Important Contributing Factors for Estimating the Active and Total Whitecap Coverage Globally Using Satellite-Derived Parameters

Aaron C. Paget
IMPORTANT CONTRIBUTING FACTORS FOR ESTIMATING THE ACTIVE AND TOTAL WHITECAP COVERAGE GLOBALLY USING SATELLITE-DERIVED PARAMETERS

By

AARON C. PAGET

A Dissertation submitted to the Geophysical Fluid Dynamics Institute in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Degree Awarded:
Summer Semester, 2013
Aaron C. Paget defended this dissertation on April 26, 2013.

The members of the supervisory committee were:

Mark A. Bourassa
Professor Directing Dissertation

Guosheng Liu
University Representative

Doron Nof
Committee Member

Allan J. Clarke
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.
Dedicated to my family.
ACKNOWLEDGEMENTS

A. Callaghan provided the tabulated whitecap data, shown in Figures 5 and 7 represented by the MAP symbol. Images for these data were collected by Prof. Gerrit de Leeuw (Climate Change Unit, Finnish Meteorological Institute, Helsinki and Department of Physics, University of Helsinki, Finland, and TNO Built Environment and Geosciences, Utrecht, Netherlands) and Leo Cohen (TNO Defence and Security, Hague, Netherlands) during the Marine Aerosol Production (MAP) campaign headed by Prof. Colin O'Dowd (National University of Ireland, Galway) and G. de Leeuw. Magdalena Anguelova at the Naval Research Laboratory in Washington, D.C. developed and provided the Whitecap Database and the Historical Database of whitecap observations. Funding was provided from the following grants: NGI award #191001-363558-01, NOAA prime award #NA06OAR4320264, NRL award #N00173-06-2-C902, NGI award #191001-363405-03, NOAA prime award #NA11OAR4320199
# TABLE OF CONTENTS

List of Tables .......................................................................................................................... vii  
List of Figures ........................................................................................................................ viii  
Abstract .................................................................................................................................... x  

1. INTRODUCTION ................................................................................................................... 1  

2. COMPARING INSITU AND SATELLITE-BASED OBSERVATIONS ..................................... 7  
   2.1 Introduction .................................................................................................................... 7  
   2.2 Data .................................................................................................................................. 13  
      2.2.1 In Situ Data ........................................................................................................... 13  
      2.2.2 Satellite Whitecap Database ................................................................................ 15  
   2.3 Methodology .................................................................................................................. 17  
      2.3.1 Converting $U_{10}$ to $U_{10EN}$ ........................................................................... 18  
      2.3.2 Fitting the Power Law ......................................................................................... 18  
   2.4 Results ........................................................................................................................... 19  
      2.4.1 In Situ $U_{10}$-reported ....................................................................................... 19  
      2.4.2 Conversion functions of In Situ ........................................................................... 22  
      2.4.3 Comparison of In Situ and Satellite .................................................................... 24  
   2.5 Conclusions .................................................................................................................... 27  

3. SATELLITE-BASED OBSERVATIONS OF WHITECAPS USING $U_{10EN}$ ......................... 30  
   3.1 Introduction .................................................................................................................. 30  
   3.2 Data ............................................................................................................................ 32  
      3.2.1 Whitecap Database ............................................................................................. 32  
   3.3 Methodology ................................................................................................................ 32  
      3.3.1 PDF Evaluation ................................................................................................... 33  
      3.3.2 Data Preparation .................................................................................................. 33  
      3.3.3 Function Fitting .................................................................................................... 33  
   3.4 Results ........................................................................................................................ 34  
   3.5 Conclusions ................................................................................................................ 38  

4. SATELLITE-BASED OBSERVATIONS OF WHITECAPS USING WIND AND OTHER  
   VARIABLES ......................................................................................................................... 40  
   4.1 Introduction ................................................................................................................ 40  
   4.2 Data .......................................................................................................................... 40  
   4.3 Methodology ............................................................................................................... 44  
   4.4 Results ....................................................................................................................... 46  
   4.5 Conclusions ............................................................................................................... 53  

5. CONCLUSIONS ................................................................................................................... 56  

APPENDICES .............................................................................................................................. 61  
   A. DATA SAMPLING ........................................................................................................... 61
B. PARAMETER NAMES RANGES AND FORMULAS ......................................................67
C. FUNCTION FITTING ..........................................................................................................71
D. RANKING.............................................................................................................................75
REFERENCES ..............................................................................................................................78
BIOGRAPHICAL SKETCH .........................................................................................................83
LIST OF TABLES

1 Data sets contributing to the combined data set (CDS) used for this study. Sources, reported whitecap type, and number of contributing values are shown. ...........................15

2 Constants used for (2 - 4)...................................................................................................19

3 Coefficient values from power law fits for in situ observations for whitecap type (W_t and W_o), wind type (U_{10} and U_{10EN}), range (r_{26} and r_{15}), R^2, and LSE .....................................20

4 Mean coefficients for W_{37} and W_{10} over the ranges r_{22EN} and r_{15EN}........................35

5 Variables with PDFs that shift or widen............................................................................47

6 Orders of functions for coefficient functions α(γ) and β(γ). Cubic is represented by 3, quadratic by 2, linear by 1, and constant by 0 .................................................................48

7 Final rankings for U_{10EN−γ} pairs and U_{10EN}..................................................................49

B1 Parameter names, ranges, and parameters used for the 37 GHz Channel (W_{37}) of the Whitecap Database for r_{18EN} and r_{22EN}.................................................................69

B2 Parameter names, ranges, and parameters used for the 10 GHz Channel (W_{10}) of the Whitecap Database for r_{18EN} and r_{22EN}.........................................................................70

C1 Coefficients for fitting functions........................................................................................74
LIST OF FIGURES

1. Whitecap formation. (Left) Wind blows water, entraining air at wave crest and forming active whitecaps. Residual whitecaps are on windward side of wave as wave propagates to the right. (Right) Wave crest toppling over, plunging, and crashing, entraining air into the water column. Bubbles in column occur when wave plunges into the water.

2. The differences between $U_{10}$ and $U_{10EN}$ after $U_{10}$ was converted $U_{10EN}$ using equations from Kara et al. [2005] from -8°C < $dT$ < 7°C and over the range of 1 < $U_{10}$ < 40 ms$^{-1}$.

3. The differences of the percentage of whitecap coverage, $W(U_{10})$ minus $W(U_{10EN})$ [Kara et al., 2005] from -8°C < $dT$ < 7°C and over the range of 1 < $U_{10}$ < 40 ms$^{-1}$ using whitecap equation adjusted to percentage whitecap coverage from Monahan [1980].

4. Whitecap coverage for $W_t$ for 5 Jan 2006 from Whitecap Database. $W_t$ is in percent.

5. $W_t$ as a function of $U_{10}$ for $r_{26}$ (red solid line) and $r_{15}$ (blue solid line) with mean of observations (black dashed line) and the estimated error of $W_t$ (green vertical lines).

6. As in Figure 5 except for $W_a$.

7. As in Figure 5 except for $U_{10EN}$.

8. As in Figure 6 except for $U_{10EN}$.

9. Density plot of $W_{37}$ from Whitecap Database shown in the color rectangles: color values indicated by the log color bar scale. $W_t$ observations from the in situ data are shown in black dots. $W_t$ functional fits for the power law are shown using $U_{10}$ (light gray) and $U_{10EN}$ (dark gray) over the ranges for $r_{26}$ (solid) and $r_{15}$ (dashed).

10. Density plot of $W_{10}$ from Whitecap Database shown in the color rectangles: color values indicated by the log color bar scale. $W_a$ observations from the in situ data are shown in black dots. $W_a$ functional fits for the power law are shown using $U_{10}$ (light gray) and $U_{10EN}$ (dark gray) over the ranges for $r_{26}$ (solid) and $r_{15}$ (dashed).

11. PDFs of $W_{37}$ from $U_{10EN}$ for 1- ms$^{-1}$ bins. Shift or broadening of the PDFs indicates a possible dependence of $W_{37}$ on $U_{10EN}$.

12. As in Figure 11 but for $W_{10}$.

13. Coefficient pairs $a$ and $b$ for $W_{10}$ $r_{22EN}$. Calculated values from the 20 sample sets (blue circles). Mean value of $a$ and $b$ (red square).

14. Density plot of $W_{37}$ from Whitecap Database with the functional fits for $r_{22EN}$ (solid black line) and $r_{10EN}$ (dashed gray line) overlaid.
As in Figure 14 but for $W_{10}$. .................................................................38

PDFs of $W_{37}$ for 11 $< U_{10EN} < 13$ m s$^{-1}$ and $T_{air}$ in 2 $^\circ$C bins from -6 to 22 $^\circ$C. .................46

Performance of the short list of parameters and $U_{10EN}$ for all cases. Lower values indicate better performance. ...........................................................................................................50

Mean active whitecap coverage ($W_{10}$) for 2006 from Whitecap Database. Whitecap coverage is in percent.................................60

Mean total whitecap coverage (W$_{37}$) for 2006 from Whitecap Database. Whitecap coverage is in percent.................................60

Flow chart for sampling one independent and one dependent parameters. ......................64

Flow chart for sampling two independent and one dependent parameters. ......................64

Histogram of $U_{10EN}$ data from Whitecap Database in 1–ms$^{-1}$ bins. .......................................65

Density plot of $U_{10EN}$ and $W_t$ from Whitecap Database. Wind colors of density on a log scale. Black dots represent the 30 samples taken from each 1–ms$^{-1}$ bin; the samples maintain the same statistical mean and variance of all the values in their respective bins.65

A density plot $W_t$ observations by bins of SST and $U_{10EN}$. Solid grey line shows range of values for which each bin contains at least 30 data points for sampling with a maximum $U_{10}$ of 22 ms$^{-1}$. Dashed black line shows range of values for which each bin contains at least 30 data points for sampling with a maximum $U_{10EN}$ of 18 ms$^{-1}$. .............................66

A larger view of the bin from Figure A5 with the ranges 16°C $<$ SST $<$ 18°C and 17 ms$^{-1}$ $< U_{10EN} < 18$ m s$^{-1}$ with black dots representing the observations available and the white squares representing the 30 sampled values which maintained the same statistical mean and variance as all the parameters contained within the bin..............................................66

Function fits to $\alpha$(SST) (above). Function fits to $\beta$(SST) (below). .................................74
ABSTRACT

This study identifies the major contributing factors in estimating whitecap coverage globally from satellite observations. Power law functions for estimating both the active and total whitecap coverage from in situ observations are derived using $U_{10}$ (in situ winds) and $U_{10EN}$ (satellite-reported equivalent neutral winds). Whitecap coverage estimates using $U_{10EN}$ and a power law reproduced in situ whitecap estimates better than using $U_{10}$ and a power law, but, when compared with the satellite-based whitecap observations from the Whitecap Database, neither can adequately estimate the whitecap. New power law coefficients using $U_{10EN}$, available in the Whitecap Database, are presented and estimates are compared to the active and total whitecap coverage from the Whitecap Database. An additional 17 parameters, independent of $U_{10EN}$, are tested to determine their roles in influencing whitecap formation and duration using a modified power law; the most influential parameters are identified. The most influential parameters for the active whitecap coverage include SST, orbital velocity, air temperature, fetch, and the along SST gradient wind; the most influential parameters for the total whitecap coverage include SST, orbital velocity, air temperature, air-sea temperature differential, and the along SST gradient wind. Sampling and performance ranking techniques to manage large data sets are presented along with function fitting techniques for a modified power law. The improved understanding of whitecap coverage aids in estimating whitecap coverage globally using satellite products and in determining the global effects of whitecaps on air–sea processes and remote sensing of the surface.
CHAPTER ONE

INTRODUCTION

Whitecaps are the bubbles and foam generated by breaking and plunging waves. Ongoing efforts to understand whitecap formation and to estimate the percentage of the ocean’s surface covered by whitecaps, or whitecap coverage, have produced both theoretical and empirical results. Disagreements between these theoretical and empirical results have caused discord and confusion about which factors should be considered in the formation and duration of whitecaps and whitecap coverage. This study identifies the important parameters that should be considered in estimating whitecap coverage globally using satellite-based observations.

The wind imparts stress to the ocean surface, which forms waves. As waves form, the crests of the waves begin to break and entrain air, forming bubbles or foam (Figure 1). The foam can be continually generated at the wave crest. This foam, known as active whitecaps, is not advected with the wave. As the wave propagates away, the remaining foam diminishes over time. The diminishing foam is called residual whitecaps. Additional air–sea characteristics apart from wind stress affect whitecap formation. Changes in topography and wave–wave interactions can steepen the slope of the wave and promote breaking and plunging. A plunging wave breaks and topples over, crashing into the water. Air is entrained into the water column, forming air bubbles that rise to the surface as residual whitecaps. The percentage of the surface covered by active and residual whitecaps is called the total whitecap coverage.
Figure 1. Whitecap formation. (Left) Wind blows water, entraining air at wave crest and forming active whitecaps. Residual whitecaps are on windward side of wave as wave propagates to the right. (Right) Wave crest toppling over, plunging, and crashing, entraining air into the water column. Bubbles in column occur when wave plunges into the water.


Whitecap observations are taken from ships, buoys, piers, towers, platforms, aircraft, and satellite. Various techniques are used to determine whitecap coverage from these observations. Early estimates of whitecap coverage, such as those made with the Beaufort scale, come from visual inspection and are referenced with a corresponding wind speed.
Estimates of whitecap coverage from photo or video observations often use a brightness or intensity threshold to determine whitecap coverage [Nordberg et al., 1971; Ross and Cardone, 1974; Monahan et al., 1984; Asher and Wanninkhov, 1998; Xu et al., 2000; Asher et al., 2002; Stramaska and Petelski, 2003; Kraan et al., 1996; Suguhara et al., 2007]. Monahan et al. [1984] projected images onto tracing paper; trained and skilled researchers traced the whitecaps, then cut out and weighed the paper to determine the fraction of whitecaps within each image.

Automated whitecap extraction techniques use similar threshold methods [Lafon et al., 2004, 2007; Massouh and Le Clave, 1999]. Callaghan and White [2009] demonstrated a varying threshold technique for determining whitecap coverage from digital images. Satellite estimates of whitecap coverage use emissivity from the ocean surface to determine the portion of the total observed signal attributable to whitecaps [Anguelova and Webster, 2006].

Many individual data sets have been published from the sea surface images and other observations of whitecap coverage since the late 1960s. These data sets, which include observations from around the world, incorporate coastal and open ocean observations from a large range of temperatures, wind speeds, and sea states. Global in situ observations of whitecaps in every possible condition are not feasible. Because of this, individual data sets from in situ observations cannot represent the globe on the whole.

From the information available at that time, Blanchard [1963, 1983] estimated the global average whitecap coverage at 1–4%. Other estimates of whitecap coverage around the globe vary two to four orders of magnitude [Anguelova and Webster, 2006]. Monahan [2012] found over
262 different equations for estimating whitecap coverage; the majority of the equations are
power laws with the 10-m wind speed ($U_{10}$) as the argument. Aside from the theory that
whitecap coverage should follow a cubic power of the $U_{10}$ [Wu, 1979, 1988, 1992], little
consensus has been reached in determining whitecap coverage globally. These problems affect in
situ and satellite estimates of whitecap coverage and need to be resolved.

Visible, infrared, and microwave satellite-based observations of the ocean surface must account
for the contributions of whitecaps to the total signal. The algorithm to remove the whitecap
signal from the total visible signal for the CZCS, SeaWiFS, and MODIS used the equation from
Monahan [1980]; however, the calculated whitecap signal using $U_{10}$ overestimated the actual
signal [Gordon and Wang, 1994; Gordon and Voss, 1999]. The problem was noted and a
correction was made, but the cause of this difference was not fully explored. As a further
complication, satellite-based winds are reported as a 10-m equivalent neutral wind ($U_{10EN}$), not
$U_{10}$ as reported from in situ observations [Kara et al., 2008; May and Bourassa, 2011].
Application of $U_{10EN}$ to equations developed for $U_{10}$ adds to the error and uncertainty. A better
understanding of global whitecap coverage is required to account for the differences observed in
the estimated and the observed whitecap signal contribution to the total visible signal.

Whitecaps affect satellite observations of the Earth’s surface and air–sea turbulent fluxes, CO$_2$
exchange, momentum transfer, and aerosol production, and surface albedo. Correctly
determining the contribution of whitecaps to these processes is necessary to fully understand the
Earth system, but an understanding of the whitecap coverage, formation, and duration in
required.
The purpose of this study is to globally determine the whitecap coverage and to identify the major contributing factors to whitecap formation and duration using currently available satellite products. This study highlights the appropriate uses of $U_{10}$ and $U_{10EN}$ and the application of whitecap coverage equations to emissivity-based whitecap estimates using previously published in situ whitecap observations and satellite-based whitecap observations from the Whitecap Database [Anguelova and Webster, 2006]. Appropriate coefficients for a power law equation are developed using $U_{10EN}$ to estimate the active and total whitecap coverage. Additional parameters, independent of $U_{10EN}$, are tested to determine their roles in influencing whitecap formation and duration. The results lead to a better understanding of whitecap coverage, which aids in estimating whitecap coverage globally using satellite products and in determining the global effects of whitecaps on air-sea processes and remote sensing of the surface.

Chapter 2 discusses the derivation of equations to estimate whitecap coverage from in situ observations. The differences between $U_{10}$ and $U_{10EN}$ are reviewed, including the basic role of stability and the potential errors in using the inappropriate winds in estimating whitecap coverage. The in situ observations and new equations are compared to the Whitecap Database. Differences are identified and suggestions for uniformity and improvement are offered.

Chapter 3 covers the use of the Whitecap Database to determine appropriate power law coefficients for estimating the mean active and total whitecap coverage using emissivity-based whitecap coverage and $U_{10EN}$ from satellites and models. Techniques for reducing bias and
computational requirements are presented in the context of managing large amounts of data with an uneven distribution of geophysical conditions.

Chapter 4 includes an analysis of 17 different parameters used in conjunction with $U_{10EN}$ to determine the role of each parameter in estimating both the active and the total whitecap coverage. Each parameter using PDFs identifies potentially important parameters for further inquiry. A shortened list of parameters is further investigated and the performance is ranked to determine the parameters that have the strongest influence on whitecap formation and duration. The possible roles of these top-ranking parameters in influencing whitecap coverage are discussed.

Whitecaps can play a role in all air–sea processes and in surface and near-surface satellite remote sensing of the Earth’s surface. This study does not attempt to provide a full explanation of the physical processes that affect whitecap formation and duration; rather, the study identifies the factors most important in estimating the mean whitecap coverage globally. These contributing factors can be used to estimate whitecap coverage around the world using available satellite products. New estimates will aid in correcting problems found in previous formulations. Whitecap coverage estimates developed for global use can aid in research and in developing new instruments and tools for observing the Earth system.
CHAPTER TWO

COMPARING IN SITU AND SATELLITE-BASED OBSERVATIONS

2.1 Introduction

For the last 50 years, scientists have studied whitecap coverage around the globe in an attempt to relate the coverage to some other measureable quantity, such as wind speed. Functions derived from these measurable quantities are used for estimating whitecap coverage locally and globally, and then these estimates are used to determine whitecap formation mechanisms and the effects of whitecaps on satellite surface measurements and geophysical processes such as surface albedo and CO₂ exchange. Reducing the error of whitecap coverage estimates increases the ability to understand the potential role of whitecaps in these processes.

Historically, whitecap coverage was estimated as a function of the 10-m wind speed \( (U_{10}) \), a frequently recorded value. Although various types of functions have been used, a power law equation has been the predominant choice for estimating whitecap coverage (1):

\[
W = aU_{10}^b. \tag{1}
\]

One source of error is using the wrong type of wind in the functions to estimate whitecap coverage. Over 262 different functions exist [Monahan, 2013]; most of these functions depend on \( U_{10} \) and follow the form of a power law. Satellite-based winds are reported as equivalent neutral 10-m winds \( (U_{10EN}) \) and differ from \( U_{10} \) [Kara et al., 2008; May and Bourassa, 2011;
Applying $U_{10}$ to functions for estimating whitecap coverage in place of $U_{10}$ increases the error of the whitecap coverage estimates. In this study, appropriate power law coefficients are determined using available in situ whitecap observations with $U_{10}$ and $U_{10EN}$. These new functions are compared to global satellite-based observations of local whitecaps and $U_{10EN}$ to assess differences.

Total whitecap coverage ($W_t$) is broken into two parts: the active whitecaps ($W_a$) formed by breaking waves, and the residual or passive whitecaps. In situ observational studies of whitecaps focused on relating the wind speed to the percentage of water covered by whitecaps. Observations came from video, photography, visual inspection, and remote sensors from observing platforms. These platforms include land, ship, pier, buoy, towers, aircraft, and satellites, which collect observations around the globe and in many bodies of water during different observing missions. Observations of whitecap coverage were fit to functions using wind speed or other recorded quantities.

To observe the whitecaps with camera or video, a series of sequential observations are taken of the whitecap and the surrounding area. These images resolve the fractal shapes of the whitecaps against the non-whitecap-covered area; sea spray and individual bubbles from whitecaps are not resolved. Various analysis techniques determine the ratio of the area containing whitecaps to the total area using brightness thresholds or other techniques [Callaghan et al., 2008]. Although various techniques are used to calculate whitecap coverage, typically only one technique is applied for the observations from each mission or series of sequential observations.
For photographs, the whitecap coverage values from each set of the sequential series of individual images are averaged. Between 4 and 1000 individual images are averaged for one whitecap coverage data point depending on the technique, available images, and time frame. Each averaged value is reported as an individual whitecap observation and the individual whitecap observations are paired with the corresponding wind, temperature, and other recorded information. The averaged values are combined into data sets, some of which have been developed and published.

Empirical equations are derived using individual data sets, portions of data sets, or medleys of data sets. Wu [1979, 1988, 1992] suggested that whitecap coverage follows a cubic power law. This suggestion has fueled some of the known equations for determining whitecap coverage (Anguelova and Webster, 2006).

Equations typically are derived using in situ observations. Values of $U_{10}$ from in situ observation data sets typically ranged from 0 through 25 ms$^{-1}$. The majority of in situ observations were for wind speeds below 12 ms$^{-1}$; observations with $U_{10}$ above 15 ms$^{-1}$ were rare and sparse. The data used to derive the empirical equations do not represent the full range of $U_{10}$, but rather a smaller range [e.g., Ross and Cardone, 1971; Monahan and O’Muircheartaigh, 1986; Wu, 1992] additional observations are needed to calculate whitecap coverage globally.

Satellite-based winds are a likely source for additional wind observations, but they are reported as $U_{10EN}$, not $U_{10}$ as in in situ observations. Converting $U_{10}$ in situ observations to $U_{10EN}$ is possible using an air temperature ($T_{air}$) and sea surface temperature (SST) differential ($dT$, such
that $dT = T_{aw} - SST$) and the surface motion [Kara et al., 2008]. In situ observations of $U_{10}$ and $dT$ with corresponding $W_i$ and $W_o$ whitecap observations already exist; satellite-based observations of $U_{10EN}$ and whitecaps also exist.

Satellite-based observations of whitecaps, like those of the Whitecap Database [Anguelova and Webster, 2006], provide long-term, continual global observations useful in studying and estimating whitecap coverage. The whitecap signal notably contributes to the total signal in portions of the microwave, infrared, and visible bands. Satellites retrieve a signal, microwave, infrared, or visible; a portion of the total signal comes from whitecaps in surface observations. When the portion of the signal from whitecaps can be identified, it is used to determine whitecap coverage.

Satellites do not resolve individual whitecaps as photographs do; rather the whitecap signal from the footprint of the satellite determines the whitecap coverage. The whitecap observations are effectively spatially averaged for the entire footprint. Spatially averaged whitecap observations are comparable to the temporally averaged in situ whitecap observations following principles proposed by Taylor [1938].

Efforts to use estimated whitecap coverage from $U_{10}$ to predict whitecap signal contributions to the total signal as seen by satellites have not always been successful. For example, CZCS, SeaWiFS, and later MODIS estimated the visible signal contribution with Monahan’s equation [1980], which uses the $U_{10}$ to estimate the whitecap coverage in a power law, but the equation strongly overestimated the actual signal contribution from the whitecaps [Gordon and Wang,
The predicted whitecap signal contribution was reduced to correctly account for the whitecap signal contribution, but the cause of the difference between the predicted whitecap signal and the actual whitecap signal was not fully addressed.

The overestimation might have been caused by using the wrong wind in Monahan’s equation. The wind applied to the equation for the MODIS estimates might have been a satellite-based 10-m equivalent neutral wind ($U_{10EN}$), which is surface relative and presents a value adjusted to an equivalent neutral boundary layer profile [Kara et al., 2008], with the goal of estimating the stress with a neutral drag coefficient. These winds are not the same and their differences should be considered when estimating whitecap coverage.

Satellite winds are a valuable source of global wind speed distribution, but the differences between $U_{10EN}$ and $U_{10}$ can be large (Figure 2). Stability plays a large role in the difference between $U_{10}$ and $U_{10EN}$ at the lower wind speeds because of atmospheric boundary layer processes. At the upper wind speeds, mechanical mixing is the dominant force, and the difference is lessened. Even the small differences between the two types of wind make large differences in the whitecap estimates (Figure 3). The discrepancy in whitecap estimates requires investigation, the focus of this study.

The whitecap coverage and the equations for estimating whitecap coverage have been frequently compared for in situ observations, but comparing the in situ whitecap observations and satellite-based whitecap observations from the Whitecap Database using $U_{10EN}$ has not been attempted until this study.
Figure 2. The differences between $U_{10}$ and $U_{10\text{EN}}$ after $U_{10}$ was converted to $U_{10\text{EN}}$ using equations from Kara et al. [2005] from -8°C < $dT$ < 7°C and over the range of $1 < U_{10} < 40$ ms$^{-1}$.

Figure 3. The differences of the percentage of whitecap coverage, $W(U_{10})$ minus $W(U_{10\text{EN}})$ [Kara et al., 2005] from -8°C < $dT$ < 7°C and over the range of $1 < U_{10} < 40$ ms$^{-1}$ using whitecap equation adjusted to percentage whitecap coverage from Monahan [1980].
2.2 Data

2.2.1 In Situ Data
The majority of published whitecap data sets are based on in situ observations. Various techniques are used to determine the percentage of whitecap coverage for active ($W_a$), passive, and the total combined fields ($W_t$). Calibration between these techniques is rare, but the basic principles of identifying whitecaps from the surrounding water remain the same. Because of this, the observations can be compared with the caveat that the techniques do have unique biases.

Observations for each data set tend to occur in or near the same region. These regions tend to be coastal and in the Northern Hemisphere. Some more recent observational efforts have occurred in the open ocean. Principles and equations proposed by regional studies and combinations of regional studies are often used for global applications, though they are often inappropriate because the regional studies are verified only over the small area. Data sets can be combined to approximate local observations and to represent a small sampling of the local whitecap coverage globally. Combining data sets assumes that all observations and techniques for determining whitecap coverage are approximately intercalibrated. Data sets normally include wind speed and other geophysical parameters with the whitecap observation.

A Historical Database [Anguelova and Webster, 2006] of whitecap observations was compiled for active, passive, and total whitecap coverage observations with their time, location, and available geophysical attributes. The whitecap observations were made from ship, buoy, tower, and aircraft platforms. The Historical Database now covers from 1971 –2010. A total of 998, 610, and 20 data points were available for the combined active and passive fields, the active field, and the passive field, respectively. Nearly all of the data points contained a corresponding
$U_{10}$. Other parameters such as the 20-m or 19.5-m wind speed, fetch, $SST$, $T_{air}$, and stability, along with statistics of standard deviation and error estimates, were occasionally included.

A combined data set (CDS) compiled for this study includes 30 years of data from many parts of the world. Many of the observations can be found in the Historical Database. Additional observations are from the 2006 Marine Aerosol Production (MAP) campaign. The MAP campaign occurred in the northeast Atlantic and provided 44 Whitecap observations with corresponding $U_{10}$, $SST$, $T_{air}$, position, and time. Observations of both $W_t$ and $W_a$ are included in the CDS (Table 1).

In the CDS, used for this study (see Table 1), the distribution of the wind speeds is skewed toward the lower wind speeds; this distribution is an artifact of the observations and does not represent the global wind speed distribution. Observations show the variability of the $W_a$ and $W_t$ increases with increasing wind speed. Although the number of observations at the higher wind speeds is often insufficient to understand the behavior of $W_a$ and $W_t$ at the higher wind speed ranges, the observations are still valuable.

### 2.2.1.1 Active Whitecap Coverage ($W_a$).

For $W_a$, 204 observations include $U_{10}$, and $dT$ or $SST$ and $T_{air}$ for calculating $dT$, necessary for calculating $U_{10EN}$ for comparison with satellite winds (Table 1). A lower limit for the wind speed range of 3 ms$^{-1}$ is enforced with an upper wind speed limit enforced by available observations, around 25 ms$^{-1}$. These observations, which can be directly compared to satellite winds and whitecap observations, are referred to as the $W_a$ comparison data set.
Table 1. Data sets contributing to the combined data set (CDS) used for this study. Sources, reported whitecap type, and number of contributing values are shown.

<table>
<thead>
<tr>
<th>Source</th>
<th>In plots</th>
<th>$W_t$</th>
<th>$W_a$</th>
<th>Number of Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norberg et al., 1971</td>
<td>Norberg71</td>
<td>X</td>
<td>X</td>
<td>4</td>
</tr>
<tr>
<td>Monahan, 1971</td>
<td>Monahan71</td>
<td>X</td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>Ross and Cardone, 1974</td>
<td>RossCardone74</td>
<td>X</td>
<td>X</td>
<td>13</td>
</tr>
<tr>
<td>Monahan et al., 1983</td>
<td>JASIN83</td>
<td></td>
<td>X</td>
<td>63</td>
</tr>
<tr>
<td>Doyle, 1984</td>
<td>STREX83</td>
<td>X</td>
<td></td>
<td>85</td>
</tr>
<tr>
<td>Monahan et al., 1984</td>
<td>MIZEX83-film</td>
<td></td>
<td>X</td>
<td>21</td>
</tr>
<tr>
<td>Monahan et al., 1984</td>
<td>MIZEX83-video</td>
<td>X</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>Monahan et al., 1985</td>
<td>MIZEX84-film</td>
<td>X</td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>Monahan et al., 1985</td>
<td>MIZEX84-video</td>
<td>X</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Monahan et al., 1985</td>
<td>HEXOS84</td>
<td></td>
<td>X</td>
<td>27</td>
</tr>
<tr>
<td>Sugihara et al., 2007</td>
<td>Sugihara07</td>
<td>X</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>Callaghan et al., 2008</td>
<td>MAP</td>
<td>X</td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>Mironov and Dulov, 2008</td>
<td>MD08</td>
<td>X</td>
<td>X</td>
<td>71</td>
</tr>
</tbody>
</table>

2.2.1.2 Total Whitecap Coverage ($W_t$). Of the in situ observations, 484 include $W_t$, $U_{10}$, and $dT$ or $SST$ and $T_{air}$ for calculating $dT$, necessary for calculating $U_{10EN}$ for comparison to satellite winds (Table 2.1). A lower limit for the wind speed range of 3 ms$^{-1}$ is enforced with an upper wind speed limit enforced by available observations, around 25 ms$^{-1}$. These observations, which can be directly compared to satellite winds and whitecap observations, are referred to as the $W_t$ comparison data set.

2.2.2 Satellite Whitecap Database

*Anguelova and Webster [2006]* developed a technique to use passive microwave observations from WindSat to determine whitecap coverage on the ocean surface using the emissivity properties of whitecap foam bubbles. Whitecap coverage was calculated using the WindSat observations and a forward model ran with input data for wind speed and direction, $SST$, water vapors, cloud liquid water, and a constant salinity. The microwave channels respond differently
to $W_a$ and $W_t$ [Anguelova and Gaiser, 2012]. As a usable product, the Whitecap Database was formed with daily global whitecap coverage estimates. Whitecap values principally depend on the surface emissivity instead of the geophysical parameters such as wind, which is used for in situ-derived whitecap coverage functions.

The Whitecap Database used for this study is a daily $0.5^\circ \times 0.5^\circ$ global gridded product covering 2006 (Figure 4). Values of $W$ are available for both the ascending and descending passes of the WindSat instrument. This is the first continuous global whitecap data set that covers all seasons; makes actual surface layer observations from a satellite platform; and includes the additional parameters of $U_{10EN}$, wind direction, SST, $T_{air}$, significant wave height ($H_s$), and significant wave period ($T_p$). The geophysical processes associated with coastal regions are not explicitly addressed with the data set. The near-daily complete global coverage for an entire year with consistent spatial coverage makes the Whitecap Database an effective resource for determining which parameters affect whitecap coverage.

**Figure 4.** Whitecap coverage for $W_t$ for 5 Jan 2006 from Whitecap Database. $W_t$ is in percent.
2.2.2.1 Whitecaps from WindSat. The Whitecap Database uses the 10 GHz and 37 GHz frequencies from the passive microwave sensor Windsat [Gaiser et al., 2004; Anguelova and Webster, 2006]. The microwave signal emitted from the whitecaps in the 37 GHz channel \( W_{37} \) is comparable to \( W_t \); the signal in the 10 GHZ channel \( W_{10} \) is comparable to \( W_a \) [Anguelova and Gaiser, 2012]. Differences exist between \( W_t \) and \( W_{37} \) and between \( W_a \) and \( W_{10} \), but those differences are only a small source of error. Both channels are independently related to other parameters used in this study, such as the wind.

2.2.2.2 Wind. The 10-m wind from the Whitecap Database contains both the wind magnitude and direction. The wind comes from QuikSCAT and SSM/I as equivalent neutral winds \( U_{10EN} \), and when those are not available, from the Global Data Assimilation System (GDAS). The wind from GDAS is considered an equivalent wind although winds over the ocean appear to be closer to \( U_{10EN} \). WindSat winds are not used at all in the Whitecap Database and were not assimilated in GDAS during 2006. Winds are collocated temporally within ±60 minutes (for QuikSCAT) or ±180 minutes (for GDAS) of the WindSat observations.

2.3 Methodology

Satellite winds and in situ winds, though both reported at a height of 10-m, are not the same. The in situ wind, considered a true wind, is Earth relative and responds to the effects of stability in the atmospheric boundary layer. Satellite winds are surface relative and are given assuming an equivalent neutral profile consistent with the parameterized stress [Tang and Liu, 1996]. Satellite
winds and in situ winds can be related when the stability is known, and mean surface motion is ignored.

2.3.1 Converting $U_{10}$ to $U_{10EN}$

The conversion of in situ winds to equivalent neutral winds for this study uses the actual wind ($u$), roughness length ($z_0$), friction velocity ($U*$), von Karman Constant ($k$), drag coefficient ($C_D$), gravity ($g$), and Charnock’s constant (2 - 4). Charnock’s constant, $g$, and $k$ are assumed to be constant (Table 2); Charnock’s constant is known to vary, but exactly how the constant changes is not well understood, therefore a constant is assumed to limit error in the analysis. A non-constant drag coefficient is used as suggested by Powell [2003]. The drag coefficient is calculated using the formulas from Kara et al. [2005] and has a third order dependence on $U_{10}$ and a second order dependence on $dT$. The drag coefficient equation is valid for $u$ values from 1 to 40 ms$^{-1}$ and $dT$ values from -8°C to 7°C and has the distinction of being continuous over the entire range of valid values. Through these equations, the 10-m in situ wind is adjusted to $U_{10EN}$ and is directly comparable to satellite winds assuming zero mean surface motion since no observation-specific current and wave information is available.

2.3.2 Fitting the Power Law

The data for both $U_{10}$ and $U_{10EN}$ are used to derive the coefficients of the power law equation (1). The coefficients are chosen by minimizing the least squares error, resulting in four sets of coefficients: a set of coefficients for each $U_{10}$ and $U_{10EN}$ for $W_t$ and a set of coefficients for each $U_{10}$ and $U_{10EN}$ for $W_a$. 
\[
U_{10EN} = \frac{U_\ast \ln \left( \frac{10}{z_0} \right)}{k} \quad (2)
\]
\[
U_\ast = U\sqrt{c_D} \quad (3)
\]
\[
z_0 = 0.015 \frac{U_\ast}{g} \quad (4)
\]

**Table 2.** Constants used for (2 - 4).

<table>
<thead>
<tr>
<th>Constants used for (2 - 4)</th>
<th>0.015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charnock's Constant</td>
<td></td>
</tr>
<tr>
<td>Gravity</td>
<td>9.81 ms(^{-2})</td>
</tr>
<tr>
<td>von Karman Constant</td>
<td>0.4</td>
</tr>
</tbody>
</table>

2.4 Results

2.4.1 In Situ \(U_{10}\)-reported

A power law describes the relationship between \(W_o\) and \(W_i\) and \(U_{10}\). Two ranges for \(U_{10}\) are used: \(3 < U_{10} < 26\) ms\(^{-1}\) (\(r_{26}\)) and \(3 < U_{10} < 15\) ms\(^{-1}\) (\(r_{15}\)). Coefficients to the power law are derived using both the \(W_o\) and \(W_i\) comparison data sets over the two wind speed ranges (Table 3). The \(R^2\) statistic is provided as a measure of variability explained. The change in the least square error (LSE) from \(U_{10}\) to \(U_{10EN}\) is provided as a measure of improvement; positive values indicate reduced error and improved fit of the function to the data.

For the \(W_i\) comparison data set, the function fits for \(r_{26}\) and \(r_{15}\) are easily contained within the data (Figure 5). Both fits perform very well from 9 ms\(^{-1}\) and higher. The \(r_{15}\) fit is within the estimated error of the data from 9 through 15 ms\(^{-1}\) but underestimates the observed \(W_i\) value for the 3 to 9 ms\(^{-1}\) range. Similarly, the \(r_{26}\) fit also underestimates the observed \(W_i\) values for the same range. The \(r_{26}\) fit is close to the estimated error above 9 ms\(^{-1}\) and captures \(W_i\) from the uppermost wind speed observations quite well.
Table 3. Coefficient values from power law fits for in situ observations for whitecap type ($W_t$ and $W_a$), wind type ($U_{10}$ and $U_{10EN}$), range ($r_{26}$ and $r_{15}$), $R^2$, and LSE.

<table>
<thead>
<tr>
<th>Whitecap</th>
<th>Wind</th>
<th>Range</th>
<th>$a$</th>
<th>b</th>
<th>$R^2$</th>
<th>LSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_t$</td>
<td>$U_{10}$</td>
<td>$r_{26}$</td>
<td>1.26E-04</td>
<td>3.92</td>
<td>0.6126</td>
<td></td>
</tr>
<tr>
<td>$W_t$</td>
<td>$U_{10EN}$</td>
<td>$r_{26}$</td>
<td>1.73E-04</td>
<td>3.79</td>
<td>0.6182</td>
<td>1.7</td>
</tr>
<tr>
<td>$W_t$</td>
<td>$U_{10}$</td>
<td>$r_{15}$</td>
<td>1.05E-04</td>
<td>3.81</td>
<td>0.2908</td>
<td></td>
</tr>
<tr>
<td>$W_t$</td>
<td>$U_{10EN}$</td>
<td>$r_{15}$</td>
<td>1.42E-04</td>
<td>3.69</td>
<td>0.2983</td>
<td>-0.26</td>
</tr>
<tr>
<td>$W_a$</td>
<td>$U_{10}$</td>
<td>$r_{26}$</td>
<td>3.84E-04</td>
<td>3.07</td>
<td>0.4674</td>
<td></td>
</tr>
<tr>
<td>$W_a$</td>
<td>$U_{10EN}$</td>
<td>$r_{26}$</td>
<td>4.36E-04</td>
<td>3.01</td>
<td>0.4762</td>
<td>1.7</td>
</tr>
<tr>
<td>$W_a$</td>
<td>$U_{10}$</td>
<td>$r_{15}$</td>
<td>4.20E-06</td>
<td>4.69</td>
<td>0.1639</td>
<td></td>
</tr>
<tr>
<td>$W_a$</td>
<td>$U_{10EN}$</td>
<td>$r_{15}$</td>
<td>2.90E-06</td>
<td>4.83</td>
<td>0.1804</td>
<td>2</td>
</tr>
</tbody>
</table>

For the $W_a$ comparison data set, both fits are contained within the data, but the $r_{26}$ fit mostly overestimates the $W_a$ values in the 3 to 10 ms$^{-1}$ range (Figure 6). The $r_{26}$ fit is easily contained within the data and mostly within the estimated error above 12 ms$^{-1}$. The $r_{15}$ fit is contained within the data and mostly within the estimated error above 5 ms$^{-1}$. It also closely follows the mean of the data.

The power law coefficient, $b$, for both cases of $W_t$ and $W_a$ for $r_{26}$ and $r_{15}$ are similar to the coefficients found in the literature. Values from 2 through 4 are common; most values approximate a cubic value [Anguelova and Webster, 2006], but values above 5 are reported, too [Hanson and Phillips, 1999]. Although the functions fit the observations from the data, higher wind speeds have fewer observations and fit points. Considering the available data, these fits are reasonable for $W_t$ and $W_a$ using $U_{10}$, but the comparison data sets do not show the true variability that would likely occur in nature. The data from the CDS are intended to serve as a proxy for a global data set, but only a small part of the Earth’s oceans is sampled. Global coverage would require millions of ships and buoys or, alternatively, a satellite platform that measures $U_{10EN}$ instead of $U_{10}$.  

20
Figure 5. $W_t$ as a function of $U_{10}$ for $r_{26}$ (red solid line) and $r_{15}$ (blue solid line) with mean of observations (black dashed line) and the estimated error of $W_t$ (green vertical lines).

Figure 6. As in Figure 5 except for $W_a$. 
2.4.2 Conversion functions of In Situ

Since $U_{10}$ and $U_{10EN}$ describe different quantities, their relationship with $W$ will also change. Though the difference between $U_{10}$ and $U_{10EN}$ might appear small, these differences should not be ignored when using satellite-based winds. For the $W_t$ and $W_a$ comparison sets, the difference is as great as 2.45 and 1.15 ms$^{-1}$, respectively, and could have been larger if currents had been available. These differences in $U_{10}$ and $U_{10EN}$ are magnified when raised to a power. Two ranges for $U_{10EN}$ are used: $3 < U_{10EN} < 26$ ($r_{26EN}$) and $3 < U_{10EN} < 15$ ($r_{15EN}$). The functional fits for $U_{10EN}$ differ from those for $U_{10}$ as shown in Table 3.

The function fits for $r_{26EN}$ and $r_{15EN}$ are contained within the $W_t$ data (Figure 7). The $r_{15EN}$ fit is within the estimated error of the data from 9 through 15 ms$^{-1}$ but underestimates the mean observed $W_t$ value for the 3 to 9 ms$^{-1}$ range. Similarly, the $r_{26EN}$ fit also underestimates the mean observed $W_t$ values for the same range. Neither the $r_{26EN}$ nor the $r_{15EN}$ is contained within the error at the lower wind speed ranges. It is possible that including other variables such as currents, swell, or orbital velocity would reduce the error. The $r_{26EN}$ fit is close to the estimated error above 9 ms$^{-1}$ and captures $W_t$ from the uppermost wind speed observations quite well.

Both fits for $r_{15EN}$ and $r_{26EN}$ for the $W_a$ comparison data set are contained within the data, but the $r_{26EN}$ and $r_{15EN}$ fits mostly underestimate the mean $W_a$ values in the 3 to 9 ms$^{-1}$ range (Figure 8). The $r_{26EN}$ fit is easily contained within the data and approaches the estimated error for $U_{10EN} > 9$ ms$^{-1}$. The $r_{15EN}$ fit is close to the estimated error above 9 ms$^{-1}$ and also closely follows the mean of the data.
Figure 7. As in Figure 5 except for $U_{10EN}$.

Figure 8. As in Figure 6 except for $U_{10EN}$. 
The coefficients from the power law using $U_{10EN}$ are different from those corresponding to $U_{10}$ for both $W_t$ and $W_a$, as they should be since $U_{10EN}$ and $U_{10}$ are different quantities and have different properties (Table 3). The functions from $r_{26EN}$ and $r_{15EN}$ capture the overall basic shape of the data, but fail to represent the mean of the data through the data ranges. The fits are not within the standard error of the available data, but the coefficients found for $U_{10EN}$ should be used when the wind is $U_{10EN}$.

Converting $U_{10}$ to $U_{10EN}$ only slightly changes the $R^2$ values, sometimes reducing and sometimes increasing the values; no clear improvements in capturing the variability are evident. Including currents and swell, assumed to be zero for this study since observations did not exist, in the calculation of $U_{10EN}$ from $U_{10}$ might aid in improving $R^2$. Likely, the techniques used in determining whitecap coverage are also large contributors to the $R^2$ value, but this cannot be addressed here. The change of the LSE from $U_{10}$ to $U_{10EN}$ shows a reduction of about 2% in most cases. This reduction of error indicates $U_{10EN}$ is an improvement over $U_{10}$.

### 2.4.3 Comparison of In Situ and Satellite

The satellite-based Whitecap Database provides a wind and corresponding whitecap variable. The Whitecap Database wind is assumed to be $U_{10EN}$. The total whitecap field, $W_t$, is comparable to $W_{37}$. The active whitecap field, $W_a$, is comparable to $W_{10}$.

For $W_{37}$, the in situ-based coefficients from the $W_t$ fit are used over the ranges described by $r_{26EN}$ and $r_{15EN}$ (Figure 9). In situ observations are included to show the differences between the satellite observations and the in situ observations. Once more, as with the in situ case, the
coefficients for both $r_{26EN}$ and $r_{15EN}$ underestimate the mean value of $W_{37}$ for $3 < U_{10EN} < 12 \text{ ms}^{-1}$. 

The $r_{26EN}$ fit exponentially overestimates all values of $W_{37}$ for $U_{10EN} > 15 \text{ ms}^{-1}$, failing to capture any values of $W_{37}$ for above $U_{10EN} > 14 \text{ ms}^{-1}$. The $r_{15EN}$ fit resides near the mean value $W_{37}$ close to 15 ms$^{-1}$, but would also exponentially overestimate $W_{37}$ if extended to higher wind speeds. The change of the LSE from the coefficients derived from $U_{10}$ to the coefficients for $U_{10EN}$ shows a reduction of 17% and 3% for $r_{26EN}$ and $r_{15EN}$, respectively. This LSE improvement verifies that the $U_{10EN}$-based fits are more appropriate for $W_{37}$ from the Whitecap Database.

For $W_{10}$, the coefficients from the $W_a$ fit are used over the ranges described by $r_{26EN}$ and $r_{15EN}$ (Figure 10). The coefficients for both $r_{26EN}$ and $r_{15EN}$ underestimate the mean value of $W_{10}$ for $3 < U_{10EN} < 15 \text{ ms}^{-1}$. In fact, the $r_{15}$ and the $r_{15EN}$ fits underestimate over 99.99% of the values of the $W_{10}$ for $3 < U_{10EN} < 15 \text{ ms}^{-1}$. The $r_{26EN}$ and $r_{26}$ fits lie within $W_{10}$ for $16 < U_{10EN} < 22 \text{ ms}^{-1}$ but overestimate for $U_{10EN} > 23 \text{ ms}^{-1}$. The $r_{26}$ fit overestimates $W_{10}$ at lower wind speeds than the $r_{26EN}$ fit. The $r_{26EN}$ fit performs remarkably well compared to the $r_{15}$ fit, but still poorly estimates $W_{10}$. The change of the LSE from the coefficients derived from $U_{10}$ to the coefficients for $U_{10EN}$ shows an increase of 6% and 2% for $r_{26EN}$ and $r_{15EN}$, respectively. This LSE increase does not discount the usefulness of $U_{10EN}$-based fits. Rather, combined with the overall poor performance of fitting the data, the LSE increase suggests that none of the coefficients derived from the in situ data capture the mean $W_{10}$ from the Whitecap Database.
Figure 9. Density plot of $W_{37}$ from Whitecap Database shown in the color rectangles: color values indicated by the log color bar scale. $W_{t}$ observations from the in situ data are shown in black dots. $W_{t}$ functional fits for the power law are shown using $U_{10}$ (light gray) and $U_{10EN}$ (dark gray) over the ranges for $r_{26}$ (solid) and $r_{15}$ (dashed).

Figure 10. Density plot of $W_{10}$ from Whitecap Database shown in the color rectangles: color values indicated by the log color bar scale. $W_{a}$ observations from the in situ data are shown in black dots. $W_{t}$ functional fits for the power law are shown using $U_{10}$ (light gray) and $U_{10EN}$ (dark gray) over the ranges for $r_{26}$ (solid) and $r_{15}$ (dashed).
For both the $W_a$ and the $W_t$ cases, the actual in situ observations of the whitecaps with their corresponding $U_{10}$ and $U_{10EN}$ at times vary orders of magnitude when compared to other whitecap observations of similar $U_{10}$ or $U_{10EN}$. Additional geophysical parameters such as currents, swell, or orbital velocity might provide one possible explanation for these differences, but not entirely. The errors in observations of $U_{10}$ and $dT$ also cannot contribute enough to explain such differences. A more probable explanation is that the differences result from the techniques used to determine $W_t$ and $W_a$, and the other possible errors contribute as well. A calibration between techniques is merited in the future but is not included as part of this research.

2.5 Conclusions

The difference between satellite winds and in situ winds is large enough to merit new coefficients of the power law using $U_{10EN}$ as the argument. Whitecap coverage from satellites is predicted better using coefficients for $U_{10EN}$ than coefficients for $U_{10}$. This is the case for both $W_t$ and $W_a$. Though the changes in the coefficients may appear small, their application to the correct wind to the correct equation reduces overall error.

For the functional fits to the in situ $W_t$ and $W_a$, different coefficients are derived for $r_{26EN}$ and $r_{15EN}$. The coefficients are intended to be valid over their given ranges without extrapolation. Even then, the coefficients overestimate or underestimate the values of the mean and the error most of the time. Extremes where few observations exist make verification difficult.
Additional observations of both $W_t$ and $W_a$ would provide a better understanding of the variance and mean of the data over more ranges. Additionally, a calibration between the various techniques used for determining $W_t$ and $W_a$ would better help to determine the true mean and variance of the data and allow for better functions for determining $W_t$ and $W_a$.

Coefficients from $W_t$ with $U_{10EN}$ predict $W_{37}$, with limitations. Coefficients from $W_a$ should not be used to predict $W_{10}$ from the Whitecap Database. The parameters that influence $W_t$, $W_a$, $W_{10}$, and $W_{37}$, such as sea spray with visible observations and the change over point from active to residual whitecaps for both visible and microwave observations, also influence the coefficients and are a potential source of error. Better coefficients for $W_{10}$ and $W_{37}$ might be found using satellite-based whitecap observations with the accompanying $U_{10EN}$ values provided in the Whitecap Database.

Even with the coefficients derived for $U_{10EN}$, the estimated whitecap signal contribution to the visible spectrum would strongly overestimate the actual signal contributions for CZCS, SeaWiFS, and MODIS. The application of these coefficients to visible and microwave satellite observations would overestimate the signal contribution from whitecaps. Though whitecaps respond differently in the microwave and visible spectrums, the uniform overestimation of the whitecap signal contribution using in situ-derived functions encourages further investigation into the actual local whitecap values using global satellite observations, which are potentially independent of previous in situ observations.
Visible observations of whitecaps, sea spray, and sea state using the Beaufort scale have been shown to reliably predict wind speed within a 2 m/s, similar to potential differences in $U_{10}$ to $U_{10EN}$. The Beaufort scale cannot be directly translated to a specific amount of whitecap coverage, but the ability to predict the wind based on the Beaufort scale supports the conclusion that the larger sources of error in estimating whitecap coverage come from the techniques used to determine whitecap coverage rather than the type of wind used to estimate. Visible whitecap observations should be relatable to $U_{10}$ or to $U_{10EN}$. The differences exist in the parameters that contribute to visible and microwave observations, but the differences merit further exploration in a different study.

In situ whitecap coverage can be estimated by satellite-based winds using the $U_{10EN}$ coefficients, but errors exist in estimating $W_{10}$ and $W_{37}$ from the satellite-based Whitecap Database. The different techniques used to obtain $W_t$ and $W_a$ from the in situ observations possibly would benefit from calibration to the Whitecap Database to establish a common metric. The variability in both the actual whitecaps and the observational techniques used to calculate the whitecaps along with the accuracy in determining and recording the wind speed, SST, and $T_{air}$ might explain the differences between the Whitecap Database and the in situ observation values. Until the differences are accounted for, problems and unexplained discrepancies will continue to exist, especially in global application. A more global data set with more observations using a standard platform for observing and a uniform technique for determining $W_t$ and $W_a$ is necessary for determining a better global function for estimating whitecap coverage. The Whitecap Database has these traits and offers necessary values for determining a global function for $W_t$ and $W_a$ using satellite winds.
3.1 Introduction

Satellite-based observations of whitecaps made from the surface emissivity and backscatter measurements extend back more than four decades [Wilheit, 1979; Pandey and Kakar, 1982]. Unlike photographs from in situ data, satellites do not resolve the actual whitecap field. Instead, the portion of the retrieved signal that comes from the whitecaps is used in estimating the percentage of whitecap coverage. Satellite-based observations, along with corresponding wind speeds, can be used to better understand the functional dependence of whitecap coverage on the wind speed.

In chapter 2, it was shown that a power law can be fit to in situ data to estimate whitecap coverage, but, at higher wind speeds, that fit exponentially overestimates the $W_{37}$ and $W_{10}$ from the satellite-derived Whitecap Database. Even though the Whitecap Database was developed with in situ observations, the ability to equate in situ whitecap observations and observations from the Whitecap Database is limited. Inconsistencies between the analysis techniques used to determine whitecap coverage from in situ observations prohibit the calibration of in situ and satellite-observed whitecaps.
The Whitecap Database uses emissivity from the WindSat satellite for calculating the whitecap coverage. These whitecap observations were matched with corresponding $U_{10EN}$ values; $U_{10EN}$ values are not from WindSat. The Whitecap Database contains sufficient observations to determine whether whitecap coverage depends on $U_{10EN}$ and whether it is appropriate to use a power law in estimating whitecap coverage. In this study, appropriate power law coefficients are determined to estimate $W_{10}$ and $W_{37}$ using $U_{10EN}$ to show that whitecap coverage can be estimated using satellite winds. Ample observations of $W_{37}$ and $W_{10}$ exist to determine and verify the new power law coefficients.

Power law coefficients derived from in situ observations fail to capture the mean whitecap coverage from the Whitecap Database. Therefore, new coefficients are required for estimating $W_{10}$ and $W_{37}$. The power law coefficients for estimating mean $W_{10}$ and $W_{37}$ using $U_{10EN}$ alone over large wind speed ranges to estimate whitecap coverage with satellite-based winds were not determined until this study.

The analysis of $W_{37}$, $W_{10}$, and $U_{10EN}$ begins with PDF evaluations to determine possible dependencies [Milliff et al., 2004; Lary and Lait, 2006]. The data are sampled over multiple wind speed ranges and are fit to the power law function (1) following the technique presented by Roberts et al. [2010]. The new coefficients of the functional fits are tested against the coefficients presented in chapter 2 to show reduced error.
3.2 Data

3.2.1 Whitecap Database

The whitecap database, as described in chapter 2, is a global database of whitecap coverage, both $W_{37}$ and $W_{10}$ (comparable to $W_t$ and $W_a$, respectively), with corresponding $U_{10EN}$ values on a 0.5° x 0.5° global grid. Whitecap values are calculated from the 10 and 37 GHz channels of WindSat and were determined using a forward model that accounts for the emissivity of the surface, wind speed and direction, SST, water vapors, cloud liquid water, and a constant salinity. The Whitecap Database includes daily values for 2006, from both ascending and descending passes of the satellite. The number of valid gridded values, considered observations, consist of $1.9 \times 10^7$ for $W_{37}$ and $1.8 \times 10^7$ for $W_{10}$ and their corresponding $U_{10EN}$ values. Details and characteristics of $W_{37}$, $W_{10}$, and $U_{10EN}$, the values used for this study, are as previously described.

3.3 Methodology

The number of whitecap observations from the Whitecap Database greatly exceeds the number of available in situ observations. The distribution of the $W_{37}$ and $W_{10}$ observations is heavily skewed toward the lower values. To manage the large number of values computationally and statistically, techniques are borrowed from satellite remote sensing validation applications, namely performing evaluations based on the probability distribution function (PDF) of the data, binning the data, and sampling the data. Then, as with in situ observations, the data are fit to a power law function.
3.3.1 PDF Evaluation

The PDFs of the data are evaluated by binning the data into 1- ms\(^{-1}\) bins, then calculating the PDF of the whitecap values for each of these bins. A shift in the peak location or a change in the distribution of the data indicates a possible dependence of the whitecap values on the wind speed. This PDF evaluation is performed for both \(W_{37}\) and \(W_{10}\). Shifting or broadening PDFs indicate a possible dependence of \(U_{10EN}\) on whitecap coverage, indicating need for further investigation for the dependence of \(U_{10EN}\) on whitecap coverage.

3.3.2 Data Preparation

After the PDF evaluation, the data are binned and sampled. Further descriptions of data sampling are found in Appendix A, and the case of \(W_{37}\) and \(U_{10EN}\) is described therein. For both \(W_{37}\) and \(W_{10}\), \(U_{10EN}\) is binned into 1- ms\(^{-1}\) bins; N=30 samples are selected from each bin, creating a training data set. This sampling technique maintains the statistical mean and variance of the data. The range for \(U_{10EN}\) is based on available data and valid wind speeds from the GDAS, QuikSCAT, and SSM/I sources. Two cases are presented: \(3 < U_{10EN} < 22\) (\(r_{22EN}\)) and \(3 < U_{10EN} < 18\) (\(r_{18EN}\)). This sampling is repeated 20 times for each \(r_{22EN}\) and \(r_{18EN}\) to verify accuracy and decrease error due to sampling. This sampling methodology reduces bias resulting from the uneven distribution of the data and greatly reduces computational costs.

3.3.3 Function Fitting

A power law equation (1) is used to fit the data following the methodology described in chapter 2. The coefficients of the power law equation are chosen by minimizing the least squares error. Coefficients are found for \(r_{22EN}\) and \(r_{18EN}\) over all 20 sets of samples for both \(W_{37}\) and \(W_{10}\). The
mean values of $a$ and $b$ for each case ($W_{37\, r_{22\text{EN}}}$, $W_{37\, r_{18\text{EN}}}$, $W_{10\, r_{22\text{EN}}}$, and $W_{10\, r_{18\text{EN}}}$) provide composite coefficients for the four possible cases. Estimates made using the remainder of the data, the validation data set, are then compared to the $W_{10}$ and $W_{37}$.

### 3.4 Results

A shift in the position or a broadening of the distribution of data indicates a probable dependency of one variable on another. The PDFs of the $W_{37}$ and $W_{10}$ from the 1- m s$^{-1}$ wind speed bins show both a shifting and a broadening of the distribution of the data (Figures 11 and 12, respectively) as a function of the wind speed range. The shift in the peaks indicates a dependency between $U_{10\text{EN}}$ and $W_{10}$, and between $U_{10\text{EN}}$ and $W_{37}$. The broadening of the PDFs with increasing wind speed

![PDFs](image)

**Figure 11.** PDFs of $W_{37}$ from $U_{10\text{EN}}$ for 1- m s$^{-1}$ bins. Shift or broadening of the PDFs indicates a possible dependence of $W_{37}$ on $U_{10\text{EN}}$. 

34
speed indicates a larger variability of $W_{10}$ and $W_{37}$ values at higher wind speed. This dependence requires further investigation.

Consistent with the literature, $U_{10EN}$ is considered an independent variable and $W_{10}$ and $W_{37}$ are dependent variables. The data are sampled and fit to the power law function and mean values for $a$ and $b$ for each of the four cases are determined (Table 4). These coefficients are considered

---

**Figure 12.** As in Figure 11 but for $W_{10}$.

**Table 4.** Mean coefficients for $W_{37}$ and $W_{10}$ over the ranges $r_{22EN}$ and $r_{15EN}$.

<table>
<thead>
<tr>
<th>Whitecap</th>
<th>Range</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{37}$</td>
<td>$r_{22EN}$</td>
<td>4.36E-02</td>
<td>1.5495</td>
</tr>
<tr>
<td>$W_{37}$</td>
<td>$r_{18EN}$</td>
<td>3.78E-02</td>
<td>1.6056</td>
</tr>
<tr>
<td>$W_{10}$</td>
<td>$r_{22EN}$</td>
<td>5.74E-03</td>
<td>2.1719</td>
</tr>
<tr>
<td>$W_{10}$</td>
<td>$r_{18EN}$</td>
<td>5.24E-03</td>
<td>2.2858</td>
</tr>
</tbody>
</table>
valid only over the range of values in their derivation. From each case, a range exists over which
\(a\) and \(b\) appear to be valid (Figure 13). For \(W_{10\,r_{22EN}}\), there appears to be some dependence of the
coefficients \(a\) and \(b\) on each other. This dependency exists for the other three cases as well (not
shown).

![Figure 13. Coefficient pairs \(a\) and \(b\) for \(W_{10\,r_{22EN}}\). Calculated values from the 20 sample sets
(blue circles). Mean value of \(a\) and \(b\) (red square).](image)

The differences between the coefficients for each case are non-negligible. Both the ranges of the
\(U_{10EN}\) values used and the type of whitecap coverage estimated influence the coefficients. The
ranges of the coefficients do not overlap; each case is determined for a specific range of \(U_{10EN}\)
and whitecap type. Using the coefficients from $r_{18\text{EN}}$ to estimate $W_{10}$ or $W_{37}$ above 18 m s$^{-1}$ introduces error unnecessarily.

When compared to the Whitecap Database, the calculated $W_{37}$ from the validation data set using the derived coefficients from the power law fit closely follows the mean composite line for both the $r_{22\text{EN}}$ and $r_{18\text{EN}}$ fits (Figure 14). The fits for $W_{10}$ similarly follow the mean (Figure 15). All fits are contained within the data. There are still regions where the functions over- or underestimate the mean value, but the fit follows the mean quite well. In contrast to the coefficients presented in chapter 2 using the in situ observations, the total error from a least squares fit is reduced by over 95\% for all cases— not a trivial improvement. The spread of the data are still not explained by a single dependent variable.

![Figure 14. Density plot of $W_{37}$ from Whitecap Database with the functional fits for $r_{22\text{EN}}$ (solid black line) and $r_{18\text{EN}}$ (dashed gray line) overlaid.](image)
Figure 15. As in Figure 14 but for $W_{10}$.

3.5 Conclusions

The Whitecap Database’s values of $W_{37}$ and $W_{10}$ can be considered dependent on $U_{10EN}$. A power law function using $U_{10EN}$ can reproduce $W_{37}$ and $W_{10}$ reasonably well from the Whitecap Database. Over the appropriate ranges, the coefficients reproduce the mean $W_{37}$ and $W_{10}$ values with 95% reduced error, compared to the coefficients in chapter 2. Functional representations of $W_{37}$ and $W_{10}$ are possible.

The data sampling method produces consistent, repeatable results; the sampling is representative of the mean and variance of the data set and reduces biases related to the uneven distribution of the data for both the whitecap (seen in the broadening of the PDFs with higher wind speed) and
$U_{10EN}$ values. The repetition of the sampling reduces sampling error, confirms ranges of valid values of $a$ and $b$ for each case, and provides composite mean values of $a$ and $b$.

The coefficients $a$ and $b$ are codependent: the choice of $b$ greatly influences the value of $a$ to minimize error. Choosing $b$ to be a cubic, for example, also changes $a$, increases the overall error, and does not follow the mean values well. The coefficients in Table 4 are determined to be the best choice for estimating whitecap coverage over their given ranges, minimizing the least square error and reducing bias caused by an uneven distribution of the data.

This study is valid for ranges of $U_{10EN}$ up to 22 ms$^{-1}$; any extrapolation of the equations is inappropriate. The data were fit to over 95% of the total available data. The validity of wind speeds exceeding 22 ms$^{-1}$ from the Whitecap Database might be called into question because of the limitations of their sources. To prevent error, the threshold of 22 ms$^{-1}$ is introduced. The additional threshold of 18 ms$^{-1}$ is useful for considering additional variables in further studies.

The simple power law is used here to capture the mean whitecap coverage and cannot explain all of the variability of the data. The repeatability of the sampling and function fitting and the large number of observations in the Whitecap Database make it possible to extend the work using $W_{37}$, $W_{10}$, and $U_{10EN}$. Additional parameters available and derivable from the Whitecap Database can be studied for their possible influence on $W_{37}$ and $W_{10}$. 
CHAPTER FOUR

SATELLITE-BASED OBSERVATIONS OF WHITECAPS USING WIND AND OTHER VARIABLES

4.1 Introduction

Whitecap coverage can be estimated using a power law function and $U_{10EN}$ as the independent parameter. The power law function, as determined in chapter 3, captures the mean whitecap coverage structure, but a power law function using $U_{10EN}$ cannot account for the variability of $W_{10}$ or $W_{37}$. Additional parameters have been suggested as being important in whitecap formation and persistence [see Anguelova and Webster, 2006 and references therein]. This study pairs $U_{10EN}$ and a second independent parameter ($\gamma$) to determine the role of $\gamma$ in estimating whitecap coverage (both $W_{10}$ and $W_{37}$) using a modified power law equation (C1). The overall performances of 17 $U_{10EN}$–$\gamma$ pairs are considered.

4.2 Data

In addition to the parameters explored in chapter 3, the Whitecap Database also includes other gridded parameters useful for investigating factors contributing to $W_{37}$ and $W_{10}$. Parameters from GDAS include SST and $T_{air}$, and a portion of the $U_{10EN}$ data. Wind direction from the $U_{10EN}$ sources is also included. $U_{10EN}$ can be converted into u and v components. The significant wave height ($Hs$) and the significant wave period ($Tp$) from a Wave Watch III model are included as part of the data set. Vectors relating to $Hs$ and $Tp$ are not available; $Hs$ and $Tp$ are magnitude
only. Latitude (Lat), Longitude (Lon), and day of the year (Date) are included in the Whitecap Database.

Bathymetry is taken from the Hybrid Coordinate Ocean Model (HYCOM). The native 1/25° resolution is regridded onto the coarser 0.5° x 0.5° resolution of the Whitecap Database.

Other parameters are derived from the previously mentioned parameters. The parameters for this study are chosen because (1) they were previously suggested [Monahan and O’Muircheartaigh, 1986; Stramska and Petelski, 2003; Anguelova and Webster, 2006 and references therein], (2) the parameter could affect the wave dynamics or the stress, or (3) they highlight possible regional and seasonal influences. These parameters are also chosen because of the availability of the data. The independent parameters, independent of whitecap coverage, not each other, included in this study are SST, T_{air}, Lon, Bath, dT, Wave, Date, Season, Fetch, U_{orb}, dSST_{x}, dSST_{y}, dSST_{xy}, U_{10ENDSST}, NH, SH, and Slope. Only U_{10ENDSST} is uses of U_{10EN} specifically in calculations, and many geophysical parameters are closely related to or previously calculated with U_{10EN}, but, for convenience, all parameters are considered independent. Further details regarding calculation of these parameters and the ranges over which they are valid are found in Appendix B. A short summary of each variable follows:

- **SST** – The sea surface temperature from the GDAS first 10 cm of the water. Suggested in the literature. SST directly influences water density and viscosity and indirectly influences animal and plant life and potential sources of biological surfactants.

- **T_{air}** – The 2 m air temperature. Suggested in the literature. T_{air} is inversely proportional to air density; air density is proportional to stress, so stress is inversely proportional to T_{air}. 
Additionally, $T_{\text{air}}$ is related to the air viscosity and has a general latitudinal dependence (i.e., $T_{\text{air}}$ increases at lower latitudes).

- **Lon** – The longitude. Included to highlight possible regional influences. Lon identifies water bodies such as the Atlantic, Pacific, and Indian oceans.

- **Bath** – Included because of wave dynamics. Bathymetry influences shallow and deep water waves.

- **$dT$** – The air–sea temperature difference. Suggested in the literature. As a proxy for stability, the air–sea temperature differential affects the vertical wind profile. Additionally, unstable conditions are more likely to lead to whitecapping at lower wind speeds.

- **Wave** – The non-dimensional wave height. Included because of wave dynamics. The non-dimensional wave height aids to identify potential wave breaking from ocean dynamics and not just cause by wind.

- **Date** – Day of the year, daily resolution. Included to highlight possible seasonal influences. The day of the year indicates distance from the sun and, depending on the hemisphere, the seasons. Date would also indicate events on the monthly time scale.

- **Season** – The seasonally adjusted date, daily resolution. Included to highlight possible seasonal influences. The seasonally adjusted date indicates the season of the year, which relates to $T_{\text{air}}, \text{SST}$, and seasonal shifts in global circulation patterns and wind events.

- **Fetch** – The fetch of the wind. Suggested in the literature. Limited Fetch is related to less developed seas and to fewer whitecaps.
• $U_{\text{orb}}$ – The magnitude of the orbital velocity. Included because of stress. $U_{10EN}$ is considered surface relative but does not account for $U_{\text{orb}}$; $U_{\text{orb}}$ is arguably [Bourassa, 2006] necessary to accurately relate $U_{10EN}$ to stress.

• $dSST_x$ – The longitudinal $SST$ gradient. Included because of stress. The $SST$ gradient can affect viscosity and stability, both related to stress. The longitudinal gradient accounts for zonal changes.

• $dSST_y$ – The latitudinal $SST$ gradient. Included because of stress. The $SST$ gradient can affect viscosity and stability, both related to stress. The latitudinal gradient accounts for meridional changes.

• $dSST_{xy}$ – The magnitude of the 2-D $SST$ gradient. Included because of stress. The $SST$ gradient can affect viscosity and stability, both related to stress. The magnitude of the gradient accounts for regional changes.

• $U_{10EN}dSST$ – The along $SST$ gradient wind (dot product of $U_{10EN}$ and the 2-D $SST$ gradient). Included because of stress. Wind changes speed and direction moving from one body of water to another. The vertical wind profile in the boundary layer changes, and $dT$ and the orientation of the wind to the waves can change. The stress is modified as a result of these changes.

• $NH$ and $SH$ – The Northern Hemisphere and Southern Hemisphere latitude, respectively. Divided for individual investigations. Included to highlight possible regional influences. $SST$ decreases with increasing latitude, in general. $NH$ and $SH$ might be used as a proxy for $SST$. 


• **Slope** – The slope of the wave assuming deep water waves. Included because of wave dynamics. Steepening waves are more likely to break. The slope of the wave is used to try to identify the breaking point.

### 4.3 Methodology

Whitecap coverage can be calculated with more than one independent parameter, such as $U_{10EN}$, used in chapter 3. It is not clear, however, whether including a second independent parameter ($\gamma$) would improve whitecap coverage estimates. For a preliminary investigation, $U_{10EN}$ is paired with $\gamma$ and a PDF analysis is performed for each $U_{10EN}$–$\gamma$ pair. A 2-D binning is performed using $U_{10EN}$ and $\gamma$. The intervals of the bins are as indicated in Tables B.1 and B.2. The PDF for the whitecap coverage in each of the 2-D bins is calculated. For a constant $U_{10EN}$ interval, changes in the PDFs of the whitecap coverage are observed. Shifting or broadening of the PDFs within the same $U_{10EN}$ bin indicates a possible dependence of $\gamma$ on the whitecap coverage. This analysis is performed for $W_{37}$ and $W_{10}$ over the ranges for each $U_{10EN}$–$\gamma$ pair according to Tables B.1 and B.2. The shifting or broadening of the PDFs only indicates possible dependence, not how $U_{10EN}$ and $\gamma$ together relate to the whitecap coverage.

To identify possible dependence of whitecap coverage on a parameter, $\gamma$, functions for $W_{37}$ and $W_{10}$ are derived for two independent variables [i.e., $W_{37}(U_{10EN}, \gamma)$ and $W_{10}(U_{10EN}, \gamma)$]. A modified power law is used (C1). The functions $\alpha(\gamma)$ and $\beta(\gamma)$ are allowed to take on the form of a constant, linear, quadratic, or cubic function.
Because of the size and distribution of the data, the data are sampled for two independent and one dependent parameters according to Appendix A, creating a training data set for each $U_{10EN}$–$\gamma$ pair. The first independent variable in every case is $U_{10EN}$. The other independent variables are paired with $U_{10EN}$ one at a time and act as $\gamma$ in the analysis. The ranges used for $\gamma$ are contained in Tables B.1 and B.2. The number of samples enforced for each bin is $N=30$. As before, sampling is performed 20 times for each $U_{10EN}$–$\gamma$ pair to ensure repeatability and to reduce biases due to the distribution of the data.

For a single $U_{10EN}$–$\gamma$ pair, the sampled data are fit to (C1), resulting in four functions of $\alpha(\gamma)$ and four functions of $\beta(\gamma)$, for 16 possible combinations. Extensive details for fitting each function are included in Appendix C. The validation data set, the remainder of the unsampled data, is applied to all the combinations, which then receive an ordinal rank (see Appendix D). The best performing combinations of $\alpha(\gamma)$ and $\beta(\gamma)$ are reserved for further comparisons. This ranking is done for both $W_{37}$ and $W_{10}$ for $r_{22EN}$ and $r_{18EN}$. The process is repeated for all $U_{10EN}$–$\gamma$ pairs so that each pair has one modified power law equation for $W_{37} (U_{10EN}; \gamma)$ and $W_{10} (U_{10EN}; \gamma)$ over the ranges $r_{22EN}$ and $r_{18EN}$.

The remaining equations, each representing one $U_{10EN}$–$\gamma$ pair, and the corresponding $U_{10EN}$ only equation from chapter 3 receive an ordinal rank once more following the procedure in appendix D. The ranks are compared to the $U_{10EN}$ rank to indicate dependencies of the whitecap coverage on $\gamma$. 

45
4.4 Results

The PDF analysis of whitecap coverage using two independent variables indicates possible
dependence of whitecaps on those two variables. For $W_{37}$ with $r_{22EN}$ and $T_{air}$ in the 11–13 ms$^{-1}$
$U_{10EN}$ range, the PDFs shift to lower $W_{37}$ with increasing $T_{air}$ (Figure 16); $T_{air}$ is inversely
proportional stress, so the shifting of the PDFs to lower $W_{37}$ corresponds to decreasing stress.

This signal has been identified in scatterometer winds [May and Bourassa, 2010]. A similar shift
in the PDFs is observed over all wind speed ranges (not shown). This shift in PDFs indicates that
$W_{37}$ might have a dependency on $T_{air}$ as well as on $U_{10EN}$. This same test is performed for all
combinations of $W_{37}$ and $W_{10}$ for $r_{22EN}$ and $r_{18EN}$ for all 17 independent variables. Those
combinations that show a shift or broadening of the PDFs require further investigation. Of the
original 17 variables, 9 remain for the $r_{18EN}$ cases and 8 remain for the $r_{22EN}$ cases. Fetch for $r_{22EN}$
cannot be analyzed because of binning requirements.

![Figure 16](image.png)

**Figure 16.** PDFs of $W_{37}$ for 11 < $U_{10EN}$ < 13 ms$^{-1}$ and $T_{air}$ in 2°C bins from -6 to 22°C.
The number of independent variables to investigate is reduced to create a “short list.” Although 

\( NH \) and \( SH \) met the requirements for the short list, the limited observations of \( U_{10EN} > 15 \text{ ms}^{-1} \) at 

lower latitudes do not meet the continuity requirements necessary for this global study. The 

remaining parameters meet the continuity requirements as described in Appendix A and are 

included in the short list.

Each \( U_{10EN} - \gamma \) pair from the short list is sampled and is analyzed independently to determine the 

best functional form for power law coefficients, which depend on \( \gamma \). The coefficient functions 

depend only on \( \gamma \) and not on each other. The 16 possible coefficient pairs of \( \alpha(\gamma) \) and \( \beta(\gamma) \) are 

ranked to indicate the best performing coefficient pair for each \( U_{10EN} - \gamma \) pair for the \( W_{37} \) and \( W_{10} \) 

over \( r_{18EN} \) and \( r_{22EN} \).
For the case of \( \alpha(SST) \) and \( \beta(SST) \) with \( W_{37, r_{18EN}} \), if \( \alpha \) and \( \beta \) were taken to be constants, there would be no dependence of whitecap coverage on \( SST \), and (C1) would become (1). The different functions capture different features of the functions \( \alpha(SST) \) and \( \beta(SST) \) (see Figure C1). A visual inspection indicates a possible \( SST \) dependence, but a visual inspection alone does not indicate the best functional fit and cannot indicate the quality of fit of the functions to (C1). A similar dependence is found for the other independent variables (not shown). The best combination of \( \alpha(\gamma) \) and \( \beta(\gamma) \) pairs is found through a ranking. For \( SST \), the best combinations of \( \alpha(SST) \) and \( \beta(SST) \) with \( W_{37, r_{18EN}} \) are a cubic and a quadratic function, respectively. The best coefficient function pairs for the entire short list are determined for \( \alpha(\gamma) \) and \( \beta(\gamma) \) (Table 6).

### Table 6. Orders of functions for coefficient functions \( \alpha(\gamma) \) and \( \beta(\gamma) \). Cubic is represented by 3, quadratic by 2, linear by 1, and constant by 0.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>( W_{10} )</th>
<th>( r_{18EN} )</th>
<th>( r_{22EN} )</th>
<th>( W_{37} )</th>
<th>( r_{18EN} )</th>
<th>( r_{22EN} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( SST )</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>( T_{air} )</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>dT</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Wave</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fetch</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( U_{orb} )</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>( dSST_{y} )</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>( U_{10EN}dSST )</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Slope</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

The short list of the 17 original \( U_{10EN}–\gamma \) pairs shows a potential coefficient dependence on \( \gamma \) for all coefficient pairs except \( Wave \) for \( W_{37} \) and \( r_{10EN} \). The best coefficients for each \( U_{10EN}–\gamma \) pair with \( U_{10EN} \) from chapter 3 are ranked in the union of their valid ranges using the validation data.
set. The short list of equations and $U_{10EN}$ are ranked for the values within the ranges over which the functions are valid. For three of the cases, 31% of the total whitecap observations are included in the analysis; for the case $W_{37}$, only 29% of the total whitecap observations are included. The functions are ranked against one another starting with 1 as the best overall performing function, and so on (Table 7).

**Table 7. Final rankings for $U_{10EN}$ pairs and $U_{10EN}$.**

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>$W_{10}$</th>
<th>$W_{37}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{10EN}$</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>SST</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$r_{18EN}$</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>$r_{22EN}$</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>$dSST$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$dT$</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Wave</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>$F_t$</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$U_{orb}$</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>$U_{10EN}dSST$</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>$dSST_y$</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Slope</td>
<td>10</td>
<td>5</td>
</tr>
</tbody>
</table>

A ranking better than $U_{10EN}$ indicates that the $U_{10EN}$ function captures the mean, variance, or both, better than $U_{10EN}$ alone. Different physical mechanisms are likely responsible for the formation of $W_{10}$ and the persistence of $W_{37}$, so different rankings between $W_{37}$ and $W_{10}$ are expected. The ranges, $r_{18EN}$ and $r_{22EN}$, also capture different characteristics of the parameters; $r_{18EN}$ captures larger ranges of most non-$U_{10EN}$ parameters (see Figure A5 for SST), and $r_{22EN}$ captures higher wind speeds but reduces ranges of most non-$U_{10EN}$ parameters. Different rankings are also expected. Ranking indicates the order in which parameters should be explored.
for possible dependence; the normalized performance before ranking shows how each parameter compares to the others (Figure 17).

**Figure 17.** Performance of the short list of parameters and $U_{10EN}$ for all cases. Lower values indicate better performance.

For $W_{10r18EN}$, $T_{air}$ ranks the best overall, followed by $U_{10EN}dSST$; $SST$, $Fetch$, and $U_{orb}$ tie for third; seven functions rank better than $U_{10EN}$; $dT$ and $Slope$ rank worst. For $W_{10r22EN}$, $U_{orb}$ and $dSST_y$ tie for the best rank followed closely by $SST$. Again, only $dT$ and $Slope$ rank worse than $U_{10EN}$. Since $U_{10EN}$ is an equivalent neutral wind and already accounts for stability [Kara, 2006],
$dT$ ranking worse than $U_{10EN}$ is understandable and supports the theory that the surface stress, not the wind, is the important formation mechanism for actively breaking whitecaps.

Overall for $W_{10}$, $U_{orb}$, SST, $T_{air}$, and $U_{10EN}dSST$ rank the best. The exact reasons for the performance of these additional parameters are not determined in this study, but some possible reasons for the improved performance over $U_{10EN}$ alone are provided here. For $U_{orb}$, $U_{10EN}$ is considered a surface relative wind, but does not account for the orbital velocity of the ocean surface. Since $U_{orb}$ in this study does not have a vector direction for comparison with the $U_{10EN}$ components, specifics regarding the exact changes to the surface relative wind are unknown, but differences of two ms$^{-1}$ for the surface relative wind are not unreasonable. Additionally, $U_{orb}$ is a parameterized value and does not directly adjust $U_{10EN}$ values. Still, $U_{orb}$ improves the $W_{10}$ estimate the most.

The parameters SST and $T_{air}$ might affect density or viscosity of the water or air. These parameters appear to affect $W_{10}$ and have been suggested before in previous in situ studies [e.g. Monahan and O’Muircheartaigh, 1986; Stramska and Petelski, 2003]. A $W_{10}$ dependence on SST is expected since SST is included in the forward model used to calculate $W_{10}$, but the value of this parameter in estimating $W_{10}$ cannot be ignored. A change in the ranges of $T_{air}$ from the $r_{18EN}$ to the $r_{22EN}$ cases might provide some explanation regarding the drop in overall rank; $T_{air}$ might have an increased influence on $W_{10}$ at higher air temperatures, but this could also be a residual of normally lower wind speeds near the equator where values of $T_{air}$ are high. Alternatively, the air density, closely related to $T_{air}$ and essential in converting from stress to $U_{10EN}$, could account for errors in converting stress to $U_{10EN}$ from satellite observations [May and Bourassa, 2010].
The wind blowing along the SST gradient, $U_{10\text{EN}}dSST$ is not actually independent of $U_{10\text{EN}}$, but the $U_{10\text{EN}}-U_{10\text{EN}}dSST$ pair ranks well for both cases of $r\text{E}_{18\text{EN}}$ and $r\text{E}_{22\text{EN}}$. Discrepancies are observed in $u10$, $U_{10\text{EN}}$, and surface stress across SST fronts [Small et al., 2008; Sweet et al., 1981; Hayes et al., 1989; Liu et al., 2000, 2007; Chelton et al., 2001, 2004, 2007; O’Neill et al., 2003, 2005, 2010, 2012; Tokinaga et al., 2005]. Multiple mechanisms have been suggested for the physical processes that occur when wind crosses a SST front. Positive $U_{10\text{EN}}dSST$, wind moving from cold to warm SST, might induce more momentum transfer from the wind to the water on the warmer side [Sweet et al., 1981; Hayes et al., 1989; Wallace et al., 1989; Liu et al., 2000; Tokinaga et al., 2005; Spall et al., 2007] or produce small-scale pressure gradients that would accelerate the wind [Small et al., 2003, 2005; Song et al., 2006; Spall et al., 2007; O’Neill et al., 2010]. The wind direction would turn and a thermal wind would be created going from cold to warm water. The changes in the total wind would change the vertical profile of the wind and the stress. A specific physical process for the $U_{10\text{EN}}-U_{10\text{EN}}dSST$ pair is not identified in this study, but this pair performs consistently well.

Other parameters combined with $U_{10\text{EN}}$ perform better than $U_{10\text{EN}}$ alone also. Wave slightly improves performance. Fetch ties for third for $r\text{E}_{18\text{EN}}$, but $r\text{E}_{22\text{EN}}$ is not available because of restrictions in ranges with higher wind speeds. For $W_{10}$, $U_{\text{orb}}$, SST, $T_{\text{air}}$, and $U_{10\text{EN}}dSST$ perform best overall, suggesting a possible dependence of $W_{10}$ on these parameters.

For $W_{37\text{EN}}$ $r_{18\text{EN}}$, $T_{\text{air}}$ ranks first followed closely by SST. For $W_{37\text{EN}}$ $r_{22\text{EN}}$, $U_{10\text{EN}}dSST$ ranks the best, strongly outperforming $dT$ and Slope, the next closest ranking variables. For both cases of $W_{37\text{EN}}$,
U_{orb} and dSST, perform notably worse than U_{10EN}. The change in ranges of r_{18EN} to r_{22EN} change T_{air} from ranking first to worse than U_{10EN} alone; this change indicates a potential dependence of W_{37} on the T_{air} range or a weak dependence of W_{37} on T_{air} at higher wind speeds.

The parameters U_{10EN}dSST, dT, and SST rank best overall for W_{37}. Since W_{37} represents the total whitecap coverage, which includes the residual field, these parameters might affect the persistence of the residual field. Only dT from this list is excluded from the possible W_{10} dependent parameters, indicating that dT might play a role in the persistence of whitecaps beyond the influence of dT to the wind at the surface. The inclusion of dT as a possible dependent parameter for W_{37} supports previous research efforts for determining W_{t} using U_{10} and dT, though possibly for different theoretical reasons.

Overall, whitecap coverage for both W_{37} and W_{10} depends most on SST and U_{10EN}dSST, but T_{air} should be considered for the r_{18EN} range. Neither SST nor U_{10EN}dSST is completely independent of U_{10EN}, but their performance and ranking always surpass U_{10EN} alone. They are the only parameters that outperform U_{10EN} alone for all cases. For W_{10}, U_{orb} should also be considered; for W_{37}, dT should also be considered.

4.5 Conclusions

Modified power-law functions using two independent variables can estimate W_{37} and W_{10} better than power law functions using U_{10EN} alone. The modified power law function more closely captures the mean value and/or reduces the variance of the whitecaps compared to a power law function. Different parameters are shown to be important for estimating W_{10} and W_{37}. These
parameters indicate physical mechanisms possibly important to whitecap formation and persistence. The parameters that show notable improvement over $U_{10EN}$ alone and that should be considered are as follows:

- For $W_{10}$ — $U_{orb}$, SST, and $U_{10EN}dSST$, $T_{air}$ and Fetch for $U_{10EN} < 18$ ms$^{-1}$
- For $W_{37}$ — SST, $U_{10EN}dSST$, and $dT$. $T_{air}$ for $U_{10EN} < 18$ m s$^{-1}$
- For both $W_{10}$ and $W_{37}$ — SST and $U_{10EN}dSST$. $T_{air}$ for $U_{10EN} < 18$ ms$^{-1}$

The parameter $U_{10EN}dSST$ is not independent of $U_{10EN}$, but the $U_{10EN}$—$dSST$ pair consistently ranks better than $U_{10EN}$ alone. The forward model uses SST and $U_{10EN}$ in determining $W_{10}$ and $W_{37}$ from the surface emissivity observed by WindSat. Only $T_{air}$ is independent of the forward model since $dT$ uses SST, but $T_{air}$ might have an unknown dependence on SST from GDAS. Regardless of the relationship to the forward model, these parameters outperform $U_{10EN}$ and should be considered.

This study is not intended to exhaust the possible parameters affecting whitecaps nor is it intended to explain physical processes behind formation and persistence of whitecaps. Rather, this study provides parameters potentially important in estimating the local mean whitecap coverage using global satellite winds. This study also provides a benchmark for parameters that should be considered and parameters that can be eliminated from further consideration. Those parameters available from satellite platforms can be used to produce estimates of $W_{10}$ and $W_{37}$ globally independent of model or in situ observations.
Whitecap coverage estimates using available satellite products such as $U_{10EN}$, $T_{air}$, and $SST$ can be used to estimate $W_{10}$ and $W_{37}$ from the Whitecap Database. This independent method allows for calculations of whitecaps globally and provides a research tool for estimating whitecaps and their contributions to other research areas such as CO$_2$ exchange, surface albedo, turbulent fluxes, and satellite surface measurements. Consistent, verified, and repeatable results show whitecap coverage can be estimated and other parameters besides $U_{10EN}$ might play a role in the formation of and persistence in whitecap coverage.
CHAPTER FIVE

CONCLUSIONS

The objective of this study was to determine equations appropriate for estimating the whitecap coverage globally and to identify the major contributing factors to whitecap formation and duration using currently available satellite products. First, appropriate power law coefficients were determined using available in situ (visible) whitecap observations with $U_{10}$ and $U_{10EN}$. These new functions were compared to global satellite-based observations of microwave (10 GHz and 37 GHz) local whitecaps and $U_{10EN}$ to determine consistency. Second, appropriate power law coefficients were determined to estimate $W_{10}$ and $W_{37}$ using $U_{10EN}$ to show that whitecap coverage can be estimated using satellite winds. Finally, 17 $U_{10EN}$–$\gamma$ pairs are analyzed to determine the role of non-wind variables ($\gamma$) in estimating whitecap coverage (both $W_{10}$ and $W_{37}$) using a modified power law equation. The most influential parameters in whitecap formation and duration are identified.

First, a notable difference was shown between $U_{10}$ and $U_{10EN}$. This difference and resulting potential error in whitecap coverage were accentuated when the winds were applied to a power law. New coefficients for the power law equation using $U_{10EN}$ were derived using in situ observations. Although the new equations using $U_{10EN}$ improved whitecap estimates, they still failed to reasonably estimate the whitecap coverage from the Whitecap Database, strongly suggesting that the different techniques for analyzing in situ whitecap observations are a larger source of error than the change in wind type. A standardization or metric to equate the multiple analysis techniques is necessary to definitively equate in situ to satellite-based whitecap
observations. Otherwise, only a single technique with sufficient observations to statistically represent global whitecap coverage should be used for determining a function appropriate for estimating global whitecap coverage.

Second, the Whitecap Database provided satellite-based global whitecap coverage from a single platform using a single analysis technique to provide over $1.8 \times 10^7$ observations matched with $U_{10EN}$ values for $W_{10}$ and $W_{37}$. Estimates of whitecap coverage calculated with a power law closely followed the mean whitecap coverage from the Whitecap Database for $U_{10EN}$ up to 22 ms$^{-1}$. The power law alone could not capture the variance of the data and further improvements to capturing the mean of the data were not possible with this function of a single variable.

Finally, a modified power law using $U_{10EN}$ and each of 17 different parameters was tested to account for the variance and to capture the mean of the data better. A PDF analysis eliminated many of the parameters, leaving a short list of 8 to 9 parameters to compare against $U_{10EN}$ alone. The parameters in the short list were ranked according to their ability to estimate whitecap coverage. The best performing and most consistent parameters for estimating whitecap coverage were noted as follows:

- $W_{10} - U_{orb}$, SST, and $U_{10EN}dSST$. $T_{air}$ and $Fetch$ for $U_{10EN} < 18$ ms$^{-1}$
- $W_{37} - SST$, $U_{10EN}dSST$, and $dT$. $T_{air}$ for $U_{10EN} < 18$ ms$^{-1}$
- $W_{10}$ and $W_{37} - SST$ and $U_{10EN}dSST$. $T_{air}$ for $U_{10EN} < 18$ ms$^{-1}$
A modified power law using these parameters provides an improvement over a standard power law using $U_{10EN}$ alone. Of the available satellite parameters, $U_{10EN}$ is the single most important factor to determining whitecap coverage. These additional parameters are identified as secondary contributing factors. However, further investigation of the exact effects of each of these parameters and the physics of whitecap formation and duration they contribute to is beyond the scope of this study. These findings are consistent with the literature, which indicates SST, $T_{air}$, $dT$, and Fetch likely play a role in whitecap formation and duration. The parameter $U_{10EN}dSST$, presented for the first time in this study, shows merit and should be included in further studies.

The techniques used to manage large amounts of unevenly distributed data reduced both bias and computational requirements. The sampling, function fitting, and ranking were shown to be repeatable. Equations derived were verified against all the data available within the union of the individual parameter’s ranges. All parameters were considered equally with no a priori knowledge of the results.

The overall behavior of the power law functions for both the active and the total whitecap coverage estimates using the Whitecap Database did not follow a cubic. The whitecap coverage from the Whitecap Database, calculated using surface emissivity in the microwave channels observed by WindSat, would have been largely overestimated by the power law equation presented by Monahan [1980]. Similar problems are noted in CZCS, SeaWiFS, and MODIS, in the visible channel, though the surface reflective processes are different; it is possible that whitecap coverage does not follow a cube of the wind. The focus on finding a cubic function should be replaced with efforts to determine the coefficient pair for the power law. Current
publicly available in situ whitecap observations with corresponding \( U_{10EN} \) values are insufficient to make this determination with in situ observations alone.

*Blanchard* [1963, 1983] estimated the global whitecap coverage to be 1–4%. The Whitecap Database roughly agrees with this estimate. The global whitecap coverage for the active whitecap coverage is 0.62% and for the total whitecap coverage, 1.13% (Figures 18 and 19, respectively).

In the future, as additional parameters become available, this work could be extended to other physical factors such as currents, vector orbital velocity, salinity, and surfactants. This research could be used to improve surface and near-surface satellite retrievals (e.g., to correct ocean color). The findings could also be applied to areas such as turbulent flux parameterizations, atmosphere–ocean \( \text{CO}_2 \) exchange, wave energy dissipation from breaking waves, ocean acoustics, and ocean topography changes.

This study builds upon the invaluable theories, methodologies, and observations from the research community. Further understanding and calibration between methods to determine whitecap coverage are necessary. The future of this work depends on the community coming together to work on understanding the differences in the whitecap estimates for in situ and remote-sensed observations, making all the previous efforts truly useful. By working together with different theories, backgrounds, and ideas, scientists can further understand whitecap formation and duration.
Figure 18. Mean active whitecap coverage ($W_{10}$) for 2006 from Whitecap Database. Whitecap coverage is in percent.

Figure 19. Mean total whitecap coverage ($W_{37}$) for 2006 from Whitecap Database. Whitecap coverage is in percent.
APPENDIX A

DATA SAMPLING

Processing large amounts of data with uneven distributions can create issues in biasing results and computational power and time. These issues are not unique to this study. A common technique to address these issues is data sampling. Sampling data to create a working data set from a much larger data set is used frequently in working with remotely sensed data. Sampling can reduce bias in the results and greatly reduce computational requirements, while still maintaining many of the desired statistical properties even with repeated use. An additional benefit to the sampling is that the data not chosen for the sample can be used as a separate validation set. The sampling technique presented follows closely to those used by Roberts et al. [2010] but is presented with more detail.

In situ observations of whitecaps are generally sparse globally, fragmented in time, and they are often parameterized in terms of some physical parameter such as $U_{10EN}$, SST, or $T_{air}$. Satellite-based observations provide near daily global coverage. The Whitecap Database, is described in chapters 2 - 4, contains satellite-based observations of whitecap coverage with other corresponding geophysical parameters. This data sampling technique is applied to manage and interpret the Whitecap database’s more than 19 million whitecap observations with corresponding geophysical parameters for 2006. Specifically, to address sampling this large amount of data and multiple parameters or parameters, details of the data sampling technique are provided for sampling two parameters and for sampling three parameters.
The goal of this data sampling technique is to obtain a sample of the data from a larger pool of data maintaining the same statistical mean and variance for both the sample and the larger pool of data. When applied properly, this method reduces biases due to unevenly distributed data, a common geophysical problem. This smaller sample will maintain the same statistical characteristics of mean and variance while using a much smaller number of the available data points.

The data are considered to have one dependent parameter and one or more independent parameters. The ranges of the independent parameter(s) must be continuous. Each of the independent parameters is divided up into evenly spaced bins based on the range over which they are valid to meet or exceed the requirements as suggested by Panofsky and Brier [1953]. The ranges over which the data are considered valid are determined by physical limitations or instrumentation to limit non-physical results. In the case of two independent parameters, the data is binned by both parameters in even bins for the ranges over which each parameter is valid. A constant number of values, $N$, is sampled from each bin such that the sampled values represent the mean and variance of their source bin. To insure repeatability and validity of the sampling, these rules for sampling must be followed. The specific requirements for the dependent and independent parameters for this technique are listed below.

For the independent parameter(s)

1. The mean and variance of sampled data and the larger pool of data must be the same for each independent parameter within a given tolerance, 5% for this work.
2. Bins of data must be continuous over the range of interest.
3. Data must be considered valid.

For the dependent parameter

1. The mean and variance of sampled data and the larger pool of data must be the same for each dependent parameter within a given tolerance, 5% for this work.

2. Data must be considered valid.

Flow charts are provided for the one independent parameter and two independent parameter cases (Figures A1 and A2, respectively).

An example of how to use this process for one independent parameter $W(U_{10EN})$ shows the binning of the $U_{10EN}$ data (Figure A3) and the location of the samples with respect to the entirety of the binned data (Figure A4).

An example of how to use this process for two independent parameters $W(U_{10EN}, SST)$ shows the binning of the $U_{10EN}$ and SST data (Figure A5) and the location of the samples with respect to the entirety of the binned data (Figure A6).
Flow chart for one independent variable

Determine range of valid, continuous independent variable

Bin data into evenly spaced bins

Select N samples from each bin maintaining mean and variance for both independent and dependent variables

Figure A1. Flow chart for sampling one independent and one dependent parameters.

Flow chart for two independent variables

Determine ranges of valid, continuous independent variables

Bin data into evenly spaced bins for both independent variables

Adjust usable ranges to ensure that each bin has at least N values for sampling

Select N samples from each bin maintaining mean and variance for both independent and dependent variables

Figure A2. Flow chart for sampling two independent and one dependent parameters.
Figure A3. Histogram of $U_{10EN}$ data from Whitecap Database in 1–ms$^{-1}$ bins.

Figure A4. Density plot of $U_{10EN}$ and $W_t$ from Whitecap Database. Wind colors of density on a log scale. Black dots represent the 30 samples taken from each 1–ms$^{-1}$ bin; the samples maintain the same statistical mean and variance of all the values in their respective bins.
**Figure A5.** A density plot of observations by bins of SST and $U_{10EN}$. Solid grey line shows range of values for which each bin contains at least 30 data points for sampling with a maximum $U_{10}$ of 22 ms$^{-1}$. Dashed black line shows range of values for which each bin contains at least 30 data points for sampling with a maximum $U_{10EN}$ of 18 ms$^{-1}$.

**Figure A6.** A larger view of the bin from Figure A5 with the ranges 16°C < SST < 18°C and 17 ms$^{-1}$ < $U_{10EN}$ < 18 ms$^{-1}$ with black dots representing the observations available and the white squares representing the 30 sampled values which maintained the same statistical mean and variance as all the parameters contained within the bin.
APPENDIX B

PARAMETER NAMES, RANGES, AND FORMULAS

The following are the formulas for the calculated parameters:

\[ dT = T_{\text{air}} - SST \quad (B1) \]

\[ \text{Wave} = \frac{H_s}{B_{\text{ath}}} \quad (B2) \]

\[ \text{Season} \quad \text{The Southern Hemisphere date shifted by 6 months} \]

\[ \text{Fetch} = 9.8 \left( \frac{T_p}{U_{\text{igen}}} \right)^2 \quad (B3) \]
\( dSST_x \) – The longitudinal gradient of SST

\[
dSST_x = \nabla_{\text{Lon}} SST
\]  

(B4)

\( dSST_y \) – The latitudinal gradient of SST

\[
dSST_y = \nabla_{\text{Lat}} SST
\]  

(B5)

\( dSST_{xy} \) – The magnitude of the 2-D SST gradient

\[
dSST_{xy} = \sqrt{(dSST_x)^2 + (dSST_y)^2}
\]  

(B6)

\( U_{10END}dSST \) – The along SST gradient wind

\[
U_{10END}dSST = \vec{U}_{10END} \cdot \nabla^2 SST
\]  

(B7)

Slope – The slope of the wave

\[
\text{Slope} = \frac{H_s}{1.56 \times T_p^2}
\]  

(B8)
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Uses</th>
<th>Units</th>
<th>Low</th>
<th>High</th>
<th>Increment</th>
<th>Low</th>
<th>High</th>
<th>Increment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{10EN}$</td>
<td>$U_{10EN}$</td>
<td>ms$^{-1}$</td>
<td>3</td>
<td>18</td>
<td>1</td>
<td>3</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>SST</td>
<td>SST</td>
<td>°C</td>
<td>0</td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>$T_{air}$</td>
<td>$T_{air}$</td>
<td>°C</td>
<td>-8</td>
<td>30</td>
<td>2</td>
<td>-6</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Lon</td>
<td>Longitude</td>
<td>Degrees</td>
<td>0</td>
<td>360</td>
<td>0.5</td>
<td>0</td>
<td>360</td>
<td>0.5</td>
</tr>
<tr>
<td>Lat</td>
<td>Latitude</td>
<td>Degrees</td>
<td>-69.5</td>
<td>70.5</td>
<td>0.5</td>
<td>-69.5</td>
<td>70.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Bath</td>
<td>Bathymetry</td>
<td>m</td>
<td>-6400</td>
<td>0</td>
<td>100</td>
<td>-6200</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Hs</td>
<td>Hs</td>
<td>m</td>
<td>0</td>
<td>14.9</td>
<td>N/A</td>
<td>0</td>
<td>14.9</td>
<td>N/A</td>
</tr>
<tr>
<td>Tp</td>
<td>Tp</td>
<td>s</td>
<td>0</td>
<td>21.5</td>
<td>N/A</td>
<td>0</td>
<td>21.5</td>
<td>N/A</td>
</tr>
<tr>
<td>$dT$</td>
<td>$T_{air}$, SST</td>
<td>°C</td>
<td>-10</td>
<td>5</td>
<td>1</td>
<td>-8</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Wave</td>
<td>Hs, Bath</td>
<td>unitless</td>
<td>-1.28E-2</td>
<td>-4.00E-3</td>
<td>2.00E-4</td>
<td>-5.60E-3</td>
<td>-4.00E-3</td>
<td>2.00E-4</td>
</tr>
<tr>
<td>Date</td>
<td>Day of year</td>
<td>Day</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
</tr>
<tr>
<td>Season</td>
<td>Day of Year</td>
<td>Day</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
</tr>
<tr>
<td>Fetch</td>
<td>Tp, $U_{10EN}$, gravity</td>
<td>m</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>$U_{orb}$</td>
<td>$Hs$, $Tp$</td>
<td>ms$^{-1}$</td>
<td>0.5</td>
<td>1.75</td>
<td>0.25</td>
<td>0.75</td>
<td>1.75</td>
<td>0.25</td>
</tr>
<tr>
<td>$dSST_x$</td>
<td>$SST$, Lon</td>
<td>°C/m</td>
<td>-2.00E-4</td>
<td>2.1E-4</td>
<td>1.00E-5</td>
<td>-9.00E-5</td>
<td>9.00E-5</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$dSST_y$</td>
<td>$SST$, Lat</td>
<td>°C/m</td>
<td>-6.00E-5</td>
<td>5.00E-5</td>
<td>1.00E-5</td>
<td>-4.00E-5</td>
<td>5.00E-5</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$dSST_{xy}$</td>
<td>$SST$, Lat, Lon</td>
<td>°C/m</td>
<td>0</td>
<td>3.30E-4</td>
<td>1.00E-5</td>
<td>1.50E-4</td>
<td>1.50E-4</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$U_{10ENDSST}$</td>
<td>$U_{10EN}$, $Udir$, $dSST_x$, $dSST_y$</td>
<td>°C/s</td>
<td>-3.00E-5</td>
<td>3.00E-3</td>
<td>1.00E-5</td>
<td>-3.00E-5</td>
<td>3.00E-3</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>NH</td>
<td>Latitude</td>
<td>Degrees</td>
<td>28.5</td>
<td>70.5</td>
<td>0.5</td>
<td>33</td>
<td>57</td>
<td>0.5</td>
</tr>
<tr>
<td>SH</td>
<td>Latitude</td>
<td>Degrees</td>
<td>-65</td>
<td>-28.5</td>
<td>0.5</td>
<td>-64</td>
<td>-35.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Slope</td>
<td>$Hs$, $Tp$</td>
<td>unitless</td>
<td>0.01</td>
<td>0.05</td>
<td>5.00E-3</td>
<td>1.5E-2</td>
<td>0.05</td>
<td>5.00E-3</td>
</tr>
</tbody>
</table>
Table B2. Parameter names, ranges, and parameters used for the 10 GHz Channel ($W_{10}$) of the Whitecap Database for $r_{18EN}$ and $r_{22EN}$

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Uses</th>
<th>Units</th>
<th>Low</th>
<th>High</th>
<th>Increment</th>
<th>Low</th>
<th>High</th>
<th>Increment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{10EN}$</td>
<td>$U_{10EN}$</td>
<td>ms$^{-1}$</td>
<td>3</td>
<td>18</td>
<td>1</td>
<td>3</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>SST</td>
<td>SST</td>
<td>°C</td>
<td>0</td>
<td>30</td>
<td>2</td>
<td>0</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>$T_{air}$</td>
<td>$T_{air}$</td>
<td>°C</td>
<td>-8</td>
<td>30</td>
<td>2</td>
<td>-6</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Lon</td>
<td>Longitude</td>
<td>Degrees</td>
<td>0</td>
<td>360</td>
<td>0.5</td>
<td>0</td>
<td>360</td>
<td>0.5</td>
</tr>
<tr>
<td>Lat</td>
<td>Latitude</td>
<td>Degrees</td>
<td>-69.5</td>
<td>70.5</td>
<td>0.5</td>
<td>-69.5</td>
<td>70.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Bath</td>
<td>Bathymetry</td>
<td>m</td>
<td>-6400</td>
<td>0</td>
<td>100</td>
<td>-6200</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>$Hs$</td>
<td>$Hs$</td>
<td>m</td>
<td>0</td>
<td>14.9</td>
<td>N/A</td>
<td>0</td>
<td>14.9</td>
<td>N/A</td>
</tr>
<tr>
<td>$Tp$</td>
<td>$Tp$</td>
<td>s</td>
<td>0</td>
<td>21.5</td>
<td>N/A</td>
<td>0</td>
<td>21.5</td>
<td>N/A</td>
</tr>
<tr>
<td>$dT$</td>
<td>$T_{air}, SST$</td>
<td>°C</td>
<td>-10</td>
<td>6</td>
<td>1</td>
<td>-10</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Wave</td>
<td>$Hs$, Bath</td>
<td>unitless</td>
<td>-1.28E-2</td>
<td>-4.00E-3</td>
<td>2.00E-4</td>
<td>-5.60E-3</td>
<td>-4.00E-3</td>
<td>2.00E-4</td>
</tr>
<tr>
<td>Date</td>
<td>Day of year</td>
<td>Day</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
</tr>
<tr>
<td>Season</td>
<td>Day of Year</td>
<td>Day</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
<td>1</td>
<td>366</td>
<td>Month Length</td>
</tr>
<tr>
<td>Fetch</td>
<td>$Tp$, $U_{10EN}$, gravity</td>
<td>m</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>$U_{orb}$</td>
<td>$Hs$, $Tp$</td>
<td>ms$^{-1}$</td>
<td>0.5</td>
<td>1.75</td>
<td>0.25</td>
<td>0.75</td>
<td>1.75</td>
<td>0.25</td>
</tr>
<tr>
<td>$dSST_x$</td>
<td>$SST$, Lon</td>
<td>°C/m</td>
<td>-2.00E-4</td>
<td>2.1E-4</td>
<td>1.00E-5</td>
<td>-9.00E-5</td>
<td>9.00E-5</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$dSST_y$</td>
<td>$SST$, Lat</td>
<td>°C/m</td>
<td>-5.00E-5</td>
<td>5.00E-5</td>
<td>1.00E-5</td>
<td>-4.00E-5</td>
<td>5.00E-5</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$dSST_{xy}$</td>
<td>$SST$, Lat, Lon</td>
<td>°C/m</td>
<td>0</td>
<td>3.30E-4</td>
<td>1.00E-5</td>
<td>0</td>
<td>1.50E-4</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>$U_{10EN}dSST$</td>
<td>$U_{10EN}$, $U_{dir}$, $dSST_x$, $dSST_y$</td>
<td>°C/s</td>
<td>-3.00E-5</td>
<td>3.00E-3</td>
<td>1.00E-5</td>
<td>-3.00E-5</td>
<td>3.00E-3</td>
<td>1.00E-5</td>
</tr>
<tr>
<td>NH</td>
<td>Latitude</td>
<td>Degrees</td>
<td>28.5</td>
<td>70.5</td>
<td>0.5</td>
<td>34</td>
<td>64</td>
<td>0.5</td>
</tr>
<tr>
<td>SH</td>
<td>Latitude</td>
<td>Degrees</td>
<td>-66.5</td>
<td>-28.5</td>
<td>0.5</td>
<td>-64</td>
<td>-35.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Slope</td>
<td>$Hs$, $Tp$</td>
<td>unitless</td>
<td>0.01</td>
<td>0.05</td>
<td>5.00E-3</td>
<td>1.5E-2</td>
<td>0.05</td>
<td>5.00E-3</td>
</tr>
</tbody>
</table>
APPENDIX C

FUNCTION FITTING

The percentage of whitecap coverage ($W$) is most often considered to primarily depend on the wind, and for standardization purposes, the 10m wind above the water surface ($U_{10}$). The majority of the efforts to create an equation to calculate $W$ have focused on a power law equation using $U_{10}$ as the argument starting with Blanchard [1963]. Various efforts have been used to determine additional dependencies in the whitecap field other than $U_{10}$ [Pandey and Kakar, 1982; Monahan and O’Muircheartaigh, 1986; Spillane et al., 1986; Bortokovshii, 1987; Monahan and Woolf, 1989; Monahan, 1993; Villarino et al., 2003; Monahan, 2012] or, for use with satellite-based winds like those used in this study, $U_{10EN}$. Additional terms multiplied a power law, modified the argument of a power law, or modified the coefficients of the power law. While maintaining the basic structure of a power law equation, one method for determining additional dependencies of the whitecap coverage using $U_{10EN}$ and an additional parameter (e.g. $\gamma$) allows for the coefficients of the power law to depend on $\gamma$ (C1). A similar method has shown interesting results using $\gamma = \text{SST}$ (Monahan, 2012), but $\gamma$ can represent any parameter in this method.

$$W = a(\gamma)u_{10}^{b(\gamma)} \quad \text{(C1)}$$

Allowing the coefficients of a power law equation to be a function of a given parameter can provide valuable insight regarding the dependence of whitecap coverage on the parameter and help to determine the influence of the parameter on the actual whitecap coverage. In this appendix, an example is given of the fitting the coefficients of a power law equation using $U_{10EN}$.
as the argument of the power law and an additional parameter as the argument of the coefficient functions. The coefficient functions can be applied and used in further research for predicting whitecap coverage and for determining the dependence of whitecaps on $U_{10EN}$ and other parameters.

The following case is presented as a sample using the Whitecap Database and assigning $\gamma = SST$. This method can be applied to other parameters with the appropriate ranges and constraints described in Appendix B.

The ranges over which the functional fit will be valid are provided through the techniques described in Appendix A. For this case the independent variables of interest are $U_{10EN}$ and $SST$ over the ranges $3 \text{ ms}^{-1} < U_{10EN} < 18 \text{ ms}^{-1}$ and $0 \degree C < SST < 30 \degree C$ with a $SST$ increment step of $2 \degree C$ totaling 15 increments. A total of 20 sets of sampled values are taken from each of the increment steps totaling 15*20 sets of sampled values. The purpose of using 20 sets instead of a single set is to reduce any possible error introduced through the sampling technique (see Appendix A).

The coefficients $a$ and $b$ of (C2) are fit with the first set of sampled values by minimizing the least squared error, the standard statistical approach to fitting coefficients for this equation. The values of $a$ and $b$ are constant and recorded for later use along with the mean $SST$ value of sampled values. This is repeated for each set of the sampled values until 15*20 $a$ and $b$ pair constants with their respective mean $SST$ are found.

$$W = au_{10}^b$$

(C2)
Each of the coefficients, $a$ and $b$, along with their respective mean SST values are then fit to the functions $\alpha(SST)$ and $\beta(SST)$ following equation (C1) (Figure C1) The functions $\alpha$ and $\beta$ are allowed to take on any functional form, but for the purposes of this work, the functional forms included are constant, linear, quadratic, and cubic (Table C1). Previous efforts have also included logarithmic and exponential functions [Monahan, 2012]. Visual inspection of the data dictated the functional forms used here. The functions $\alpha$ and $\beta$ are found independently of one another and can be used in any combination, ie. a constant $\alpha$ can be used with a quadratic $\beta$. The functions $\alpha$ and $\beta$ are valid for the ranges that their arguments are valid based on the sampled values. Extrapolation of these functions over larger ranges is not recommended and highly discouraged.

**Table C1. Coefficients for fitting functions.**

<table>
<thead>
<tr>
<th></th>
<th>$\alpha(\gamma)$</th>
<th>$\beta(\gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>$\alpha = \alpha_1$</td>
<td>$\beta = \beta_1$</td>
</tr>
<tr>
<td>Linear</td>
<td>$\alpha = \alpha_1 \gamma + \alpha_2$</td>
<td>$\beta = \beta_1 \gamma + \beta_2$</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$\alpha = \alpha_1 \gamma^2 + \alpha_2 \gamma + \alpha_3$</td>
<td>$\beta = \beta_1 \gamma^2 + \beta_2 \gamma + \beta_3$</td>
</tr>
<tr>
<td>Cubic</td>
<td>$\alpha = \alpha_1 \gamma^3 + \alpha_2 \gamma^2 + \alpha_3 \gamma + \alpha_4$</td>
<td>$\beta = \beta_1 \gamma^3 + \beta_2 \gamma^2 + \beta_3 \gamma + \beta_4$</td>
</tr>
</tbody>
</table>

This method determines equations for $W$ as a function of two independent variables. This method has been described for the case of $\gamma =\text{SST}$, but the method can be applied independently to other parameters including all those used within this study. Additionally, this method is not only limited to $U_{10EN}$ as the argument. For in situ winds, $U_{10EN}$ can be replaced by $U_{10}$. The performance of the functional fits using the same parameter or multiple independent parameters is left to Appendix D.
Figure C1. Function fits to $\alpha(SST)$ (above). Function fits to $\beta(SST)$ (below).
APPENDIX D

RANKING

Having already used the least squares statistic in previous steps to determine the best empirical fit for functions to calculate the whitecap coverage based on a variety of independent variables, choosing a closely related statistic such as $R^2$ to determine the performance of these equations to predict the whitecap coverage might introduce unnecessary error and uncertainties. To address this problem, a simple ordinal ranking system is presented to use in determining the performance of two or more empirically fit equations using the same validation data. This ranking is independent of $R^2$ and other previous and similar statistics used. The result of this technique is to compare the overall performance of each of the equations against one another in representing the validation data using ordinal ranking for each validation point. It does not predict error or deviations. This is a simple statistic with a simple interpretation of finding the best performing empirically fit equation.

For this approach to be valid for ranking the equations, all data used for validation must be contained within the union of the ranges over which all of the equations are valid. To avoid biasing due to the uneven distribution of $U_{10EN}$, a binning process is applied during the process (Appendix A). This method provides performance for each of equations over the entire valid $U_{10EN}$ range as well as over smaller ranges. The “truth” values are called validation points and values calculated from the empirically fit equations being tested are called test points; $M$ represents the number of equations being tested. The steps for this ranking method follow.
1. For one validation point, the absolute difference between the validation point value and the test point value for each equation is determined. The smallest difference is assigned the value of one and the next smallest difference is assigned the value of two and so forth for all M test values corresponding to the one validation point receive an ordinal rank. This is repeated for each validation point until each test value for each equation receives an ordinal rank according to the performance compared to other test values for the given validation point.

2. The data are divided into 1–ms\(^{-1}\) \(U_{10}\) bins. The ranked values for each of the equations for each point in each bin are then summed and normalized by the number of points in each bin. This creates M values per bin each representing the M equations.

3. The new values for each bin receive an ordinal rank as before with one set of ranks per bin.

4. The newly ranked values from each bin for each equation are then summed and normalized leaving M values each representing one equation.

5. The new values are ranked once.

The equation corresponding to the final rank of one performed the best with the equation with the final rank of two performing the next best and so forth. This only indicates the best performing equations compared to the entire validation data set. For smaller ranges of \(U_{10EN}\), the results from step 3 provide a similar result over each of the bins. For limiting ranges further,
modifying the requirements on the validation points is recommended. For limiting $U_{10EN}$ only, step 4 can be modified to sum only over the $U_{10EN}$ range of interest.

A qualitative analysis of how each of these equations performed compared to one another can be done by analyzing the values of the summed and normalized ranks before re-ranking (steps 2 and 4). For a basic interpretation, the values closer to each other indicate a closer overall performance for those values. Further qualitative interpretations depend more on the data being considered and are not included as part of this appendix.
REFERENCES


BIOGRAPHICAL SKETCH

Aaron C. Paget attended St. Louis Community College, Ricks College (now Brigham Young University – Idaho), and completed a Bachelor’s degree in Applied Physics from Brigham Young University. After completing the Bachelor’s degree, he came to Florida State University for graduate school in Meteorology, graduating with a Master’s Degree in Fall 2009. He continued his graduate education seeking a PhD in Geophysical Fluid Dynamics at Florida State University.