the observed wave height. An along track averaging of the observations is therefore performed. Before the altimeter data are averaged, the data are collocated against the model results. Due to the high resolution of the altimeter data, the model result in a grid-box may be collocated against more than one observation. This group of observations are then averaged. The maximum time span between model and observation is set to +/- 30 min. The coverage of the collocated altimeter is displayed in Fig.(2). The blue contours represent the coverage when

[1]HIRLAM12 = Atmospheric model with 12km horizontal resolution,[Unden (2002)]
[2]HIRLAM8 = Atmospheric model with 8km horizontal resolution,[Unden (2002)]
[3]HIRLAM4 = Atmospheric model with 4km horizontal resolution,[Unden (2002)]
[4]http://www.esa.int/esaEO/SEMGWH2VQUD_index_0_m.html

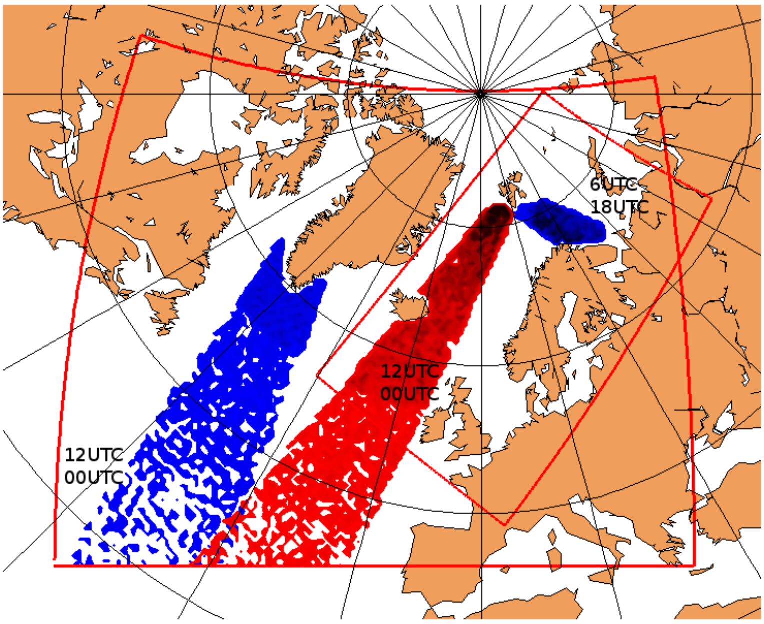


Figure 2: Displayed is the density of the collocated EnviSat RA-2 observations. Red contours shows the coverage when the satellite is descending, while the blue contours shows the coverage when the satellite is ascending. The hours are the approximate time for the given paths, with a time span of +/- 30 min. The outer red area is the domain of WAM50, while the inner red area is the domain of WAM10.

the satellite is ascending (from south to north), while the red contours represent the coverage when the satellite is descending (from north to south). The figure shows a repeating pattern for the different satellite paths, which is due to the cyclic pattern of the satellite. The EnviSat RA-2 is continuously providing measurements around the whole orbit with a 35 day repeating cycle. At the same time, WAM is producing a 66 hour forecast four times a day (6UTC, 12UTC, 18UTC and 00UTC), where only results from the 12UTC and 00UTC runs are stored, and therefore validated in this study. Additionally, we have only validated the model results every 6 hour. The fixed model hours together with the cyclic observation pattern gives the limited observation coverage in Fig.(2). EnviSat RA-2 data are used to validate WAM50 and WAM10. For a resonable comparison between the two models, only observations covering the WAM10 domain is applied (the inner red area shown in Fig.(2)).

Models	WAM50	WAM10	WAM4
total obs	30890	20190	6530
total buoys	50	35	24
Models to compare	WAM50	WAM10	
obs	18300	18300	
buoys	33	33	
Models to compare		WAM10	WAM4
obs		6200	6200
buoys		11	11

Table 1: Numbers of buoys and observations used to validate the models. Observations refers to the number of observations at each forecast time. Also presented is the number of buoys and observations used when comparing two models.

3.2 Buoys and Wave Radar observations

3.2.1 Observations from ECMWF

The applied buoy observations in Chapter 5.3 are shown in Fig.(1). They have been processed and quality controlled at ECMWF. Since buoys exhibit high-frequency variability not captured by the model results, the observations are averaged in a window of 4 hours centered around the verification time, see Bidlot et al. (2002). The resulting time series have a 4 hour time interval. Not averaging the data can result in a scatter between the models and observations [Janssen et al. (1997)]. For a more detailed description of the data treatment, see Bidlot et al. (2002) and Sætra et al. (2004).

A summary of the data used can be seen in Table (1). The observations used to validate WAM50 come from 50 buoys with approximately 30890 observations at analysis time, while the observations used to validate WAM10 come from 35 buoys with approximately 20190 observations. For WAM4, there are 24 buoys with 6530 observations. For comparison between models, only common observations are used.

3.2.2 Observations from met.no

The six sites in the Norwegian and North sea used to validate WAM in Chapter 5.1 are shown in Fig.(3). These in-situ observations are quality controlled at met.no but have not been averaged in a window of 4 hours. They have been averaged over each hour, and the resulting time series have a 1 hour time interval. This is the same method used in previous work on validating WAM, see Gusdal (2010). Since the results in Chapter 5.1, are extended time series from former study, we apply the same method.

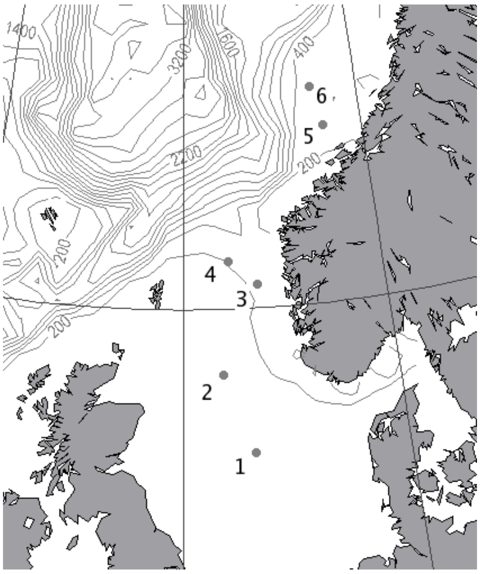


Figure 3: Displayed is the observation sites located in the Norwegian and the North Sea applied in Chapter 5.1. The sites are 1: Ekofisk, 2: Sleipner, 3: Troll A, 4: Gullfaks C, 5: Draugen and 6: Heidrun.

4 Methods

4.1 Statistics

The skill is measured using standard statistics. The Mean Square Error (*MS Error*) and bias, is defined as

$$MS\,Error_j = \frac{1}{n}\sum_{i=1}^{n}(H_i^{mod} - H_i^{obs})^2 \tag{1}$$

$$bias_j = \frac{1}{n}\sum_{i=1}^{n}(H_i^{mod} - H_i^{obs}) \tag{2}$$

where the subscript j denote the day number in a month, i represent the observation number and H_i^{mod} and H_i^{obs} is the modeled and observed wave height respectively. The monthly Root Mean Square Error (rmse) and bias are then defined as

$$rmse = \sqrt{\frac{1}{N_T}\sum_{j=1}^{N_d} MS\,Error_j \cdot N_j} \tag{3}$$

$$bias = \frac{1}{N_T}\sum_{j=1}^{N_d} bias_j \cdot N_j \tag{4}$$

$$N_T = \sum_{j=1}^{N_d} N_j \tag{5}$$

where N_j is the number of existing observations for day j and N_T is the number of observations in a month.

4.2 Categorical Statistics

Table 2: Contigency table, showing the frequency of ”yes” and ”no” forecasts and occurrences.

		Observed		
		yes	no	Total
Forecast	yes	hits	false alarm	forecast yes
	no	misses	correct negatives	forecast no
	Total	observed yes	observed no	total

hits - event forecast to occur, and did occur

misses - event forecast not to occur, but did occur

false alarm - event forecast to occur, but did not occur

correct negative - event forecast not to occur, and did not occur.

Categorical statistics are computed from the contigency table to describe particular aspects of the forecast performance. For example, the forecast skill for wave heights exceeding 7m. A large variety of categorical statistics can be computed from the table, in this study the following have been computed:

Hit Rate - measures the fraction of the observed yes events that were correctly forecasted!

$$\frac{hits}{hits + misses} \tag{6}$$

False alarm ratio - measures the fraction of the predicted yes events that did not occur

$$\frac{falsealarm}{hits + falsealarms} \tag{7}$$

Frequency bias - measures the ratio of the frequency of forecast events to the frequency of the observed events.

$$\frac{hits + falsealarm}{hits + misses} \tag{8}$$

5 Results

5.1 WAM - 1999 to 2010

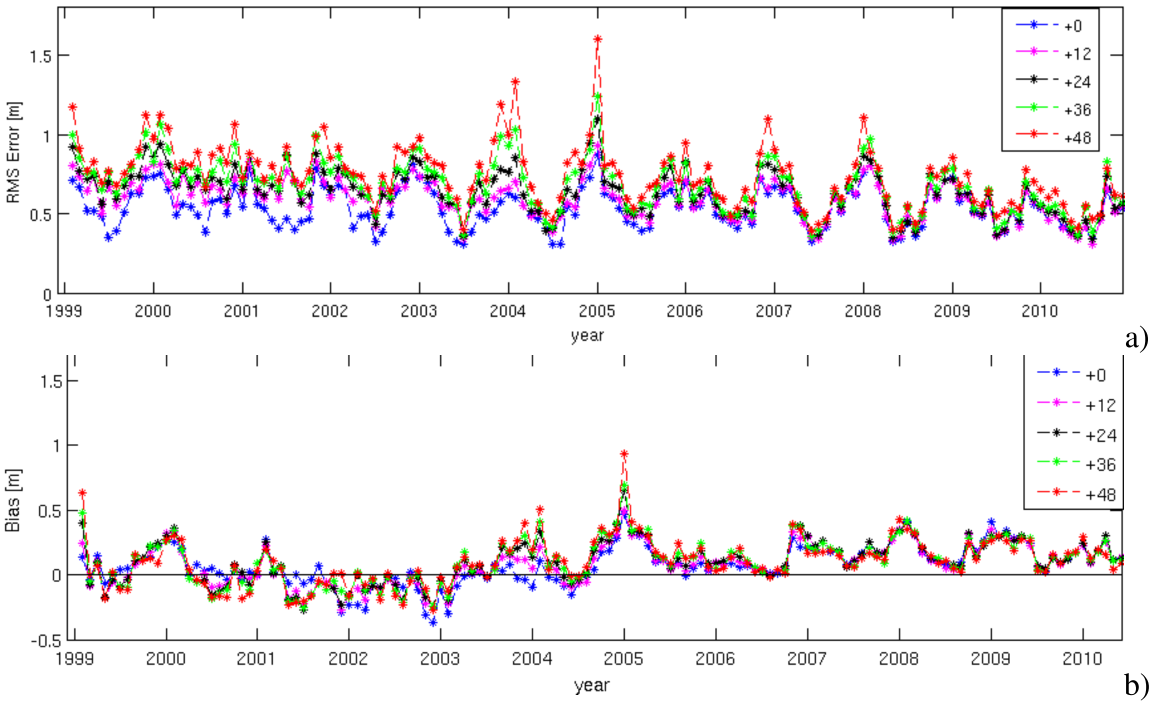


Figure 4: Time series of the rmse and bias for the forecast of Hs for the period 1999 to 2010. Note that model results from WAM50 are included for the period before March 2007, while model results from WAM10 are included for the later period. Observations from six sites in the Norwegian and the North Sea is applied. Ekofisk, Sleipner, Troll A, Gullfaks C, Draugen and Heidrun

The rmse and bias (model minus observations) for different lead times are displayed in Fig.(4), covering the period February 1999 through 2010. Before March 2007 model results from WAM50 is applied while for the later period WAM10 results are used in the comparison [Gusdal (2010)]. The results reveal no decreasing trend in the rmse for the analysis hour. However, the forecast is improved, illustrated by the decreasing deviation between the rmse for the analysis and the different lead times. In 2003 as displayed in Fig.(4b), we find a shift in the bias, as WAM starts to simulate higher waves than observed. In 2003, the resolution of HIRLAM was increased to 20km instead of 50km. The physics in WAM is not tuned due to the different changes implemented in HIRLAM, and may be the reason for the systematically overestimation of Hs.

Results from the categorical statistics are displayed in Fig.(5), where the left column shows results for the period 1999 - 2007 for WAM50 and the right column shows results for the period 2007 through 2010 for WAM10. It is not possible to compare the two models, since they represent different periods of the decade 1999 - 2010, where different improvements have been implemented in HIRLAM with an apparent high effect on the forecasted Hs. For the latest period 2007 - 2010, WAM has a higher hit rate for all exceeding wave heights compared to the period 1999 - 2007. The false alarm ratio has become lower for the long forecasts (+36 and +48) but higher for the analysis and +12 forecast. In the latest period (2007 - 2010) WAM gives a higher frequency bias for the highest waves, meaning that the number of forcasted events of high waves, are larger than the number of observed events.

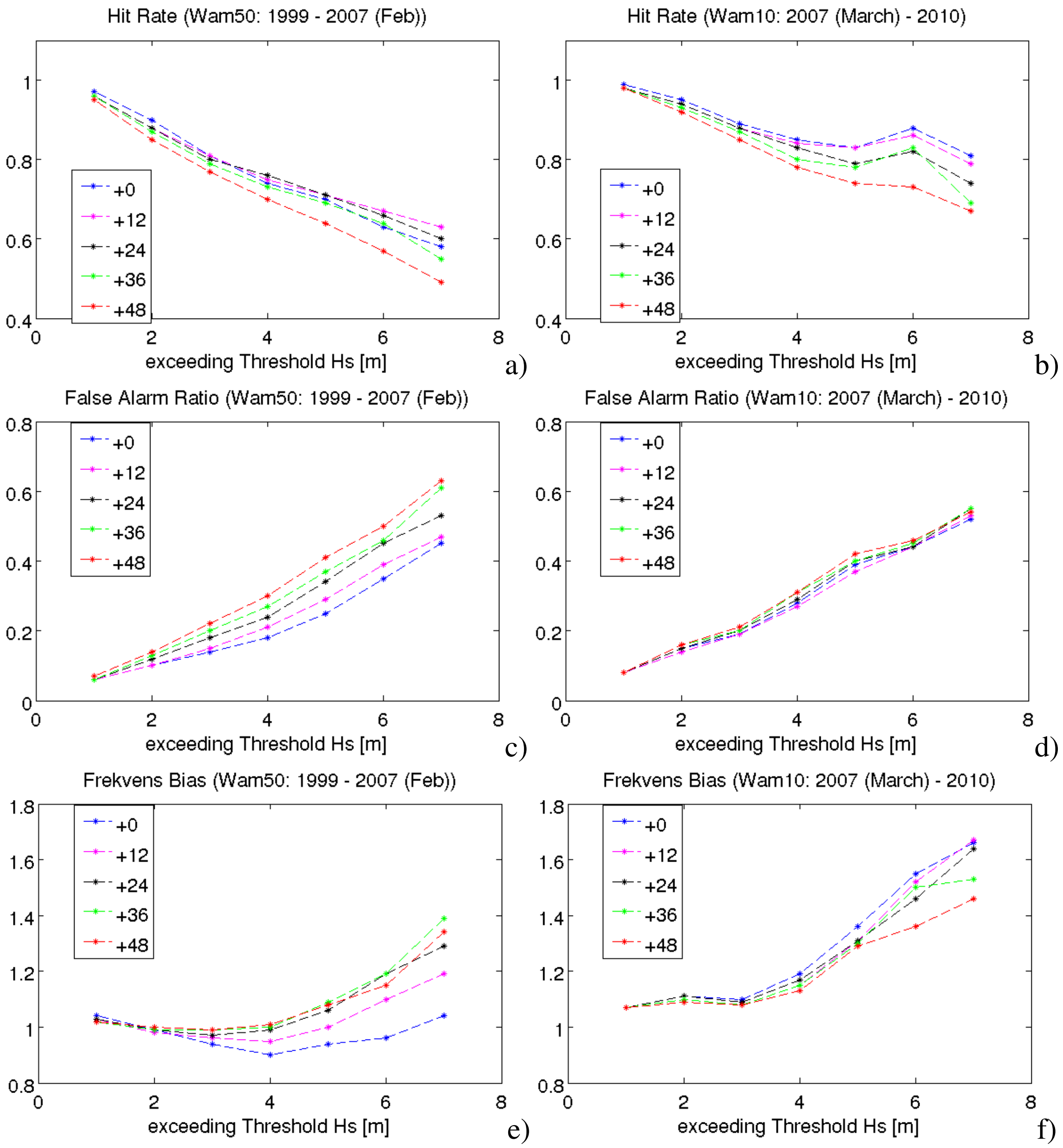


Figure 5: Shows the forecast skill of wave heights exceeding a threshold Hs. The statistics computed are Hit rate, False alarm ratio and Frequency bias. The left column shows results for the period 1999 to 2007 (February) for WAM50, while the right column show results for the period 2007 (March) through 2010 for WAM10.

5.2 EnviSat-RA2

5.2.1 WAM10 vs WAM50 2010

Due to the limited observation coverage from EnviSat RA-2, the altimeter data is only applied to validate WAM50 and WAM10. To get a reasonable comparison between the two models, observations found solely inside the WAM10 domain is applied (the inner red area shown in Fig.(2)). The collocated Hs of altimeter data and model analysis for both WAM10 and WAM50 are plotted in the scatter diagram shown in Fig.(6). As the scatter plot shows, the agreement between the observed and the modeled Hs is very good for both models with a correlation of 0.96 at the analysis time. The high correlation between the models and the altimeter data, is also seen for the other lead times.
Fig.(7) shows a quantile-quantile plot (Q-Q plot) between the models and the collocated EnviSat data for the year 2010. A Q-Q plot is a graphical method for comparing two probability distributions by plotting their quantiles against each other. If the two distributions are similar, the points in the Q-Q plot will approximately lie on the 1:1 line. As the Q-Q plot shows, the agreement between the observed and the modeled Hs is very good for the small wave heights. However, at the tail of the distribution when the wave height exceed 6m, both models gives higher waves than observed. Throughout the forecast period, except at the analysis time, WAM10 has a higher overestimation of the higher waves than WAM50. This can only be observed for the descending path (red area Fig.(2)). For lead time +6 (Fig.(7) b and e), the EnviSat-RA2 is only covering the Barents Sea and there is not an overestimation of Hs in the Q-Q plot. It is worth noting that when we validate the analysis (12UTC and 00UTC), the observations covering the Barents Sea is not included.

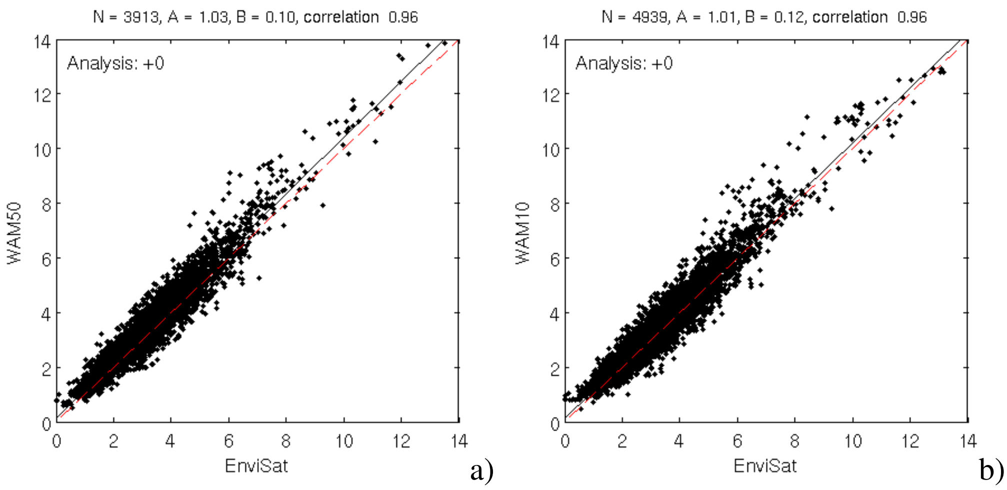


Figure 6: Shown are scatter plots between observed and modeled wave height for WAM10 and WAM50. The plots includes model results and observations from EnviSat for the year 2010. The black line is the linear regression while the red dashed line represents the perfect fit between the two data sets.

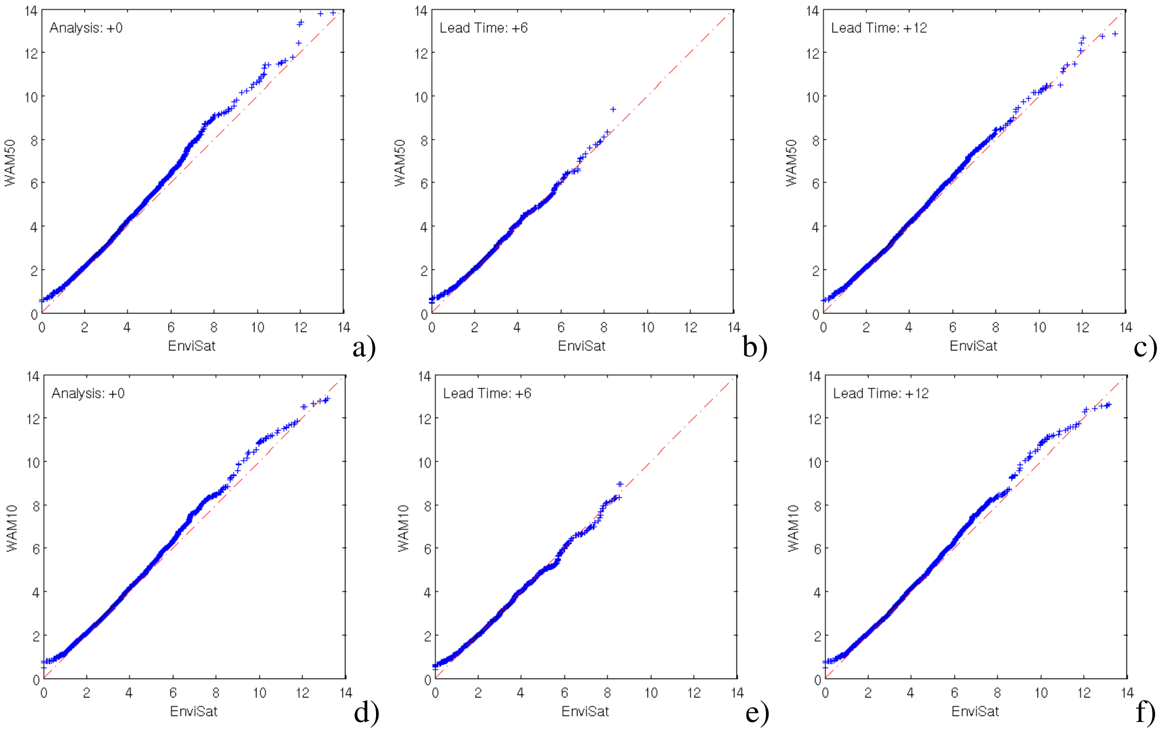


Figure 7: Shown are Q-Q plots between observed and modeled wave height for WAM10 and WAM50. The plots includes model results and observations from EnviSat for the year 2010. The black line is the linear regression while the red dashed line represents the perfect fit between the two data sets.

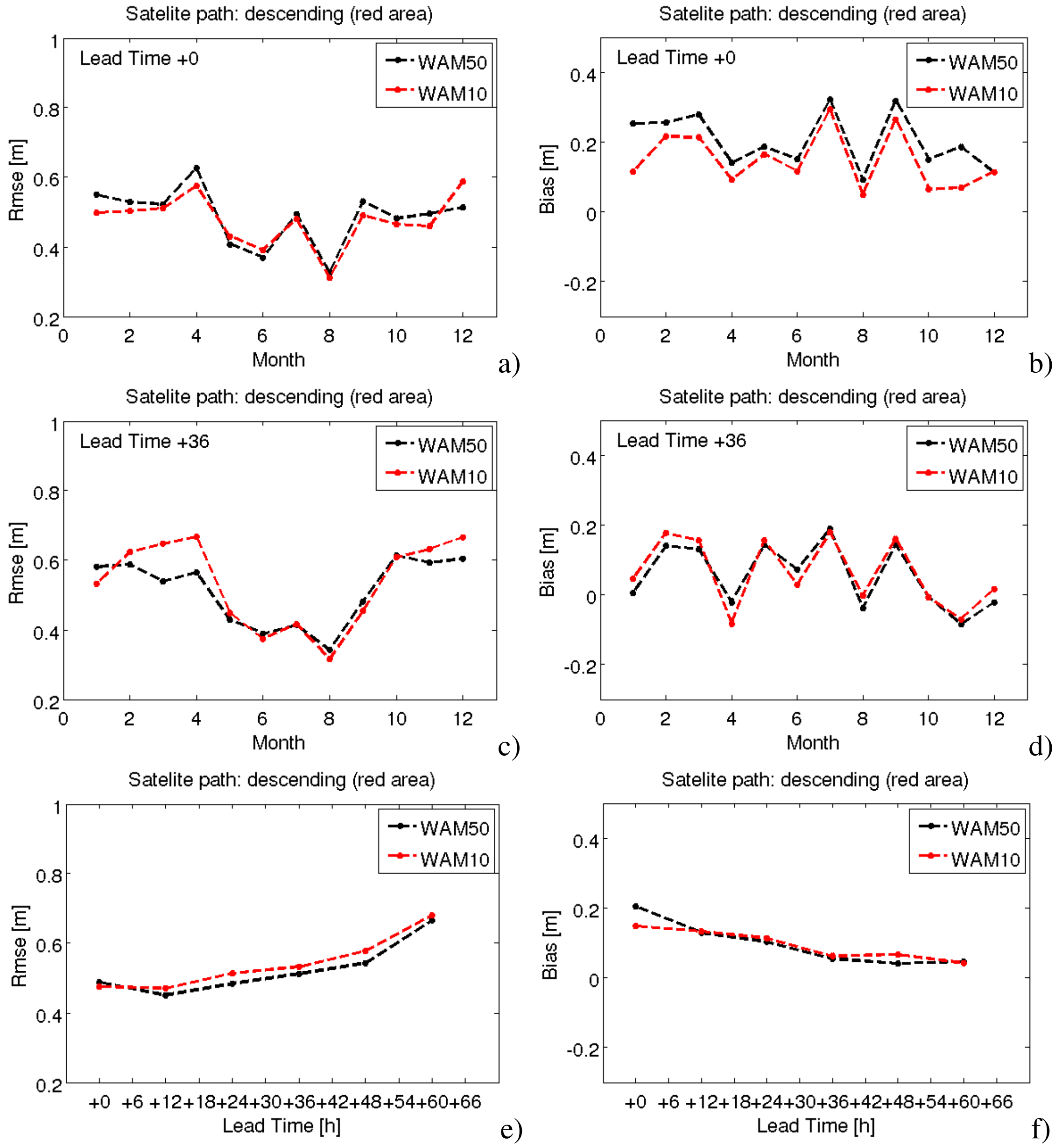


Figure 8: Displayed is a comparison of the wave height from WAM50 (black line) and WAM10 (red line) for the descending path. Fig.(a and b): shows the rmse and bias for the analysis, Fig.(c and d): shows the rmse and bias for the 36 hour forecast Fig.(e and f): shows the mean rmse and bias (model minus observation) for each lead time for both models.

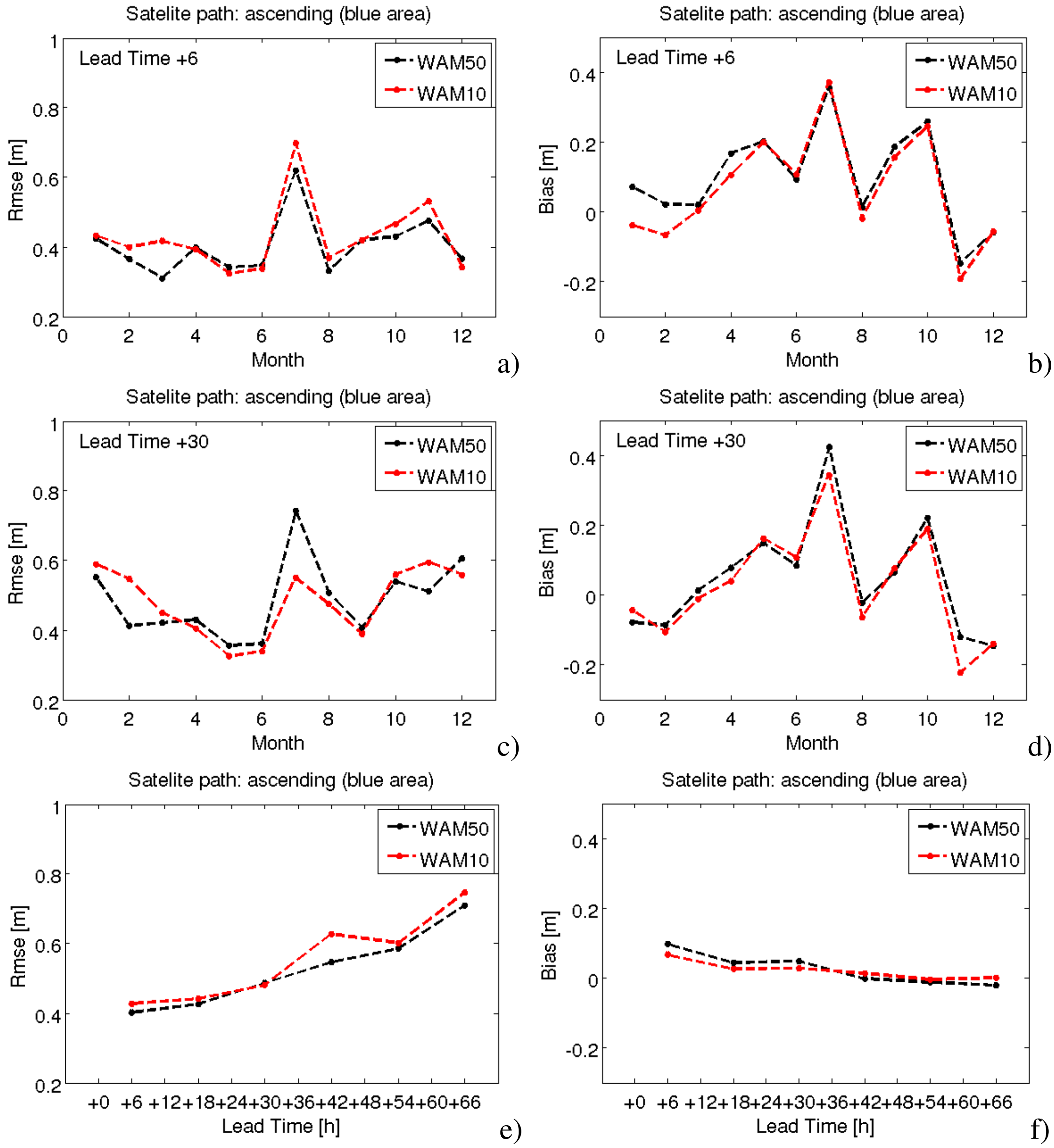


Figure 9: Displayed is a comparison of the wave height from WAM50 (black line) and WAM10 (red line) for the ascending path. Fig.(a and b): shows the rmse and bias for the 6 hour forecast, Fig.(c and d): shows the rmse and bias for the 30 hour forecast Fig.(e and f): shows the mean rmse and bias (model minus observation) for each lead time for both models.

Shown in Fig.(8) and Fig.(9) is a monthly comparison between WAM50 and WAM10 for the year 2010. The plots show rmse and bias between the modeled Hs and the collocated EnviSat Hs, where bias is model results minus observations. When comparing the two models for the analysis, only the observations from the descending path (red area in Fig.(2)) is applied. The statistical comparison shows how the models achieve approximately the same bias and rmse for the analysis, however WAM10 has a slightly better agreement against the observations than WAM50. We can see how the season affect the results with the lowest rmse achieved around June, where wave heights are low. For the 6 hour lead time, as shown in Fig.(9a) and Fig.(9b), only the observations covering the Barents Sea is included in the comparison. For this area WAM50 has a better agreement against the observations than WAM10 when looking at the rmse results, except in the summer period June to September. This may be due to the sea ice which occur during the winter months in the Barents Sea. The sea ice may lead to poor satellite observations. Due to the along track averaging of the observations, WAM50 is validated against observation averaged over more data than WAM10. If bad quality data occur, this will have a larger effect in the WAM10 results. In Fig.(8e and f) and Fig.(9e and f), a comparison for each lead time for the two different paths are shown. For the descending path, WAM10 has a lower rmse and bias for the analysis. However for the 30 hour lead time and onward WAM10 has a slightly higher rmse than WAM50, but WAM10 will probably describe the wave field in a better manner nearshore than WAM50, since the location of the EnviSat data is offshore, the nearshore Hs is not validated with the altimeter data.

5.3 Buoys in 2010

5.3.1 Each model

From the scatter and Q-Q plots at analysis time shown in Fig.(10) during 2010, we can see that the three models: WAM50, WAM10 and WAM4, behave quite well, for small wave heights, with a slightly tendency towards overestimation as the wave heights increase. The overestimation in WAM50 and WAM10 starts at around $Hs = 3.5m$ while for the higher resolution model WAM4 it starts at around $Hs = 4m$. From the scatter plots we can see that WAM50 and WAM10 present a very hight correlation coefficient, 0.95 and 0.96, with a large amount of data, 30892 and 20194 co-locations respectively. For WAM4 the correlation coefficient was 0.89 but with only 6534 data points.

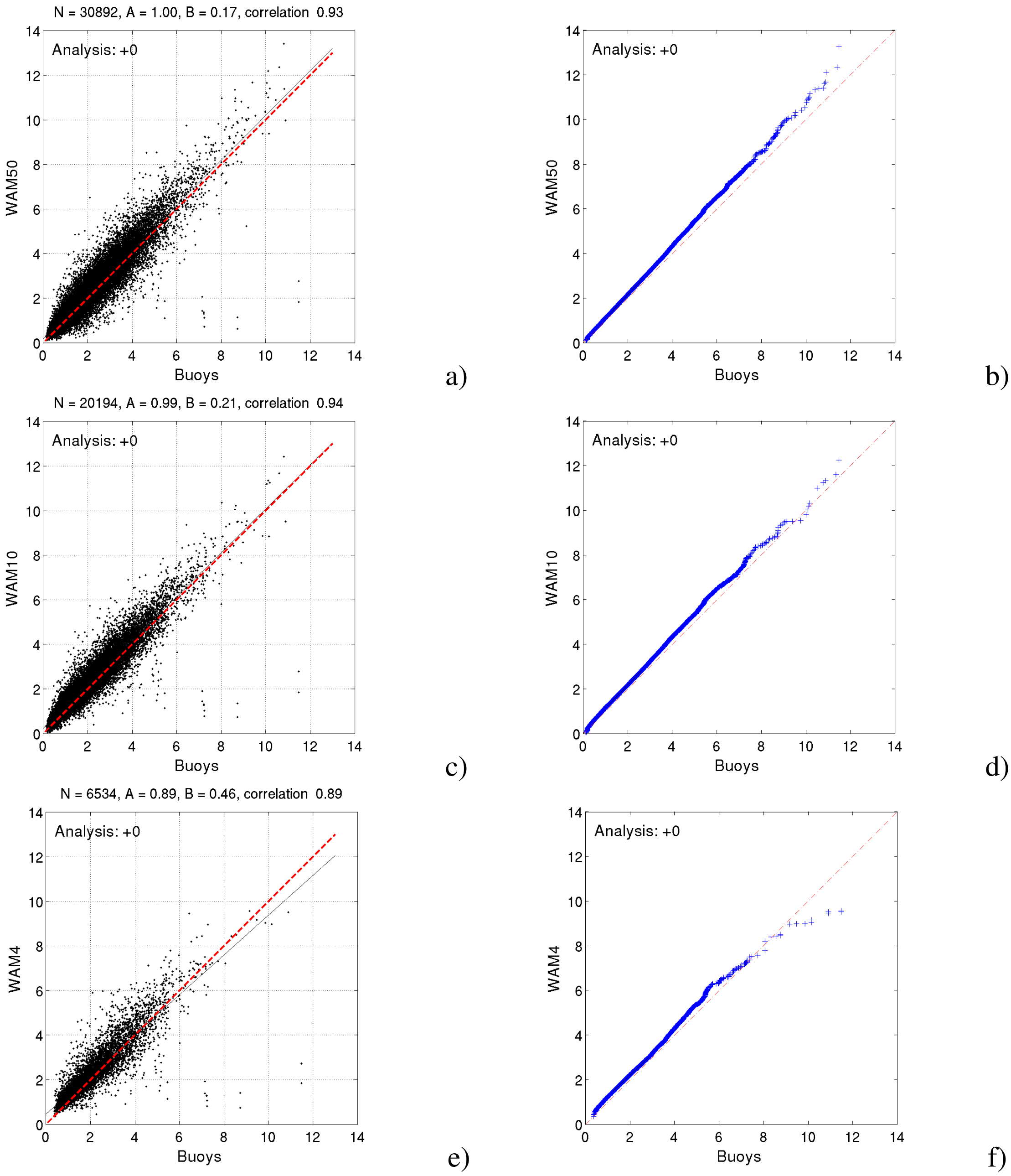


Figure 10: Scatter and quantile-quantile between observed and modeled significant wave height for WAM50 a) and b), for WAM10 c) and d) and for WAM4 e) and f). The black line is the linear regression while the red dashed line represents the perfect fit between the two data sets.

5.3.2 WAM50 vs WAM10

In order to compare the statistical performance of WAM50 and WAM10, only buoys covered by both model domains are included. This reduce drastically the number of data, see Table (1). For the comparison between WAM50 and WAM10, only data from 33 buoys, with 18300 observations at each forecast hour are used. The bias and the rmse are plotted in Fig.(11). WAM10 performs better than WAM50 with lower bias and rmse. The behavior of these two models is quite similar. At analysis time we can see that models perform better in the summer, specially in June and August, than in the winter. The bias is always positive and decreases with forecast time while the rmse increases with forecast time.

5.3.3 WAM10 vs WAM4

The stations inside the smallest domain in Fig.(1) with 11 buoys, and aprox. 6200 observations, are used to compare WAM10 with WAM4. The bias and the rmse are plotted in Fig.(12). The striking feature, as in 2009, is that the two models perform equally well, and it does not seem to be an apparent advantage of WAM4 over WAM10. This might be related to the fact

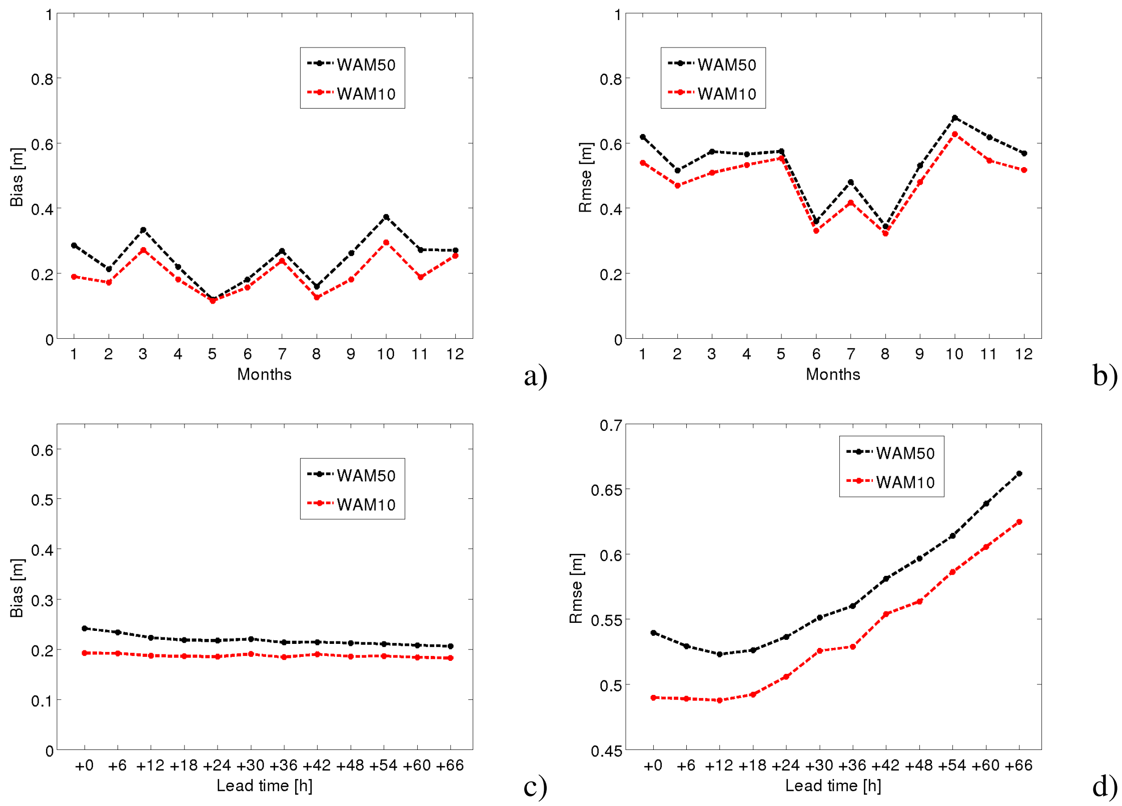


Figure 11: Comparison of the significant wave height from WAM50 and WAM10 during 2010. Fig.(a and b) show the bias and the rmse variation during the year for the analysis. Fig.(c and d) show the bias and rmse with Lead time.

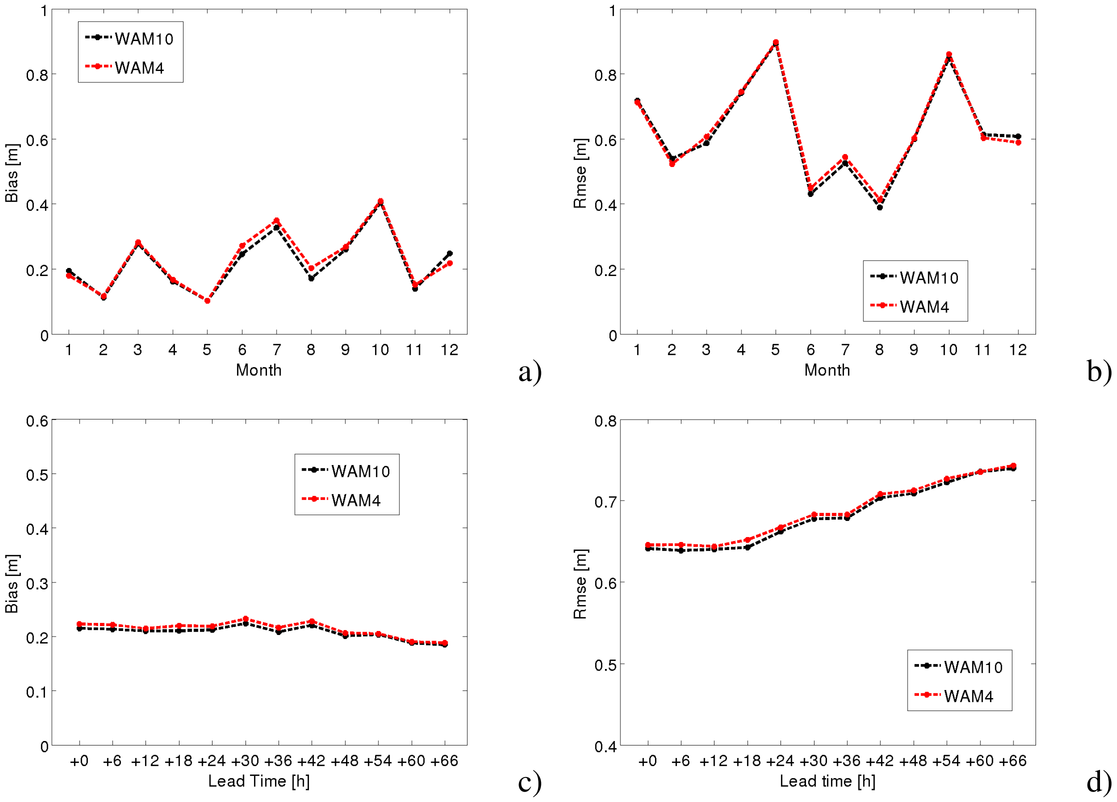


Figure 12: Comparison of the significant wave height from WAM10 and WAM4 during 2010. Fig.(a and b) show the bias and the rmse variation during the year for the analysis. Fig.(c and d) show the bias and rmse with Lead time.

that the buoys are located offshore, where the presumed advantage of running WAM4 is not noticeable. The variation of the bias and rmse in real time and in lead time, Fig.(12), have the same features as in Fig.(11).

5.4 Summary and Conclusion

We find that WAM10 performs better than WAM50 when comparing model results with in-situ observations, and only small improvement when applying Satellite Altimeter data. This may be due to the fact that the buoys are located near the coast while the satellite coverage is not. The finer resolution model WAM10 may describe the wave field in a better way nearshore than the coarse model. Since the Altimeter data is not covering the WAM4 domain, along the Norwegian coast, the high resolution model is only validated against in-situ observations. However, we only find small improvements in the higher resolution model WAM4 compared to WAM10. This may be due to the fact that the available in-situ observations are located offshore where the advantage of WAM4 can't be seen.
We find that the introduction of the higher resolution model WAM10 together with improve-

ments implemented in HIRLAM over the period 1999 - 2010, have had a positive impact on the forecast of Hs. However, due to the continually upgrade of the mesh size in HIRLAM, WAM has systematically overestimated the wave height since 2003. Higher resolution weather prediction models may give stronger winds due to finer description of the pressure field. If the model is not tuned due to the higher resolution wind fields, it may lead to higher simulated wave heights by the wave models.
As a consequence of the validation presented in this report, an artificial enhancement of the wind used in WAM at met.no, which has been systematically applied since around 1998 were removed per November 1, 2011. The enhancement was 4% for winds between 15m/s and 25m/s.

References

Abdalla, S., Global Validation of EnviSat Wind, Wave and Water Vapour Products from RA-2, MWR, ASAR and MERIS, 2005, ESA report, Available at: http://www.ecmwf.int/publications/library/do/references/list/18

Bidlot, J. R., D. J. Holmes, P. A. Wittmann, R. L. Lalbeharry, & H. S. Chen, Intercomparison of the performance of operational ocean wave forecasting systems with buoy data, 2002, Wea. Forecasting, 17, 287-310.

Booij, N., R. Ris, & L. Holthuijsen, A third-generation wave model for coastal regions 1. Model description and validation, 1999, Journal of geophysical research 104(C4), 7649-7666.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. Whilte, & N. Wood, A new dynamical core for the Met Offices global and regional modelling of the atmosphere, 2005, Q. J. R. Meteorol. Soc. 131. 1759-1782.

Gusdal, Y., Validation of the Operational Wave Model WAM - February 1999 through June 2009, 2010, met.no Rep. 3, Norwegian Meteorological Institute. 12 pp.

Janssen, P., B. Hansen, & J.-R. Bidlot, Verification of the ECMWF forecasting system against buoy and altimeter data, 1997, Wea. Forecasting, 12, 763-784.

Komen, G., J.L. Cavaleri, M. Donelan, K.Hasselmann, S.Hasselmann, & P.A.E.M. Janssen, Dynamics and Modelling of Ocean Waves, 1994, Cambridge University Press, 533 pp.

Sætra, O., & J.R. Bidlot, Potential Benefits of Using Probabilistic Forecast for Waves and Marine Winds Based on the ECMWF Ensemble Prediction System, 2004,American Meteorological Society, 673-689.

Unden, P., Hirlam-5 Scientific Documentation, (ed.) 2002, 1-144, Available at: http://www.hirlam.org

WAMDI Group: S. Hasselmann, K. Hasselmann, E. Bauer, P.A.E.M. Janssen, G.J. Komen, L. Bertotti, P. Lionello, A. Guillaume, V.C. Cardone, J.A. Greenwood, M. Reistad, L. Zambresky & J.A. Ewing, The WAM model — a third generation ocean wave prediction model, 1988, J. Phys. Oceanogr., 18, 1775–1810.