

## 8.1 VALIDATION OF SATELLITE-BASED ESTIMATES OF WHITECAP COVERAGE: APPROACHES AND INITIAL RESULTS

Magdalena D. Anguelova<sup>\*</sup>, Justin P. Bobak, William E. Asher<sup>1</sup>, David J. Dowgiallo, Ben I. Moat<sup>2</sup>, Robin W. Pascal<sup>2</sup>, Margaret J. Yelland<sup>2</sup>

Remote Sensing Division, Naval Research Laboratory, 20375, Washington DC, USA  
Telephone: 202-404-6342; E-mail: [maggie.anguelova@nrl.navy.mil](mailto:maggie.anguelova@nrl.navy.mil)

<sup>1</sup>Applied Physics Laboratory; University of Washington; Seattle, WA, USA

<sup>2</sup>National Oceanography Centre, Southampton, UK

### 1. INTRODUCTION

Sea salt aerosols (SSAs) are actively produced within the oceanic whitecaps by the evaporation of the sea spray droplets formed when whitecap bubbles burst. Large discrepancies still exist in the estimates of the emissions and concentrations of SSAs. As the SSAs establish the baseline of a background marine atmosphere, uncertainty in predicting their emission introduces large uncertainty in the aerosol and atmospheric chemistry modules of both global circulation (GC) and chemical transport (CT) models. While the parameterizations for SSA emissions currently used in GC and CT models produce acceptable results, the need to update them with better constrained parameterizations is continuously increasing as the performance of the GC and CT models improve.

SSA emission is modeled with a sea salt source function (SSSF), which predicts how many sea spray droplets are produced in a given size interval depending on the wind speed. The modeling uncertainty of the SSA emissions stems from the fact that the current parameterizations do not capture well (1) the full range of particle sizes and (2) the high temporal and spatial variability of the flux magnitude. Updated parameterizations, promoted recently from the oceanographic community, address these issues. Some of these updates had found their way in the GC and CT models. For example, some models (e.g., GEOS-CHEM<sup>1</sup>) have updated the size distribution dependence to reflect closer improved constraints established by recent laboratory and field data. Yet others (e.g., MODIS aerosol algorithm, Remer et al., 2005) continue to model the SSAs over a limited size range.

The wide held notion among air-sea interaction oceanographers that a suite of meteorological and oceanographic factors, in addition to wind speed, influence the whitecap coverage of the ocean, and thus the magnitude of the sea spray and SSA production, has only recently been addressed. As a result virtually all current GC and CT models continue to use the

conventional parameterization involving the wind speed only.

A dedicated project within the framework of the WindSat mission at the Naval Research Laboratory (NRL) addresses the high variability of whitecap coverage ( $W$ ) and the sea salt source function. A major task of this project is retrieving  $W$  on a global scale from satellite-based measurements. This will allow us to build a database representing  $W$  under a wide range of meteorological and environmental conditions. Such an extended database for whitecap coverage will allow, in turn, improvement of existing or development of new parameterizations for  $W$ , which better account for the effects of factors additional to wind speed.

After proving the feasibility (Anguelova and Webster, 2006), NRL now has the capability to obtain  $W$  globally using WindSat data and an improved version of the algorithm. The study presented here focuses on the effort to validate the satellite-based estimates of whitecap coverage obtained with this improved algorithm.

### 2. FOAM FRACTION FROM SATELLITES

#### 2.1 Algorithm changes

The algorithm for estimating  $W$  from satellite measurements relies on changes of ocean surface emissivity at microwave frequencies (6 to 37 GHz) due to presence of sea foam on a rough sea surface. These changes in the emissivity due to foam are registered by space-based radiometers as changes in the brightness temperature ( $T_B$ ) of the ocean.

Anguelova et al. (2006) correct the major shortcomings of the feasibility-study algorithm discussed by Anguelova and Webster (2006). These include use of independent sources for the input variables, use of physically based models for the emissivity of rough sea surface ( $e_r$ ) and emissivity of foam ( $e_f$ ), improved rain flag, and improved atmospheric model necessary for the atmospheric correction.

The physically-based algorithm, described in the work of Anguelova et al. (2006), has been further improved recently by continuing work on the foam emissivity model (Anguelova, 2008) and the calibration

<sup>1</sup> [www-as.harvard.edu/chemistry/trop/geos/geos\\_versions.html](http://www-as.harvard.edu/chemistry/trop/geos/geos_versions.html)

of the WindSat brightness temperature (Bettenhausen, 2007).

In addition, minimization of errors due to modeling, calibration, and matching of various datasets is sought by running the algorithm with two different surface emissivity models. The foam fraction values are first obtained using the WindSat measured  $T_{BS}$  and  $T_{BS}$  from rough-surface-only model, similarly to the way described by Anguelova et al. (2006). Then the errors of the atmospheric model, the calibration, and the match-ups are evaluated together (their separation is not possible) by running the algorithm with all elements the same but the modeled  $T_{BS}$ , which this time are obtained with a model for a composite surface involving both rough and foamy areas (Bettenhausen et al., 2006). Differencing the results of the two runs minimizes, though does not fully remove, the error in obtaining the foam fraction values. The remaining errors are those of the various emissivity models, namely the rough-surface-only model ( $e_r$ ), the composite surface model ( $e = e_r + e_f$ ), and the foam emissivity model ( $e_f$ ). Work on characterizing these errors is on-going.

The foam fraction estimates used in this study are obtained with this latest version (v1.9.6) of the foam-retrieval algorithm from Windsat measurements. Details for this improved algorithm will be given in forthcoming paper (Anguelova et al., in preparation<sup>1</sup>).

## 2.2 Specifics of the satellite values

There are 80 pixels within the WindSat swath (width about 350 km) with an approximate spacing of 12.5 km across the swath and along the spacecraft track (Bettenhausen et al., 2006). Due to the averaging and sampling procedures of the WindSat data processing at the lowest level, each WindSat pixel represents a value averaged over an area of 50 km  $\times$  71 km. WindSat retrievals at higher resolutions (i.e., pixel value averaged over an area of 35 km  $\times$  53 km or 25 km  $\times$  35 km) are also available, but the work here uses whitecap coverage retrievals at the lowest resolution. Figure 1 present daily global map of foam fraction from WindSat data for 1 March, 2007 (orbits 21479 to 21492).

The comparison in Figure 2 of satellite-based whitecap coverage as a function of wind speed at three frequencies (10, 18, and 37 GHz), horizontal polarization (H-pol.), illustrates a major feature of the satellite-based estimates of foam, namely there is frequency dependence of the  $W(U_{10})$  relationship.

In the initial algorithm, we justified the use of one frequency (19 GHz) and one polarization (H-pol.) of the SSM/I data as more appropriate for obtaining whitecap coverage (Anguelova and Webster, 2006, section 3.3.1). By the same token,  $W$  estimates from Windsat

data at 18.7 GHz, H-pol should be used. However, WindSat provides measurements at more frequencies suitable for remote sensing of  $W$  than SSM/I.

In addition, our recent work on the electromagnetic properties of the sea foam (Anguelova and Gaiser, in preparation<sup>2</sup>), including the penetration depth of sea foam at various microwave frequencies, shows that each frequency has different sensitivity to the sea foam in different stages. While all WindSat frequencies would react to foam thicker than 1 cm, as the frequency decreases from 37 to 6 GHz, its sensitivity to thinner foam decreases. The lower limit of detectable foam thickness for 37 GHz is around 1 mm; for 10 GHz it is 4 mm; and for 6 GHz it is around 1 cm. Since thick foam is associated with the active wave breaking (stage A whitecaps), while thin foam characterizes decaying foam patches (stage B whitecaps), we infer that 6 GHz will detect predominantly stage A whitecaps. At the other extreme, 37 GHz will detect even the smallest presence of sea foam on the surface, be it young or mature. At 10 GHz, stage A will be always detected, but stage B whitecaps will be detected only partially (foam thinner than 4 mm will not be detected).

The frequency sensitivity to different types of foam implies that satellite-based estimates of whitecap coverage at various frequencies can be used to study different processes (e.g., turbulent mixing, gas exchange, spray production with different rates, and albedo) thus  $W$  should be reported at different frequencies (analogously to reporting aerosol optical depth at several wavelengths). At the same time there must be one value for the whitecap coverage from oceanographic point of view. Thus, a combination of the  $W$ -retrievals, i.e., some "effective"  $W$ , needs to be considered. While work on such an oceanographically-representative  $W$  is on going, in this study we use separate frequencies.

## 3. VALIDATION APPROACHES

A major task, before the building of the foam-fraction database and the development of improved parameterizations, is validation of the satellite-based estimates of  $W$ . Anguelova and Webster (2006) briefly discussed the difficulty of validating satellite-based estimates of whitecap coverage,  $W_{sat}$ . Direct validation and assessment of the quality of  $W_{sat}$  on a global scale is not straightforward for at least two reasons.

The first is the fundamental difference in the principle of measurement used, optically for virtually all available *in situ* data and radiometrically for satellite-based data. As explained in section 2.2, whether at some specific frequencies or as a combined quantity at all available frequencies, the microwave radiometers have the potential to sense well both types of foam

---

<sup>1</sup> Anguelova, M.D., M.H. Bettenhausen, P.W. Gaiser, W.F. Johnston, and R.M. Bevilacqua, Improved satellite-based estimates of whitecap coverage from passive microwave measurements.

---

<sup>2</sup> Anguelova, M.D., and P.W. Gaiser, Dielectric and radiative properties of sea foam at microwave frequencies: Conceptual understanding of the high emissivity of foam.

(thick, young and thin, decaying). Thus, the radiometric measurements are prone to see more foam on the surface. Meanwhile, optical measurements see well stage A whitecaps because their reflectivity is high, but tend to underestimate the less bright and less reflective decaying foam in stage B whitecaps, especially at low winds.

The second is the lack of sufficient and well-constrained ground truth values representing wide range of conditions. Plans for new whitecap measurements are sporadic and non-systematic, let alone coordinated with satellite measurements. This impedes the creation of a database of *in situ* whitecap coverage values,  $W_{ins}$ , co-located in time and space with satellite-based estimates and covering wide range of conditions.

To compensate for these difficulties with the ground truth data, we use different approaches for validation, each of which presents different issue. In the following we briefly describe these approaches and list the problems associated with each of them.

### 3.1 Validation with historical *in situ* data

Though lacking temporal and spatial matching with the satellite-based values, the historically available *in situ* measurements of whitecap coverage provide order of magnitude verification. The *in situ* database used by Anguelova and Webster (2006, Table 2) for the same purpose is now updated to include new measurements (e.g., Lafon et al., 2007; Sugihara et al., 2007) and is currently totaling 1251 points.

Besides the time/space mismatch, another major issue with the historical database is that for different datasets the photographic or video records used to extract  $W$  values are processed differently. Though the general principle of obtaining whitecap coverage from photographs or video records is well established (Asher and Wanninkhof, 1998), each group reporting *in situ*  $W$  data may implement it somewhat differently. Furthermore, different groups main impose the criteria for the intensity threshold separating foam from the surrounding water with different stringency. The  $W$ s thus evaluated would differ in the degree of representing the different stages of the whitecaps: mostly A, mostly B, the total A+B, or any combination of total and partial representations of A and B. Furthermore, the effect due to the use of photographs at oblique angle may be accounted for or not.

Finally, a systematic investigation of the reasons behind the wide spread in the *in situ* data is currently missing. It is commonly accepted to attribute about a third of this spread to experimental errors. But we do not have proven measure of how much of the spread in the *in situ* data is due to experimental/processing error(s) as opposed to varying environmental conditions.

Considering that, depending on the frequency, the satellite-based estimates of whitecap coverage represent either stage A whitecaps or the fractional area covered by A+B whitecaps, we separated the available

*in situ* data by the type of whitecaps they represent ( $W_A$ ,  $W_B$ , or  $W_A+W_B$ ). Eleven data sets are specifically reported to represent stage A whitecaps solely (507 points out of the total 1251 points). Only 3 data sets report stage B whitecaps separately (20 points). In all remaining cases (724 points from 22 data sets), the reported values for the whitecap coverage represent stage A+B foam. When the type of the whitecaps is not clearly stated (4 cases), we assumed that the reported values represent  $W_A+W_B$ . The two databases, for  $W_A+W_B$  and  $W_A$ , are shown in Figure 4b with black and blue diamonds, respectively.

### 3.2 Validation with wind-speed formula

Validation with existing wind-speed-dependent parameterizations of whitecap coverage is a convenient way to look at the global distribution of the whitecap coverage. Considering that, when needed in GC models, a global distribution of whitecap coverage is obtained by combining some  $W(U_{10})$  relationship with global satellite-based fields of  $U_{10}$ , such comparison can show what a directly measured field of  $W$  from satellites can offer. Such a comparison can illustrate similarities due to wind-speed dependence and differences due to other meteorological and oceanographic factors.

The main issue in this approach is the choice of the  $W(U_{10})$  relationship to be used for the comparison. Figure 1 in Anguelova and Webster (2006) shows the vast number of relationships that have resulted from the *in situ* data in the past. It shows that the data are not just noisy, but that interpretation, thus parameterization, of the data varies enormously too. As for the *in situ* data on which the  $W(U_{10})$  models are based (section 3.1), there is not a reliable measure presently making one or another  $W(U_{10})$  relationship preferable.

In this work we choose to work with two  $W(U_{10})$  relationships: the  $W$  parameterization Monahan and O'Muirchaertaigh (1980) obtained with robust biweight fitting (RBF) because it is widely used in GC models; and the Bortkovskii (1987) parameterization for moderate (3°C to 15°C) and warm (around 27°C) temperatures of the seawater, because these are appropriate for the ship-borne (section 3.3) and the air-born (section 3.4) measurements, respectively. These two models are plotted in Figure 4.

### 3.3 Validation with ship-borne data

High Wind Air-Sea Exchanges (HiWASE) experiment aims to obtain continuous flux measurements and concomitant meteorological and sea-state variables, including whitecap coverage (Yelland et al., 2007). The HiWASE instrumentation is mounted on the Ocean Weather Ship *Polarfront* occupying Station Mike at 66N, 2E. Data have been collected routinely since September 2006 with plans for continuous operation until September 2009. In addition, we used the preliminary image records gathered from June to September, 2006. HiWASE long-term measurements of whitecap coverage present unique opportunity to **directly** validate satellite-based retrievals

of whitecap coverage obtained from the WindSat observations.

Two cameras are collecting photographic records at pre-determined time intervals during the daylight hours of HiWASE experiment. In accompanying poster, Moat et al. (2009) present details on the measurements and explanation of the image processing method used to extract *in situ* whitecap coverage values.

For the period from June 2006 to August 2008, temporal and spatial match-ups between *Polarfront* and WindSat data have been made (4048 points in total). After various quality control screenings, total of 1787 satellite-*in situ* data pairs remain available for validation. This raw data set is further examined and then processed with the so-called spatial-temporal approach in order to obtain the satellite-*in situ* pairs ultimately compared.

The spatial-temporal approach (Ichoku et al., 2002) is widely used to validate satellite retrievals of various space-based missions with *in situ* data (e.g., Remer et al., 2002; Bailey and Werdell, 2006). The essence of this approach is that at each validation point spatially averaged satellite values are compared to temporally averaged *in situ* values.

We use three statistical parameters to judge how well the two data sets relate to each other: absolute difference (bias), correlation coefficient, and the coefficients (slope and intercept) of a linear model fitted to the paired data. The absolute bias is defined as the difference between the satellite and *in situ* values ( $W_{sat} - W_{ins}$ ). The correlation coefficient between two parameters,  $x$  and  $y$ , is defined as  $r = \sigma_{xy} / (\sigma_x \sigma_y)$ , where  $\sigma_{xy}$  is the co-variance of the two parameters and  $\sigma_x$  and  $\sigma_y$  are their standard deviations.

### 3.3.1 Processing of the satellite data

For the satellite data points of the raw data set the processing involves evaluation and decision how to average them spatially. Following the Ichoku et al (2002) approach, for each *in situ* location (latitude/longitude) an area covered by a box of  $N^\circ \times N^\circ$  and centered on this location is formed and the WindSat pixels falling within the  $N^\circ \times N^\circ$  box are taken. The average of the pixels falling in the box provides one spatially-averaged satellite-based value of whitecap coverage for the validation point in the center of this box.

Acknowledging that each WindSat pixel represents already intrinsically averaged value (section 2.2), an alternative to the averaging over an  $N^\circ \times N^\circ$  box is to use the WindSat pixel closest to the considered *in situ* location. To decide how to obtain the raw  $W_{sat}$  values for comparison—as the nearest pixel or as an average over  $N^\circ \times N^\circ$  box—it is important to evaluate the spatial stability of the satellite data.

To evaluate the spatial stability of the satellite data, we constructed three data sets using the station Mike position (66N, 2E) as a fixed *in situ* point. One data set takes all the variables (whitecap coverage, wind speed, and auxiliary information) at the WindSat pixel nearest to the fixed *in situ* point. In the other two data sets, all variables are averaged over the number of Windsat pixels falling within  $N^\circ \times N^\circ$  box centered on the fixed *in situ* point. The two considered averaging areas cover  $0.5^\circ \times 0.5^\circ$  (longitude/latitude) box (23 km x 56 km) and  $2^\circ \times 1^\circ$  (90 km x 111 km).

Examining the statistics of these three data sets, such as means, standard deviations, standard errors, biases, and relative differences, we observed that: (1) The geophysical variability for each of the box-averaged data sets, quantified by the standard deviation of the set, is comparable to the variability presented by the 'nearest' data; (2) With higher number of pixels available for averaging, the overall standard error for '2x1' data set decreases compared to that of '0.5x0.5' data set; (3) The mean standard deviations of the  $N \times N$ -box data sets behave conversely as the standard deviation for '2x1' data set increases compared to that of '0.5x0.5' data set most probably because by increasing the area for averaging we introduce wider range of geophysical variability; (4) The changes due to different averaging areas affect the 37H retrievals more than the 10H retrievals.

To avoid the additional effects on the  $W_{sat}$  values caused by the averaging over an area around a validation point and to keep the uncertainties of the satellite-based variables at their minimum, we decided to work with the nearest WindSat pixel.

### 3.3.2 Processing of the *in situ* data

For the *in situ* data points in the raw data set the processing involves evaluation and decision how to average them temporally. To find a time window that would compare the *in situ* data fairly to the spatially-averaged satellite data, we construct sets of satellite-*in situ* pairs for a range of time windows  $\Delta T$  and for each paired set we obtain the mean bias, slope and correlation coefficient. The range of time windows considered is from 5 to 180 minutes with a step of 5 minutes.

We expect narrower time windows (e.g., 10 min. or 20 min.) to provide less variations of the wind speed and whitecap coverage averaged in it. But, the restriction of the data to fall in narrower time window decreases the number of samples available for temporal averaging of the *in situ* data, which degrades the comparison statistics. At the other extreme is the use of a large time window, e.g., 180 min. While the availability of more data to average improves the comparison statistics, the longer time period includes wider range of geophysical variability, which may negate the improvements achieved with the use of more samples.

We find that these two opposing tendencies, which affect the comparison statistics, are well balanced when

averaging the *in situ* data in a time window of 60 min. We note however, that using a time window of 180 minutes does not ruin the comparison statistics significantly. On this basis, our current decision is to use this larger time window in order to maximize the number of averaged samples. Further analyses may change this initial decision.

In the time window chosen for temporal averaging, usually (60% of all cases) one and the same satellite pixel corresponds to each of the raw *in situ* values falling in this window. For example, if there are 10 different *in situ* values falling in the 180-minute window, the satellite values for all 10 points is the same. This is the satellite value compared to the temporally-averaged *in situ* value. There are cases, however, when more than one satellite pixels (2 to at most 4) correspond to the *in situ* values falling in the time window. While temporally these few satellite pixels represent changes in the surface scenery over 2-5 seconds, spatially they cover areas from approximately  $12.5 \times 12.5 \text{ km}^2$  to about  $37 \times 37 \text{ km}^2$ . In these cases, the satellite value compared to the temporally-averaged *in situ* value is the average of the available pixels. As a result of this processing, only a few of the satellite values used in the compared satellite-*in situ* pairs have error measures (standard deviation and standard error). When only one satellite pixel (the chosen nearest one, section 3.3.1) corresponds to the temporally-averaged *in situ* values, no error measure is available because (i) no spatial averaging beyond the intrinsic spatial averaging of the WindSat sampling is made; and (ii) the error due to modeling is still not fully characterized.

### 3.3.3 Binning data by wind speed

In addition to the rigorous spatial-temporal approach, we processed the raw satellite and *in situ* data in relation to their corresponding wind speed values. The rationale to look at this way of processing the data is the strong non-linearity of the  $W(U_{10})$  relationship.

When grouping the  $W$  data in a time window (section 3.3.2) without regard for the corresponding wind speed values, we may end-up averaging  $W$  data related to a wide range of wind speeds that may result from fast increase or decrease of the wind or wind gusts happening during the investigated time period. Introducing wide geophysical variability in this way, the  $W$  data used for comparison may have large standard deviations.

When grouping (binning) the  $W$  data by wind speed (satellite-based  $U_{10}$  values for  $W_{sat}$  and *in situ*-based  $U_{10}$  values for  $W_{ins}$ ), no matter what are the times and locations of the data, their wind speeds are similar. Thus, binning by wind speed would lead to less standard deviation of the averaged  $W$  data due to geophysical variability, but the spatial and temporal mismatch may still affect the data.

The weakness of this way of processing the data is that it leads, in effect, to the type of comparison we do

with the historical database (section 3.1). All the advantages that we may gain by having the satellite and *in situ* data co-located and simultaneous are lost. But the differences in the satellite-*in situ* comparisons obtained with the two processing approaches (spatio-temporal versus by wind speed, section 4) invoke some ambivalence as to which of these two approaches is more appropriate for whitecap coverage. The question is: What affects our data more, the geophysical variability, especially  $U_{10}$ , which we know is in strongly non-linear relation to  $W$ , or the spatial-temporal mismatch?

### 3.4 Validation with air-borne data

The field campaign Radiometry and Sea Surface Imagery (RASSI), conducted in August 2007, involves measurements of breaking waves with the NRL's Airborne Polarimetric Microwave Imaging Radiometer (APMIR) and the University of Washington's high resolution FoamCam both mounted in a U.S. Navy P3 aircraft. Data were collected over the Atlantic coast during low wind speed conditions on several days and over the Gulf of Mexico while heading towards Hurricane Dean on 21 August. In this study we use the data for higher wind speed conditions collected in the Gulf of Mexico.

The data set involves recorded microwave and video images of the sea surface in the Gulf of Mexico as a function of distance from the eye of Hurricane Dean (Figure 3). At seven measuring stations (green squares in the figure), from a working altitude of 6.1 km, the FoamCam sea surface optical imagery determined sea state, including presence of whitecaps, while APMIR measured the ocean surface brightness temperature  $T_B$  at frequency bands from 6.6 to 37 GHz. Collocated, nearly cotemporaneous data from both WindSat and SSM/I satellite radiometer overpasses were available, as were oceanographic and meteorological data from National Buoy Data Center buoys and ships (red squares and yellow triangles, respectively, in Figure 3).

Bobak et al. (2008) reported initial comparisons between satellite and aerial observations of  $T_B$  as well as the close relationship between APMIR-measured  $T_B$  and the whitecap coverage measured with the FoamCam. Further details on this experiment will be provided in a forthcoming paper (Bobak et al., in preparation<sup>1</sup>). In this report we present the first comparison between the whitecap coverage obtained from the video records of the FoamCam and the satellite-based whitecap coverage obtained from WindSat data.

The spatial resolution of the FoamCam was calculated to be 0.11 m per pixel at an aircraft altitude of 6.6 km giving a total field of view (FOV) of 159 m by 119 m when looking at the sea surface at nadir. Images of the sea surface within this FOV were taken along a

---

<sup>1</sup> Bobak, J.P., W. E. Asher, D. J. Dowgiallo, and M. D. Anguelova, Aerial radiometric and video measurements of whitecap coverage.

circular flight pattern around each measuring station. The whitecap coverage for each station was determined as the average of the values obtained along 3 to 8 flying circles.

The satellite-based whitecap coverage for each RASSI station was calculated from WindSat swath data as the average of all available WindSat pixels falling within  $0.5^\circ \times 0.5^\circ$  box (51 km x 56 km) centered at the latitude/longitude position for each station. In this analysis we work with box-averaged values for  $W_{sat}$ , instead of  $W$  at the nearest pixel, as for *Polarfront* analysis (section 3.3.1), because in the RASSI experiment we use and compare data from many different sources. To smooth to some extent the differences due to different sources, all used data are averaged over a  $0.5^\circ \times 0.5^\circ$  box. The preliminary results for the comparison statistics obtained with the two different processing—box-averaged versus the nearest pixel—differ by 1% to at most 6%.

Ultimately, spatially averaged whitecap coverage obtained along the circular flight pattern of the RASSI experiment with the FoamCam is compared to spatially averaged WindSat estimates of  $W$ . For each sensor, therefore, we have statistically averaged value, which can be considered representative for the whitecap variability within the area of measurement. In this way we reconcile the large disparity between the footprints of the two measurements, in order of hundreds of meters for the video data and in order of kilometers for the satellite data.

The main difficulty in obtaining the whitecap coverage from the RASSI observations is the high altitude of the measurements, chosen so to minimize the effect of the atmosphere on the APMIR records. This altitude does not affect the expectation that the radiometric measurements will register foam in both stages A and B, fully or partially depending on the frequency as discussed earlier (section 2.2 and the beginning of section 3). But for the video data this altitude will further compound the underestimation of stage B whitecaps and perhaps affect even the extraction of stage A whitecaps. We expect, therefore, FoamCam to provide values only for stage A whitecaps (more details in Bobak et al.).

#### 4. RESULTS AND DISCUSSION

Figure 4a compares the daily (1 March, 2007) satellite retrievals of whitecap coverage from WindSat measurements at all 5 frequencies, H pol. (colored clusters in the figure) to the historical database of *in situ* whitecap data representing stages A and B,  $W_A+W_B$  (black diamonds). Despite the spread of the *in situ* data, their trend is well delineated by the models of Monahan and O’Muirchaertaigh (1980) (purple line, MOM80) and Bortkovskii (1987) for moderate seawater temperature (magenta line, B87mod).

The tighter clustering of the satellite estimates, as compared to those shown in the feasibility study (Figure 8 in Anguelova and Webster, 2006), is due to the

minimization of some of the errors in the latest version of the retrieval algorithm. Anguelova and Webster (2006) did evaluate errors of the  $W$  estimates, but did not attempt to remove them.

Figure 4a demonstrates that  $W_{sat}$ , as computed with the revised algorithm, are order of magnitude comparable with  $W_{ins}$ . As expected, there are specific differences in this comparison for each of the satellite frequencies.

For most of the wind speed range, the satellite retrievals at 6 GHz are lower than the general trend of the data shown with the two models and usually comparable with the lower values of the *in situ*  $W_A+W_B$  (black diamonds). One possible reason for this observation is that this frequency is expected to be sensitive predominantly to stage A whitecaps and would not detect decaying foam effectively. To illustrate this statement, we show in Figure 4b the coverage of stage A whitecaps,  $W_A$  (blue diamonds) in addition to that of both stages,  $W_A+W_B$  (black diamonds). In magnitude, the 6 GHz satellite-based values are comparable to the upper half of the  $W_A$  values. This supports qualitatively the notion that 6 GHz retrievals represent predominantly stage A whitecaps.

The placement of the satellite-based values at 10 GHz relative to the *in situ*  $W_A$  and  $W_A+W_B$  in Figure 4b and the other frequencies in Figure 4a confirms the expectation that the representation of stage B whitecaps increases with the frequency. Using the models as a reference, we see that for higher frequencies (10 to 37 GHz), the satellite data are lower than most of the *in situ* data for wind speed above around 10-12  $m s^{-1}$ . For winds below this limit,  $W_{sat}$  at 10 GHz are closest to the trend of  $W_{ins}$  delineated by the MOM80 model. This could be anticipated because  $W_{sat}$  at 10 GHz most probably represent stage A whitecaps fully and stage B whitecaps partially, which is perhaps true for most of the  $W_{ins}$  points. Retrievals at frequencies above 18 GHz overestimate the *in situ* data and the models significantly.

This overestimation can be partially explained with the higher sensitivity of these frequencies to mature (thin) foam patches and strips. Since 37 GHz is the frequency most sensitive to the thinnest foam layers, its  $W_{sat}$  values at low winds are high. The satellite-based values  $W_{sat}$  at 18 GHz should be placed somewhere between the clusters of the 10 and 37 GHz. Instead, we see the 18 GHz estimates to be the highest and equal to those of 23 GHz. This hints that another reason for the overestimation at low winds is deficiencies in the satellite retrieval algorithm.

Generally, radiometric (passive) remote sensing at 23 GHz is used to obtain atmospheric variables (water vapor and cloud liquid water). The similarity of the  $W_{sat}$  estimates at 23 and 18 GHz therefore implies that the most plausible reason for the overestimation of  $W_{sat}$  at 18 GHz is shortcoming of our atmospheric correction. Difficulties in modeling the sea surface roughness and inadequate choices made in the foam emissivity model

may add to the overestimation of the whitecap coverage at low winds.

Figure 5 presents a quantitative measure of the differences between satellite-based estimates  $W_{sat}$  and the historical *in situ* data  $W_{ins}$ . Plotted as a function of the wind speed are the biases  $\Delta W = W_{sat} - W_{ins}$  for all WindSat frequencies (H-pol.) obtained with *in situ* data for  $W_A + W_B$  (panel a) and  $W_A$  (panel b). Lacking temporal and spatial matching, we pair the satellite and *in situ* data by binning them in wind speeds bins of  $1 \text{ ms}^{-1}$  (satellite-based  $U_{10}$  values for  $W_{sat}$  and *in situ*-based  $U_{10}$  values for  $W_{ins}$ ). The biases are plotted against the binned *in situ*  $U_{10}$  values.

Figure 5a, representing the case for  $W_A + W_B$  whitecap coverage, shows that the biases for 10 GHz (green squares) are the lowest up to about  $13 \text{ m s}^{-1}$ , while in Figure 5b, the 10 GHz estimates increasingly overestimate  $W_A$  for winds above  $4 \text{ m s}^{-1}$ . This supports the notion that 10 GHz frequency estimates stage A whitecaps fully and stage B whitecaps partially. This also confirms the qualitative observation in Figure 4 that the 10 GHz satellite retrievals of the A+B coverage is the closest to what the *in situ* database represents.

For 6 GHz (cyan squares), Figure 5a shows slight underestimation of the  $W_A + W_B$  coverage, while the biases for the stage A only in Figure 5b are the smallest. This confirms, therefore, the expectation that 6 GHz frequency is the best one to evaluate predominantly stage A whitecaps. Because most *in situ* data and models developed on their basis are representative of  $W_A + W_B$  values (section 3.1), we do not consider  $W_{sat}$  at 6 GHz further.

At higher frequencies (18 and 37 GHz, blue and red squares, respectively in Figure 5), the biases for A+B coverage (panel a) are smaller than the respective biases for  $W_A$  (panel b), lending credibility to the notion that these frequencies estimate well both young and decaying foam. While acknowledging that the satellite retrieval algorithm needs more improvements, the observation that the biases for 18 and 37 GHz for A+B whitecaps (Figure 5a) are larger than the A+B biases of 10 GHz is a reasonable support of the interpretation that physically frequencies of 18 and 37 GHz may estimate stage B whitecaps better than the *in situ* measurements and the 10 GHz retrievals are capable.

Figure 6 shows monthly (March, 2007) global distribution of whitecap coverage as obtained from WindSat data at 10 GHz, H pol. (upper panel) and as predicted with Monahan and O'Muirchaertaigh (1980) model (lower panel). The model is run with  $U_{10}$  values available from various sources (e.g., NASA scatterometer QuikSCAT or weather prediction analysis GDAS) matched temporally and spatially with the WindSat measurements. Similar comparison with the Bortkovskii (1987) model is not given because, being formulated for a specific seawater temperature, none of his three models is applicable globally. Meanwhile, a combination of Bortkovskii (1987) parameterizations for

cold, moderate and warm seawater temperatures produces discontinuities on a global map.

Figure 7 shows difference maps between whitecap coverage estimates at 10 and 18 GHz (H pol.) and the model,  $\Delta W = W_{sat} - W_{mod}$ . For 10H, there is slight overestimation of  $W_{sat}$  over the model at lower latitudes (global mean  $\Delta W$  of 0.04%, light pink areas) and more significant underestimation at higher latitudes (global mean  $\Delta W$  of -0.44%, bluish areas). In accord with Figure 4,  $W_{sat}$  at 18H overestimates the model for almost all global conditions (global mean  $\Delta W$  of 0.6%).

Compared to the feasibility study, these global distributions preserve the main feature discussed by Anguelova and Webster (2006, Figure 5) that the satellite-based whitecap coverage is more uniform latitudinally than the whitecap coverage from the model. The difference is that the  $W_{sat}$  values obtained with the latest retrieval algorithm present, albeit diminished compared to the model, the high latitudes (not the Trade winds zones) as the places with the largest whitecap coverage. As anticipated (Anguelova and Webster, 2006, §69), this is a result of the improved modeling and, more likely, the effort to minimize some of the errors in the algorithm (section 2.1). Overall, improvements in the algorithm do bring  $W_{sat}$  and  $W_{mod}$  (and by association  $W_{ins}$ ) closer, yet continue to exhibit differences, which can be attributed to the effects of various factors on  $W$  in addition to wind as discussed by Anguelova and Webster (2006, section 5.2).

Figure 8 plots the raw *in situ* data for whitecap coverage from the HiWASE experiment on *Polarfront* (red triangles) and the WindSat data matched with them (green squares). The comparison with the historical database (black and blue diamonds for  $W_A + W_B$  and  $W_A$ , respectively) shows two traits of the *Polarfront* data. First, the *Polarfront* data set contributes significantly to the lower limit of  $W_{ins}$  at winds below  $17 \text{ m s}^{-1}$ . The most probable reason for the lack of more data at higher winds is that at the *Polarfront* position (66N, 2E) the daylight availability restricts the collection of the video records to spring and summer thus emphasizing relatively lower winds. The lower temperature of the seawater (from 5 to 12°C) may also contribute to the presence of more points with low  $W$  value, but this suggestion needs further investigation. Second, by magnitude, the whitecap coverage values of the *Polarfront* data set are closer to the historical  $W_A$  values than to  $W_A + W_B$ . This suggests that, despite the effort to extract whitecap coverage for both stage A and B whitecaps, in reality we obtain mostly  $W_A$  because either the images collected on *Polarfront* represent predominantly active whitecaps or the intensity threshold in the image processing excludes some of the whitecaps at stage B.

The analysis of the raw data with the spatio-temporal approach, applied with a time window of 180 minutes for the *in situ* data and the nearest pixel for the satellite data, produces 61 pairs of *in situ*-satellite values, one point for each day when match-ups

between WindSat and *Polarfront* data are identified. To remove outliers, we impose a limit of  $\pm 1\%$  on the bias of the paired satellite-*in situ* data. The remaining 58 points are compared in Figure 9. Heeding the results revealed in Figure 4 that the satellite estimates at 10 GHz, H pol., are closest to the historical *in situ* database, we compare the 10H  $W_{sat}$  values to the  $W_{ins}$  values. The error bars in the figure represent the standard error resulting from the temporal averaging of the *in situ* data falling within the 180-min window. As noted previously, the longer time window allows more data points to be averaged, which results in relatively low standard errors. For the satellite data, error bars are available and shown in only a few cases (recall the last paragraph in section 3.3.2). The dotted line in the figure represent the 1:1 comparison between the *in situ*-satellite pairs.

Figure 9 shows that  $W_{sat}$  and  $W_{ins}$  values compare well for  $W \geq 0.2\%$ . It is worth noting that the recent addition of just a few data points at higher winds brought the comparison significantly closer to the 1:1 line as compared to an earlier version when we lacked high-wind points completely. The similarity of higher  $W_{sat}$  and  $W_{ins}$  values is in contrast to the underestimation at higher winds seen in Figure 4. For  $W$  values below 0.2%, the overestimation of the satellite data noted with the other validation approaches (Figure 4 and Figure 7) is confirmed.

The comparison statistics for the  $W_{sat}$  (10H) and  $W_{ins}$  pairs shows mean bias ( $W_{sat} - W_{ins}$ ) of 0.33%; the data are highly correlated with a correlation coefficient of 0.92; and fit well with a straight line. With a slope of 0.85 and an intercept of the y-axis of +0.39%, the straight line fitted to the data veers away from the 1:1 comparison in clockwise direction thus quantifying the  $W_{sat}$  overestimation at low values observed in Figure 9. The comparison statistics for 18H and 37H is similar with larger mean bias and y-axis intercept, quantifying their larger overestimation at low whitecap coverage values.

Figure 10 compares satellite and *in situ* data this time paired by wind speed bin (section 3.3.3), not by temporal/spatial matching. The error bars are the standard error for the averaging within each wind-speed bin. The blue symbols in Figure 10 compare the binned  $W_{ins}$  to binned modeled whitecap coverage obtained with the Bortkovskii (1987) model for moderate temperature run with the *in situ* wind speeds. We add this comparison to the figure only to demonstrate that the *in situ* wind speed values have some noise.

Though it is not immediately clear if pairing by wind speed bins (instead of by time-space matching) is legitimate, we make this comparison to gain some insights about the question which variability is more influential for whitecap coverage. The first observation in Figure 10 is that the standard errors in wind-speed bins are smaller than those for time-space matching seen in Figure 9. Next, the 1:1 comparison between  $W_{sat}$  and  $W_{ins}$  looks slightly better than that in Figure 9. It is conceivable then that for the case of the whitecap

coverage, its strongly non-linear relation to the wind speed trumps the variability caused by temporal and spatial mismatch. In other words, having wide wind speed variations in a time window is worse (Figure 9) than having time-space variations in a wind-speed bin (Figure 10).

Figure 11 plots, as a function of wind speed, the whitecap coverage values obtained with FoamCam during the RASSI experiment along the flight path (red triangles) and the whitecap coverage values obtained from WindSat data at 10 GHz, H pol., for each station where APMIR made measurements (green squares). Once again, for comparison we give the *in situ* historical database (black and blue diamonds for  $W_A + W_B$  and  $W_A$ , respectively) and the models of Monahan and O'Muirchaertaigh (1980) (purple line) and Bortkovskii (1987) for warm seawater temperature (magenta line). As expected, FoamCam values are strongly biased low even at the relatively high wind speed conditions of the RASSI experiment because they account only for type A whitecaps. To obtain representation of A+B whitecaps, the values plotted in Figure 11 are the original FoamCam data multiplied by a wind-speed dependent factor ranging from 11 to 15 (Monahan and Woolf, 1989, Table 1).

Figure 12 compares the temporally and spatially matched data from FoamCam and WindSat. We see features similar to those observed for the direct validation with the *Polarfront* data: relatively good comparison at high values of the whitecap coverage and overestimation of the *in situ* values by the satellite-based values at lower whitecap coverage. The shift of the paired data to the left from the 1:1 line implies that the multiplicative factor currently applied to obtain A+B whitecap representation is low. For example, a factor of 18 places the paired data exactly at the 1:1 line. Such nudging raises the question of how best to determine the factor converting  $W_A$  values to  $W_A + W_B$  values.

Setting aside this systematic bias toward lower values, we observe that the 1:1 comparison of the RASSI data in Figure 12 is somewhat better than that in Figure 9. One possible explanation is that with the RASSI data whitecap coverage at both high and low winds speeds is quantified with approximately the same number of points. Work on the statistical analysis of the comparison between whitecap coverage from WindSat and the RASSI experiment is under way.

## 5. CONCLUSIONS

Within the framework of the WindSat mission at Naval Research Laboratory (NRL), we have a dedicated project to address the high variability of whitecap coverage ( $W$ ) and the sea salt source function. A major task of this project is retrieving  $W$  on a global scale from satellite-based measurements. This will allow us to build a database representing  $W$  under a wide range of meteorological and environmental conditions.

The algorithm for estimating  $W$  from satellite measurements relies on changes of ocean surface

emissivity at microwave frequencies (6 to 37 GHz) due to presence of sea foam on a rough sea. After proving the feasibility (Anguelova and Webster, 2006), NRL now has the capability to obtain  $W$  globally using WindSat data and an improved version of the algorithm. Besides the use of physically-based models and independent data sets reported by (Anguelova et al., 2006), the latest improvements of the algorithm include improved calibration of the WindSat measurements, better understanding of and choices in the foam emissivity model, and minimization of some of the errors.

Direct validation and assessment of the quality of satellite-based whitecap coverage on a global scale is not straightforward due to lack of well constrained ground truth values under sufficiently varying conditions. Plans for new whitecap measurements are sporadic and non-systematic, making the creation of a database of *in situ*  $W$  values co-located in time and space with satellite-based estimates difficult. To compensate for insufficient ground truth data, we use different approaches for validation.

1) Validation with the historical *in situ* database gives an order of magnitude verification of  $W_{sat}$ .

2) Validation with existing wind-speed-dependent parameterizations of whitecap coverage illustrates similarities and differences of the global distribution of whitecap coverage obtained from models and from satellite-based data.

3) Validation with ship borne data collected during the High Wind Air-Sea Exchanges (HiWASE) experiment on ship *Polarfront* positioned at Station Mike (66N 2E) offers direct comparison of temporally and spatially matched *in situ* and satellite data for whitecap coverage. The *in situ* data are temporally averaged over 180 min time interval. These are compared to the nearest Windsat pixel, which represents spatial averaging over area of 50 km x 70 km. Comparison statistics, such as mean bias, correlation coefficient, and fit of a linear model to the data are used to judge how well  $W_{sat}$  and  $W_{ins}$  relate to each other. The same data are also compared after binning them by wind speed.

4) Validation with air borne data is another way to directly compare temporally and spatially matched values for whitecap coverage obtained from optical and radiometric observations. Data was collected in August 2007 during the field campaign Radiometry and Sea Surface Imagery (RASSI), which involves measurements of breaking waves with radiometric system (from NRL) and high-resolution video camera (from University of Washington).

The validation with the *in situ* historical database and existing  $W(U_{10})$  models confirms that the of  $W_{sat}$  are in the ball park of the expected values. There is underestimation of the models at high winds, and overestimation of the models and the *in situ* data at low winds.

There are three plausible reasons for the overestimation of  $W_{sat}$  at lower whitecap coverage

values. First,  $W_{sat}$ , by the virtue of their principle of measurement, are more sensitive to both type A and type B whitecaps. Thus, a satellite-based method to measure whitecaps is prone to register more foam in places where optical observation may underestimate decaying foam. This sensitivity to the thickness (thus type) of the foam depends on the frequency, thus the magnitude of the overestimation varies, being smaller at 10 GHz and increasing for 18 and 37 GHz.

Second, expected deficiencies in the retrieving algorithm would most likely add to the overestimation at low whitecap coverage values. These include shortcomings of the atmospheric correction, difficulties in the modeling of the rough sea surface, and inadequate choice in the foam emissivity model. The results show that algorithm deficiencies affect the  $W_{sat}$  estimates at 18 GHz the most.

The third reason for the overestimation of  $W_{sat}$  at low values is that we cannot consider the ground truth values perfect, free of error. Differences in implementing the extraction of whitecap coverage from video records and processing errors are expected to influence the low (less than  $10^{-3}\%$ ) values the most. Thus, the overestimation of  $W_{sat}$  may not be due solely to the satellite-based method.

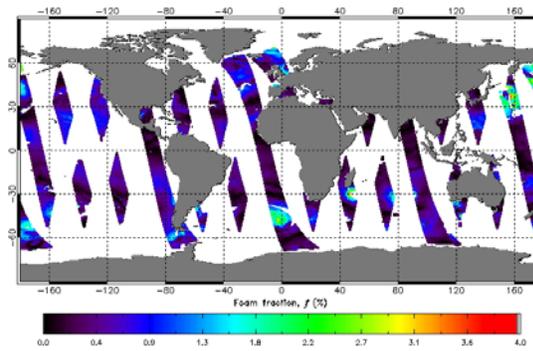
Direct comparison of temporally and spatially matched *in situ* and satellite values for whitecap coverage obtained from ship-borne and air-borne data converge on similar result. There is relatively good comparison between  $W_{sat}$  and  $W_{ins}$  at higher winds (high values of the whitecap coverage). The main problem for now is the scarcity of validation points at high winds. At lower winds (i.e., low whitecap coverage values),  $W_{sat}$  systematically overestimate  $W_{ins}$ , most probably for the same reasons listed above.

Future work can continue in two aspects. One is to patiently collect *in situ* data for direct validation using temporal-spatial matching and averaging. Another is further work on the retrievals algorithm. The latter should first involve tuning of some of the input parameters in the foam emissivity model  $e_f$ . We do not expect, however, that the overestimation for low whitecap coverage values will be removed fully with such a tuning. The main culprit for the overestimation at low  $W$  values is most probably a deficiency in our atmospheric correction and then the model for rough sea surface ( $e_r$ ).

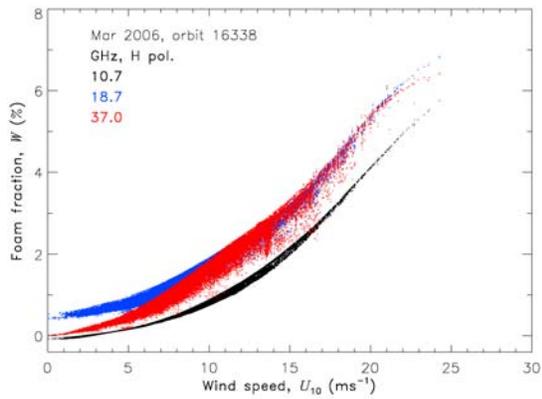
## 6. REFERENCES

- Anguelova, M. D., 2008: Complex dielectric constant of sea foam at microwave frequencies. *J. Geophys. Res.*, 113, C08001, doi:10.1029/2007JC004212.
- Anguelova, M.D., M.H. Bettenhausen, and P.W. Gaiser, 2006: Passive remote sensing of sea foam using physically-based modes. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS'06)*, Proceedings, 7, 3676 - 3679.

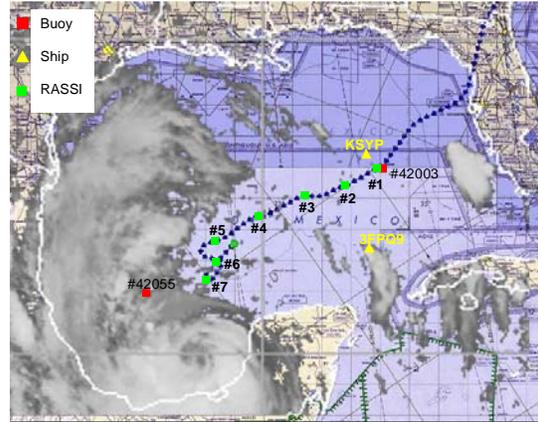
- Anguelova, M.D., and F. Webster, 2006: Whitecap coverage from satellite measurements: A first step toward modeling the variability of oceanic whitecaps. *J. Geophys. Res.-Oceans*, 111, C03017, doi:10.1029/2005JC003158.
- Asher, W., and R. Wanninkhof, 1998: The effects of bubble-mediated gas transfer on purposeful dual-gaseous tracer experiment. *J. Geophys. Res.*, 103, 10, 555–10,560.
- Bailey, S.W., P.J. Werdell, 2006: A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Rem. Sens. Env.*, 102, 12-23.
- Bettenhausen, M.H., 2007: On-Orbit Calibration of WindSat Brightness Temperatures. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS'07)*, July, Barcelona, Spain, (<http://www.grss-ieee.org/Resources/Audio>).
- Bettenhausen, M.H., C.K. Smith, R.M. Bevilacqua, N.-Y. Wang, P.W. Gaiser, and S. Cox, 2006: A nonlinear optimization algorithm for WindSat wind vector retrievals. *IEEE Trans. Geosci. Remote Sens.*, 44, 597-610.
- Bobak, J.P., W. E. Asher, D. J. Dowgiallo, and M. D. Anguelova, 2008: On the Correlation of Area-Extensive Measurement of Fractional Area Whitecap Coverage with Microwave Brightness Temperatures. 10th Meeting on Microwave Radiometry and remote sensing of the environment (*MicroRad*), March, Florence, Italy ([http://www.microrad2008.org/public/presentations/2008\\_0314\\_11.30\\_Bobak.pdf](http://www.microrad2008.org/public/presentations/2008_0314_11.30_Bobak.pdf))
- Bortkovskii, R., 1987: Air-Sea Exchange of Heat and Moisture during Storms. 193 pp., Springer, New York.
- Ichoku, C., D.A. Chu, S. Mattoo, Y.J. Kaufman, L.A. Remer, D. Tanre´, I. Slutsker, and B.N. Holben, 2002: A spatio-temporal approach for global validation and analysis of MODIS aerosol products. *Geophys. Res. Lett.*, 29, 10.1029/2001GL013206.
- Lafon C, Piazzola J, Forget P, Despiau S., 2007: Whitecap coverage in coastal environment for steady and unsteady wave field conditions. *J. Mar. Systems.*, 66, 38-46.
- Moat, B. I., M. J. Yelland, and R. W. Pascal, 2009: Oceanic whitecap coverage measured during UK-SOLAS cruises. AMS 16th Conference on Air-sea interaction, January 12-15, Phoenix, Arizona.
- Monahan, E., and I. G. O'Muircheartaigh, 1980: Optimal power-law description of oceanic whitecap coverage dependence on wind speed. *J. Phys. Oceanogr.*, 10, 2094–2099.
- Monahan, E., and D. Woolf, 1989: Comments on "Variations of whitecap coverage with wind stress and water temperature." *J. Phys. Oceanogr.*, 19, 706–709
- Remer, L.A., et al., 2002: Validation of MODIS aerosol retrieval over ocean. *J. Geophys. Res.*, 29(12).
- Remer, L.A., Y.J. Kaufman, D. Tanre´, S. Mattoo, D.A. Chu, J.V. Martins, R.R. Li, C. Ichoku, R.C. Levy, R.G. Kleidman, T.F. Eck, E. Vermote, and B.N. Holben, 2005: The MODIS Aerosol Algorithm, Products, and Validation. *J. Atmos. Sci.*, 62, 947–973.
- Sugihara, Y, H Tsumori, T Ohga, H Yoshioka, S Serizawa, 2007: Variation of whitecap coverage with wave-field conditions. *J Mar Sys*, 66, 47-60.
- Yelland, M.J., R.W. Pascal, P.K. Taylor, B.I. Moat, I. Skjelvan, C. Neill, 2007: High Wind Air-Sea Exchanges (HiWASE)—continuous air-sea flux measurements at station Mike. *15th Conference on Air-Sea Interaction*, August, Portland, OR.



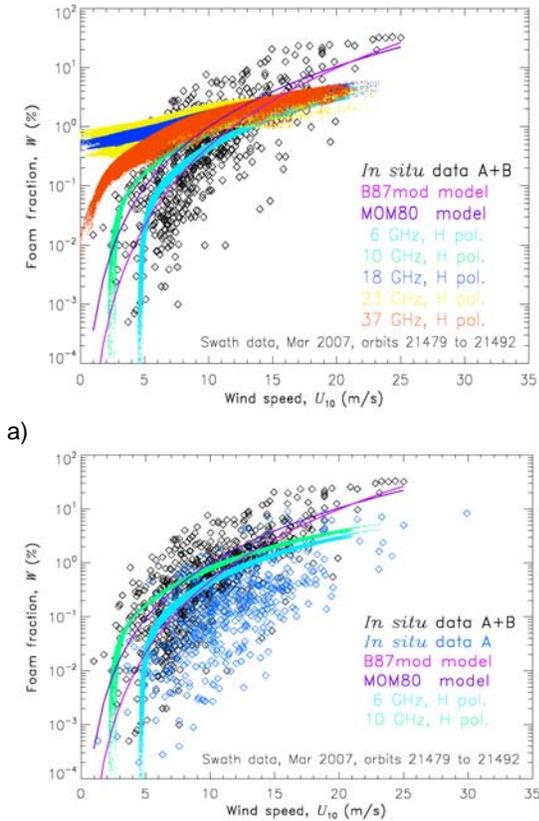
**Figure 1** Foam fraction  $W$  (in %) from satellite data at 10 GHz, H pol. Daily map for 1 March 2007 (orbits 21479 to 21492), swath resolution.



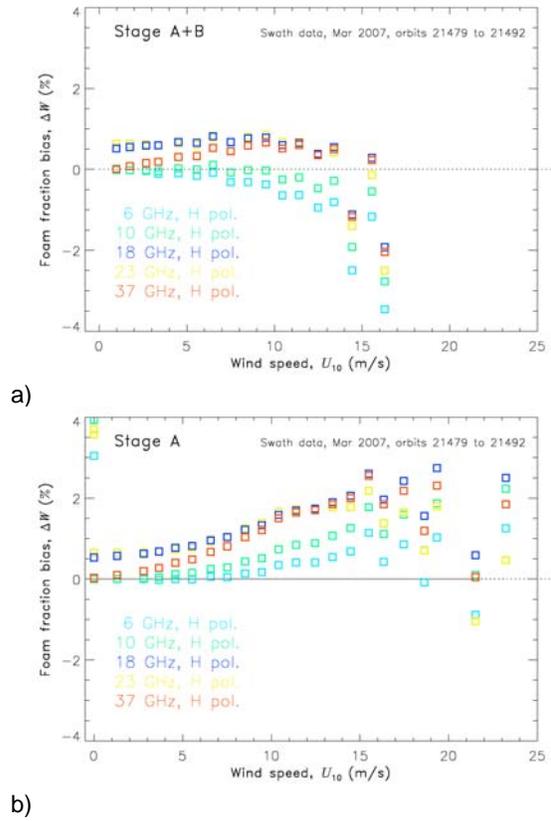
**Figure 2** Satellite-based estimate of foam fraction  $W$  (in %) vs wind speed obtained at different frequencies (H polarization).



**Figure 3** Flight track on August 21, 2007 from 2057 UTC to 2338 UTC overlaid on map of the flight region. Superimposed on map is an IR image of Hurricane Dean in the region taken at 2345 UTC on August 21, 2007. RASSI measuring sites are marked with green symbols. Red and yellow symbols show the positions of buoys and ships providing relevant auxiliary information.



**Figure 4** Whitecap coverage, as a function of wind speed, from satellite and *in situ* measurements: a) Historical *in situ* data for stage A+B whitecaps (black diamonds); satellite data at 5 frequencies (H polarization);  $W(U_{10})$  models of Monahan O’Muirchaertaigh (1980) (MOM80, purple line) and Bortkovskii (1987) for moderate seawater temperature (B87mod, magenta line); b) Historical *in situ* data for stage A whitecaps (blue diamonds) added and compared to satellite estimates from 6 GHz and 10 GHz.



**Figure 5** Whitecap coverage biases between satellite and *in situ* data ( $\Delta W = W_{\text{sat}} - W_{\text{ins}}$ , in %) as a function of *in situ* wind speed. Satellite data for 5 frequencies; *in situ* data for whitecap coverage  $W_A + W_B$  (a) and  $W_A$  (b). Data are paired by wind-speed binning of all available points.

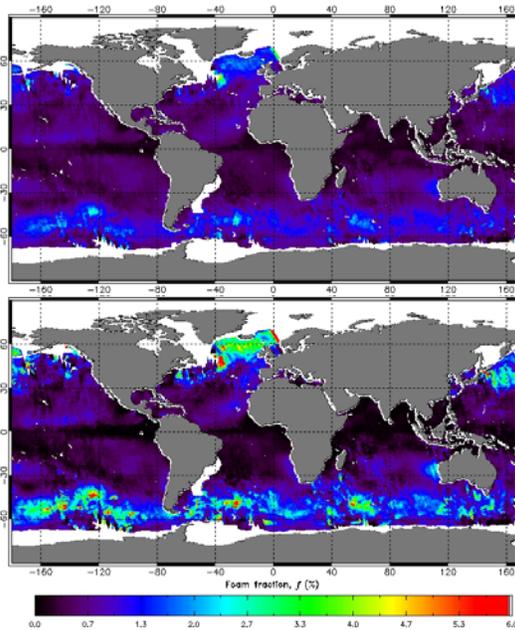


Figure 6 Global monthly (March, 2007) distribution of whitecap coverage from WindSat measurements at 10 GHz, H pol. (10H, upper panel) and  $W(U_{10})$  model of Monahan and O’Muirchaertaigh (1980) (lower panel).

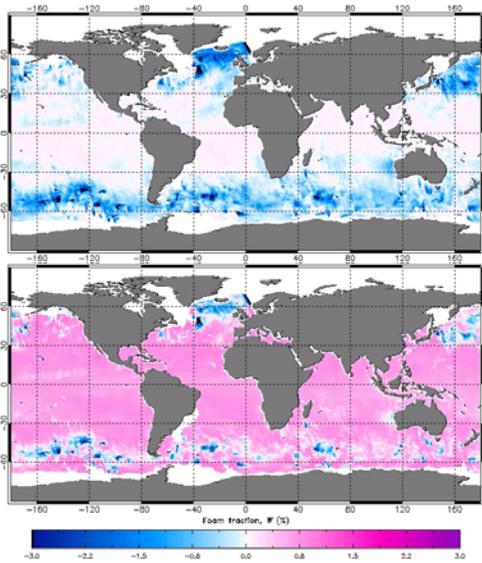


Figure 7 Difference maps  $\Delta W = W_{sat} - W_{mod}$  (in %) between whitecap coverage obtained from WindSat data ( $W_{sat}$ ) and Monahan and O’Muirchaertaigh (1980)  $W(U_{10})$  model. Upper panel:  $\Delta W$  for  $W_{sat}$  at 10 GHz (H pol.); Lower panel:  $\Delta W$  for  $W_{sat}$  at 18 GHz (H pol.)

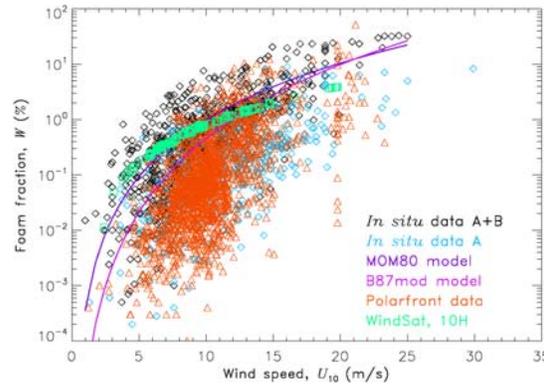


Figure 8 Raw whitecap coverage as a function of wind speed from *in situ* measurements at ship *Polarfront* (red triangles) and satellite data from WindSat (green squares). Shown for comparison are historical *in situ* data (black and blue diamonds for stages A+B and A, respectively) and two  $W(U_{10})$  models (as in Figure 4).

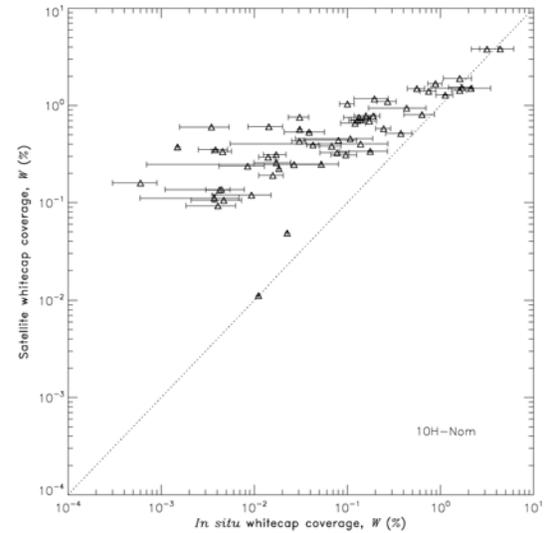


Figure 9 Pairs of *in situ* and satellite values for whitecap coverage matched in time and space. *In situ* data are averaged within 180 min time window; satellite data are at the nearest pixel.

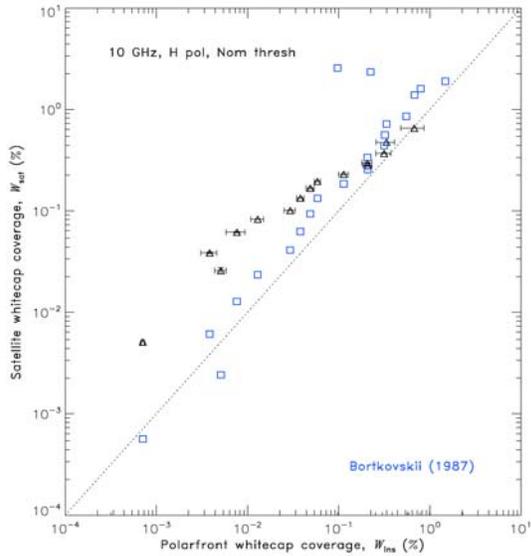


Figure 10 *In situ*-satellite pairs for whitecap coverage obtained via binning of data by wind speed (black triangles). The satellite data are for 10 GHz, H pol. Wind-speed-binned *in situ*-modeled pairs of whitecap coverage (blue squares). Bortkovskii (1987) model for moderate seawater temperature is used with *in situ* wind speed values.

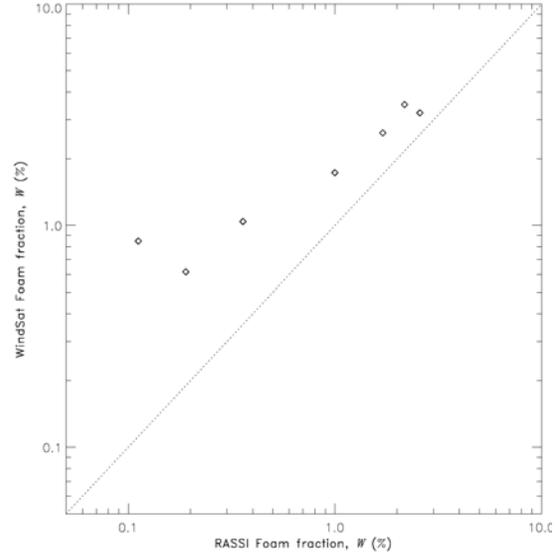


Figure 12 Pairs of air-born (FoamCam) and satellite (WindSat) values for whitecap coverage matched in time and space for the 7 stations of the RASSI experiment. A wind-speed dependent multiplicative factor ranging from 11 to 15 is applied to the whitecap coverage values from FoamCam to obtain representation of A+B foam (details in the text).

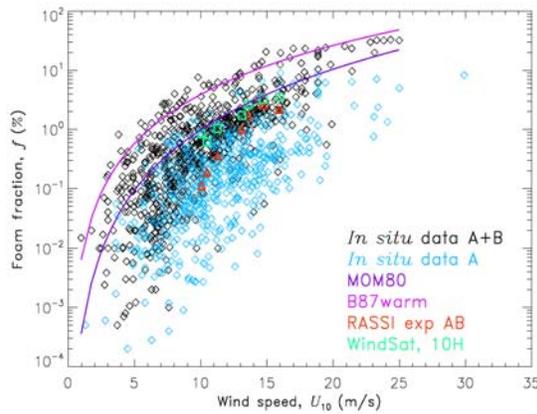


Figure 11 Whitecap coverage as a function of wind speed from *aerial* measurements with FoamCam during the RASSI experiment (red triangles) and satellite data from WindSat (green squares). Shown for comparison are historical *in situ* data (black and blue diamonds for A+B and A stage whitecaps, respectively) and two  $W(U_{10})$  models (as in Figure 4).