

Igor R. Ivić⁽¹⁾⁽²⁾, and Sebastian M. Torres⁽¹⁾⁽²⁾

- (1) Cooperative Institute for Mesoscale Meteorological Studies (CIMMS), The University of Oklahoma
 (2) NOAA/OAR National Severe Storms Laboratory, Norman, Oklahoma

1. INTRODUCTION

Proper censoring of data on the National Weather Surveillance Radar – 1988 Doppler (i.e., WSR-88D) is essential for the forecasters and automated algorithms. Presently, spectral moments at each range location are censored (i.e., labeled not useful) if the Signal-to-Noise Ratio (SNR) is insufficient, or the echoes from the subsequent trips are overlaid. Current censoring uses power measurements to determine if the SNR is above predetermined threshold relative to the noise power (typically 2 dB for reflectivity, and 3.5 dB for velocity measurements). In addition to SNR, Doppler radar offers coherency (i.e., autocorrelation) measurement in sample-time as a variable suitable for censoring (Keeler 1990, SIGMET 2006). It is likely that using the combination of the two measurements one can design a detection scheme that improves the detection rate as opposed to using either of these two measurements separately; thus, effectively increasing the sensitivity of the radar. In this paper such possibility is investigated.

2. COHERENCY BASED DETECTION

The main idea of the approach investigated in this work is to utilize the weather signal coherency in sample-time to improve its detection. Each approach essentially consists of comparing the output of some function of the time series data (i.e., $f(\mathbf{V})$ where $\mathbf{V} = [V(0, \tau_s), \dots, V(M-1, \tau_s)]^T$) against a threshold to decide if significant signal is present at range location given by τ_s . Complex random variables $V(m, \tau_s)$ are obtained by sampling voltage echoes at times mT (where T is the Pulse Repetition Time or PRT). Clearly, these samples can be combined in many ways to obtain the function f . However, a good combination is defined by the one that can emphasize signal features in white Gaussian noise. For example, the measured power is the sum of signal and noise powers, thus the total power reflects the notion that its estimate is higher if a signal is present. Moreover, the autocorrelation function of white noise at lags other than zero is zero but increases if signal is present. Hence, the autocorrelation coefficient estimate is a measure of the coherency in sample-time. Consequently, it seems reasonable that we formulate function f as a sum of the powers and the autocorrelation estimates. This has an added convenience since the calculation of these functions is needed for generating spectral moments and the computational impact on existing resources is thus minimal.

The general form of the weighted sum is

$$WS = \left(\hat{P} + \alpha \left| \hat{R}(T) \right| \right) \geq THR, \quad (1)$$

where⁽¹⁾

$$\begin{aligned} \hat{P} &= \frac{1}{M} \sum_{m=0}^{M-1} |V(m)|^2 \\ \hat{R}(T) &= \frac{1}{M} \sum_{m=0}^{M-1} V^*(m) V(m+1), \end{aligned} \quad (2)$$

and the THR is the threshold set so that the preset rate of false detections is not exceeded. The weight α provides the degree of freedom to optimize the rate of detection (i.e., probability of detection or POD) for a particular application. This can be done by setting α to maximize POD given specific values of signal parameters. Generally, the expected value of the power estimate is the sum of the signal and noise powers. In the case of the weather signal, the modulus of the autocorrelation is (Doviak and Zrnić 1993)

$$\begin{aligned} R(T) &= S \exp \left[-(\pi \sigma_v / v_a)^2 / 2 \right] \\ &= S \exp \left[-(\pi \sigma_{vn})^2 / 2 \right], \end{aligned} \quad (3)$$

where σ_v is the signal spectrum width, and v_a is the unambiguous velocity. The ratio of these two $\sigma_{vn} = \sigma_v / v_a$ is the normalized spectrum width. Clearly, when only the white Gaussian noise is present, we can assume that $\sigma_{vn} \rightarrow \infty$ and the contribution of the autocorrelation estimate to the WS is fairly small. This contribution is inversely proportional to the number of samples M because (Ivić 2009):

$$E \left\{ \left| \hat{R}(T) \right|^2 \right\} = \frac{N^2}{M-1}, \quad (4)$$

where N is the noise power. Thus, the shape of the WS *pdf* depends only on the number of samples M and the weight α . When signal and noise are present the parameters that determine the probability density function (i.e., *pdf*) shape are M , α , the SNR, and σ_{vn} . Given specific values of all these parameters (except α) the optimal weight is the one that minimizes the portion of the *pdf* that falls under the threshold (thus maximizing the POD). Thus, for any threshold value, we want the weight that minimizes the overlap between the *pdf*s of noise and signal+noise. By varying the weight value one changes both the mean value and the spread of the WS *pdf*. Because the mean of the autocorrelation estimate modulus is higher when signal is present the difference

* Corresponding author address: Igor R. Ivić, CIMMS/NSSL, 120 David L. Boren Blvd., Norman, OK 73072; email: Igor.Ivic@noaa.gov

¹ Note that the range designation τ_s has been dropped as it is understood that the same procedure is repeated for each set of M samples at all range locations

in the means of the noise and signal+noise *pdfs* is directly proportional to the value of α . The farther apart these two means are the less overlap should occur between the two *pdfs*. On the other hand, the larger the weight, the more spread are the *pdfs* which has the potential to increase the overlap. These two opposing effects need to be balanced by the proper choice of the weight. As shown in Ivić (2009), the SNR has no bearing on the choice of the weight. This is so because given the specific value of σ_{vn} , the same value of α maximizes the POD at all SNR levels. Naturally, the POD is directly proportional to the SNR. As for the number of samples M , the spread of the *pdf* when noise or signal+noise are present is inversely proportional to it, whereas the mean stays the same. At the same time, the performance of the sum degrades as the difference between the normalized spectrum width of samples in \mathbf{V} and the σ_{vn} value that the sum is optimized for grows larger. We can, however, attempt to find a value of σ_{vn} for which to optimize the weighted sum that would minimize this performance degradation. We speculate that such value is its median. Fang et al. (2004) show that 2 m s^{-1} is the median of σ_v for most weather events of interest; thus, we would like to adjust the α weight accordingly. Another possible approach would be to find such a combination of weights that would produce relatively balanced detection rate for a wide range of signal parameters of interest.

3. WEIGHT OPTIMIZATION

Let us now analyze the approach needed to find the weight α given the value of σ_{vn} and the desired PFA. The most straightforward way to achieve this would be, of course, to find the weighted sum *pdf* in a closed form which would be a function of signal parameters as well as the weight. Assuming we are able to do this we would essentially have two functions. One would be the weighted sum *pdf* when only noise is present and the other one when both signal+noise are present. The first we shall denote as $f(W_{S_N})$ and the latter one as $f(W_{S_{S+N}})$, where the W_{S_N} and $W_{S_{S+N}}$ are random variables that equal the weighted sum when only noise and both signal and noise are present, respectively. Notice that both functions are designated with the same letter and the difference is in the subscript of the random variable that is the function input. This is done purposely because these two functions, in their general form, are the same except that when only noise is present we set $\text{SNR} = -\infty \text{ dB}$, and $\sigma_{vn} = \infty \text{ m s}^{-1}$. Now we can define our goal as:

Find a value for the weight α such that:

$$P(W_{S_{S+N}} \geq THR | \alpha) \geq P(W_{S_{S+N}} \geq THR | \alpha')$$

OR

$$\int_{THR}^{\infty} f(W_{S_{S+N}}; \alpha) dW_{S_{S+N}} \geq \int_{THR}^{\infty} f(W_{S_{S+N}}; \alpha') dW_{S_{S+N}}$$

For $\forall \alpha' \in \mathbb{R}$, under the condition

$$P(W_{S_N} \geq THR | \alpha) = \int_{THR}^{\infty} f(W_{S_N}; \alpha) dW_{S_N} \leq \varepsilon \quad (5)$$

where ε is the desired PFA, and THR is the threshold.

Clearly, both W_{S_N} and $W_{S_{S+N}}$ are functions of elements of the observations vector $\mathbf{V} = [V(0), \dots, V(M-1)]^T$. Hence, we are dealing with the transformation of a random variable and to derive f we would need to start with the joint *pdf* for the elements of \mathbf{V} (Miller 1969):

$$pdf(\mathbf{V}) = \pi^{-M} |\det \mathbf{C}^{-1}| \exp(-\mathbf{V}^H \mathbf{C}^{-1} \mathbf{V}), \quad (6)$$

where \mathbf{C} is the covariance matrix of size $M \times M$ defined as:

$$\mathbf{C} = E\{\mathbf{V}\mathbf{V}^H\}. \quad (7)$$

Brief examination of the involved functions reveals that a closed form solution is not possible. All is not lost however. Given the exact expression for joint *pdfs* using (5) we can state the goal as:

Find values for the weight α such that:

$$\int_{W_{S_{S+N}} \geq THR} \pi^{-M} |\det \mathbf{C}^{-1}| \exp(-\mathbf{V}^H \mathbf{C}^{-1} \mathbf{V}) d\mathbf{V} \geq \int_{W_{S_{S+N}} > THR} \pi^{-M} |\det \mathbf{C}^{-1}| \exp(-\mathbf{V}^H \mathbf{C}^{-1} \mathbf{V}) d\mathbf{V}$$

for $\forall \alpha' \in \mathbb{R}$

under the condition

$$\int_{W_{S_N} \geq THR} (\pi^2 N)^{-M} \exp\left(-\frac{1}{N} \sum_{m=0}^{M-1} |V(m)|^2\right) d\mathbf{V} \leq \varepsilon \quad (8)$$

where ε is the desired PFA, and THR is the threshold.

The condition $W_{S_{S+N}} \geq THR$ corresponds to the region in the M dimensional sample space for which the weighted sum is greater than the threshold when signal and noise are present. The condition $W_{S_N} \geq THR$ corresponds to the region in M dimensional sample space for which the weighted sum is greater than the threshold when only noise is present. Thus, we desire to maximize the region $W_{S_{S+N}} \geq THR$ while keeping the region $W_{S_N} \geq THR$ constant. Short of the ability to derive the *pdf* analytically we resort to approximate methods.

The first that comes to mind is the plain Hit or Miss Monte Carlo simulation method. Let us analyze what steps we would have to perform to adjust the weight. Clearly, to find the optimal set of weights using simulation one would need to inspect the resulting rate of detection for large number of weight combinations and choose the one having the highest and/or the most balanced POD. Because the detection rates being estimated are on the order of $\sim 10^{-1}$, for a known weight and a given threshold, a POD evaluation does not require huge number of trials to achieve the desired accuracy; hence, it can be done quickly using the plain Monte Carlo approach. Yet, before estimating POD, an appropriate threshold value that preserves the desired probability of false alarm (PFA) must be found. So far,

the only tool we have at our disposal is the plain Monte Carlo. Due to the low values of the desired PFA (on the order of 10^{-5} to 10^{-6}), evaluating each candidate THR value requires large number of trials because the *pdf* is being evaluated at its tail; thus, making it computationally impractical if numerous thresholds are to be considered. Consequently, it becomes imperative to introduce an optimization that will reduce the number of trials required to achieve the desired accuracy. When simulating the occurrence of random rare events (as this is clearly the case) the simulation technique known as Importance Sampling (IS) has the potential to dramatically reduce the number of trials required to meet the accuracy requirements (Mitchell 1981). For the problem at hand the IS technique known as the Cross-Entropy (CE) method described in de Boer et al. (2005) is used. Detailed explanations of the derivations which customize the CE technique are beyond the scope of this paper. Interested readers can find similar derivation but for the dual-polarized case in Ivić (2009). The final product that follows the algorithm 3.1 in de Boer et al. (2005) is presented below:

1) Define $\hat{N}_0(m) = N$ for $m = 1, \dots, M$. Set $t = 1$ (iteration = level counter)

2) Generate a sample $\mathbf{V}_1, \dots, \mathbf{V}_K$ from the density:

$$g(\mathbf{V}) = \left(\pi^M \prod_{m=0}^{M-1} \hat{N}_{t-1}(m) \right)^{-1} \exp \left(- \sum_{m=0}^{M-1} \frac{|V(m)|^2}{\hat{N}_{t-1}(m)} \right). \quad (9)$$

Calculate $WS(\mathbf{V}_i)$ for all and sort them in ascending order. Compute the sample $(1-\rho)$ -quantile \hat{T}_t of the performance according to $\hat{T}_t = WS \left(\lceil (1-\rho)K \rceil \right)$, provided \hat{T}_t is less than THR . Otherwise set $\hat{T}_t = THR$.

3) Use the same sample $\mathbf{V}_1, \dots, \mathbf{V}_K$ to solve the stochastic equation as given by.

$$N(m) = \frac{\sum_{k=0}^{K-1} D[WS(\mathbf{V}_k) > \hat{T}_t] \times \dots \sum_{k=0}^{K-1} D[WS(\mathbf{V}_k) > \hat{T}_t] \times \exp \left(- \sum_{l=0}^{M-1} \left(\frac{1}{N} - \frac{1}{\hat{N}_{t-1}(l)} \right) |V(l,k)|^2 \right)}{\exp \left(- \sum_{l=0}^{M-1} \left(\frac{1}{N} - \frac{1}{\hat{N}_{t-1}(l)} \right) |V(l,k)|^2 \right)}. \quad (10)$$

Denote the solution by $\hat{\mathbf{N}}_t$.

4) If $\hat{T}_t < THR$, set $t = t+1$ and reiterate from step 2) else proceed with step 5.

5) Estimate the PFA using:

$$PFA = \frac{1}{K} \sum_{k=0}^{K-1} D[WS(\mathbf{V}_k) > \hat{T}_t] \frac{\prod_{l=0}^{M-1} \hat{N}_t(l)}{N^M} \times \exp \left(- \sum_{l=0}^{M-1} \left(\frac{1}{N} - \frac{1}{\hat{N}_t(l)} \right) |V(l,k)|^2 \right). \quad (11)$$

The above algorithm provides an efficient avenue for obtaining the PFA given a threshold value. However, we are dealing with reversed case where we have the desired PFA and we aim to find the corresponding threshold. With the tools available so far, the process for threshold finding is as follows:

1. Choose initial threshold arbitrarily.

2. Estimate the corresponding PFA.

3. If the estimated PFA is within the acceptable range of the desired one, accept the threshold. If the estimated PFA is greater than the desired one, increase the threshold by some small value Δ , and go back to step 2. In case, the estimated PFA is smaller than the desired one decrease the threshold by some small value Δ , and go back to step 2.

Hopefully the above algorithm would work reliably, but even with the optimization techniques introduced, it could still prove very time consuming because the further the initial threshold is from the final one, the longer it will take for the process to converge. It is apparent that it would be particularly helpful if we could find some rough approximation to the *pdf* in a closed form. This approximation would be used to choose the initial threshold which would be then used as a starting point for the iterative process.

We choose the generalized gamma distribution (Stacy 1962), as the approximation function and the first three moments as the matching criteria. This model function is:

$$f(x) = \begin{cases} x^{p-1} e^{-\left(\frac{x}{a}\right)^q} & \text{if } x > 0. \\ 0 & \text{if } x < 0 \end{cases} \quad (12)$$

The fitting is performed by finding the parameters a , p , and q so that the first three moments of the approximation are the same as the ones estimated using

$$\hat{m}_n = \frac{1}{K} \sum_{k=0}^{K-1} \left(\hat{P}_k + \alpha |\hat{R}_k(T)| \right)^n. \quad (13)$$

The procedure for finding the parameters of the generalized gamma distribution is described in Ivić (2009). The initial threshold calculation using the approximation as well as an iterative scheme to find the true threshold value is given in Ivić (2009). The *THR* found for the unity noise power is adjusted by simply multiplying it by the measured system N .

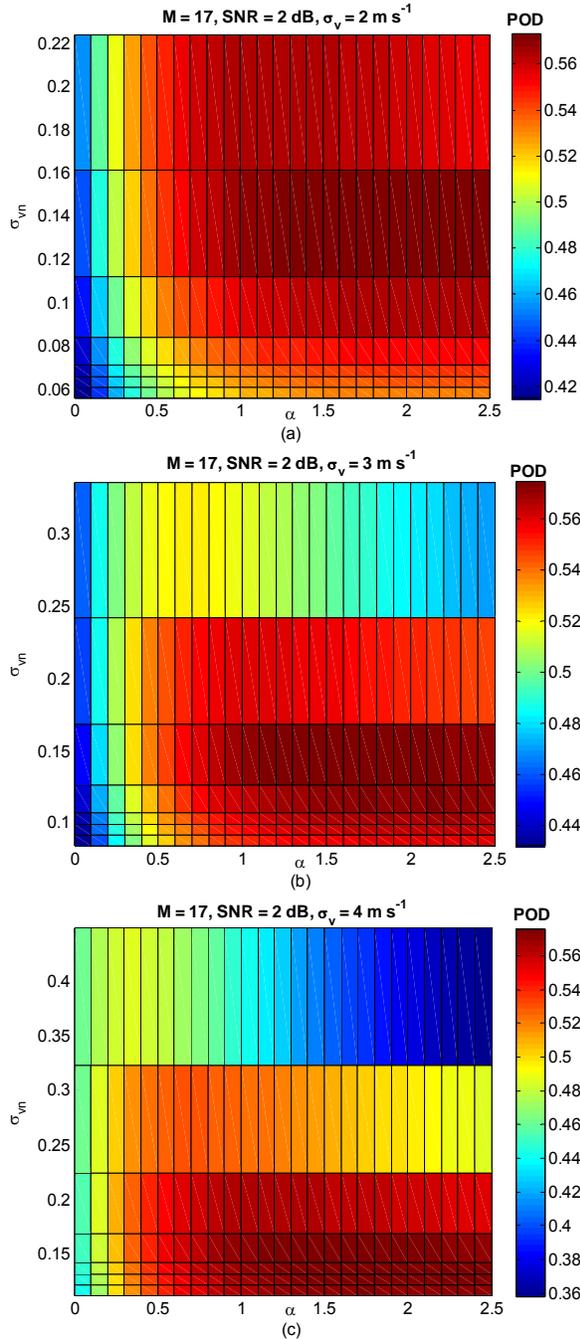


Figure 1. POD of the function $P + \alpha|R(T)|$ versus the value of the weight α and σ_{vn} for the signal spectrum widths of 2 (a), 3 (b), and 4 (c) m s⁻¹, in case when M is 17, and SNR is 2 dB.

Using the tools described, a comprehensive analysis was performed to establish how the POD behaves as function of signal parameters. Figure 1 shows the color plot of the POD as function of α and the σ_{vn} . The range of normalized σ_v values is obtained by taking the following unambiguous velocities: 8.92, 12.4, 17.85, 23.76, 28.1, 30.36, 32.75, and 35.55 m s⁻¹; and dividing them by the σ_v of 2, 3 and 4 m s⁻¹ to obtain Figure 1 (a), (b) and (c). The unambiguous velocities correspond to the operational values used in the NEXRAD network.

We notice that for all three settings of σ_v , the POD is directly proportional to the σ_{vn} for the weight values up to about 0.3. This can be because the power estimate dominates and its variance is directly proportional to the normalized spectrum width. Hence, as σ_{vn} increases, the spread of the weighted sum *pdf* decreases effectively diminishing the area that falls below the threshold and vice versa. At the same time though, the increase in the expected value of the *WS*, as the σ_{vn} decreases, is not large enough to outdo the spread increase in the *pdf*; hence, the POD reduces with σ_{vn} . Clearly, the latter effect (i.e., mean increase) is more pronounced with larger weight values. For the weights larger than 0.3 the maximum POD appears to be achieved at σ_{vn} around 0.15 for all weight values. The maximum value, however, is not the same for all. For instance, in Figure 1 (a) the maximum obtained for α of 0.5 is smaller than the one at 1. Both maxima occur between σ_{vn} values of 0.12 and 0.16. Similar trends are observed in Figure 1 (b) and (c) as well. In general it appears that the weight value at which maximum occurs, for each σ_{vn} , increases as the normalized spectrum width decreases. This is logical because as the signals become more coherent, it makes sense to accentuate the coherency measurement (i.e., autocorrelation). At the same time, increasing α beyond certain value appears to have negative effect on the detection rate. One can speculate that this occurs when an increase in the spread of the *pdf*'s becomes more prominent than the mean separation; thus, the further weight increase only augments the overlap between the noise and the signal+noise *pdf*'s.

Figure 1 (a) shows that setting α to unity ought to give satisfactory results in case we desire to optimize the POD for the σ_v of 2 m s⁻¹. On the other hand, if the weighted sum is to be optimized for the detection of signals with spectrum widths of 3 m s⁻¹, the autocorrelation estimate should be weighted down by about half (i.e., 0.5) for smaller v_a (i.e., larger σ_{vn}) but increased back to unity for higher v_a values. The same trend is observed in Figure 1 (c) for the signal σ_v of 4 m s⁻¹. In Figure 2 the POD behavior for the σ_{vn} values near 0.15 for several weights is shown. Notice that the performance does not change significantly for α larger than one. Also, given the system unambiguous velocity the maximum POD occurs at spectrum widths of $v_a \times \max(\sigma_{vn})$ (where $\max(\sigma_{vn})$ is the σ_{vn} value for which the maximum POD occurs). From Figure 2 the $\max(\sigma_{vn})$ is approximately 0.18 for the 0.5 weight and 0.15 for others. For instance, if v_a is 8.92 m s⁻¹ and the weight is

0.5 or 1 the best detection is achieved for signals with spectrum widths of 1.6 or 1.34 m s⁻¹, respectively. If v_a is 35 m s⁻¹ the weighted sum detects best the signals with spectrum widths of 6.3 or 5.25 m s⁻¹.

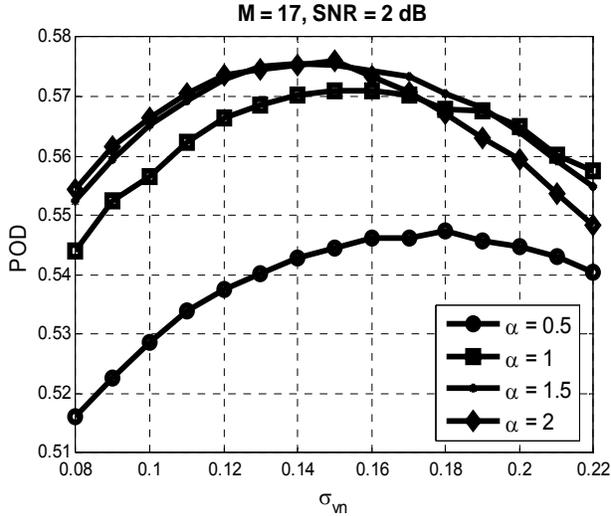


Figure 2. POD of the function $P+\alpha|R(T)|$ versus σ_{vn} for several values of the weight α , in case when M is 17, and SNR is 2 dB.

It is desirable to optimize the overall POD over the range of signal spectrum widths that are of meteorological interest. Thus, another aspect to consider is the detection rate dependence on the spectrum width. This is presented in Figure 3. In both cases when α is 1 and 0.5 we notice that the POD maxima shift towards higher spectrum widths as v_a increases. This is so because the product $v_a \times \max(\sigma_{vn})$ increases. Also, we notice that for each v_a after the POD reaches maximum it starts decreasing as the signal spectrum width increases further. This happens because the spread of the weighted sum *pdf* becomes so significant that after such point further increase in σ_v augments the area of the *pdf* that falls below the threshold. In Table 1 the mean overall PODs for each v_a across the spectrum widths of 0.5 to 5 m s⁻¹ are presented. It shows that setting the weight to 0.5 yields better mean POD only for the unambiguous velocity of 8.92 m s⁻¹ (even though the maximum POD achieved for the unity α is higher). The results in Table 1 imply that setting the weight to unity ought to provide better overall detection across a range of v_a values. Also, one could use a particular weight setting for each unambiguous velocity. Hence, in this case we could have α of 0.5 if v_a is 8.92 m s⁻¹, and unity otherwise. In either case, using the weighted sum improves the detection rate compared to power only based detector. By setting the weight lower one can adjust the detector to favor less coherent signals thus making it more similar to the power based detector. By the same token, setting the weight higher the detector would favor those low SNR signals that are more correlated.

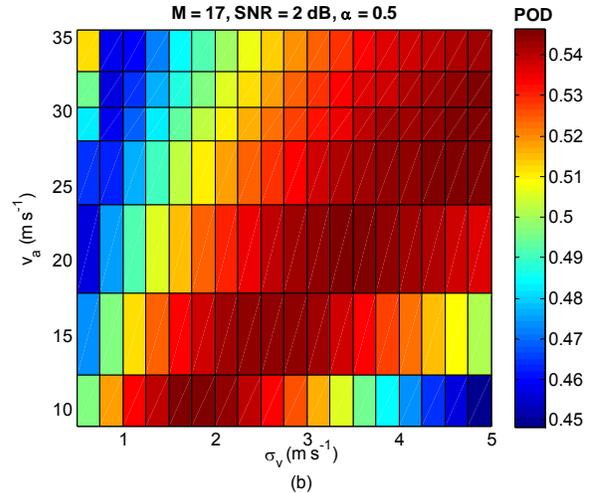
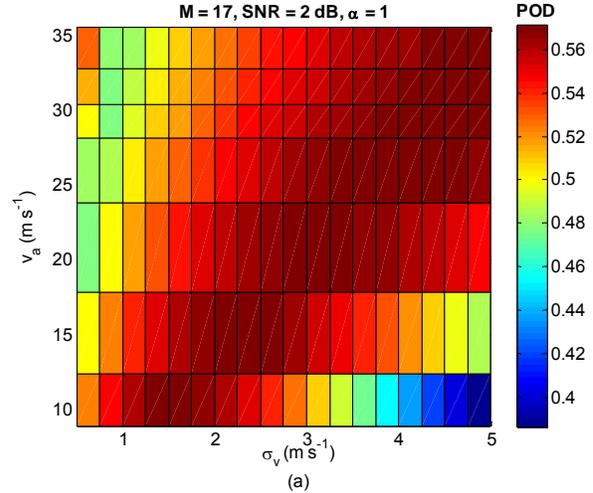


Figure 3. POD of the function $P+\alpha|R(T)|$ versus the σ_v and v_a for the weight values of 1 (a), and 0.5 (b), in case when M is 17, and SNR is 2 dB.

v_a (m s ⁻¹)	POD for $\alpha = 0.5$	POD for $\alpha = 1$
8.92	0.49442	0.47749
12.4	0.51710	0.52327
17.85	0.52580	0.54604
23.76	0.52195	0.54696
28.1	0.51848	0.54412
30.36	0.51641	0.54233
32.75	0.51489	0.54066
35.55	0.51307	0.53926
Mean POD	0.51527	0.53252

Table 1. The mean overall PODs for the range v_a values across the spectrum widths of 0.5 to 5 m s⁻¹.

4. TIME-SERIES IMPLEMENTATION

Real data evaluation of the detector sum was performed using a set of time series data. This set was obtained with the KOUN research radar in Norman, OK which operates at the wavelength of 11.09 cm. The data

shown was collected at a PRT of 3.1 ms (v_a of 8.92 m s⁻¹), with $M = 17$, and at elevation of 0.48 deg. In standard operation, this corresponds to surveillance scan with the threshold set to 2 dB above the noise yielding the PFA of 1.2×10^{-6} in this case. Consequently, this threshold will be used in subsequent analysis. Threshold for the weighted sum was set so that the PFA is also 1.2×10^{-6} . The original reflectivity field obtained using the legacy threshold is plotted in Figure 4.

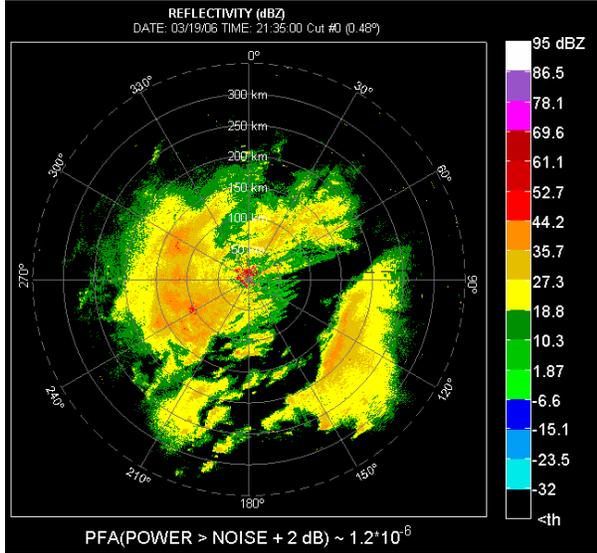


Figure 4. The reflectivity field of the surveillance scan data collected by the KOUN.

To compare the performance of the coherency detector against the SNR based one, various ratios of detections are given in Table 2. These ratios are explained next. Let us view the field (in polar coordinates) as a matrix of size $NAZ \times NRL$, where NRL stands for the *Number of Range Locations* and NAZ for the *Number of Azimuths* in the scan. Let MZ stand for the original reflectivity matrix where each matrix entry is power value at a given location. Let $MN(\alpha)$ stand for the matrix where each matrix entry is 1 if the decision that the signal is present is positive, otherwise it is 0. The weighted sum is used to determine the matrix entries, and α in the brackets denotes the weight value used in the sum. Then the values in the row termed as the *Ratio of total detections* are calculated as:

$$\frac{\text{num}([\text{MZ} \geq \text{NOISE} + 2\text{dB}] .* \text{MN}(\alpha))}{\text{num}([\text{MZ} \geq \text{NOISE} + 2\text{dB}])} \quad (14)$$

Greater than operator is binary (1 if true, and 0 if false) and is applied to each matrix entry yielding new matrix with 0, 1 entries. The operator $.*$ acts as an element-wise matrix multiplication (same as in MATLAB). The num operator gives the total number of 1s in the matrix. This ratio reflects how much of the data classified as useful signal using the SNR based detector are also not censored by the coherency detector. *The Ratio of additional detections* is:

$$\frac{\text{num}([\text{MZ} < \text{NOISE} + 2\text{dB}] .* \text{MN}(\alpha))}{\text{num}([\text{MZ} \geq \text{NOISE} + 2\text{dB}])} \quad (15)$$

This ratio gives the portion of data that are originally censored by the SNR detector, but are detected as signals using the novel approach. *The Ratio of missed detections* is:

$$\frac{\text{num}([\text{MZ} \geq \text{NOISE} + 2\text{dB}] .* \text{not}[\text{MN}(\alpha)])}{\text{num}([\text{MZ} \geq \text{NOISE} + 2\text{dB}])} \quad (16)$$

where *not* operator stands for Boolean negation. This ratio gives the portion of data that are not censored by the legacy detector, but are classified as not useful by the weighted sum approach. *Coherent detections* is the sum of the total and the additional detections ratios.

α	0.5	1
<i>Ratio of total detections</i>	0.997503	0.993242
<i>Ratio of additional detections</i>	0.009721	0.011137
<i>Ratio of missed detections</i>	0.002497	0.006758
<i>Coherent detections</i>	100.72%	100.44%

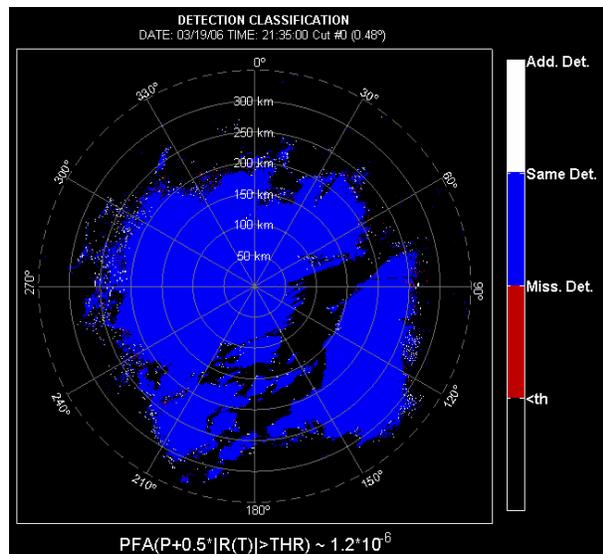
Table 2. Time-series statistics for the KOUN data.

The statistics in the Table 2 shows that in both cases the weighted sum picks up close to 100% of all legacy detections. There is, however, a slight difference of about 0.3% in favor of the sum with the weight of 0.5. Such statistics is in agreement with the numbers in Table 1 as the mean overall POD is slightly higher for α of 0.5 than for the unity if v_a is 8.92 m s⁻¹. Additional detections are very moderate percentage wise in both cases but are slightly higher for the unit weight. Expressed in numbers the total number of detections using the SNR detector for this particular phenomenon is 298017. If the coherency detector with the 0.5 weight is used, it adds 2897 detections and misses 744. Thus, it increases the overall number of detections by 2153. When the weight is unity, the number of the additional and the missed detections is 3319 and 2014, yielding the overall increase of 1305 detections. Table 3 presents mean spectrum widths and the SNRs of the additional and the missed detections. As expected, when the weight is higher the detector favors the more coherent signals. Additionally, the mean spectrum width for the lower weight is closer to the weather signal median of 2 m s⁻¹, which may be another reason why the lower weight performs slightly better in this case.

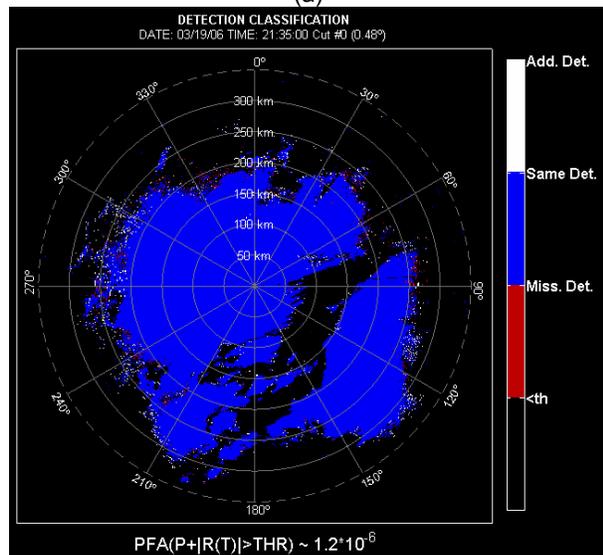
α	0.5	1
<i>Mean σ_v of additional detections</i>	1.42 m s ⁻¹	1.1 m s ⁻¹
<i>Mean σ_v of missed detections</i>	4.8 m s ⁻¹	4.47 m s ⁻¹
<i>Mean SNR of additional detections</i>	1.47 dB	1.426 dB
<i>Mean SNR of missed detections</i>	1.66 dB	1.75 dB

Table 3. Mean spectrum widths, and the SNRs of the additional and the missed detections for the KOUN data.

The classification of detections using α value of 0.5, and 1 is shown in Figure 5 (a) and (b). Visual inspection reveals that the number of missed detections is higher when using the unity weight. This is in agreement with the statistics shown in Table 2.



(a)



(b)

Figure 5. Classification of detections using the weights of 0.5 (a) and 1 (b) for data collected by the KOUN.

The coherency detector is evaluated on an additional two sets of data both collected by the National Weather Radar Testbed (NWRT) Phased Array (PAR) radar in Norman, OK (Forsyth et. al 2007). This radar operates at wavelength of 9.369 cm. The first data set was collected with 15 pulses at PRT of 2.664 ms (v_a of 8.8 m s⁻¹). The SNR based threshold was set to 2 dB corresponding to the PFA of 4.4×10^{-6} . The threshold for the WS was set to yield the same false detection rate. The reflectivity field is shown in Figure 6, and the classification of detections in Figure 7. For this particular weather phenomenon the coherency based detector

picks up 98.73% of the total SNR based detections (i.e., *Ratio of total detections*). The ratios of additional and the missed detections are 0.07 and 0.0127, resulting in the total of 5.73% more detections. The mean spectrum width of all additional and the missed detections is 0.468 m s⁻¹, and 4.18 m s⁻¹. The mean SNR of all additional and the missed detections is 1.47, and 2.41 dB.

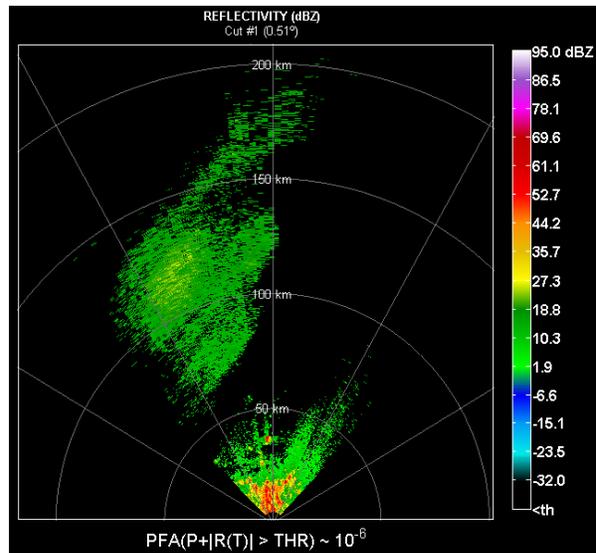


Figure 6. Reflectivity field of the uniform PRT data collected with the NWRT.

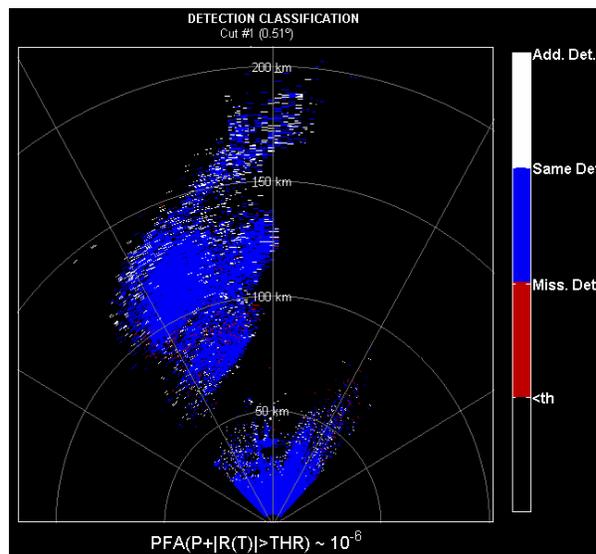


Figure 7. Classification of detections for the uniform PRT data set collected with the NWRT.

The second set was acquired using the range unfolding mode where each azimuth position was first scanned using surveillance PRT yielding the v_a of 7.55 m s⁻¹ with $M = 15$. After that, data from 44 Doppler pulses, with v_a of 26.14 m s⁻¹, were acquired at the same position. In this mode the NWRT signal processor (Torres et. al 2008) combines the power estimates from the two PRTs by using the power measurements from the short PRT at all ranges where the signal processor

determines that signal is not overlaid. As a result, the reflectivity estimates are improved. Also, so obtained power measurements are used as an input to the SNR based detector. The same is done for the autocorrelation measurements for the purpose of using those in the coherency detector. The thresholds for both the SNR and the coherency based detector were set to yield the PFAs of 4.4×10^{-6} , and 10^{-6} for reflectivity estimates obtained from the surveillance and the Doppler PRT set, respectively. For velocity measurements, these are both set to 10^{-6} . The reflectivity and the velocity fields are presented in Figure 8 (a) and (b). The classification of detection is given in Figure 8. The statistics is listed in Table 4 and Table 5.

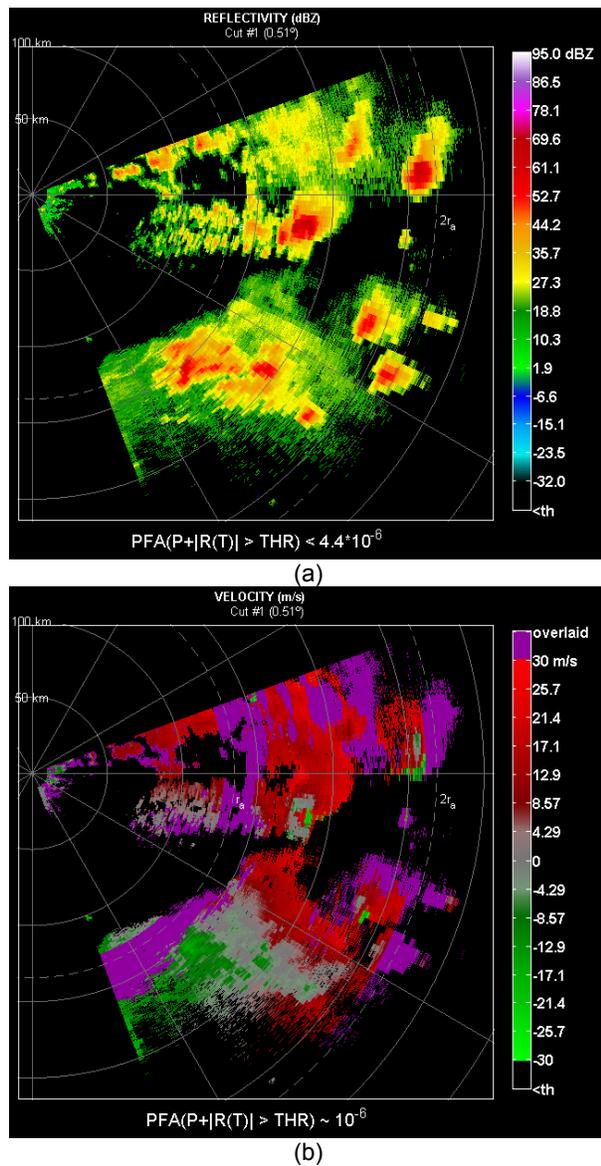


Figure 8. Reflectivity (a) and the velocity (b) field of the dual PRT data collected with the NWRT.

The statistics for the NWRT data shows similar trends as the one obtained from the KOUN radar. It is

evident that in all cases the coherency based detector misses some of the valid data detected by the legacy detector. Nonetheless, the ratio of additional detections is always somewhat larger than that of the missed ones. Hence, using coherency gains more detections on the average resulting in the increased sensitivity. Also, the average spectrum width of additional detections is significantly lower than that of the missed ones, which can have positive effect on the velocity estimates. On the other hand, the average SNR appears to be higher for the missed detections. Finally, observe that the additional detections are predominantly located at the rim of the weather system. This gives an additional assurance that the majority of these detections are indeed valid weather returns.

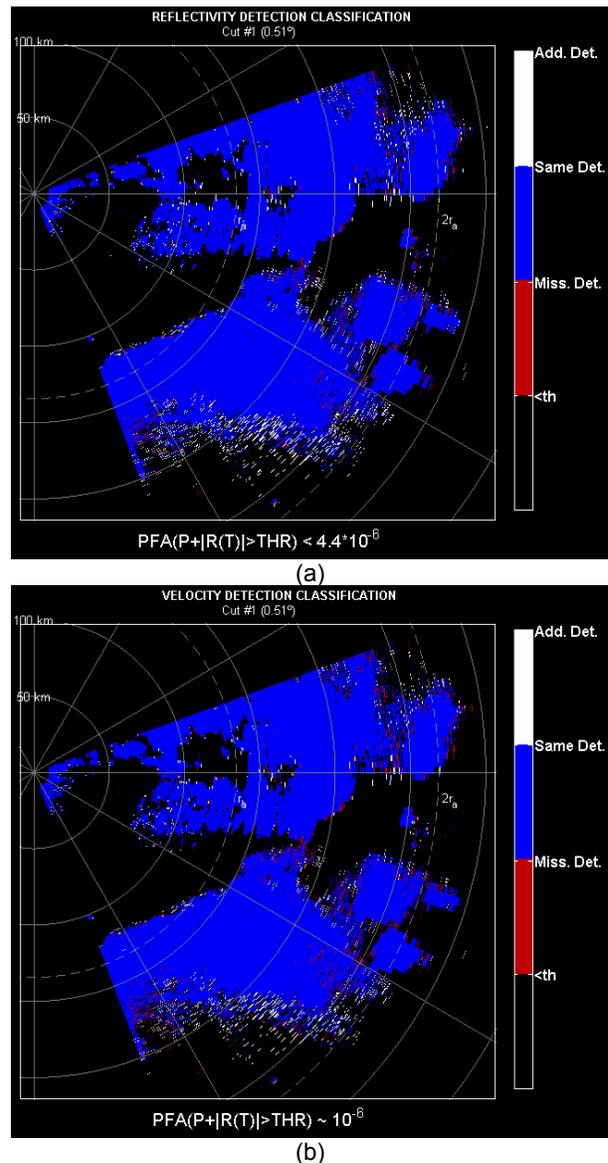


Figure 9. Classification of detections for reflectivity (a), and velocity (b) of the dual PRT data set collected with the NWRT.

Product	Reflectivity	Velocity
Ratio of total detections	0.98582	0.978184
Ratio of additional detections	0.031316	0.023729
Ratio of missed detections	0.01418	0.021816
Coherent detections	101.71%	100.2%

Table 4. Time-series statistics for the dual PRT data collected with the NWRT.

Product	Reflectivity	Velocity
Mean σ_v of additional detections	2.37 m s ⁻¹	2.36 m s ⁻¹
Mean σ_v of missed detections	3.88 m s ⁻¹	4.25 m s ⁻¹
Mean SNR of additional detections	1.98 dB	2.42 dB
Mean SNR of missed detections	3.56 dB	4.1 dB

Table 5. Mean spectrum widths, and the SNRs of the additional and the missed detections for the dual PRT data collected with the NWRT.

5. SUMMARY

Method to improve censoring of weather radar data using the combination of the SNR and coherency based signal detection was investigated. Motivation comes from the possibility to improve the sensitivity in remote sensing devices. A function that linearly combines the power and the modulus of the autocorrelation was chosen for evaluation. It is termed the “weighted sum” because the autocorrelation modulus is scaled by a coefficient. By adjusting the value of the weight, users can fine-tune the sum for particular applications. Clearly, by setting the weight to a value much smaller than one, the sum functions more like a SNR based detector, while having the weight set to a large value the sum favors more coherent signals. In case of a weather signal, we are interested in detection of partially coherent signals. Thus, a value for the scaling factor (i.e., weight) had to be found that optimizes the sums effectiveness in detecting the signals of interest. The evaluation indicates that setting the weight to unity yields balanced results in case of weather signals.

The sum estimate is compared to a threshold to perform the detection. The threshold value is chosen so that the rate of false detections does not exceed some predetermined value. In most cases this value is rather small requiring evaluation of the weighted sum *pdf* at its tail in order to calculate the threshold value. An instructional description of a procedure for threshold calculation is also presented. The performance of the sum was evaluated using both simulated and the radar data. The real data analysis shows satisfactory performance. When used on a radar data the coherency based detection produced improvements in all investigated cases. This shows that the novel approach has the potential to successfully increase the sensitivity of weather radars.

ACKNOWLEDGEMENT

Funding for part of this research was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA17RJ1227, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

REFERENCES

- De Boer P-T., D.P Kroese, S. Mannor and R.Y. Rubinstein, 2005: A Tutorial on the Cross-Entropy Method, *Annals of Operations Research*, **134**, 19-67.
- Doviak, R.J., and D.S. Zrnić, 1993: *Doppler radar and weather observations*. Academic Press, 562 pp.
- Fang M., R. J. Doviak, and V. Melnikov, 2003: Spectrum Width Measured by WSR-88D: Error Sources and Statistics of Various Weather Phenomena. *J. Atmos. Oceanic Technol.*, **21**, 888-904.
- Forsyth, D. E., J. F. Kimpel, D. S. Zrnic, R. Ferek, J. F. Heimmer, T. McNellis, J. E. Crain, A. M. Shapiro, R. J. Vogt, and W. Benner, 2007: Update on the National Weather Radar Testbed (Phased-Array). *23rd Conf. on Interactive Information Processing Systems, for Meteorology, Oceanography, and Hydrology*, San Antonio, TX, Amer. Meteor. Soc., CD-ROM 7.4.
- Miller, K. S., 1969: Complex Gaussian Processes. *SIAM Review*, **11**, 544-567.
- Ivić I. R., 2009: *Detection Thresholds for Spectral Moments and Polarimetric Variables*. VDM Verlag Dr. Müller AG & Co., 199 pp.
- Keeler, R. J., and R. E. Passarelli, 1990: Signal Processing for Atmospheric Radars *Radar in Meteorology*. D. Atlas, Ed.. Amer. Meteor. Soc., 287-315.
- Mitchell R. L., 1981: Importance Sampling Applied to Simulations of False Alarm Statistics, *IEEE Transactions on Aerospace and Electronic Systems*, **17**, 15-24.
- SIGMET, 2006: Processing algorithms. RVP8 Digital IF Receiver User’s Manual, 63 pp.
- Stacy, E. W., 1962: A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics*, **33**, 1187-1192.
- Torres, S. M., C. D. Curtis, E. Forren, and D. Priegnitz, 2008: Signal processing upgrades for the National Weather Radar Testbed. Preprints, *23rd International Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, New Orleans, LA, Amer. Meteor. Soc.
- Zrnić, D. S., 1975: Simulation of Weatherlike Doppler Spectra and Signals. *J. Appl. Meteorol.*, **14**, 619-620.