Investigations of temperature and backscatter correlation in the dry snow zone of the Greenland ice sheet

Kevin R. Moon, David G. Long

***Abstract*— Due to system degradation, satellite-borne scatterometers require post-launch calibrations to maintain accuracy. The dry snow zone of the Greenland ice sheet has been used for calibration due to its relatively constant backscatter properties. However, we recently discovered that some of the variation in the dry snow zone backscatter is seasonal. This paper uses correlation analysis to investigate the relationship between temperature and backscatter in the dry snow zone. The correlation coefficient is found to be significant, especially after spatially averaging the backscatter. However, an analysis and simulation demonstrate that spatial averaging can artificially increase the correlation coefficient.**

Satellite-borne scatterometers are radars designed to estimate the wind velocity over the earth's oceans. These wind estimates are useful for studying the earth's climate and for forecasting and monitoring weather patterns including hurricanes. Satellite-borne scatterometers are ideal for wind vector estimation because of their global coverage, high resolution, and frequent sampling [1]. Other applications of satellite-based scatterometers include tracking icebergs, mapping the sea ice extent [2], measuring deforestation in the Amazon [3], and monitoring important indicators of the global climate such as the Greenland and Antarctic ice sheets [4]. The accuracy of these applications requires accurate calibration of the scatterometer [5].

Although scatterometers are calibrated prior to launch, system degradation requires the scatterometer to be calibrated after launch as well [6]. Accurate post-launch calibration can be achieved by using radar backscatter data from natural land targets with temporally constant and isotropic backscatter. Currently, post-launch scatterometer calibration is performed using data collected from regions such as the Amazon rainforest, the Greenland and Antarctic ice sheets, and the Sahara Desert, with calibration most often performed using data from the Amazon [7], [8], [9]. However, in [6], the authors find that the backscatter in parts of the Greenland ice sheet varies less than in other regions. This suggests that the Greenland ice sheet, specifically the dry snow zone, may be better suited for scatterometer calibration than other regions of the earth.

Our studies indicate that some of the variation in backscatter (σ^0) in the dry snow zone of the Greenland ice sheet is seasonal. This variation is characterized by a slight decrease in backscatter during the summer months followed by a return to winter backscatter levels. We define this behavior as seasonal variation. Figure 1 shows an example of seasonal variation.

Some of our previous work has indicated that temperature may be causing the seasonal variation.This paper explores the relationship between temperature and backscatter in the dry snow zone using Pearson's correlation coefficient. An analysis of the effects of spatial averaging on the correlation coefficient is included.

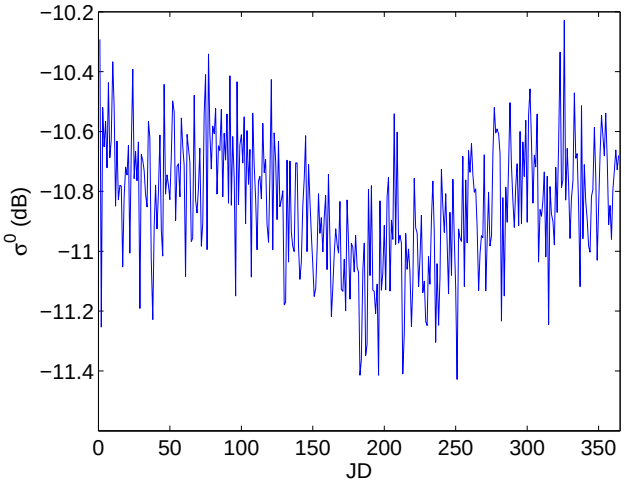

Fig. 1. An example of seasonal variation in backscatter. Taken from the Greenland summit in 2000. JD stands for Julian Day.

I. Data

The data used in this paper is generated by the QuikSCAT scatterometer. QuikSCAT is a 13.4 GHz pencil beam scatterometer with an inner and outer beam. The inner beam is horizontally polarized with a fixed incidence angle of $46°$ while the outer beam is vertically polarized with an incidence angle of $54°$.

Two types of normalized backscatter images are created using the Scatterometer Image Reconstruction (SIR) algorithm: egg and slice [10]. Egg images have coarser resolution but are less sensitive to noise than the slice images. For this study, we use vertically polarized egg images.

Only pixels classified as dry snow are used. The dry snow zone is defined as the region where little to no melt occurs throughout the year. To classify the dry snow zone, we use the Q-α melt detection algorithm [11]. Pixels that experience no melt according to the melt detection algorithm throughout the entire year are flagged as dry snow.

Air temperature data are taken from Automated Weather Stations (AWS) sponsored by the Greenland Climate Network

(GC-Net) [12]. We use data from stations 1-23 that coincide with the duration of the QuikSCAT mission.

The AWS record temperature hourly while the QuikSCAT SIR images are created for a single day. To facilitate easier comparison of the two data sets, the average temperature is calculated for each station for each day.

II. Correlation of Backscatter and Air Temperature

We calculated the Pearson correlation coefficient using the backscatter in the pixel closest to the AWS and by spatially averaging backscatter throughout the entire dry snow zone. The results are given in the following.

A. Correlation With a Single Pixel and Station

In general, the magnitude of the correlation coefficient of σ^0 and temperature is relatively low using a single pixel of backscatter data. Throughout all years of data, the correlation coefficient had a mean of -0.516 with a variance of 0.0058. While this does suggest that there is a relationship between temperature and backscatter at temperatures below freezing, it is not particularly strong.

There are several possible reasons for the low correlation coefficient. First, both the temperature records and the backscatter are noisy. This may result in a reduced correlation coefficient. Second, the true relationship between temperature and backscatter may not be linear which is what the Pearson correlation coefficient measures. Snow accumulation can also influence backscatter which may result in a reduced correlation. Third, due to the lack of reliable snow temperature records, we use air temperature as a substitute for snow temperature. Snow temperature is less variable than air temperature and may be more highly correlated with the backscatter.

The non-linearity can be taken into account by using other definitions of the correlation coefficient such as Spearman's rank correlation coefficient which is beyond the scope of this paper. To mitigate the effects of noise, we explore the potential of spatial averaging in the following.

B. Correlation With Spatial Averaging

Since we are interested in the dominant behavior in the dry snow zone, we use the spatial average of σ^0 calculated using all pixels classified as dry snow. Although the dry snow zone does not melt, σ^0 may drop a few dB on certain days in some pixels near the edges of the classified dry snow zone. This spike is followed by a return to the previous backscatter levels within a few days. Since this backscatter behavior is consistent with melting, these days are removed from the data set.

Prior to spatially averaging σ^0, interpolation is performed to fill in the missing data and the pixel time average of σ^0 is removed. Figure 2 gives an example of spatially averaged σ^0 and temperature.

Table I gives the correlation coefficient of temperature and the spatial average of σ^0 for the available stations. In general, the correlation coefficient using the spatial average is

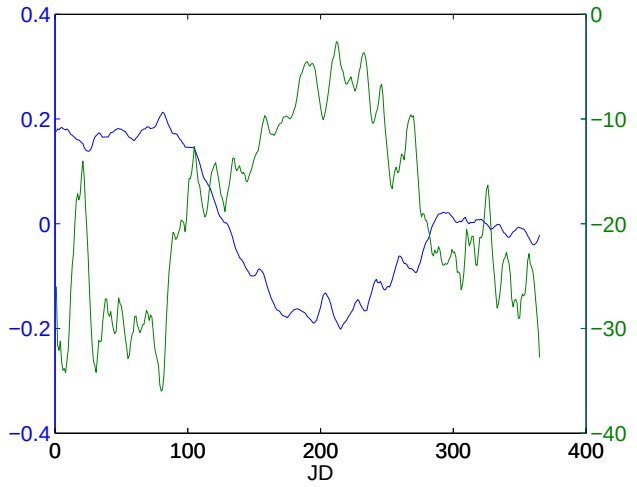

Fig. 2. Smoothed spatial averages of air temperature (green, in C) and σ^0 (blue, in dB) in 2000. The smoothing is performed with a moving average filter using a 7-day window.

significantly higher than that using a single pixel. Interestingly, the correlation coefficient of the spatial averages of σ^0 and air temperature is higher than the average of the correlation coefficients of the spatial average of σ^0 and the individual temperature records. Similarly, the correlation coefficient of air temperature and the spatial average of σ^0 increases with the number of pixels used to calculate the spatial average. While this could be attributed to noise mitigation, the analysis in the following section suggests that may not always be the case.

III. Correlation Coefficient Analysis

Let the number of pixels that are spatially averaged be N. We examine the case where $N = 2$ and where N is arbitrary.

A. Analysis for $N = 2$

Let X_1 and X_2 be positively correlated random variables. Let Y also be a random variable correlated with both X_1 and X_2. The correlation coefficient of the random variables X_k and Y is

$$\rho_{X_k Y} = \frac{\text{Cov}(X_k, Y)}{\sigma_{X_k}\sigma_Y} \tag{1}$$

where σ_{X_k} and σ_Y are the standard deviations of X_k and Y. If $Z = \frac{X_1+X_2}{2}$, then

$$\begin{aligned}\sigma_Z^2 &= \text{Var}\left(\frac{X_1+X_2}{2}\right)\\ &= \frac{1}{4}(\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2\text{Cov}(X_1, X_2))\\ &= \frac{1}{4}(\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2\sigma_{X_1}\sigma_{X_2}\rho_{X_1X_2}).\end{aligned} \tag{2}$$

TABLE I

CALCULATED CORRELATION COEFFICIENT OF TEMPERATURE AND THE SPATIAL AVERAGE OF σ^0 FOR VARIOUS AWS AND YEARS. THE BOTTOM ROW GIVES THE CORRELATION COEFFICIENT FOR THE SPATIALLY AVERAGED TEMPERATURE AND THE SPATIAL AVERAGE OF σ^0.

Station	2000	2001	2002	2003	2004	2005	2006	2007	2008
1			-0.657						
2			-0.663	-0.756			-0.622		-0.736
5			-0.703	-0.901	-0.826	-0.795			
6		-0.718		-0.804				-0.840	-0.809
7							-0.895	-0.911	
8	-0.728		-0.662		-0.776	-0.719	-0.820	-0.775	-0.707
9			-0.648		-0.765		-0.833	-0.695	-0.827
10	-0.711		-0.664		-0.806		-0.714	-0.791	-0.707
11	-0.764	-0.714	-0.657		-0.762	-0.688		-0.780	
12			-0.723		-0.852	-0.803			
13	-0.730								
14					-0.827			-0.848	-0.824
15	-0.714	-0.687	-0.656	-0.758	-0.785	-0.742			
16	-0.738								
17		-0.758	-0.616	-0.806	-0.746				
19			-0.607	-0.810					
21				-0.927					
Average	-0.768	-0.780	-0.710	-0.904	-0.859	-0.839	-0.897	-0.875	-0.830

The covariance of Z and Y is then

$$\begin{aligned}\text{Cov}(Z,Y) &= E[ZY] - E[Z]E[Y] \\ &= \frac{1}{2}(E[X_1Y] + E[X_2Y] \\ &\quad - E[X_1]E[Y] - E[X_2]E[Y]) \\ &= \frac{1}{2}(\text{Cov}(X_1,Y) + \text{Cov}(X_2,Y)) \\ &= \frac{1}{2}\sigma_Y(\rho_{X_1Y}\sigma_{X_1} + \rho_{X_2Y}\sigma_{X_2}). \end{aligned} \tag{3}$$

The correlation coefficient of Z and Y then becomes

$$\rho_{ZY} = \frac{\rho_{X_1Y}\sigma_{X_1} + \rho_{X_2Y}\sigma_{X_2}}{\sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2 + 2\sigma_{X_1}\sigma_{X_2}\rho_{X_1X_2}}}. \tag{4}$$

If we assume that $\sigma_{X_1} \approx \sigma_{X_2}$, then the correlation coefficient becomes

$$\rho_{ZY} = \frac{\rho_{X_1Y} + \rho_{X_2Y}}{\sqrt{2(1+\rho_{X_1X_2})}}. \tag{5}$$

This assumption could be valid for nearby pixels or stations. Since X_1 and X_2 are positively correlated, $0 < \rho_{X_1X_2} < 1$, and gives bounds on ρ_{ZY} can be written as

$$\frac{|\rho_{X_1Y} + \rho_{X_2Y}|}{2} < |\rho_{ZY}| < \frac{|\rho_{X_1Y} + \rho_{X_2Y}|}{\sqrt{2}}. \tag{6}$$

Equation (6) shows that if σ^0 is averaged for two adjacent pixels and then correlated with air temperature, the resulting correlation coefficient will be greater than the average of the individual correlation coefficients.

B. Analysis for Arbitrary N

This can be extended to $N \geq 2$. Define $Z_N = \frac{1}{N}\sum_{n=1}^{N} X_n$ where each X_n is a random variable positively correlated with the others. Then the variance of Z_N is

$$\begin{aligned}\sigma_{Z_N}^2 &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\text{Cov}(X_n, X_m) \\ &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\rho_{X_nX_m}\sigma_{X_n}\sigma_{X_m} \end{aligned} \tag{7}$$

and the covariance of Z_N and Y is

$$\begin{aligned}\text{Cov}(Z_N, Y) &= \frac{1}{N}\sum_{n=1}^{N}\text{Cov}(X_n, Y) \\ &= \frac{\sigma_Y}{N}\sum_{n=1}^{N}\sigma_{X_n}\rho_{X_nY}. \end{aligned} \tag{8}$$

The correlation coefficient is then

$$\rho_{Z_NY} = \frac{\sum_{n=1}^{N}\sigma_{X_n}\rho_{X_nY}}{\sqrt{\sum_{n=1}^{N}\sum_{m=1}^{N}\rho_{X_nX_m}\sigma_{X_n}\sigma_{X_m}}}. \tag{9}$$

If we assume that $\sigma_{X_n} \approx \sigma_{X_m}$ for all n and m and that $\rho_{X_nX_m} \approx \rho_X$ for all $n \neq m$, then ρ_{Z_NY} simplifies to

$$\rho_{Z_NY} = \frac{\sum_{n=1}^{N}\rho_{X_nY}}{\sqrt{N + (N-1)N\rho_X}}. \tag{10}$$

Since the X_n's are positively correlated with each other, then $0 < \rho_X < 1$. This means that

$$\frac{1}{N}\left|\sum_{n=1}^{N}\rho_{X_nY}\right| < |\rho_{Z_NY}| < \frac{1}{\sqrt{N}}\left|\sum_{n=1}^{N}\rho_{X_nY}\right|. \tag{11}$$

This shows that for adjacent pixels or stations, the magnitude of the correlation coefficient of the spatial average and the other parameter is greater than the magnitude of the spatial average of the correlation coefficient.

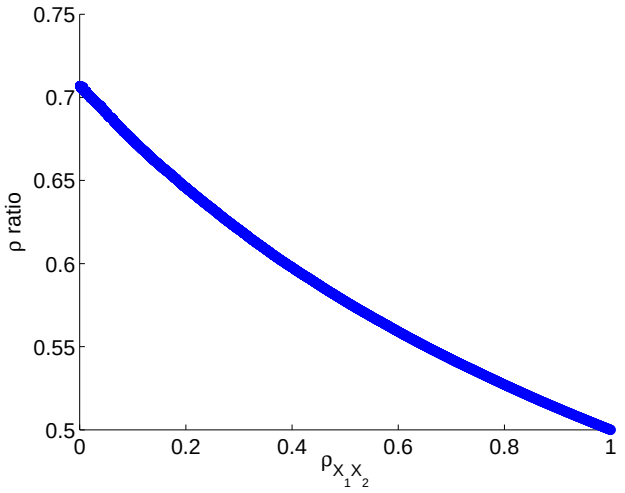

Fig. 3. Simulated results where $N = 2$. The value ρ ratio is defined in Eq. (12).

C. Verification by Simulation

To verify the analysis, we simulated the case where $N = 2$. The value ρ ratio was simulated where

$$\rho \text{ ratio} = \frac{|\rho_{ZY}|}{|\rho_{X_1Y} + \rho_{X_2Y}|}. \tag{12}$$

Then from the previous sections,

$$\frac{1}{2} < \rho \text{ ratio} < \frac{1}{\sqrt{2}}. \tag{13}$$

Figure 3 gives the results for the simulation which agree with Eq. (13).

D. Conclusion of Correlation Coefficient Analysis

The results of this section demonstrate that spatial averaging can artificially increase the correlation coefficient. According to Fig. 3, this occurs the most when the pixels are not highly correlated. This suggests that to mitigate noise effects on the correlation coefficient, a spatial average of a smaller and more correlated region would give more accurate results than using the spatial average of the entire ice sheet.

IV. Conclusion

The results of this paper suggest that backscatter is correlated with air temperature even at temperatures below freezing. We also demonstrated that spatially averaging backscatter can artificially increase the correlation coefficient between backscatter and temperature if the pixels are not highly correlated.

Future work includes exploring different definitions of the correlation coefficient to take into account any non-linear relationship between σ^0 and temperature. Additional future work includes exploring the correlation coefficient using the spatial average of smaller regions around the weather stations.

References

[1] F. Naderi, M. Freilich, and D. 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.

[2] H. Anderson and D. Long, "Sea ice mapping method for SeaWinds," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 647–657, 2005.

[3] D. Long and P. Hardin, "Vegetation studies of the Amazon Basin using enhanced resolution seasat scatterometer data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 2, pp. 449–460, 1994.

[4] D. G. Long and M. R. Drinkwater, "Greenland ice-sheet surface properties observed by the Seasat-A scatterometer at enhanced resolution," *Journal of Glaciology*, vol. 40, no. 135, pp. 213–230, 1994.

[5] W.-Y. Tsai, J. Graf, C. Winn, J. Huddleston, S. Dunbar, M. Freilich, F. Wentz, D. Long, and W. Jones, "Post-launch sensor verification and calibration of the NASA scatterometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1517–1542, May 1999.

[6] R. Kumar, S. Bhowmick, K. Babu, R. Nigam, and A. Sarkar, "Relative calibration using natural terrestrial targets: A preparation towards Oceansat-2 scatterometer," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 2268–2273, 2011.

[7] R. Kennett and F. Li, "Seasat over-land scatterometer data. ii. selection of extended area and land-target sites for the calibration of spaceborne scatterometers," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 27, no. 6, pp. 779–788, 1989.

[8] J. Zec, W. Jones, and D. Long, "NSCAT normalized radar backscattering coefficient biases using homogeneous land targets," *Journal of geophysical research*, vol. 104, no. C5, pp. 11 557–11 568, 1999.

[9] L. Kunz and D. Long, "Calibrating SeaWinds and QuikSCAT scatterometers using natural land targets," *Geoscience and Remote Sensing Letters, IEEE*, vol. 2, no. 2, pp. 182–186, 2005.

[10] D. Early and D. Long, "Image reconstruction and enhanced resolution imaging from irregular samples," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 2, pp. 291–302, 2001.

[11] I. Ashcraft and D. Long, "Comparison of methods for melt detection over Greenland using active and passive microwave measurements," *International Journal of Remote Sensing*, vol. 27, no. 12, pp. 2469–2488, 2006.

[12] K. Steffen, J. E. Box, and W. Abdalati, "Greenland Climate Network: GC-Net," *CRREL Special Report on Glaciers, Ice sheets and Volcanoes, trib. to M. Meier*, vol. 96-27, pp. 98–103, 1996.