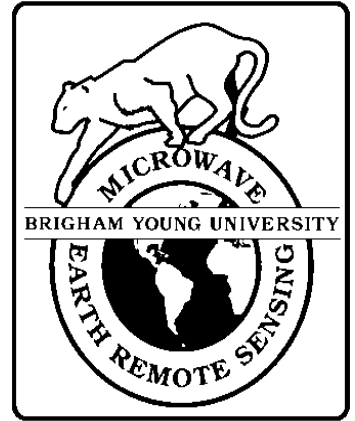




**Brigham Young University  
Department of Electrical and  
Computer Engineering**

**459 Clyde Building  
Provo, Utah 84602**



# **The Probability Distribution of NSCAT Measurements**

**Paul E. Johnson and  
David G. Long**

**20 March 1999**

**MERS Technical Report # MERS 99-05  
ECEN Department Report # TR-L110-99.5**

---

**Microwave Earth Remote Sensing (MERS)  
Laboratory**

© Copyright 1999, Brigham Young University. All rights reserved.

# The Probability Distribution of NSCAT Measurements

Paul E. Johnson and David G. Long

Microwave Earth Remote Sensing (MERS) Laboratory

Brigham Young University, Electrical and Computer Engineering Dept.

20 March 1999

## Abstract

NSCAT makes only indirect measurements of wind. The direct measurement is of the backscattered radar power. The signal power is contaminated by radiometric noise so a separate measurement of the noise power is subtracted from the signal-plus-noise measurement to estimate the backscattered power. Using the radar equation,  $\sigma^o$  is computed from the measured signal power. From multiple  $\sigma^o$  measurements made at different azimuth angles, the wind is estimated. In wind retrieval, the NSCAT  $\sigma^o$  measurements are assumed to have a Gaussian probability distribution with a variance which depends on the mean. Given this distribution model, the maximum-likelihood estimator is formed and optimized to estimate the wind.

Because of the on-board signal processing used by NSCAT, the Gaussian distribution model for the measurements is only an approximation to the actual distribution. Working from first principles and the design of the NSCAT signal processor we derive the distribution of the NSCAT measurements as a function of the surface  $\sigma^o$ , the signal to noise ratio and the cell number. The resulting distribution is skewed relative to the traditional Gaussian model. Simple compass simulations are used to compare the accuracy of winds estimated using the actual and Gaussian model distributions.

## 1 Introduction

Estimation of geophysical parameters based on their modulation of NRCS requires accurate measurements of the power scattered from the surface. Inherent in such measurements is considerable uncertainty. While the signal transmitted from the scatterometer is deterministic, the received signal is random based on the scattering from the random surface. The scattered power is accurately described as a white, Gaussian signal. Resolution of the antenna footprint into ocean cells is typically done by Doppler filtering. Further, an estimate of the noise in the receiver must be subtracted from the noisy measurement. These complications modify the statistics of the backscatter measure-

ments. For lack of a better choice, and with only a sample mean and variance of the measurements, the pdfs traditionally have been assumed Gaussian.

However, the pdf of the sum of frequency bins, based on Welch’s method with  $K$  overlapping data segments, and an arbitrary data window, is found to be clearly non-Gaussian (Johnson and Long, 1999). In this paper I apply the theory to the specific case of NSCAT processing.

Wind retrieval with NSCAT involves Doppler filtering to resolve ocean cells; the signal-plus-noise measurement is made as the sum a range of frequency bins of the estimated power spectrum (Naderi et al., 1991). Welch’s method for periodogram estimation (Welch, 1967) is used with 50% overlapping data segments and a Hann window to minimize spectral leakage (Long et al., 1988). The signal power for a cell is then computed by subtracting a noise-only power estimate from the signal-plus-noise estimate. This signal power, used in the wind estimation, has been assumed to have a Gaussian distribution. However, Welch’s periodogram estimate is not Gaussian (Johnson and Long, 1999). In this paper, the probability density function of NSCAT power estimates is found, and the effect on wind retrieval is considered. The development is made with sufficient generality to readily extend to other applications of periodogram estimation.

The pdfs of a noise-only measurement and of a signal-plus-noise measurement, appropriate for NSCAT processing, only require simple scale changes from basic pdfs; the pdf of the signal-only measurement is then found as the correlation of these two pdfs. Here I describe the scale changes necessary based on the number of overlapping data segments and frequency bins used (these vary throughout the data set). I compare these density function with Gaussian densities of the same mean and variance. I also report compass simulation results which describe the impact of using the incorrect pdf on wind retrieval.

## 2 Conditional PDF of NSCAT Measurements

For NSCAT, two power estimates are made, a signal-plus-noise measurement and a noise-only measurement. The power in the signal is estimated as the difference between these two power estimates:  $P = P_1 - P_2$ . Because  $P_1$  and  $P_2$  are independent random variables, distributed as described in (Johnson and Long, 1999), the pdf of the signal-only power is the convolution of the first pdf with that of the second pdf with a negative argument.

NSCAT employs both time and frequency filtering to resolve a wind cell. Depending on the cell location within the swath, the estimate of the signal-plus-noise,  $P_1$ , is based on either 2, 3 or 7 overlapping data segments ( $L = 256$ ) and from 5 to 20 frequency bins, and is improved by averaging 25 pulses (Long et al., 1988).

A single pulse would result in a pdf as given by

$$f_{\text{basic}}(p) = g \sum_{i=1}^D \sum_{j=1}^{h_i} A_{ij} \frac{1}{(j-1)!} p^{j-1} e^{-d_i p} U(p). \quad (1)$$

The subscript ‘basic’ indicates that this basic pdf will be scaled for given signal power. The eigenvalues and their multiplicities, for a single pulse are based on  $K = 2, 3$  or 7 data segments of length

| $K = 2$     |      | $K = 3$     |      | $K = 7$     |      |
|-------------|------|-------------|------|-------------|------|
| Eigenvalues | Mult | Eigenvalues | Mult | Eigenvalues | Mult |
| 0.5932      | 2    | 0.3968      | 2    | 0.1688      | 14   |
| 0.5860      | 2    | 0.3935      | 2    | 0.1216      | 14   |
| 0.4252      | 2    | 0.3901      | 2    | 0.0877      | 4    |
| 0.4052      | 2    | 0.2890      | 2    | 0.0727      | 2    |
| 0.2578      | 2    | 0.2795      | 2    | 0.0608      | 2    |
| 0.1495      | 2    | 0.2721      | 2    | 0.0488      | 2    |
| 0.0613      | 2    | 0.1916      | 2    | 0.0374      | 2    |
| 0.0185      | 2    | 0.1378      | 2    | 0.0226      | 4    |
| 0.0030      | 2    | 0.0849      | 2    | 0.0078      | 6    |
| 0.0002      | 2    | 0.0413      | 2    | 0.0003      | 20   |
|             |      | 0.0173      | 2    |             |      |
|             |      | 0.0051      | 2    |             |      |
|             |      | 0.0009      | 2    |             |      |
|             |      | 0.0001      | 4    |             |      |

Table 1: *Eigenvalues of the basic signal-plus-noise pdfs for a single pulse with typical NSCAT parameters:  $K = 2, 3$  and  $7$ ,  $b = 5$ ,  $L = 256$ ,  $r = 0.5$ , and a Hann window.*

$L = 256$  and  $b = 5$  to 20 frequency bins,  $r = 0.5$  for 50% and NSCAT applies a Hann data window to minimize spectral leakage. Equation (1) then provides the pdf of the generic signal-plus-noise, with mean equal to  $b$ . For example, Table 1 lists the eigenvalues corresponding to the basic signal-plus-noise for the case of 5 frequency bins and a single scatterometer pulse. The corresponding pdfs are plotted in Fig. 1, with each case having a mean equal to the number of frequency bins, 5.

Because NSCAT averages 25 signal-plus-noise pulses, it is useful to define the normalized pdf based on the self-convolution of 25 pulses (i.e., the pulses are assumed i.i.d. so that the average power can be computed by convolution) and normalized by  $b$  frequency bins, so the normalized distribution has a mean of 1:

$$f_{\text{norm}}(p) = 25b f_{\text{basic}}^{(25)}\left(\frac{p}{25b}\right), \quad (2)$$

where the superscript in parenthesis indicates self-convolution. The pdf of the signal-plus-noise, given signal and noise powers,  $P_S$  and  $P_N$ , is

$$f_{\text{SN}}(p|P_S, P_N) = \frac{25b}{P_S + P_N} f_{\text{basic}}^{(25)}\left(\frac{p}{25b}(P_S + P_N)\right). \quad (3)$$

The noise power estimate,  $P_2$ , averages 4 ‘pulses’ and sums more than 200 frequency bins. For this estimate, the scatterometer transmits nothing, but takes measurements according to the same timing as during the signal-plus-noise measurement.

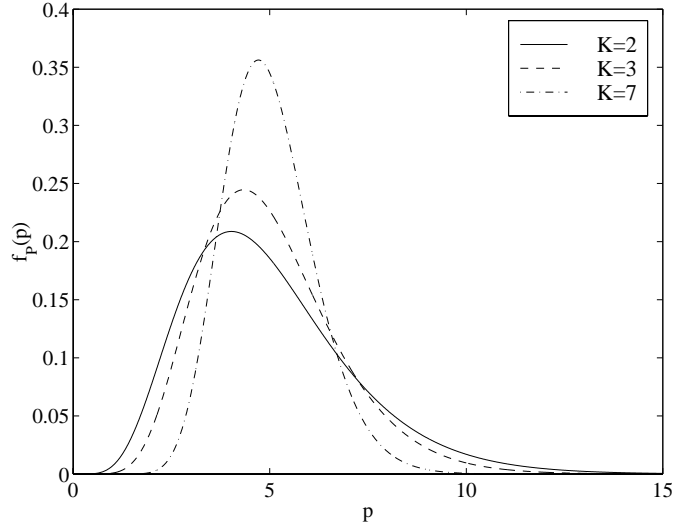


Figure 1: *Generic signal-plus-noise pdfs for a single pulse with typical NSCAT parameters*

The pdf of the signal-only power,  $f_S(p)$ , is then the convolution of  $f_{SN}(p)$  with the pdf of the noise estimate with a negative argument:  $f_S(p) = f_{SN}(p) * f_N(-p)$ . Because the noise estimate sums 200 frequency bins and 4 pulses, the noise is essentially Gaussian with a variance which becomes negligibly small. The pdf  $f_N(p)$  can be well approximated with a delta function at the noise power. The signal-only pdf is then simply

$$f_S(p|P_S, P_N) \approx \frac{25b}{P_S + P_N} f_{\text{basic}}^{(25)} \left( \frac{p}{25b} (P_S + P_N) + P_N \right) \quad (4)$$

$$= \frac{25b}{P_N (1 + SNR)} f_{\text{basic}}^{(25)} \left[ P_N \left( \frac{p}{25b} (1 + SNR) + 1 \right) \right]. \quad (5)$$

Figure 2 displays several examples of the density functions of power estimates. The near, mid and far swath cells use 7, 3 and 2 overlapping data segments, respectively, and 5 frequency bins. Note that the SNR simply scales the power estimate for each case. The dashed line indicates a Gaussian density with the same mean and variance. It is clear that the actual distribution is skewed toward low  $\sigma^o$  values. For near swath cells, the true distribution is more Gaussian.

Two useful results related to the density function are the normalized bias and the probability of negative measurements. Defining the normalized bias as the difference between the mode and the mean, normalized by the mean of the distribution, a simple measure is obtained of the distance between a Gaussian and the correct density. Also of interest is the probability of a negative power estimate. This can be computed as the integral of the pdf from  $-\infty$  to 0, or equivalently as the cumulative distribution function evaluated at zero  $F(0)$ . These measures are displayed in Fig. 3 as functions of the SNR for the three representative cell cases.

The concern is that because the actual density is skewed, the wind estimation may be biased. A compass simulation shows that this bias is small, though non-zero. Using a particular measurement

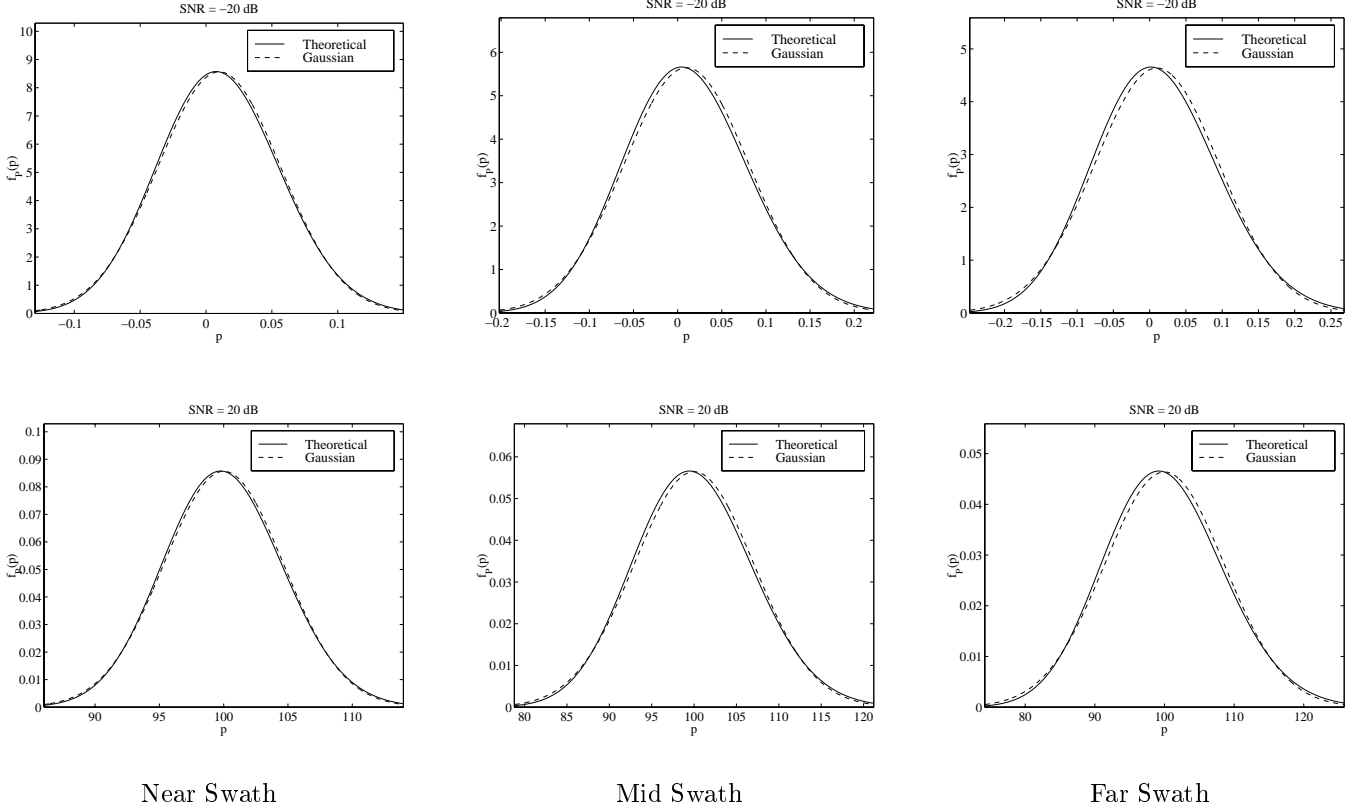


Figure 2: *Some sample pdfs for the signal-only power estimate based on NSCAT processing. Because there is less averaging for far swath cells (fewer data segments and fewer frequency bins used), the skew increases. For the different parts of the swath, the SNR scales the fundamental pdf.*

geometry and noise parameters (taken from an NSCAT L1.7 file) simulated backscatter measurements for a given wind vector cell are used to estimate the wind in a traditional compass simulation. The simulated measurements were then biased, according to Fig. 3, and the wind is estimated with this set of biased measurements. Note that a more precise method would be to use the correct pdf, rather than simply a shifted Gaussian distribution; but to obtain a reasonably accurate result with minimal code modification for this preliminary study, this approach was chosen. Figure 4 displays the results of this compass simulation for a cell in the far swath (2 overlapping data segments and 5 frequency bins). The plots show the difference between the wind estimates made without the correcting bias and the wind estimates with the correcting bias for both speed and direction. As the simulated wind speed increases, the difference between the two estimates increases, while the direction difference is minimal.

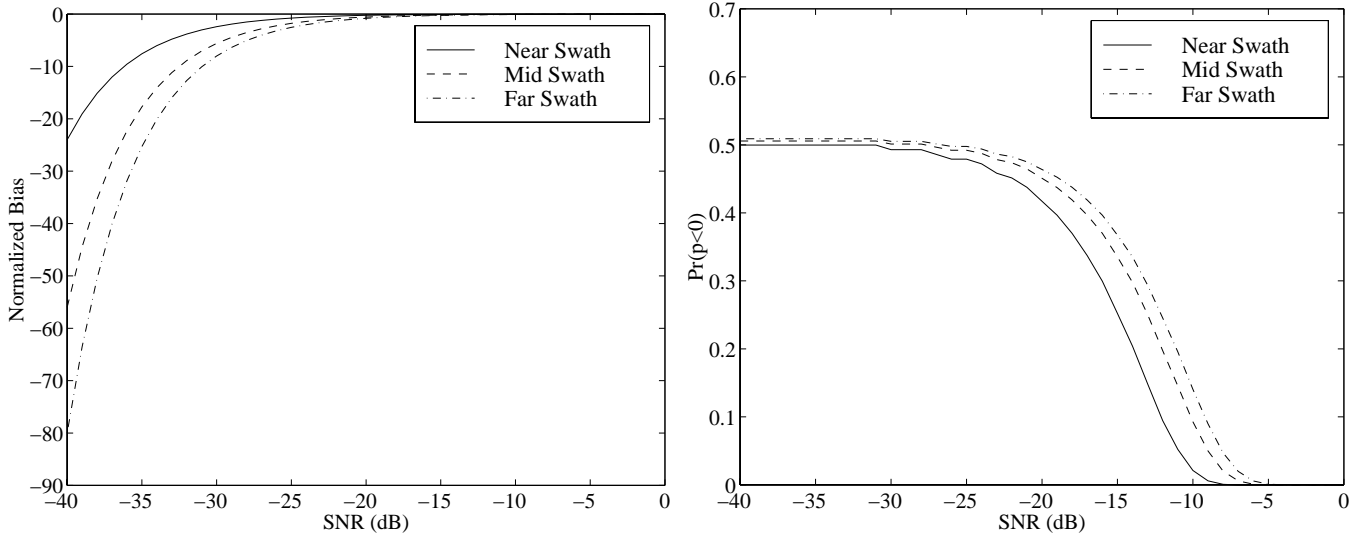


Figure 3: *Left: The normalized bias, defined as the difference in the modes of the theoretical and Gaussian densities, normalized by the mean, depends on the SNR and has similar behavior for all cells. Right: The probability of negative power estimates as a function of the signal to noise ratio.*

### 3 Impact of Wind Distribution

In the previous section, the conditional probability density function of the signal power estimate was derived for NSCAT processing as Eq. (5). In terms of wind retrieval, this is the conditional pdf which, by adjusting the wind estimate, results in the maximum likelihood estimate of the wind.

The pdf of the measurements, then, is the integral of this conditional pdf multiplied by the prior pdf of the signal power:

$$f_S(p) = \int_{-\infty}^{\infty} f_S(p|P_S) f(P_S) dP_S. \quad (6)$$

The distribution of the signal power is not known. It will depend on the distribution of the wind and on the model function relating the wind to the backscatter. In this section I develop the backscatter measurements,  $z$ , based on prior distributions of the wind and the empirical model function.

The radar equation, states that the backscatter is linearly related to the power measurement,  $\sigma^o = XP_S$ . Transforming the pdf of the measurements to backscatter space, and assuming the noise power is deterministic, yields:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{25b}{X(P_S + P_N)} f_{\text{basic}}^{(25)} \left( \frac{z}{25b} (P_S + P_N) + P_N \right) f(P_S) dP_S. \quad (7)$$

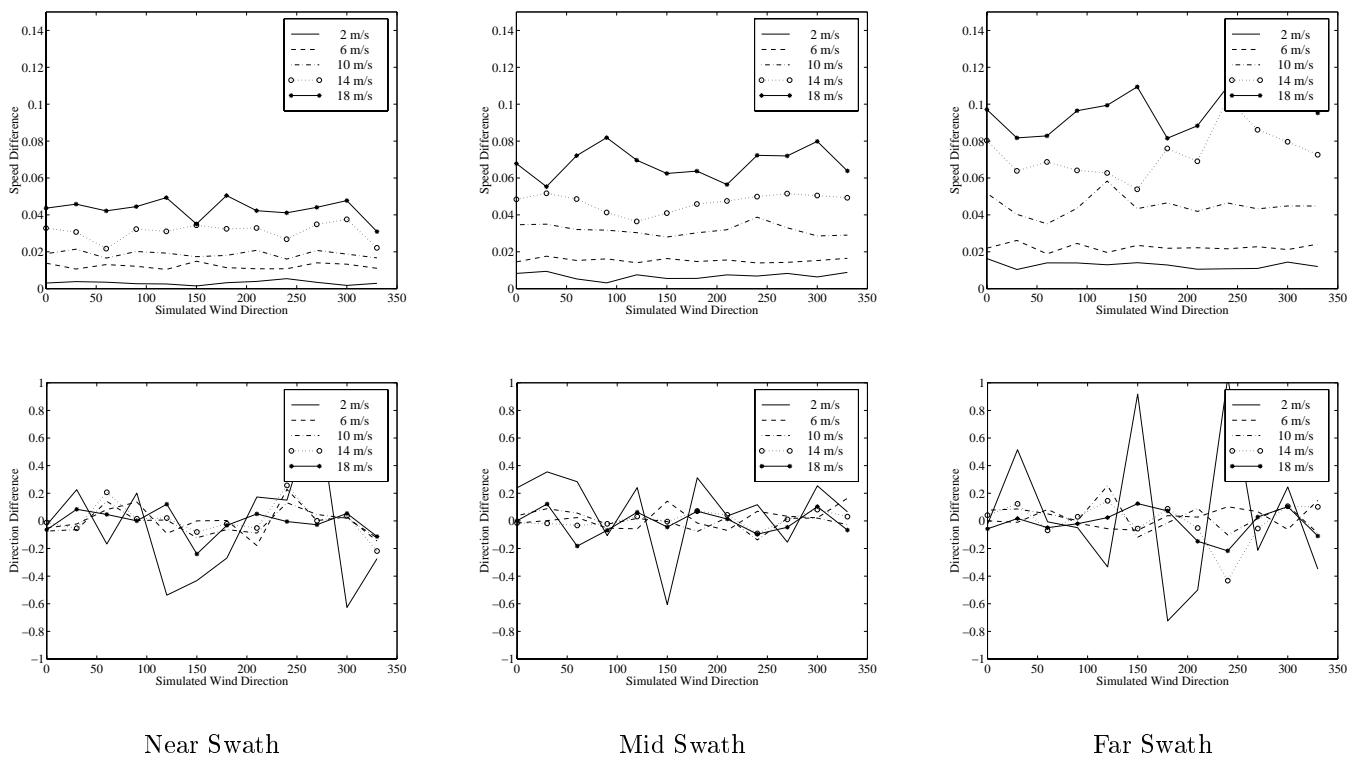


Figure 4: *Impact on wind retrieval with bias error in the pdf. For the three parts of the swath, results from compass simulations are displayed showing the difference between the wind retrieved using traditional methods and that retrieved assuming a biased Gaussian distribution for the scatterometer measurements. The latter more closely approximates the correct distribution.*

The pdf of the signal power is related simply to the pdf of the backscatter (as a function of the wind). For Fig. 5 I established a set of wind vectors with realistic statistics (plotted on the left), with a Rayleigh distribution for the wind speed and a uniform distribution for the direction. Using a typical NSCAT measurement geometry for a single beam (at near-swath) I computed the resulting backscatter according to the empirical model function. The plots on the right display the sample pdf of the backscatter. The evidence is growing for a limit on the wind speed, below which the NRCS and wind speed are only slightly correlated; the bottom plots represent this case where winds below a threshold of 3 m/s do not contribute to the backscatter.

## 4 Conclusion

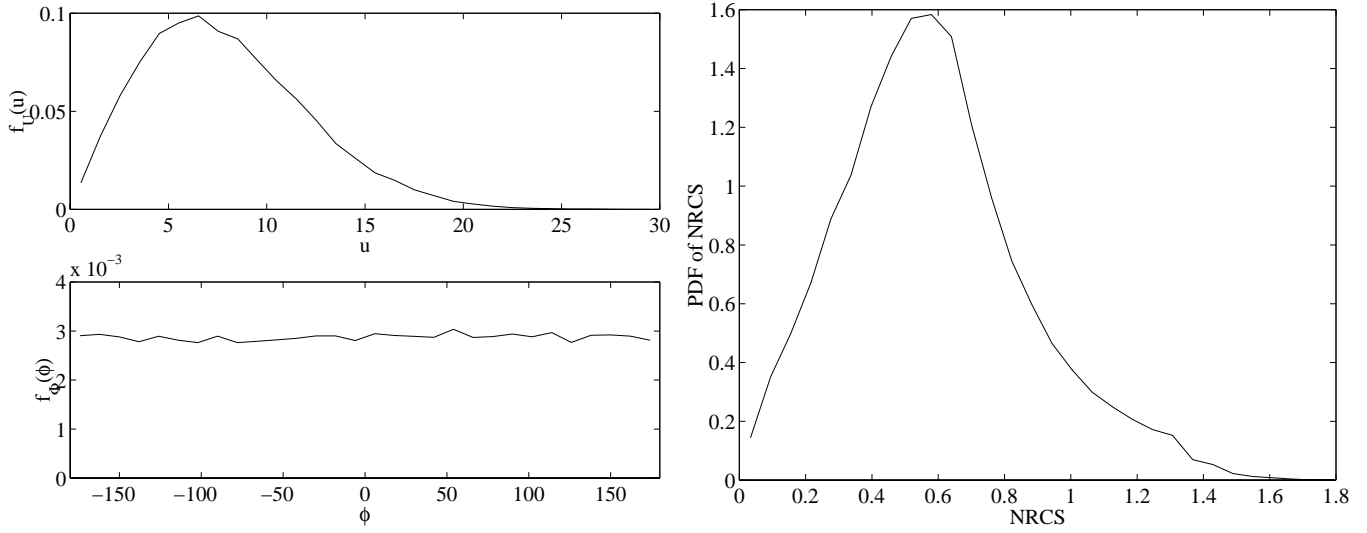
The probability structure of NSCAT signal estimates is not biased, but shows considerable skew. Initial observations suggest that this has negligible impact on the estimated wind direction, though the wind speed bias should be investigated further. The skew of the probability density function of



the power measurements causes current wind retrieval techniques (which assume measurements of Gaussian random variables) to be biased high. This effect is ameliorated to some extent by models like NSCAT1, which are tuned to the retrieved wind, as well as the considerable averaging resulting in nearly Gaussian densities.

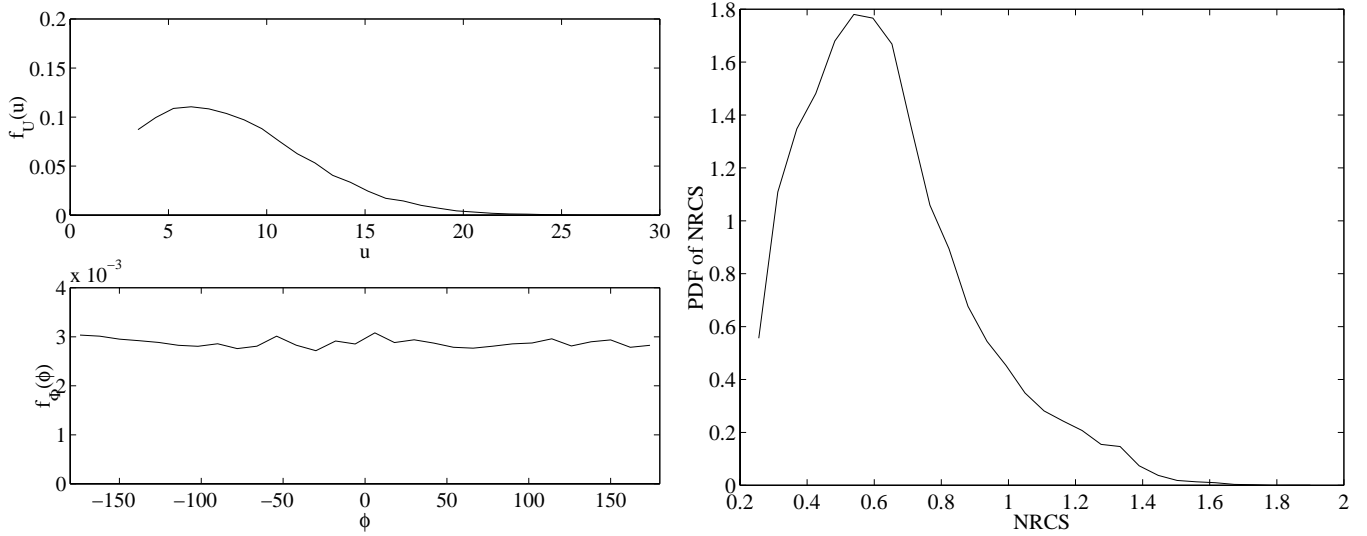
## References

- Johnson, P. E. and D. G. Long, “The Probability Density of Welch’s Modified Periodogram Estimates.” *IEEE Transactions on Signal Processing*, 1999, In Press.
- Long, D. G., C.-Y. Chi, and F. K. Li, “The Design of an Onboard Digital Doppler Processor for a Spaceborne Scatterometer.” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 6, pp. 869–878, 1988.
- Naderi, F. M., M. H. Freilich, and D. G. Long, “Spaceborne Radar Measurement of Wind Velocity Over the Ocean—An Overview of the NSCAT Scatterometer System.” *Proceedings of the IEEE*, vol. 79, no. 6, pp. 850–866, 1991.
- Welch, P. D., “The use of the fast Fourier transform for the estimation of power spectra.” *IEEE Trans. Audio Electroacoustics*, vol. AU-15, pp. 70–73, 1967.



PDF of Simulated Wind ( $U_0 = 0$ )

PDF of Backscatter ( $U_0 = 0$ )



PDF of Simulated Wind ( $U_0 = 3$ )

PDF of Backscatter ( $U_0 = 3$ )

Figure 5: *Sample estimate of the pdf of the model function backscatter.*