IMPROVED SSM/I WIND SPEED RETRIEVALS AT HIGHER WIND SPEEDS

by

Vladimir M. Krasnopolsky
General Sciences Corporation,
Laurel, MD 20707

William H. Gemmill and Laurence C. Breaker
Ocean Modeling Branch
Environmental Modeling Center
National Centers for Environmental Prediction
Washington, D.C. 20233

Submitted to the Journal of Geophysical Research

October 1995

OMB Contribution Number 111.

ABSTRACT

Limitations in making SSM/I wind speed retrievals at higher wind speeds (w > 15 m/sec) are investigated and the primary sources of error identified. Whitecaps and foam give rise to systematic changes in the empirical transfer functions which are used in making SSM/I wind speed retrievals at higher wind speeds. The addition of the 85GHz(V) channel is shown to improve retrievals at higher wind speeds from the SSM/I. A new hybrid retrieval approach has been developed which combines a modified neural network (NN) architecture (including 85GHz channel) and modified training procedures with an independent correction for a residual systematic error in the transfer function which occurs at higher wind speeds. This hybrid approach has resulted in a weighted, bias-corrected NN algorithm with five inputs (the "OMB" algorithm). Applied to matchup data used in deriving previous SSM/I wind speed algorithms, this algorithm yields a bias < 0.2 m/sec and an rms difference < 1.75 m/sec for all wind speeds and weather conditions encountered in the matchup database, and a bias of ~0.7 m/sec and an rms difference of ~2.8 m/sec for wind speeds > 15 m/sec. The OMB algorithm is capable of generating wind speeds up to 25- 27 m/sec. It also yields an average gain in coverage of ~15%, and significantly higher gains in coverage for individual synoptic events. This algorithm also reveals detailed structure in the patterns of surface wind speed not produced by other retrieval algorithms. It has been preliminarily validated using data from both the F10 and F13 SSM/I instruments. Finally, the application of surface winds retrieved using the OMB algorithm to atmospheric and oceanic forecast models is discussed.

1. INTRODUCTION

Ocean surface winds are required by operational marine forecasters to produce accurate surface weather analyses over the global oceans, and by atmospheric modelers for assimilation into global and regional weather forecast models. Since the late 1970's, both active and passive microwave radiometers aboard polar-orbiting satellites have been used to infer wind speed and, in some cases, wind direction over the ocean. Algorithms which have been developed to infer surface wind speed for these instruments such as the SMMR and the SSM/I, however, have been seriously limited in their ability to infer wind speeds under conditions of high atmospheric moisture or for wind speeds greater than ~15m/sec. Because active synoptic weather systems, which are usually characterized by relatively high levels of moisture and higher-than-average wind speeds, are of primary interest to the operational forecasting community, it is important to develop wind retrieval algorithms which are "robust" both with respect to atmospheric moisture and wind speed.

Neural networks (NNs) have recently been used to develop wind speed retrieval algorithms based on brightness temperatures (TBs) received from the SSM/I flown aboard the DMSP satellites [Stogryn et al., 1994; Krasnopolsky et al. 1994, 1995]. Krasnopolsky et al. [1995; referred to as KBG, hereafter] developed a single, extended-range NN algorithm (SER NN) together with a wind speed retrieval flag based on cloud liquid water path (LWP) which retrieved wind speeds up to ~17 m/sec for moisture levels up to ~0.5 kg/m^2 with a small bias and an rms error of less than 1.7 m/sec.

Based on our previous work in establishing an acceptable rain flag criterion for the SER NN, we are approaching the upper limit for making SSM/I wind speed retrievals at higher levels of atmospheric moisture. However, it is likely that further refinements to the threshold value of 0.5 kg/m^2 can still be made when a more extensive database is created which includes significantly more matchups under high moisture conditions.

Another limitation of existing SSM/I wind speed retrieval algorithms is that they are primarily restricted to retrieving wind speeds with acceptable accuracy only up to 15 - 20 m/sec. Often, high wind speed events are accompanied by high levels of moisture (e.g., hurricanes), which preclude the possibility of making successful retrievals because the moisture threshold is exceeded. For lower levels of moisture, however, processes at the ocean surface which directly affect the emissivity of the air-water interface establish an upper limit for retrieving surface wind speed. In particular, as wind speed increases, wave breaking occurs which produces whitecaps and foam. Eventually, when the ocean surface becomes completely obscured by foam, microwave emissions from the surface no longer change as wind speed increases. Therefore, an additional physical limitation in retrieving SSM/I wind speeds, based on the wind speed itself exists, which limits the range of SSM/I wind speed retrievals to the interval (0, w_{max}). It is difficult to estimate the value of w_{max} from theoretical considerations alone; however, observational estimates indicate that this threshold occurs in the vicinity of 30 m/sec [Bortkovskii, 1987; Monahan and Mac Niocaill, 1986; Swift, 1990]. Clearly, based on these estimates, significant improvements can still be made to existing NN algorithms by extending their range of applicability from (0,16-17) m/sec, to (0,25-30) m/sec. In what follows, we describe a new NN algorithm with extended wind speed retrieval capabilities.

The SSM/I is a seven-channel passive microwave radiometer. This instrument receives vertically-polarized (V) radiation at 22.2GHz and vertically- and horizontally-polarized (H) radiation at 19.3, 37.0, and 85.5GHz. The SSM/I has a nominal spatial resolution of 12.5 km at 85.5GHz and a resolution of 25 km at 19.3, 22.2 and 37.0GHz. For additional information concerning the SSM/I, see Hollinger et al. [1990].

In Section 2, we discuss why both empirically- and theoretically-derived algorithms encounter problems in reproducing high wind speeds. In Section 3, it is shown that TBs at 85GHz(V) are sensitive to changes in wind speed, and an improved training approach is introduced which extends the high wind speed limit of the NN algorithm which we develop. A new, weighted NN algorithm is introduced which uses five channels from the SSM/I (four previous channels plus 85GHz(V)). We then examine the bias of the new algorithm as a function of wind speed and introduce a simple bias correction which further improves the high wind speed retrieval capability of the weighted NN (WNN). In Section 4, we introduce a new matchup database based on TBs acquired from the F10 SSM/I instrument. We use this new database to compare the current operational algorithm [Goodberlet, Swift and Wilkerson, 1989; hereafter referred to as GSW] with the new WNN and the bias-corrected WNN algorithms. Finally, we examine the dependence of retrieval accuracy for the GSW and the NN algorithms on the spatial and temporal resolution of the data. Section 5 contains our conclusions.

2. OBTAINING RETRIEVALS AT HIGHER WIND SPEEDS

Several sources of error affect the accuracy of SSM/I retrieval algorithms. First, there are errors due to the buoy measurements themselves. Second, errors arise because wind speeds are not uniformly distributed. A third source of error is due to the lack of coincidence in space and time between the ground-truth and the satellite observations in creating the wind speed matchups. Also, the matchups consist of inherently different types of observations. On the one hand, buoy wind speeds correspond to point measurements averaged over time; on the other, the satellite measurements are acquired instantaneously but are essentially area-averaged. A fourth source of error arises from the physical limitations of microwave remote sensing of wind speed when the ocean surface becomes whitecap foam covered. This error manifests itself as a systematic change in the empirical transfer function that defines the wind speed -SSM/I brightness temperature (TB) relationship at higher wind speeds.

2.1 Buoy data accuracy

All SSM/I wind speed retrieval algorithms to date have used observations from ocean data buoys for algorithm development and validation. Empirically-derived algorithms approximate empirical transfer functions, i.e. the relationship between the SSM/I TBs and the buoy wind speeds [e.g., GSW; Goodberlet and Swift, 1992; Stogryn et al., 1994; KBG]. Theoretical algorithms based on the radiative transfer equations also use buoy data for certain model parameterizations [e.g., Petty and Katsaros, 1994] and validation [e.g., Wentz et al., 1991]. Thus, any problems associated with the buoy observations themselves, directly affect the

accuracy of the algorithms which are derived, especially at high wind speeds. According to Gilhousen [1986], wind speed accuracy for the anemometers deployed on National Data Buoy Center (NDBC) buoys is ±0.5 m/sec for winds less than 10 m/sec and ±5% of the wind speed for winds greater than 10 m/sec. Thus, there are significant uncertainties associated with buoy measurements of wind speed and they become particularly important at high ($w > 15$ m/sec) wind speeds. Also, because these errors are essentially random, they contribute primarily to the scatter.

Buoy wind speed observations at higher wind speeds are limited because high wind speeds are a relatively infrequent occurrence, and additionally because many of the fixed buoys (upon which the matchups are based) are located in partially-sheltered coastal areas. These problems have been recognized and discussed elsewhere [e.g., Pierson and Sylvester, 1995]. In the SSM/I/F8 matchup database, the maximum buoy wind speed is 21.2 m/sec. In the SSM/I/F10 database which we have recently assembled, the maximum wind speed is 21.4 m/sec (Section 4). In the SSM/I/F8 matchup database (3958 matchups), there are very few measurements of wind speed higher than 17-18 m/sec, and only 44 matchups (~1%) at wind speeds higher than 15 m/sec. The observed wind speed distribution (Fig. 5) is clearly nonuniform, and based on theoretical considerations, wind speed data are generally expected to follow a Rayleigh distribution [e.g., Breckling, 1989].

Nonuniform wind speed distributions create significant problems for algorithm development because they imply fewer observations at the tails of the distribution. The nonuniformity of wind speed distributions creates an additional problem for algorithm validation due to the interaction of random errors in buoy wind speed data with a nonuniform wind speed

distribution [Tolman, 1995]. This interaction may produce an additional nonrandom error which increases with the wind speed and introduces an additional positive bias at high wind speeds.

2.2 Matchup errors

Both the training and the test data consist of matchups which compare two inherently different types of observations, (i), buoy wind speeds acquired from anemometers which are point measurements at a fixed elevation above the ocean surface averaged over intervals of 8.5 minutes, and (ii), instantaneous satellite observations that cover an approximate 25 km × 25 km footprint on the ocean surface (at the four lowest SSM/I frequencies). Even for perfect matchups (where the center of the SSM/I footprint coincides precisely with the buoy location, and the time of the buoy measurement coincides precisely with the time of the satellite measurement), an additional error is introduced because of the differences in the space-time windows which must be employed. Thus, the inherent variability of wind speed within the satellite footprint over an 8.5 minute period (τ) may be an additional source of error that enters into the matchup. If this variability is truly random, no bias is expected; however, for a 25 km × 25 km footprint, we may expect an additional contribution to the observed scatter on the order of 0.5 m/sec [Overland and Gemmill,1977; Monaldo, 1988]. Additionally, because most of the NDBC buoys are located in coastal regions where small-scale, transient high wind speed events frequently occur (vs. over the open ocean far-removed from land and boundary currents), there may be higher natural variability in wind speed and, thus, higher matchup errors at the higher wind speeds.

A related problem associated with matching the buoy and satellite data arises because perfect matchups occur infrequently and, as a result, the time interval R_t and the distance R_s between the buoy and satellite measurements must be expanded in order to obtain statistically-meaningful sample sizes. Because of the finite satellite footprint dimension r_s, the minimum uncertainty in distance is equal to $0.5\ r_s$ (~6.25 km at 85GHz and ~12.5 km at the four lower SSM/I frequencies). In the general case, the uncertainty in distance is equal to $R_s \geq 0.5\ r_s$, and the uncertainty in time is equal to $R_t \geq \tau$. Both of these uncertainties introduce additional errors which are due to the natural variability in wind speed over the scales R_s and R_t. Because these errors are essentially random, they contribute primarily to the scatter.

2.3 Physical limitations of microwave retrievals at higher wind speeds

At low wind speeds, the increase in ocean surface roughness with increasing wind speed produces a TB dependence on wind speed. At higher wind speeds, waves begin to break producing whitecaps and foam (called whitecap foam from hereon). At wind speeds of ~10 m/sec [Swift, 1990] or less [Monahan and O'Muircheartaigh, 1986], whitecap foam starts to contribute significantly to TB at microwave frequencies. The surface area covered by whitecap foam increases rapidly with increasing wind speed at higher wind speeds. At some limiting value of wind speed, w_{max}, when the surface area covered by whitecap foam becomes appreciable, the TB dependence on wind speed breaks down. Although it is difficult to estimate w_{max} precisely, Bortkovskii [1987], Monahan and MacNiocaill [1986] and Swift [1990] provide estimates of the area covered by whitecap foam as a function of wind speed. Empirical relationships have been developed from these observations that express the relationship between

whitecap foam coverage as a function of wind speed.[1] They usually follow a simple power law relationship between whitecap foam coverage (S in percent) and wind speed, w,

$$S = a\, w^{\alpha} \qquad (1)$$

where a and α are the empirically-determined constants shown in Table 1. Fig. 1 shows four different parameterizations (Table 1) extrapolated up to S = 100%. Despite the uncertainties involved, these results provide at least rough estimates for w_{max}.

For 10 m/sec < w < w_{max}, surface emissions from the rough ocean surface and the foam-covered areas both contribute to the TB signals sensed by the SSM/I. Because whitecap foam acts roughly as a black body radiator, TBs from the foam-covered areas far exceed those emitted by the non-foam-covered areas. Thus, the physical mechanism which primarily contributes to the signal received by the SSM/I at higher wind speeds is quite different. When the wind speed is about 10 m/sec, less than 5% of the surface is covered by whitecap foam; thus, it may be reasonable to assume that further retrievals become difficult, or perhaps impossible, when at least 50% or more of the ocean surface is covered by whitecap foam. Table 1 shows estimates for w_{max} which fall between 25 and 32 m/sec for 50% coverage (Swift [1990] estimated a value of 40 m/sec for 100% whitecap foam coverage). This threshold, although imprecise, represents an upper limit for retrieving surface wind speeds using microwave radiometry.

---

[1]Monahan and O'Muircheartaigh(1986) have shown that although whitecap foam coverage is primarily related to wind speed, it also depends on the air-water temperature difference, the water temperature per se, and the effective wind duration.

Previous empirical SSM/I wind speed retrieval algorithms have been based on the assumption that the retrieved wind speed depends only on TB, or expressed in terms of the SSM/I transfer function (see KBG),

$$w = f(\boldsymbol{T}) \qquad (2)$$

where f is the transfer function, and $\boldsymbol{T} = \{t_1,..,t_7\}$ is a TB vector containing a maximum of 7 elements. At low wind speeds, surface roughness determines the wind speed dependence on TB. At higher wind speeds, when whitecap foam formation becomes important, the functional dependence between TB and wind speed changes, and retrievals based on (2) may become ambiguous if the wind speed itself is not taken into account. The simple representation expressed in (2) can be generalized for higher wind speeds, as

$$w = F(\boldsymbol{T},w) \qquad (3)$$

Therefore, at higher wind speeds, the SSM/I retrieval algorithm, based only on SSM/I TBs (i.e., the transfer function f, in (2)), may generate wind speeds with a systematic error due to the difference between the correct transfer function, F, and the original transfer function, f. This systematic error (hereafter refer to as the δ-error), contributes to the bias (and thus the rms error) at wind speeds > 10-15 m/sec and reflects the change in mechanisms at the ocean surface which contribute to TBs at higher wind speeds. These errors (together with the lack of sufficient

matchup data at high wind speeds) are primarily responsible for underestimating high wind speeds in previous NN algorithms.

3. A NEURAL NETWORK ALGORITHM WITH IMPROVED HIGH WIND SPEED RETRIEVAL CAPABILITY

To develop a new NN algorithm which is capable of generating higher wind speeds than our previous NN algorithm, we have initially examined the sensitivity of the different SSM/I channels to changes in wind speed, and then developed a new training strategy, followed by a correction for the expected systematic residual error. As a result, a new NN architecture has been developed which differs significantly from that used previously by Stogryn et al. [1994], or by KBG.

As in KBG, we use the same SSM/I TB/buoy wind speed matchup database that was originally compiled by GSW. This database consists of matchups from the F8 SSM/I instrument which has been used in previous algorithm development and validation by GSW, Wentz [1989], Stogryn et al. [1994], and KBG. A complete description of this database can be found in GSW or KBG.

3.1 Channel sensitivity to wind speed

KBG previously examined the utility of the 85GHz channels to improve the accuracy of SSM/I wind speed retrievals and showed that under cloudy conditions inclusion of the 85GHz channels yielded a slight reduction (~10%) in rms error. We now examine the sensitivity of the different channel outputs, including the 85GHz channels, to changes in wind speed. With respect to the SSM/I channelization, wind speed may be considered to be a function of the seven SSM/I TBs. These TBs, $\{T_i\}^N_{i=1}$, which are contained in the training matchups, produce a cluster of points that cover a finite connected domain in seven-dimensional space. The centroid of this cluster is calculated as

$$T_0=\frac{1}{N}\sum_{i=1}^{N} T_i \tag{4}$$

The mean radius of this cluster is

$$R=\sqrt{(\frac{1}{N}\sum_{i=1}^{N}(T_i-T_0)^2)} \tag{5}$$

If a continuous transfer function f exists, then this domain is a connected domain and the TBs, $\boldsymbol{T}'$ for wind speeds in any bin $w_1 \le w' < w_2$ create a connected subdomain or subcluster. The centroid of this bin cluster $\boldsymbol{T_0'}$, and its mean radius $\boldsymbol{R}'$, can be calculated using (4) and (5); however, a summation over i must now be performed over the number of TB vectors in each bin cluster, N'. Similarly, the projection of the centroid of the bin cluster on the j-th axis, corresponding to the j-th TB, t_j, and the mean radius of this projection can be calculated.

We examine these bin cluster characteristics as functions of wind speed (Fig. 2). Three channels, 22GHz(V), 37GHz(H) and 85GHz(V), are highly sensitive to changes in wind speed. In addition to the original four channels that have been used in most previous SSM/I wind speed retrieval algorithms, the 85GHz(V) channel also depends on wind speed and, consequently, has been included in the development of a new retrieval algorithm. Fig. 2 also shows a distinct change in the functional dependence at higher wind speeds (w > 15 m/sec) for all channels which may be due to significant contributions from whitecap foam (and/or the limited sample sizes for these bins).

3.2 Training strategy

Because there are very few matchups available at higher wind speeds (w > 15 m/sec), the distribution of wind speeds in the training set is highly nonuniform (Fig. 5). If we use the standard cost function for NN training (e.g., Wasserman, 1989) where all of the data are equally weighted, the few matchups that do exist at high wind speeds become completely obscured in the training process. This problem can be reduced if a weighted cost function is used instead of the standard cost function for training, where the weighted cost function can be expressed as

$$E = \frac{1}{N} \sum_{i=1}^{N} \alpha_i \, (W_i - w_i)^2 \qquad (6)$$

Here, N is the number of matchups in the training set, the W_i are the wind speeds obtained during the training process, the w_i are the observed buoy wind speeds, and the α_i are the weights. Because a uniform data distribution is clearly preferable for training (e.g., Cheng and

Titterington, 1994), the distribution of the α_i can be chosen to compensate for the inherent nonuniformity of wind speeds in our training set. A distribution which is inversely proportional to the square root of the wind speed distribution is an acceptable choice. For such a weighting scheme, the highest wind speeds are emphasized and the NN is forced to learn more from these few cases. The penalty, of course, is that any errors in these data will be amplified during the process of training.

As mentioned in Section 2, both random and systematic errors in wind speed increase with increasing wind speed. Fortunately, NNs are somewhat insensitive to random errors of high amplitude when they are properly designed [Kerlirzin and Vallet, 1993]; however, they are sensitive to systematic errors. NNs provide a valid model for the transfer function f in (2) for low-to-medium wind speeds. At higher wind speeds, the transfer function F in (3) is valid. The wind speeds generated by the weighted NN (WNN) at high wind speeds contain systematic errors which are due to the differences between f and F (see Section 3.4).

3.3 Weighted NNs

Following the approach outlined in the previous sections, we have chosen a new architecture for the SSM/I wind speed retrieval algorithm. This NN has one input layer with five inputs, the four channels which have been used in previous algorithms (19V, 22V, 37V, and 37H) and the additional 85GHz(V) channel, one hidden layer with two nodes, and an output layer containing one node. In preparation for training the NN, the weights α_i in the cost function (6) were generated using the following formula,

$$\alpha_i = \frac{C}{\sqrt{p(w_i)}} \qquad (7)$$

where $p(w)$ is the observed wind speed probability distribution and C is a normalization constant. This choice of weights allows us to assign higher values to the tails of the distribution and effectively produces a distribution which is approximately uniform. By introducing the square root in the denominator of (7), we have restricted the rate of increase in the weights and thus reducing any noise-like influences at the highest wind speeds.

Table 2 shows the summary statistics for four different algorithms, the original GSW algorithm, the SER NN (our original NN algorithm), the WNN4, the weighted NN algorithm with four inputs, and finally, the WNN5, the weighted NN algorithm with five inputs. Only the statistics for clear atmospheric conditions (i.e., for low levels of moisture) are included, i.e., the case where the GSW algorithm performs best. Table 2 shows that (1) the weighted NNs are capable of generating higher wind speeds, (2) appropriate weighting of the cost function at high wind speeds amplifies the noise (as expected), producing slightly less favorable statistics compared to the SER NN, (3) the weighted NN with five inputs (WNN5) outperforms the weighted NN with four inputs (WNN4) both in terms of extended high wind speed generation capability plus improved statistics (in this case, similar to those obtained using the SER NN), and (4) all NN algorithms outperform GSW.

Table 3 shows the same information as Table 2 but for the clear plus cloudy case. In addition, the statistics for high wind speeds only ($w > 15$ m/sec) are included (these statistics are

not presented in Table 2 because of insufficient sample size for the clear case). Because of buoy data inaccuracies and the matchup problems indicated earlier at high wind speeds, the statistics for high wind speeds should be treated with caution. We conclude that the WNN5 algorithm demonstrates the best overall performance for the clear plus cloudy case. Finally, for high wind speeds the table shows that although the GSW algorithm has the ability to generate high wind speeds, the standard deviation in this case is much higher than the standard deviation for the observed wind speeds.

An additional characteristic of NN algorithms can be introduced which is important for evaluating their ability to generate high wind speeds. This characteristic corresponds to a well-defined maximum output value which, by virtue of the weights which are derived during training, can be generated by a particular NN. For any input TBs, the output wind speed cannot exceed this value, a value referred to here as W_Ω. The equations for calculating W_Ω are given in the Appendix. Table 4 shows W_Ω for the SER NN, WNN4, and WNN5 algorithms and clearly demonstrates significant improvement in high wind speed generation capability, progressing from the SER NN to the WNN5 algorithms.

3.4 Bias correction

From Section 2.3, we expect the systematic δ-error to increase with increasing wind speed at higher wind speeds. In terms of the transfer function, f, (2) is valid only for low-to-moderate wind speeds. To include higher wind speeds, f should be generalized as shown in (3) where the generalized transfer function F now depends on the wind speed itself, as well as the TBs. We can, in principle, use the NN approach to develop a NN model for the transfer function

expressed in (3); however, the corresponding NN architecture will be more complicated and will require the addition of a feedback loop (i.e., recurrence) . As a practical alternative, we consider a simpler approach based on our expectation that the systematic wind speed-dependent δ-error will give rise to a wind speed-dependent bias. Fig. 3 shows the binned bias as a function of wind speed (calculated as the difference between the buoy and the WNN-generated wind speeds) for the training set (dashed line). This bias has a systematic trend which we approximate by the following expression,

$$b(w) = a + b\,(w - c)^3\,(1 - \exp(-dw)) \qquad (8)$$

where $a = 0.5$ m/sec, $b = 0.004$ sec^2/m^2, $c = 10$ m/sec, and $d = 0.5$ sec/m. This approximation is represented by the solid line. The bias correction (8) increases rapidly at wind speeds > 10 m/sec and is apparently related to the change in the dominant mechanism generating the TB wind speed signature which occurs in the vicinity of 10 m/sec when whitecap foam, rather than surface roughness, constitutes the major emission from the ocean surface. Fig. 2 also shows the bias for the test set (dash-dotted line). Equation (8) yields a satisfactory approximation for these data as well. The existence of such a well-defined wind speed-dependent bias may be interpreted in terms of the transfer function as follows. By decomposing the transfer function expressed in (3) into the sum of two terms, it may now be approximated as,

$$F(\boldsymbol{T},w) = f(\boldsymbol{T}) + b(w) \qquad (9)$$

where f is the transfer function given in (2) which depends only on TB, and $b(w)$ is the wind speed-dependent bias given by (8).

We now use our weighted NN (i.e., WNN5) as a model for the transfer function f in (9) to retrieve the wind speed w corresponding to the TBs, $\boldsymbol{T}$, where

$$w = f(\boldsymbol{T}) \approx WNN5(\boldsymbol{T}) \qquad (10)$$

By applying the bias correction $b(w)$ to this wind speed we obtain the wind speed w_c, which is now corrected for the systematic error,

$$w_c = w + b(w) \qquad (11)$$

In these calculations, the same training set has been used for developing the WNN algorithm and for calculating the bias correction, $b(w)$.

Equations (10) and (11) together constitute an algorithm which we call the bias-corrected WNN5 or OMB[2] algorithm. Table 5 shows summary statistics for the WNN5, the OMB, and the BWNN4 (WNN4 with four inputs including the bias correction (8)) algorithms, and separate high wind speed statistics for the GSW, OMB and BWNN4 algorithms. The OMB algorithm clearly outperforms the other algorithms. Fig. 4 shows a binned (bin size = 2 m/sec) scatter plot for the three NN algorithms. We see successive improvements in high wind speed performance,

---

[2]Ocean Modeling Branch (OMB), Environmental Modeling Center of the National Centers for Environmental Prediction.

with no degradation at the lower wind speeds in progressing from the SER NN to the WNN5, and eventually to the OMB algorithm. Both BWNN4 and OMB produce high wind speed statistics which are significantly better than those produced by GSW. That the bias correction is compatible with both the WNN5 and WNN4 algorithms indicates that this correction is independent of the NN architecture and is consistent with our assumption that it is due to a systematic error in the original transfer function. Fig. 5 shows the observed wind speed distribution together with the SER NN- and OMB-generated wind speed distributions. The wind speed distribution generated by the OMB algorithm is much closer to the observed, reproducing both the high and low wind speed tails of the distribution.

4. VALIDATION AND MATCHUP UNCERTAINTIES

All of the results described in the previous section have been obtained using the matchup database created by GSW which only contains SSM/I data from the F8 instrument. This database contains a limited number of SSM/I TB/buoy wind speed matchups with a spatial uncertainty $R_s \leq 25$ km and a temporal uncertainty $R_t \leq 0.5$ hrs. To validate the OMB algorithm with data from different SSM/I instruments, and to investigate the dependence of retrieval accuracy on matchup uncertainty, we have created a new matchup database containing approximately 27,000 SSM/I BT/buoy wind speed matchups from the F10 SSM/I instrument with a spatial matchup uncertainty $25 \leq R_s \leq 100$ km, and a temporal matchup uncertainty $0.5 \leq R_t \leq 3.0$ hour. Because the buoy data have been preprocessed with a roundoff error of ~0.5

m/sec, an additional random error of approximately 0.3 m/sec rms has been introduced. Although the SSM/I instrument on the F10 DMSP satellite has certain problems related to the ellipticity of the orbit which effects the scanning geometry and thus reduces retrieval accuracy for all algorithms, we have nonetheless used the data from this instrument to compare the performance of the various algorithms.

These data have been used to create 24 subsets with different spatial and temporal matchup uncertainties. For the spatial matchup uncertainty, we use the following four values, R_s = {25, 50, 75, 100} km, and for temporal matchup uncertainty the following six values, R_t = {0.5, 1, 1.5, 2, 2.5, 3} hours. For each subset we have calculated total and high wind speed (w > 15 m/sec) statistics (biases and rms errors) for the GSW, WNN5 and OMB algorithms. These statistics are shown in Fig. 6 as functions of R_s and R_t. Figs. 6a,b show the biases and rms errors for clear atmospheric conditions, Figs. 6c,d, for clear plus cloudy conditions, and Figs. 6e,f show the high wind speed (wind speed > 15 m/sec) statistics. We summarize the information contained in these figures below:

- For all weather conditions, both NN algorithms outperform the GSW algorithm. The total statistics for the OMB algorithm are similar to those for WNN5. Under clear plus cloudy conditions, the biases and rms errors are unacceptably high for GSW, whereas both the WNN5 and OMB algorithms yield biases and rms errors which are acceptable for operational use. Additionally, the NN-based algorithms produce an expanded retrieval domain from clear, to clear plus cloudy, conditions yielding an increase in retrieval coverage of ~15%. This result is particularly significant for obtaining more complete coverage of synoptic-scale weather systems

such as extratropical cyclones which typically have higher levels of moisture and higher wind speeds.

- At high wind speeds, the OMB algorithm performs significantly better than WNN5 or GSW. This demonstrates that the bias correction which was derived using the F8 data performs satisfactorily using the data from F10 as well, and thus tends to be instrument-independent.
- Both NN algorithms display a gradual degradation in performance with increasing temporal and spatial matchup uncertainty. They perform best on data with the lowest matchup errors. Thus, the NN algorithms provide an accurate approximation for the transfer function f; they are also robust with respect to random errors in the matchups.
- GSW demonstrates an opposite trend with respect to increasing matchup uncertainties in both time and space. The biases and rms errors decrease with increasing matchup uncertainty (the standard deviation increases, however). Thus, the GSW algorithm only approximates the transfer function, f, in a mean sense. As a result, this algorithm performs optimally when the retrievals are averaged over relatively large space and time intervals.
- The last two conclusions have direct implications for the assimilation of SSM/I wind speeds into atmospheric and ocean forecast models. For matchup uncertainties of the order of 100 km and 3 hours (typical resolutions for global models), the GSW and NN algorithms yield the smallest differences in bias and rms error. The primary advantage of the NN-based algorithms in this case is the gain in coverage due to the inclusion of cloudy (i.e., high moisture) data in the retrieval process. The NN-based algorithms display greater benefits for low matchup uncertainties (which correspond to higher model resolutions) where they provide significant improvements both in terms of accuracy and coverage.

Although the gain in coverage due to the inclusion of cloudy data in the retrieval process is estimated to be roughly 15%, this value is based on global estimates for an extended period of time. Local and short-term increases in coverage may be much more impressive and may reach 30% in some cases. Moreover, from a synoptic standpoint, when high levels of moisture are accompanied by high winds, the NN-based algorithms (especially the OMB algorithm) may reveal detailed structure in surface wind fields which could be completely missed using the GSW algorithm. In this regard, Fig. 7 shows an intense storm which occurred over the south Australian Basin on 24 August 1995 centered at 47.5°S and 122°E with an analyzed central pressure of 982 mb. The storm created winds in the 25 m/sec (50 kt) range over an extended area. Figs. 7a,b,c show the event as it was seen by the F10 instrument, and Figs. 7d,e,f, by the F13 instrument. Figs. 7a and d show the surface wind speeds retrieved using GSW, Figs. 7b and e, using SER NN, and Figs. 7c and f, using the OMB algorithm. The transition from GSW to SER NN, and then to the OMB algorithm, clearly shows the increase in coverage just described. In addition, the OMB algorithm reveals a rich and detailed structure in this synoptic field which is completely missed by the GSW algorithm and significantly smoothed by SER NN algorithm.

5. CONCLUSIONS

We have presented a new NN-based algorithm for SSM/I wind speed retrievals which demonstrates high retrieval accuracy together with the ability to generate high wind

speeds with acceptable accuracy. The 85GHz(V) channel, which has not been used in previous SSM/I wind speed retrieval algorithm development, was shown to be sensitive to changes in wind speed and thus has been incorporated in the new algorithm. This sensitivity at 85GHz may be related to the higher spatial resolution of the SSM/I at this frequency.

Previous algorithms have not performed well at high wind speeds. This problem may be due to several factors including increased buoy wind speed errors at high wind speeds, nonuniformity of the wind speed distribution, matchup uncertainties, and systematic errors which occur at high wind speeds due to the presence of whitecaps and foam. A practical upper limit for making SSM/I wind speed retrievals, based on physical considerations, may be as low as ~30 m/sec.

A new NN training strategy which includes preferential weighting at high wind speeds was introduced to compensate for the nonuniformity in the distribution of observed wind speeds. A hybrid approach has been developed which combines a weighted NN and a wind speed-dependent bias correction for systematic errors which occur at high wind speeds. This bias correction increases rapidly at wind speeds > 10 m/sec and is apparently related to the change in the dominant mechanism responsible for generating the TB wind speed signature which occurs in the vicinity of 10 m/sec. A new algorithm (OMB) based on the hybrid approach was developed for retrieving SSM/I wind speeds which produces operationally-useful retrievals up to 25 - 27 m/sec. Also, an important characteristic of NNs which determines their theoretical maximum output has been introduced and applied to our particular problem.

The OMB algorithm was tested on matchup data for the F10 SSM/I instrument and showed significant improvement both in the accuracy of the retrievals and in increased areal

coverage. The NN algorithm demonstrates the greatest improvement in retrieval accuracy based on accurate, high-resolution matchups (with low matchup uncertainty). Consequently, higher-resolution atmospheric and ocean forecast models will benefit to a greater extent than lower-resolution models from the inclusion of surface wind speed data produced using the OMB algorithm.

The OMB algorithm was applied to an extratropical cyclone in the southern hemisphere which contained both significant moisture and high wind speeds. Wind speeds from both the F10 and F13 instruments were retrieved. These retrievals, based on the new algorithm, showed major improvements in resolving details in the surface wind speed field.

Finally, the new algorithm (including both the weighted NN and the bias correction) has been developed using only the SSM/I F8 database. This database has many limitations; it does not have a sufficient number of matchups at high wind speeds, high latitudes are poorly represented, and matchup uncertainties could be improved. Taking into account these limitations, the OMB algorithm will be re-evaluated and improved following the hybrid approach presented here when a more representative matchup database is produced. The creation of such a database is presently underway.

Acknowledgments

We take this opportunity to thank D.B. Rao for a thorough review of this manuscript. We also thank Marie Colton of the Fleet Numerical Meteorology and Oceanography Center and Gene Poe of the Naval Research Laboratory for their continued encouragement in this work.

APPENDIX

The NN output can be expressed as [KBG],

$$W = b + a \tanh\left(\sum_{j=1}^{k} \omega_j z_j + \beta\right) \qquad (A.1)$$

where the ω_j are the weights and β is the bias in the output layer, a and b are positive scaling factors, k is the number of hidden nodes, and z_j is the output of the j-th hidden node, which can be expressed as

$$z_j = \tanh(X) \qquad (A.2)$$

where X is a linear combination of the inputs to the NN. For any combination of inputs the absolute value of z_j is always less than, or equal to, 1. By taking into account that $\tanh(X)$ is a monotonically increasing function in the interval (-1,1), we can estimate the theoretical upper bound W_Ω for the NN output W as,

$$W \leq W_\Omega = b + a \tanh\left(\sum_{j=1}^{k} |\omega_j| + \beta\right) \qquad (A.3)$$

REFERENCES

Bortkovskii, R.S., *Air-Sea Exchange of Heat and Moisture During Storms,* D. Reidel, Dordrecht, Holland, 1987.

Breckling, J., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, Springer-Verlag, Germany, 1989.

Cheng, B., and D.M. Titterington, Neural networks: A review from statistical perspective, *Statistical Science, 9,* 2-54, 1994.

Gilhousen, D. B., An accuracy statement for meteorological measurements obtained from NDBC moored buoys, in *Proc. MDS'86 Marine Data Syst. Int. Symp.*, Marine Tech. Soc., New Orleans, LA, 1986.

Goodberlet, M.A. , C.T. Swift, and J.C. Wilkerson, Remote sensing of ocean surface winds with the Special Sensor Microwave/Imager, *J. Geophys. Res., 94,* 14,574-14, 555, 1989.

Goodberlet, M.A., and C.T. Swift, Improved retrievals from the DMSP wind speed algorithm under adverse weather conditions, *IEEE Trans. Geosci. Remote Sens., 30,* 1076-1077, 1992.

Hollinger, J.P., J.L. Peirce, and G.A. Poe, SSM/I instrument evaluation, *IEEE Trans. Geosci. Remote Sens.,GE-28,* 781-790, 1990.

Kerlirzin, P., and F. Vallet, Robustness in multilayer perceptrons, *Neural Computation, 5*, 473 - 482, 1993.

Krasnopolsky, V., L.C. Breaker, and W.H. Gemmill, Development of a single "all-weather" neural network algorithm for estimating ocean surface wind from the Special Sensor Microwave

Imager, Technical Note, OPC contribution No. 94, National Meteorological Center, Washington D.C., 1994.

Krasnopolsky, V., L.C. Breaker, and W.H. Gemmill, A neural network as a nonlinear transfer function model for retrieving surface wind speeds from the special sensor microwave imager, *J. Geophys. Res*, *100,* 11,033-11,045, 1995.

Monahan, E.C., and G. Mac Niocaill, *Oceanic Whitecaps,* D.Reidel, Dordrecht, Holland, 1986.

Monahan, E.C., and I.G. O'Muircheartaigh, Whitecaps and the passive remote sensing of the ocean surface, *Int. J. Remote Sensing, 7,* 627-642, 1986.

Monaldo, F., Expected difference between buoy and radar altimeter estimates of wind speed and significant wave height and their implication on buoy-altimeter comparisons, *J. Geophys. Res*, *93*, 2285-2302, 1988.

Overland, J.E., and W.H. Gemmill, Prediction of marine winds in the New York Bight, *Monthly Weather Review*, *105*, 1003-1008, 1977.

Pierson, W.J. Jr and W.B. Sylvester, Remote sensing research for NASA grant NAGW-690, Remote Sensing Laboratory report, The City College of CUNY, N.Y., 1995.

Petty, G.W., and K.B. Katsaros, The response of the SSM/I to the marine environment. Part II: A parametrization of the effect of the sea surface slope distribution on emission and reflection, *J. Atmos. Oceanic Technol., 11*, 617-628, 1994.

Swift, C.T., Passive microwave remote sensing of ocean surface wind speed, in *Surface Waves and Fluxes, v.II - Remote Sensing,* ed. by G.L. Geernaert, and W.J. Plant, pp. 265-292, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990.

Stogryn, A.P., C.T. Butler, and T.J. Bartolac, Ocean surface wind retrievals from special sensor microwave imager data with neural networks, *J. of Geophys. Res., 90,* 981-984, 1994.

Tolman, H., Statistical model validation techniques applied to marine wind analyses, 1995, submitted for publication.

Wasserman, P. D., *Neural Computing,* Van Nostrand Reinhold, New York, 1989.

Wentz, F.J., Measurement of oceanic wind vector using satellite microwave radiometers, *IEEE Trans. Geosci. Remote Sens., GE-30*, 960-972, 1992.

TABLE 1. Parameters of power law representing foam and whitecaps coverage as a function of the wind speed. The threshold values w_{max} are estimated for the coverage S = 50%.

Type of coverage	a	α	w_{max}	Reference
Foam & Whitecaps Tropical regions	6.78×10^{-3}	2.76	25.2 m/sec	Bortkovskii, 1987
Foam & Whitecaps Midlatitude regions	1.71×10^{-5}	4.43	28.8 m/sec	Bortkovskii, 1987
Foam & Whitecaps	7.751×10^{-4}	3.231	30.8 m/sec	Swift, 1990
Whitecaps	3.84×10^{-4}	3.4	31.9 m/sec	Monahan and Mac Niocaill, 1986

TABLE 2. Performance of GSW, SER NN, WNN4, and WNN5 algorithms for data under CLEAR conditions. Columns 2 - 4 show statistics for the wind speeds per se, and columns 5 - 7 for the difference between buoy and algorithm-generated wind speeds. Each cell in the table shows two numbers, one (above) for training and second (below) for testing (F8 development set). CC denotes correlation coefficient, and σ_w denotes standard deviation.

	Max W	Mean W	σ_w	Bias	RMS	CC
Buoy	17.4	6.50	2.94	N/A	N/A	N/A
	19.3	6.59	2.99			
GSW	19.3	7.47	3.46	-0.97	1.99	0.86
	22.0	7.65	3.65	-1.06	2.13	0.86
SER NN	14.5	6.46	2.58	0.04	1.38	0.88
	16.1	6.59	2.69	-0.01	1.41	0.88
WNN4	16.4	6.31	3.00	0.18	1.44	0.88
	18.3	6.46	3.15	0.12	1.52	0.88
WNN5	16.6	6.43	2.73	0.07	1.38	0.88
	18.8	6.55	2.85	0.03	1.44	0.88

TABLE 3. Performance of GSW, SER NN, WNN4, and WNN5 algorithms for CLEAR plus CLOUDY case. Columns 2 - 4 show statistics for the wind speeds per se, and columns 5 - 7 for the difference between buoy and algorithm-generated wind speeds. Each cell in the table shows two numbers, one (above) for training and second (below) for testing (F8 development set). CC denotes correlation coefficient, and σ_w denotes standard deviation. High wind speed statistics (for w > 15 m/s) are shown in the lower part.

TOTAL	Max W	Mean W	σ_w	Bias	RMS	CC
Buoy	21.2 19.3	6.78 6.90	3.11 3.15	N/A	N/A	N/A
GSW	23.5 28.8	8.10 8.32	3.90 4.06	-1.32 -1.42	2.56 2.69	0.83 0.83
SER NN	15.4 16.1	6.75 6.91	2.71 2.81	0.03 -0.01	1.52 1.56	0.87 0.87
WNN4	18.0 18.3	6.68 6.86	3.22 3.62	0.10 0.04	1.62 1.70	0.87 0.86
WNN5	19.0 18.8	6.73 6.88	2.96 3.07	0.05 0.02	1.56 1.63	0.87 0.86
W > 15	**Max W**	**Mean W**	σ_w	**Bias**	**RMS**	**CC**
Buoy	21.2 19.3	16.4 16.4	1.60 1.13	N/A	N/A	N/A
GSW	25.3 23.3	17.8 17.7	2.58 3.39	-1.50 -1.32	2.84 3.66	0.38 0.08
SER NN	16.6 15.5	13.8 12.9	1.35 2.10	3.08 3.51	3.52 4.13	0.27 0.16
WNN4	18.1 18.3	14.9 14.5	1.64 2.61	1.39 1.92	2.30 3.23	0.33 0.17
WNN5	19.0 18.8	14.8 14.2	1.83 2.59	1.42 2.22	2.30 3.40	0.43 0.18

TABLE 4. Theoretical high wind speed upper limit W_Ω for three NN algorithms.

Algorithm	**SER NN**	**WNN4**	**WNN5**
W_Ω **in m/sec**	19.4	24.5	26.5

TABLE 5. Total statistics for WNN5, BWNN4 and OMB algorithms and high wind speed statistics for GSW, BWNN4, and OMB algorithms for CLEAR plus CLOUDY conditions. Columns 2 - 4 show statistics for the wind speeds per se, and columns 5 - 7 for the difference between buoy and algorithm-generated wind speeds. Each cell in the table shows two numbers, one (above) for training and second (below) for testing (F8 development set). CC denotes correlation coefficient, and σ_w denotes standard deviation.

TOTAL	Max W	Mean W	σ_w	Bias	RMS	CC
Buoy	21.2	6.78	3.11	N/A	N/A	N/A
	19.3	6.90	3.15			
WNN5	19.0	6.73	2.96	0.05	1.56	0.87
	18.8	6.88	3.07	0.02	1.63	0.86
BWNN4	20.6	6.99	3.50	-0.21	1.74	0.87
	20.5	7.19	3.62	-0.29	1.83	0.86
OMB	22.3	6.91	3.34	-0.13	1.66	0.87
	22.0	7.07	3.44	0.17	1.74	0.86
W > 15	**Max W**	**Mean W**	σ	**Bias**	**RMS**	**CC**
Buoy	21.2	16.4	1.60	N/A	N/A	N/A
	19.3	16.4	1.13			
GSW	25.3	17.8	2.58	-1.50	2.84	0.38
(w>15)	23.3	17.7	3.39	-1.32	3.66	0.08
BWNN4	20.6	15.8	2.11	0.43	2.14	0.36
(w>15)	20.5	15.4	2.95	0.98	3.04	0.20
OMB	22.3	16.0	2.48	0.28	2.27	0.43
(w>15)	22.0	15.2	3.10	1.19	3.21	0.22

FIGURE CAPTIONS

Fig. 1. Four parameterizations that express whitecap foam coverage versus the wind speed (see Table 1 also). All measurements are concentrated at the coverage between 0 and 30%. Extrapolations up to 100% coverage are shown. Solid line shows parameterization 1, dashed line - 2, dash-dotted line - 3, and dash-triple dotted line - 4.

Fig. 2. TB cluster centroid (see (4) in text) as a function of wind speed (a) and its projection on seven TB axes (b-h) - solid lines. Dashed lines show centroid plus (upper dashed line) minus (lower dashed line) the mean radius of the cluster (see (5) in the text).

Fig. 3. Binned bias versus wind speed (bin size = 1 m/sec). Dashed line shows the bias for the training set , dash-dotted line - for the test set, and the solid line is the bias correction derived from the training set (see (8) in the text).

Fig. 4. Binned scatter plot (bin size = 2 m/sec) of SSM/I wind speed versus buoy wind speed. Diamonds refer to the SER NN algorithm, crosses - WNN5, and asterisks - OMB algorithm.

Fig. 5. Distribution of the wind speed derived from the F8 matchup data base. Solid line refers to the observed buoy wind speed distribution, dotted line - distribution generated by SER NN algorithm, dashed - by WNN5, and dash-dotted - by OMB algorithm.

Fig. 6. Bias and rms error statistics for F10 matchup data base as functions of spatial (R_s in km) and temporal (R_t in hours) matchup uncertainties. (a, b) show statistics for clear conditions, (c, d) - for clear plus cloudy conditions, and (e, f) - for high wind speeds (w > 15 m/sec).

Fig.7. A synoptic event which occurred SW of Australia on 24 August 1995 as observed by the F10 (upper row) and F13 (lower row) SSM/I instruments. (a, d) show the

wind speed field retrieved by the GSW algorithm, (b, e) - by the SER NN, and (c, f) - by the OMB algorithm.

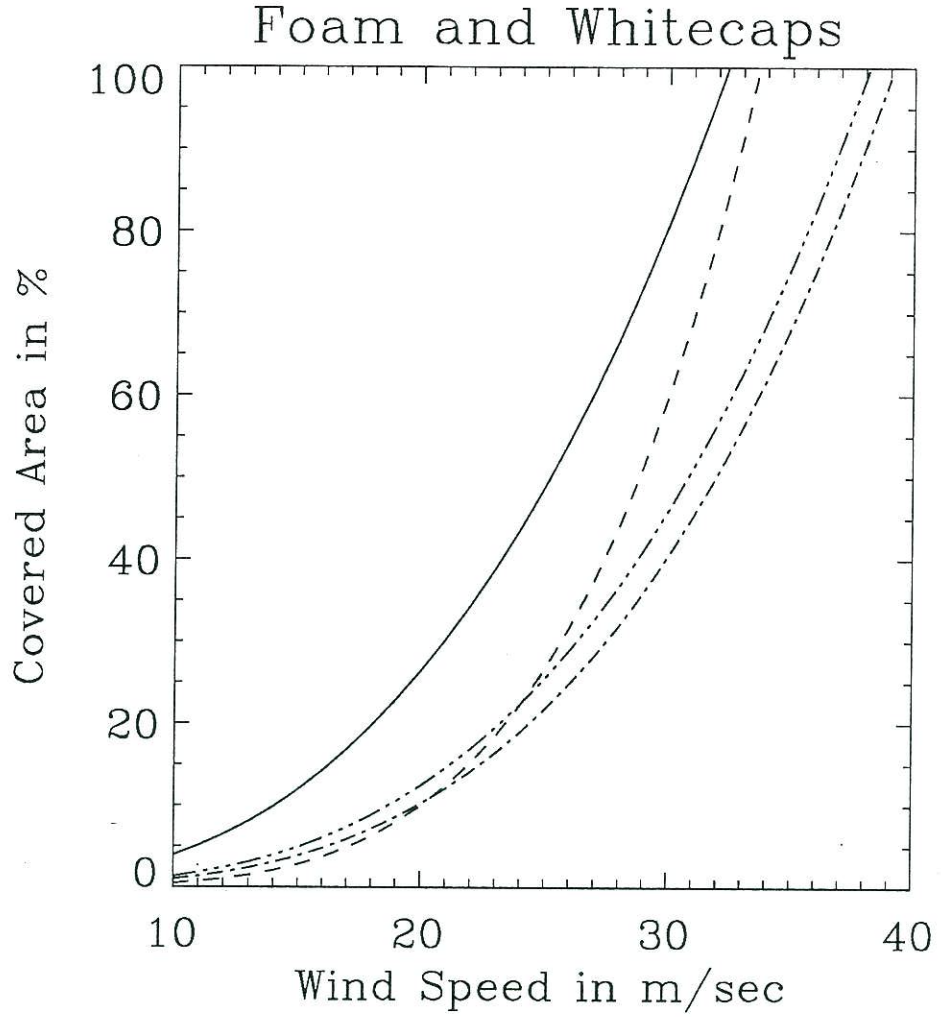
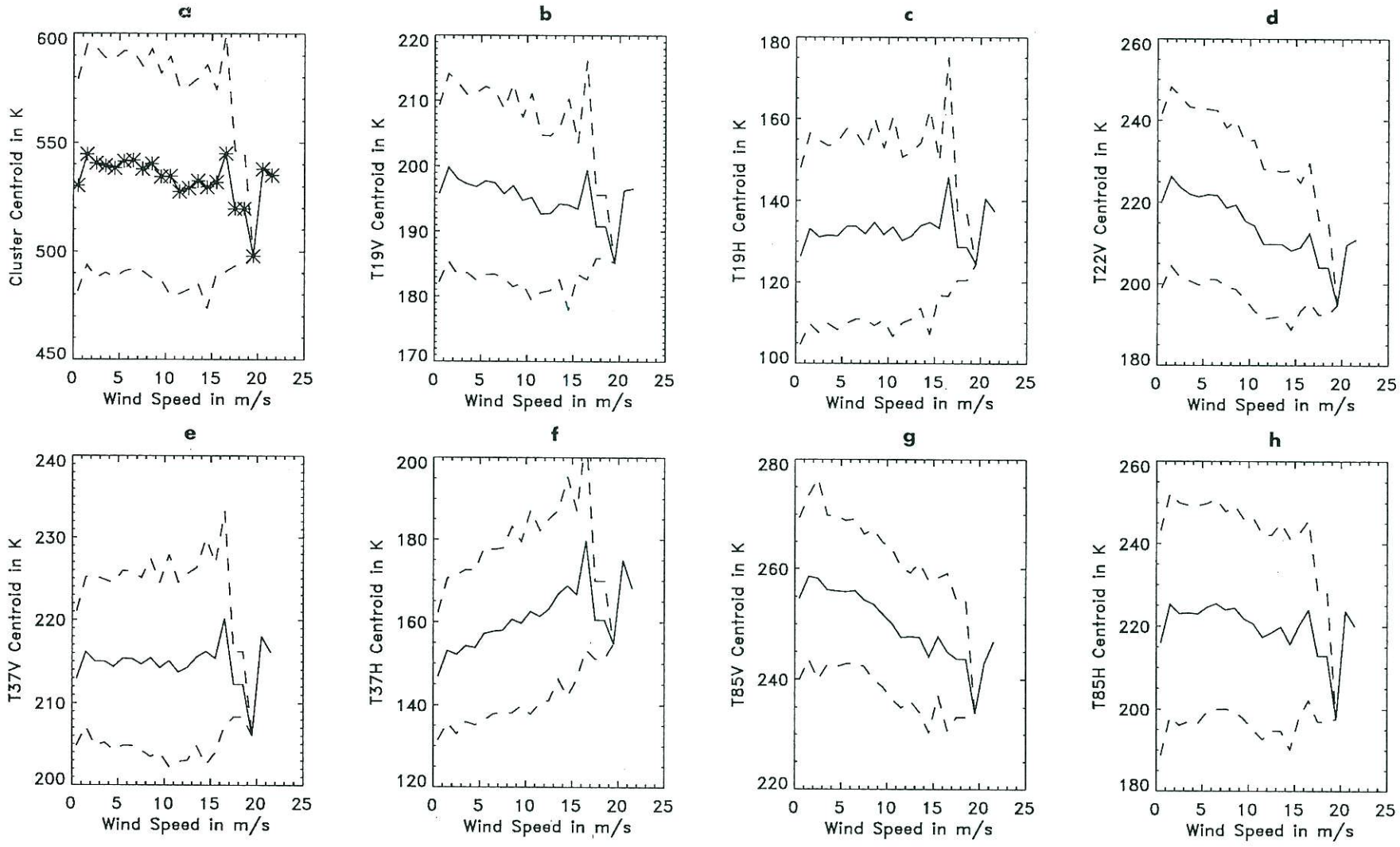

Foam and Whitecaps
Covered Area in %
100
80
60
40
20
0
10
20
30
40
Wind Speed in m/sec

Figure 1.

Figure 2.


a
Cluster Centroid in K
600
550
500
450
0 5 10 15 20 25
Wind Speed in m/s
b
T19V Centroid in K
220
210
200
190
180
170
0 5 10 15 20 25
Wind Speed in m/s
c
T19H Centroid in K
180
160
140
120
100
0 5 10 15 20 25
Wind Speed in m/s
d
T22V Centroid in K
260
240
220
200
180
0 5 10 15 20 25
Wind Speed in m/s
e
T37V Centroid in K
240
230
220
210
200
0 5 10 15 20 25
Wind Speed in m/s
f
T37H Centroid in K
200
180
160
140
120
0 5 10 15 20 25
Wind Speed in m/s
g
T85V Centroid in K
280
260
240
220
0 5 10 15 20 25
Wind Speed in m/s
h
T85H Centroid in K
260
240
220
200
180
0 5 10 15 20 25
Wind Speed in m/s

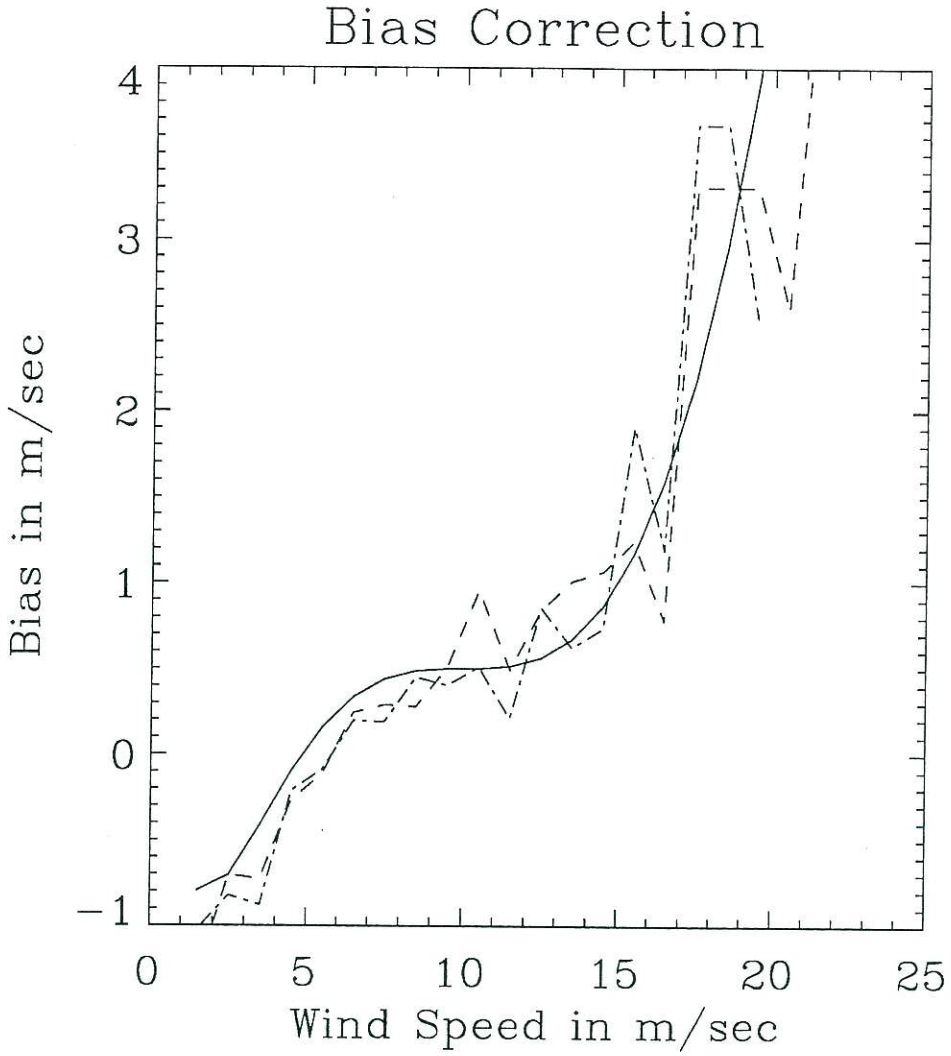

Bias Correction
Bias in m/sec
4
3
2
1
0
−1
0
5
10
15
20
25
Wind Speed in m/sec

Figure 3.

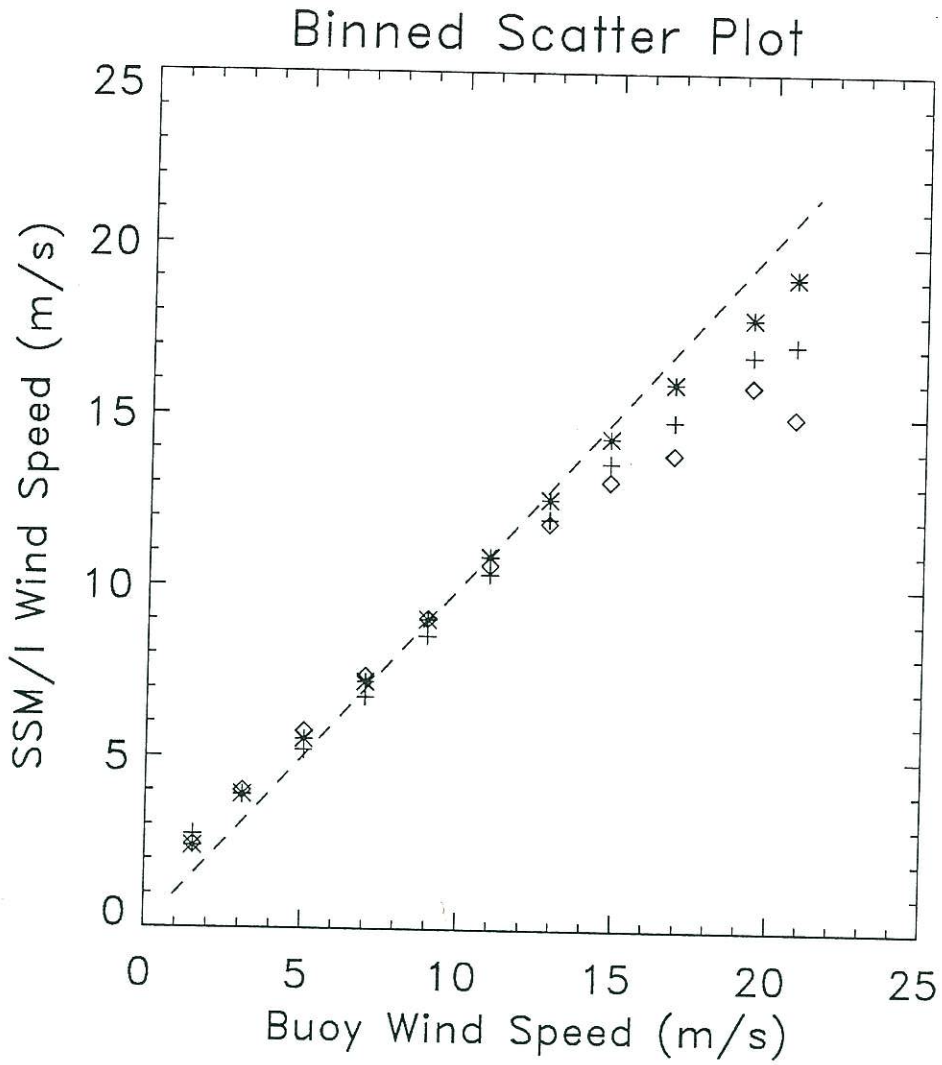

Binned Scatter Plot
SSM/I Wind Speed (m/s)
Buoy Wind Speed (m/s)
0
5
10
15
20
25

Figure 4.

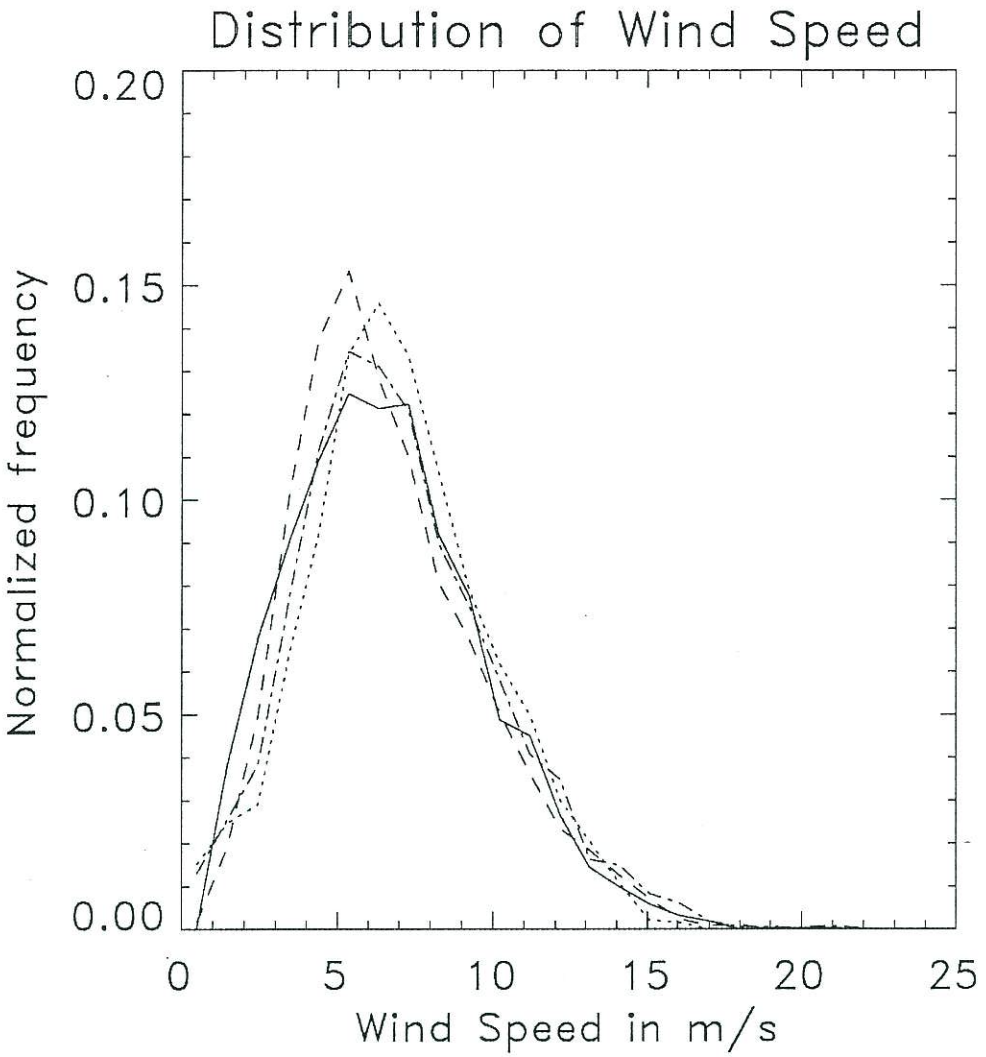

Distribution of Wind Speed
Normalized frequency
0.20
0.15
0.10
0.05
0.00
0
5
10
15
20
25
Wind Speed in m/s

Figure 5.

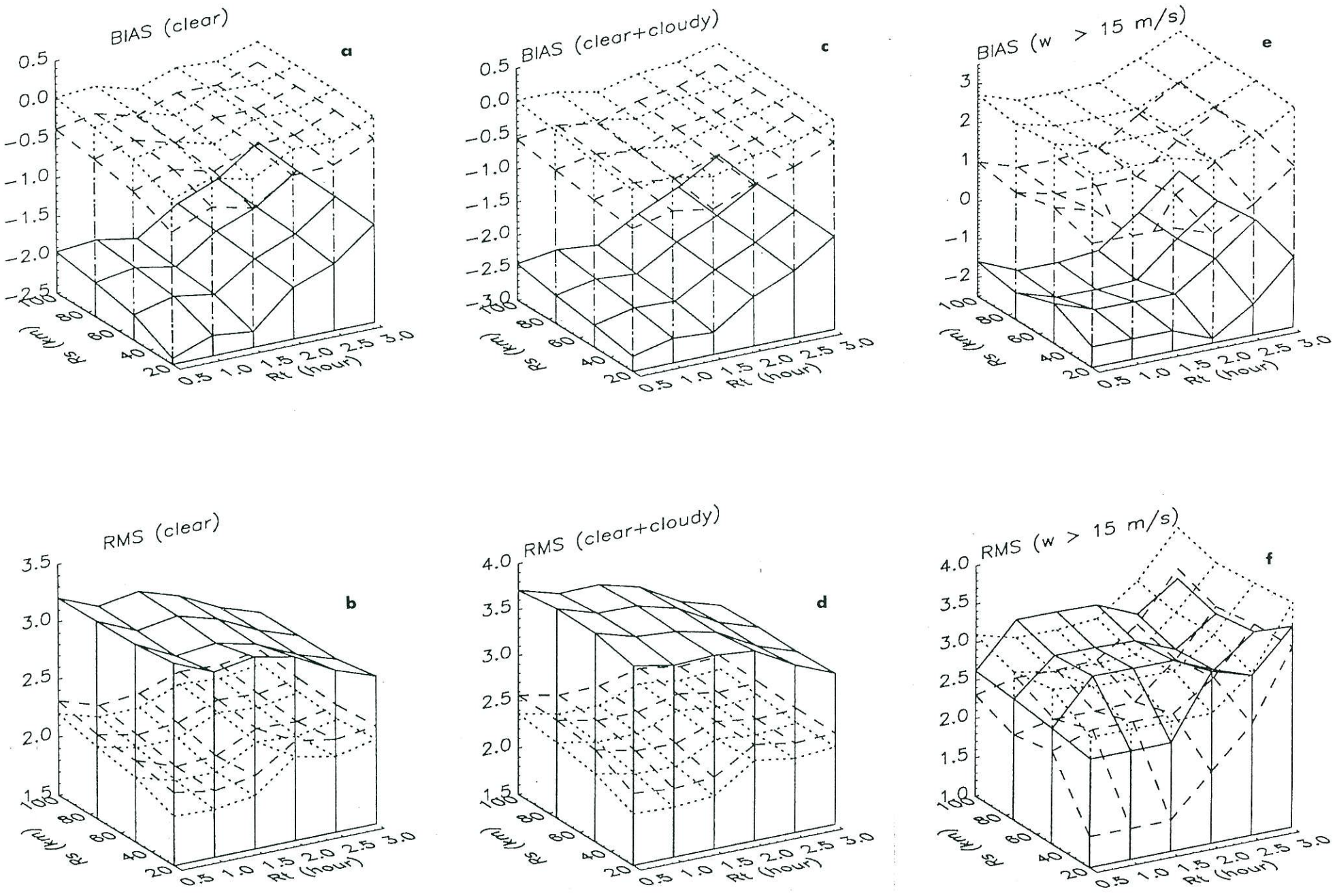

BIAS (clear)
a
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)
BIAS (clear+cloudy)
c
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)
BIAS (w > 15 m/s)
e
3
2
1
0
−1
−2
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)
RMS (clear)
b
3.5
3.0
2.5
2.0
1.5
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)
RMS (clear+cloudy)
d
4.0
3.5
3.0
2.5
2.0
1.5
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)
RMS (w > 15 m/s)
f
4.0
3.5
3.0
2.5
2.0
1.5
1.0
100
80
60
40
20
R (km)
0.5 1.0 1.5 2.0 2.5 3.0
Rf (hour)

Figure 6.

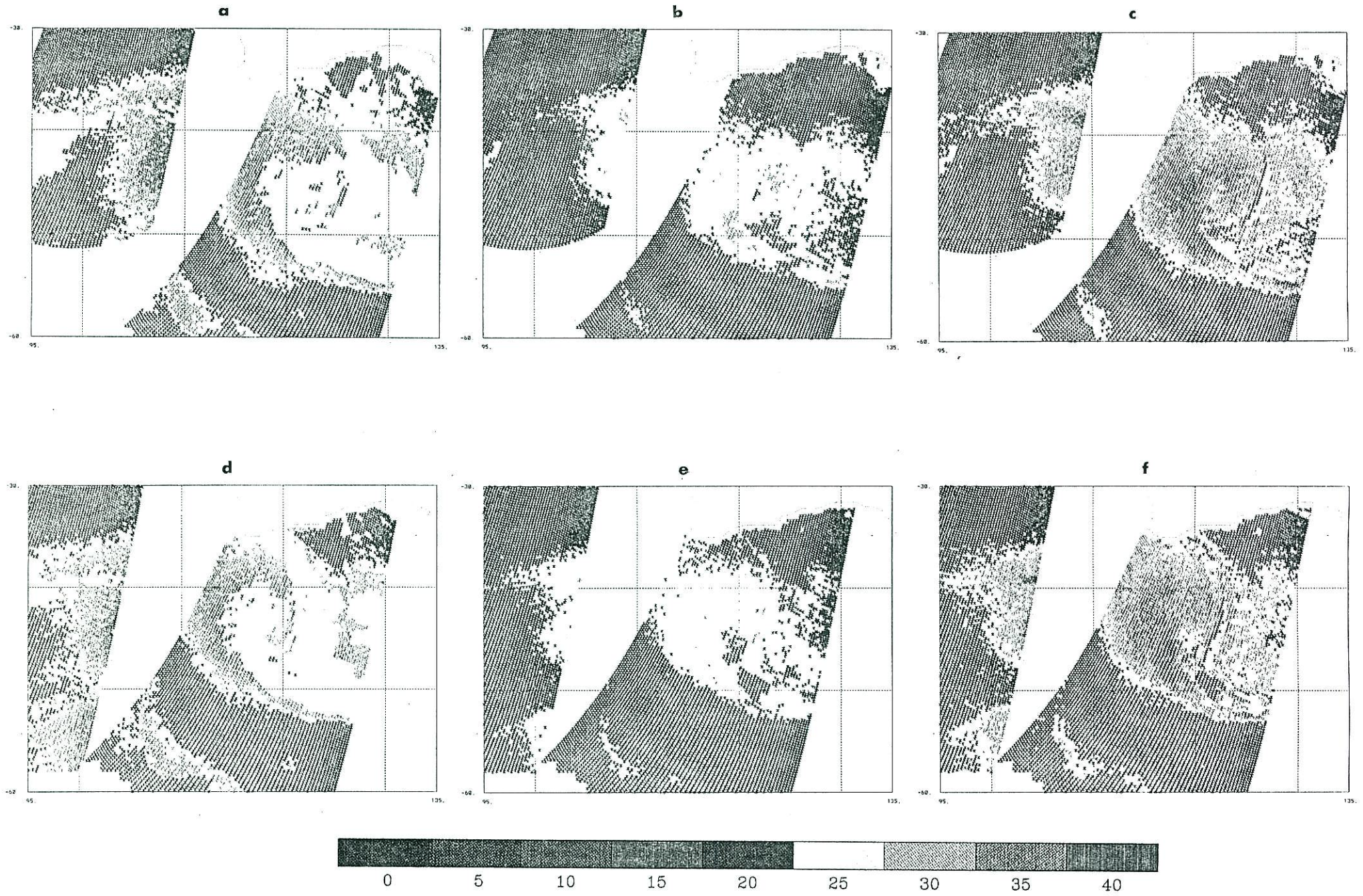


Figure 7.