Updates to the Hybrid Spectrum Width Estimator

GREGORY MEYMARIS*
National Center for Atmospheric Research, Boulder, Colorado

1. INTRODUCTION

With the advent of the Open Radar Data Acquisition (ORDA) system on WSR-88D radars and the introduction of significantly more powerful signal processing hardware comes the opportunity to improve the method used for estimating the spectrum width, a measure of the variability of radial wind velocities within a measurement pulse volume. In addition, the implementation of new operational modes for improved data quality, including SZ phase coding and staggered PRT, will involve very different signal processing techniques and hence may require novel methods to meet the WSR-88D specifications. While spectrum width has not been used extensively by radar meteorologists in the past, the NEXRAD Turbulence Detection Algorithm (NTDA), developed under direction and funding from the FAA's Aviation Weather Research Program, uses the WSR-88D spectrum width as a key input for providing in-cloud turbulence estimates (eddy dissipation rate, EDR) for an operational aviation decision support system (Williams et al. 2005). Achieving improved spectrum width estimator performance would directly benefit the accuracy of the NTDA product. The Hybrid Spectrum Width estimator (HSW), which uses three spectrum width estimators was developed with support from the NEXRAD Radar Operations Center (ROC) (Meymaris et al. 2009). The HSW estimator should start being deployed in 2012 with the build 13 ORDA update. While slightly more computationally intensive, HSW is more accurate and robust than any of the constituent estimators alone, including the standard $R0/R1$ pulse-pair estimator (Doviak and Zrnić 1993, p. 136) currently used in the WSR-88D .

Hitherto, the HSW has been developed for evenly spaced pulse schemes, which is used exclusively on the current NEXRAD system. However, staggered PRT (pulse repetition time) is currently slated to be deployed with build 14. Staggered-PRT, a popular scheme for mitigating the unambiguous range-velocity dilemma of weather radars, is a pulsing scheme in which the radar PRTs alternate between T_1 and T_2 where $T_1/T_2 = 2/3$ (Zrnić and Mahapatra 1985). The default spectrum width estimator for staggered PRT, in the literature, is the pulse pair estimator $R0/R2$, which has poor performance for narrow widths and low signal-to-noise ratios (SNR). In this paper, HSW is adapted to this pulsing scheme. Simulation statistics are presented.

2. Methodology

To evaluate and compare different spectrum width estimators we generated random complex time-series data for various true spectrum width, signal-to-noise ratio (SNR), Nyquist velocities, and number of pulses scenarios. We used an I&Q simulation technique based on the method described in Frehlich and Yadlowsky (1994); Frehlich (2000); Frehlich et al. (2001) except that the autocorrelation function is that of a weather echo as defined in Doviak and Zrnić (1993, p. 125). This is a preferred method for generating complex time-series with a given average autocorrelation function because it is not necessary to generate long time-series in order to get the correct temporal statistics unless the spectrum width is very narrow.

In what follows, the simulator input ("true") spectrum width will be denoted as W, while the estimated spectrum width will be denoted as $\hat{W}$ with a modifying subscript specifying the estimation technique used. Estimation errors were calculated by subtracting the simulator input values from the estimated values (i.e. $\hat{W} - W$). It should be noted that biases and standard deviations have different implications for turbulence detection since bias cannot be mitigated by spatial or temporal averaging while

* *Corresponding author address*:
Gregory Meymaris, RAL, NCAR
1850 Table Mesa Dr., Boulder, CO 80305
meymaris at ucar dot edu

random unbiased errors can.

Because of their ease of calculation, we used autocorrelation-based estimators. The (unbiased) auto-correlation is defined as $R_t = M_t^{-1} \sum_s V^* (s) V (s+t)$, where t is the lag in seconds, V are the complex-valued I&Q radar time-series, the sum is taken over all s such that both $V(s)$ and $V(s+t)$ are available (i.e. measured), and M_t is the number of summands in the sum. For evenly-spaced time-series, this can be written as the more familiar $R_j = (N-j)^{-1} \sum_{k=0}^{N-j-1} V^* (k\tau) V((k+j)\tau)$ where τ is the PRT. The autocorrelation (AC) of a staggered PRT time-series is not evenly sampled since the time-series is not. If T_c is defined as $T_1/2$ (and thus $T_2 = 3T_c$), then R_1 and R_4 (i.e. the ACs at time $1T_c$, $4T_c$, resp.) cannot be directly estimated. See figure 1, which shows the number of pairs (i.e. $M_j \equiv M_{jT_c}$) going into the AC estimate for each lag when the total number of pulses is 80. The reason for the large variability is the staggered nature of the pulses. For example, there are only 40 pairs going into the estimate R_2 since only for $s = k(5T_c) = k(T_1 + T_2)$, with $k = 0, \ldots, 39$, are both $V(s)$ and $V(s + 2T_c)$ available.

There are two main points to draw from this discussion. First, there are now a lot more AC-based estimators one can come up with, especially when including multi-lag estimators (for example, $R0R2R3R5$). Second, it becomes more difficult to predict how different estimators will perform since the AC lag estimator errors will fluctuate due to the fluctuating number of pairs going into the average.

3. Spectrum Width Estimators for Evenly Spaced Pulses

The assumed form for the magnitude of the average auto-correlation function of weather is generally assumed to be Gaussian (Doviak and Zrnić 1993, p. 125) given by:

$$|\mathcal{R}(t)| = \mathcal{P} \exp\left(-\frac{1}{2}\left(\frac{\pi\sigma_v t}{V_a \tau}\right)^2\right) \tag{1}$$

where $\mathcal{P}$ is the true echo power of the weather, σ_v is the spectrum width in ms^{-1}, V_a is the Nyquist velocity in ms^{-1}, τ is the PRT in seconds, and t is the lag time in seconds.

a. The Pulse Pair Estimators

The standard spectrum width estimator currently used in the WSR-88D radars, typically on short PRT data, is the $R0R1$ estimator (Doviak and Zrnić 1993, p. 136), so named because it utilizes the ratio of the first two lags of the autocorrelation function:

$$\hat{W}_{01} = \left(\sqrt{2}/\pi\right) V_a \left|\log\left(R_0/|R_1|\right)\right|^{1/2} \tag{2}$$

Here R_0 is the average power of the signal *with noise removed*, and R_1 is the first lag of the auto-correlation function. In the event that $|R_1| < R_0$, in which case the log has a negative argument, the spectrum width is set to 0 as is done on the WSR-88D. This is derived from eq. 1 by setting up two equations ($R_0 = |\mathcal{R}(0)|$ and $|R_1| = |\mathcal{R}(\tau)|$) with two unknowns ($\mathcal{P}$ and σ_v) and solving for σ_v.

In general, two-lag estimators can be written as

$$\hat{W}_{ab} = \frac{V_a\sqrt{2}}{\pi\sqrt{b^2 - a^2}} \left|\log\left(\left|\frac{R_a}{R_b}\right|\right)\right|^{1/2} \tag{3}$$

where, if R_0 is used, it is always noise corrected.

b. Multi-lag Estimators

Instead of just using two lags, one can use more lags to fit a Gaussian auto-correlation function. $R0R1R2$ is used in the traditional HSW (Meymaris et al. 2009). Hubbert et al. (2011) used a 7 lag estimator for measuring the spectrum of clutter.

While other approaches could be taken, the simplest is to note that taking the logarithm of both sides of eq. 1 yields:

$$\log\left(|\mathcal{R}(t)|\right) = \log\mathcal{P} - \frac{1}{2}\left(\frac{\pi\sigma_v t}{V_a \tau}\right)^2 \tag{4}$$

which is a quadratic equation with respect to t. This is convenient because one can find a closed form solution for a least-squared fit. This makes the spectrum width estimation nothing more than a linear combination of the logarithm of the auto-correlation function at the desired lags. One might be tempted to create an estimator that uses, say, 10 lags, or perhaps all available lags. The problem is that as t gets larger, $|\mathcal{R}(t)| \to 0$, but $|R_t| \nrightarrow 0$. This is because, while $\mathrm{E}[R_t] \to 0$, $\mathrm{E}\left[|R_t|^2\right] = \mathrm{V}[R_t]$ which depends on the signal power and M_t. Thus, the model does not well represent the data, causing the least-squares fit to be poor. The net effect is that the more lags are used, the lower the spectrum width estimator saturates. Melnikov and Zrnić (2004) observed this when examining saturation levels for $R0R1$ and $R0R2$. The family of multi-lag estimators works well, but care must be taken to take saturation levels into account when deciding which lags are used.

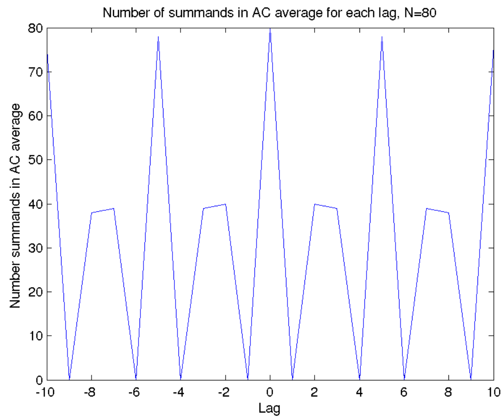


Figure 1: The number of pairs going into the AC average for each lag relative to T_c, the common time for the staggered PRT $2/3$ scheme (so $T_c = T_1/2 = T_2/3$) when the total number of pulses is 80.

c. The Hybrid Spectrum Width Estimator

For the evenly spaced pulse scheme, this estimator is discussed in detail in Meymaris et al. 2009. Briefly, the basic idea comes from the fact that different estimators have various strengths and weaknesses. For example, $R0R1$ works well for wide spectrum widths (relative to the Nyquist velocity), $R1R2$ performs well for medium spectrum widths, and $R1R3$ performs well for narrow. $R0R1$, $R0R1R2$, and $R1R3$ are used to determine whether the spectrum width is small, medium, or large, taking into account that certain types of mistakes are worse than others. The logic for this step is determined a-priori by using simulations in conjunction with decision trees. Once the size category has been determined, the appropriate estimator is used: $R1R3$, $R1R2$, or $R0R1$ for small, medium, and large (resp.). The performance of the estimator has been shown in that paper to be generally superior to the traditional $R0R1$ estimator in places of low SNR and narrow true spectrum width.

The approach to developing the HSW is to, first, identify the 3 estimators to use for small, medium, and large along with corresponding cutoffs (normalized by the Nyquist velocity). This is done by examining the performance of the various estimators. Then, using simulations and simple classification decision trees, identify the best estimators and thresholds to determine whether simulated data is small, medium, or large. The procedure is repeated for different possible number of pulses. The PRT chosen for the tuning is the largest one possible, which is the hardest case to deal with. The SNR chosen for the tuning in 8 dB.

A case study comparing the traditional $R0R1$ estimator to the hybrid spectrum width estimator for an evenly spaced pulse scheme is shown in figure 2. The hybrid spectrum width estimator has less variance and the meteorological features are much clearer.

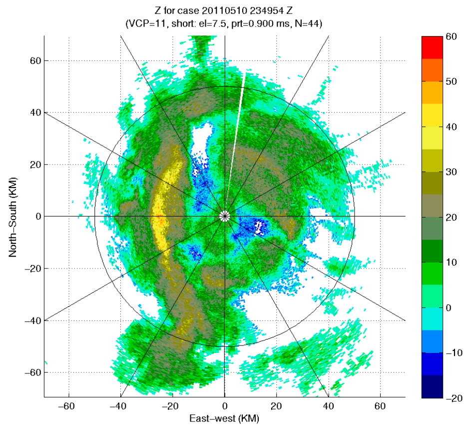
4. Proposed Staggered PRT HSW

For staggered PRT data, we chose to look at $R0R2$, $R0R3$, $R2R5$, $R2R7$, and $R3R7$ for the pulse-pair estimators and $R0R2R3$ and $R0R2R3R5$ for the multi-lag estimators. Note that when discussing staggered PRT data, the lags are relative to T_c and not T_1 or T_2. We also considered averages (various combinations of two) of the already listed estimators for use in the size determination. Our approach remained the same as for the evenly-spaced pulse case. However, the problem is simpler here due to the fact that the currently proposed NEXRAD volume control patterns (VCP) that include staggered PRT have very few options. Namely:

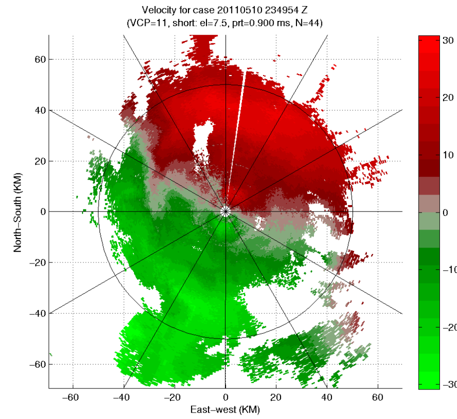
N	T_1 (μs)
46	1497
56	1251
62	1128
80	880

Table 1: Options for T_1 and N in current VCPs that include staggered PRT

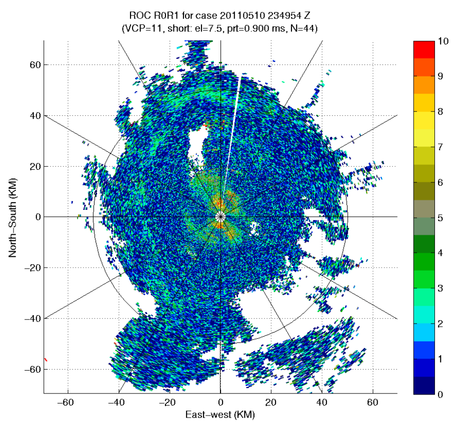
This allows for a more aggressive tuning since, in the evenly spaced pulse scheme, the tuning had to accommodate all the Doppler PRTs and range of



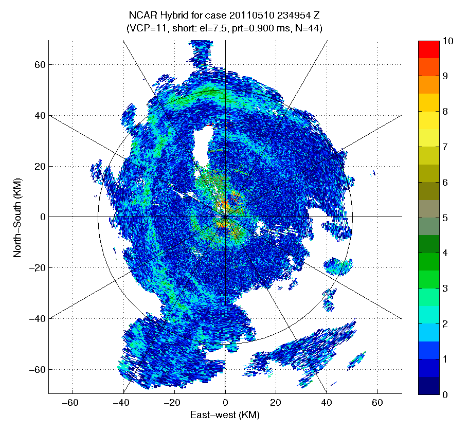
(a) Z



(b) V



(c) $R0R1$ spectrum width



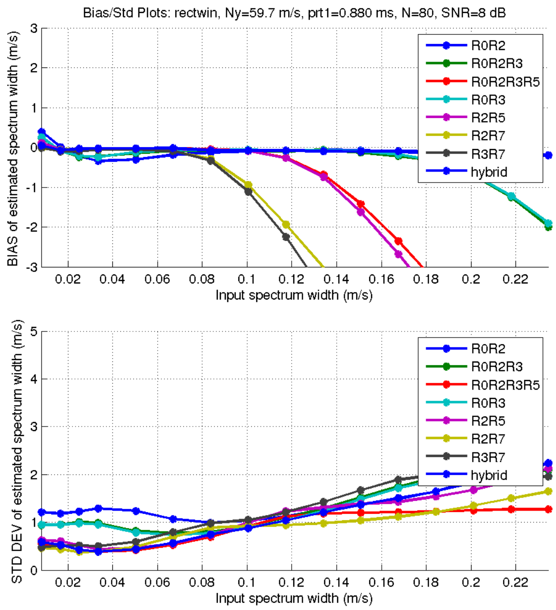
(d) Hybrid spectrum width

Figure 2: Case study of evenly spaced data from KOUN May 10, 2011 23:49:54Z. VCP 11, 7.5° elevation, $900\,\mu s$ PRT, $N = 44$.

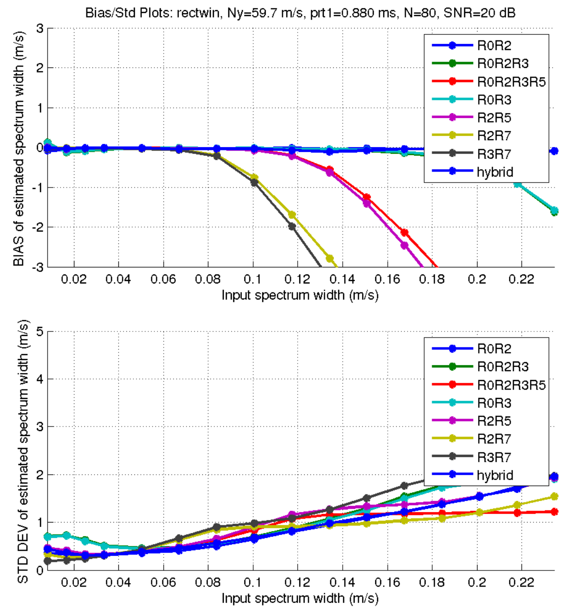
dwell times.

Figure 3 shows the bias and standard deviations of the proposed estimators for $N = 80$, $\mathrm{SNR} = 8$, 20 dB, and $T_1 = 880\,\mu s$. It was determined using this data, along with others (for example, figure 4), that $R0R2$ was the best (smallest bias and standard deviation) for wide normalized spectrum widths (larger than 0.1341), $R0R3$ was the best of medium normalized widths (between 0.0838 and 0.1341), and $R0R2R3R5$ was the best for small normalized widths (less than 0.0838). Note that the normalized widths here are normalized by the Nyquist velocity based on T_c. These cutoffs are used only in comparison to the *true* spectrum width and never with the estimated spectrum widths.

Next, we simulate 10,000 I&Q time-series for the different values and N, with corresponding PRT, for various spectrum widths (from 0.5 to 14 ms^{-1}) and compute each of the different spectrum width estimators. We then feed all the estimator outputs along with the truth field (small, medium, and large classification represented as 1, 2, and 3, resp.) into the MATLAB® decision tree software. This software generates a very deep tree and thus we trim the tree down to just 2 decisions for simplicity. It should be noted that the decision tree software is also provided with a cost matrix that allows the tree to be tuned to take into account the fact that some misclassifications are worse (more costly) than others. For example, it is better to use the large spectrum width estimator on a small spectrum width (result: poorer performance) than vice versa (result: possibly severe saturation). An example output of this process is shown in figure 5. Because the decision trees for the different values of N, with corresponding PRT, turn out to be the same except for

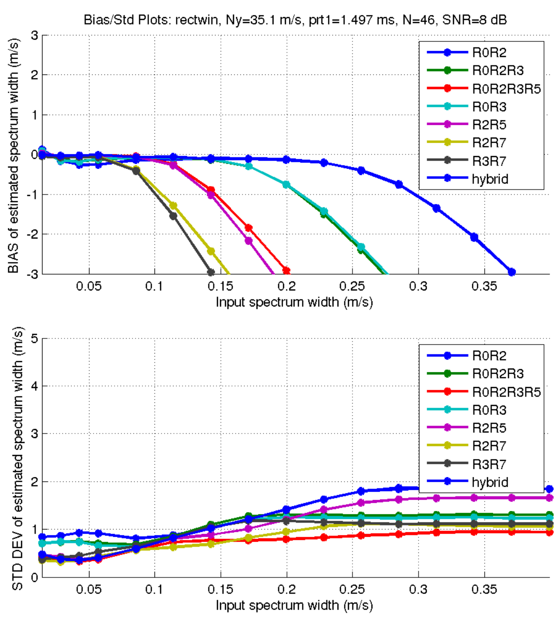


(a) 8 dB

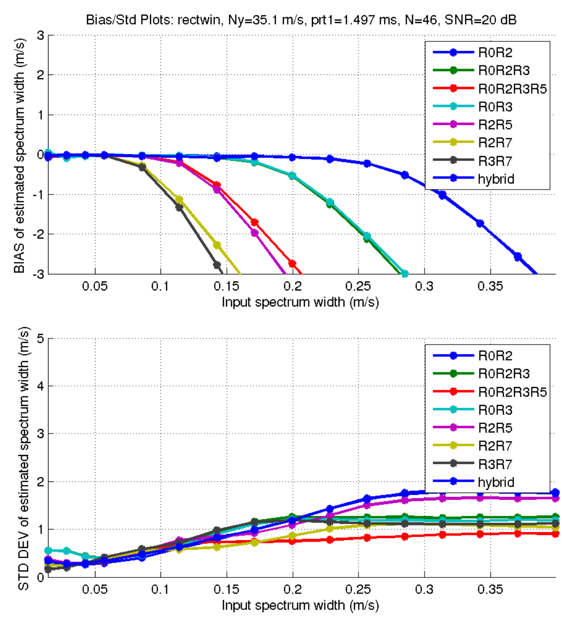


(b) 20 dB

Figure 3: Bias and standard deviation for 7 proposed estimators along with the proposed HSW for $T_1 = 880$ μs, $N = 80$, and $\mathrm{SNR} = 8,\ 20$ dB. Note that the input spectrum widths (*x*-axis) are normalized by the Nyquist velocity of corresponding to T_c.



(a) 8 dB



(b) 20 dB

Figure 4: Bias and standard deviation for 7 proposed estimators along with the proposed HSW for $T_1 = 1497$ μs, $N = 46$, and $\mathrm{SNR} = 8,\ 20$ dB. Note that the input spectrum widths (*x*-axis) are normalized by the Nyquist velocity of corresponding to T_c.

the cutoffs, a single decision logic is used with cutoffs stored in a lookup table. In the case where $N = 62$, the decision tree was unable to discriminate between small and medium with adequate skill, so in that case the lower cutoff is set to -1, thus ensuring that the narrow spectrum width estimator is not used.

5. Results

We use the same simulations (different realizations of the data, however) to evaluate the hybrid spectrum width estimator. Some results shown as 2-D histograms are shown in figures 6 and 7. The true spectrum width (*x*-axis) versus the estimator output (*y*-axis) are shown for SNRs of 5 and 12 dB, for both $R0R2$ and HSW, and for both $N = 80$ with $T_1 = 880\,\mu s$ (figure 6) and $N = 46$ with $T_1 = 1497\,\mu s$ (figure 7). As can be seen for all cases, HSW outperforms the $R0R2$ estimator when the true spectrum widths are small or medium, and have similar performance for large spectrum widths.

Bias and standard deviation as a function of true spectrum width of $R0R2$ and HSW for varying SNRs (5, 8, 12, and 10 dB) are shown for both $N = 80$ with $T_1 = 880\,\mu s$ (figure 8) and $N = 46$ with $T_1 = 1497\,\mu s$ (figure 9). As can be seen the performance of HSW and $R0R2$ are essentially the same for 20 dB SNR, but as the SNR decreases the relative improvement of HSW over $R0R2$ increases. Specifically, the improvement is for the small and medium true spectrum widths.

6. Conclusions

A hybrid approach that combines different spectrum widths shows great promise in producing improved overall performance for staggered PRT spectrum width estimation. Marked improvement is seen with low to medium SNRs (under 20 dB) and narrow to medium spectrum widths, and at least did no worse than the $R0/R2$ estimator. Computationally, the hybrid algorithm is fairly modest, requiring fewer operations than the FFT needed by a spectral technique.

Future work includes further tuning and consideration of other spectrum width estimators. Also, the algorithm needs to be evaluated on staggered PRT I&Q data collected from, say, NCAR's S-POL or NSSL's KOUN.

7. Acknowledgments

This research was supported in part by the ROC (Radar Operations Center) of Norman OK. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the ROC.

References

Doviak, R. J. and D. S. Zrnić: 1993, *Doppler Radar and Weather Observations*, Academic Press, San Diego, California. 2nd edition.

Frehlich, R., 2000: Simulation of coherent doppler lidar performance for space-based platforms. *Journal of Applied Meteorology*, **39**, 245–262.

Frehlich, R., L. B. Cornman, and R. Sharman, 2001: Simulation of three-dimensional turbulent velocity fields. *Journal of Applied Meteorology*, **40**, 246–258.

Frehlich, R. and M. J. Yadlowsky, 1994: Performance of mean-frequency estimators for doppler radar and lidar. *Journal of Atmospheric and Oceanic Technology*, **11**, 1217–1230, corrigenda, **12**, 445–446.

Hubbert, J., G. Meymaris, M. Dixon, and S. Ellis: 2011, Modeling and experimental observations of weather radar ground clutter. AMS 27th International Conference on Interactive Information and Processing Systems, Seattle, WA, 8B.6.

Melnikov, V. M. and D. S. Zrnić, 2004: Estimates of large spectrum width from autocovariances. *J. Atmos. Oceanic Technol.*, **21**, 969–974.

Meymaris, G., J. K. Williams, and J. C. Hubbert: 2009, Performance of a proposed hybrid spectrum width estimator for the nexrad orda. AMS 25th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, Phoenix, AZ.

Williams, J. K., L. Cornman, J. Yee, S. G. Carson, and A. Cotter: 2005, Real-time remote detection of convectively-induced turbulence. AMS 32nd Radar Meteorology Conference, Albuquerque, NM.

Zrnić, D. S. and P. Mahapatra, 1985: Two Methods of Ambiguity Resolution in Pulse Doppler Weather Radars. *Aerospace and Electronic Systems, IEEE Transactions on*, **AES-21**, 470–483.

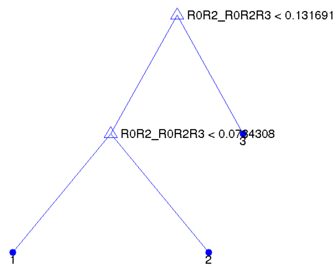
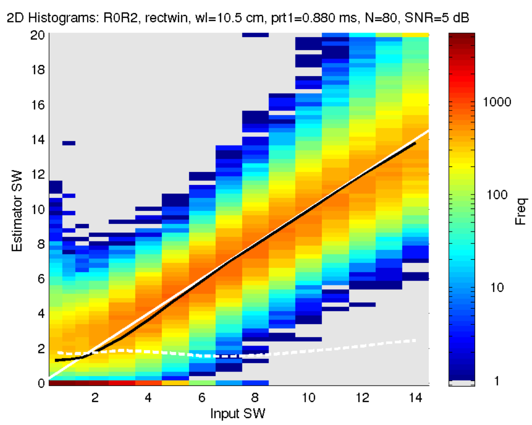
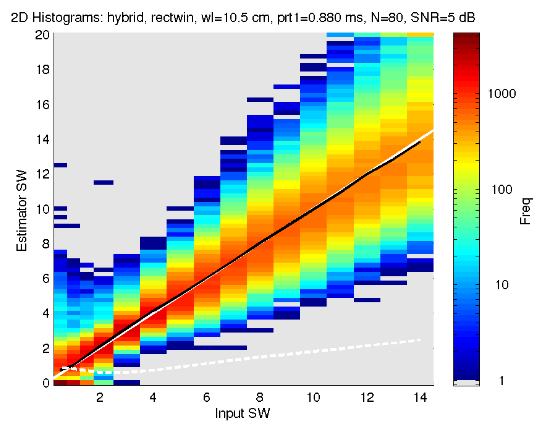


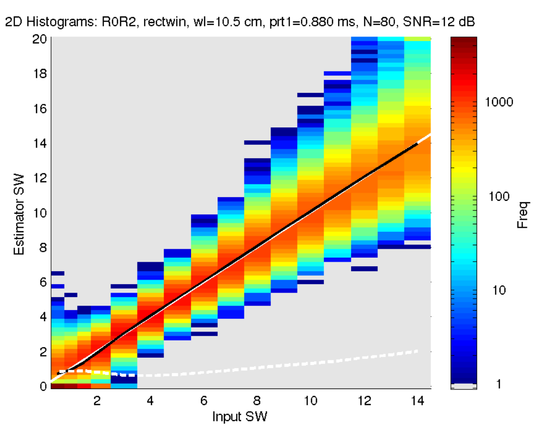
Figure 5: An example classification tree



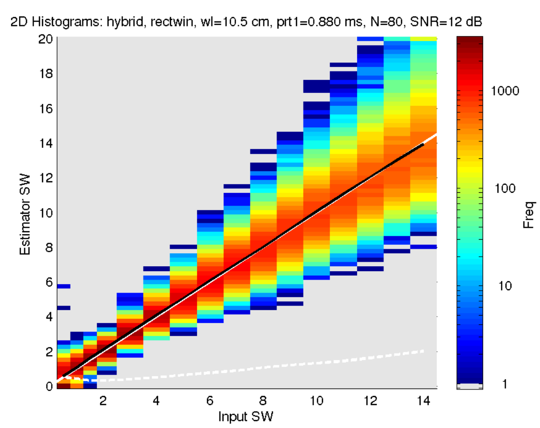
(a) $R0R2$ spectrum width for 5 dB SNR



(b) Hybrid spectrum width for 5 dB SNR

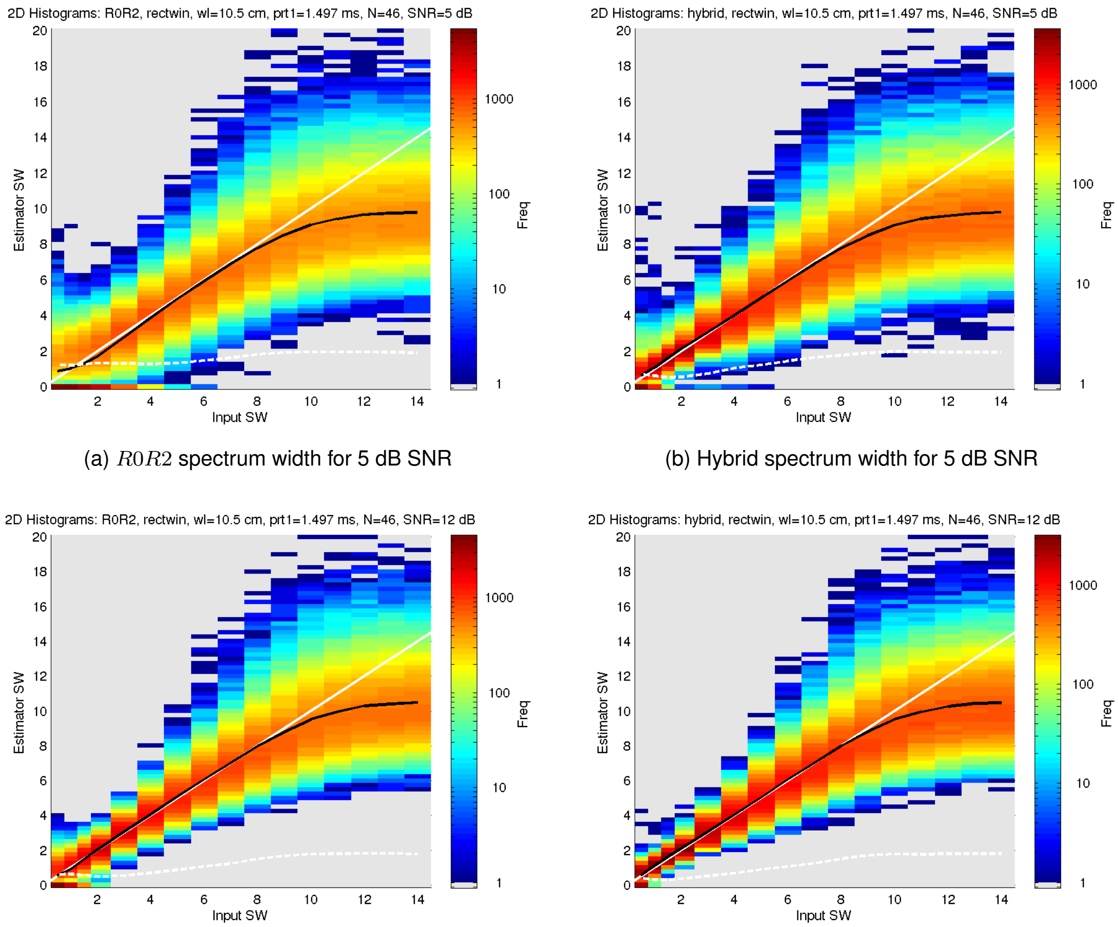


(c) $R0R2$ spectrum width for 12 dB SNR



(d) Hybrid spectrum width for 12 dB SNR

Figure 6: 2-D Histogram of true (x-axis) vs estimated spectrum widths (y-axis) for $T_1 = 880\,\mu s$, $N = 80$. Shown on the left are the results from $R0R2$ and on the right are the results from the staggered PRT hybrid estimator. SNRs of 5 dB (top) and 12 dB (bottom) are shown.



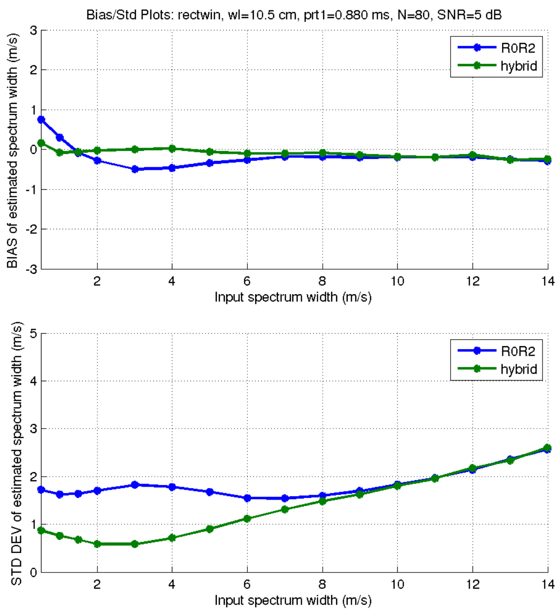
(a) $R0R2$ spectrum width for 5 dB SNR

(b) Hybrid spectrum width for 5 dB SNR

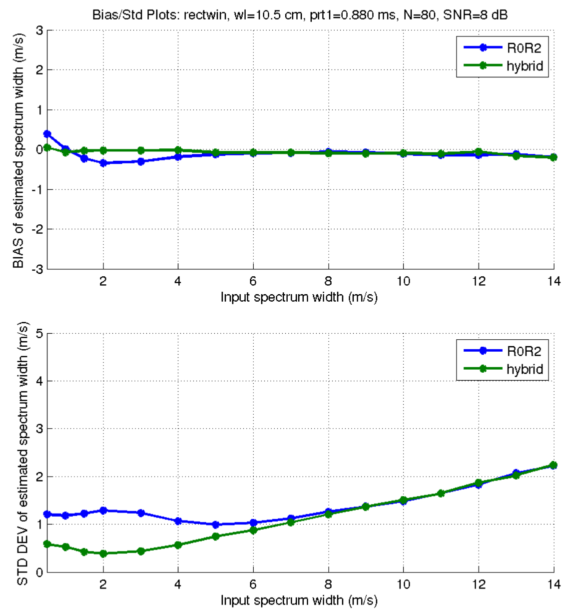
(c) $R0R2$ spectrum width for 12 dB SNR

(d) Hybrid spectrum width for 12 dB SNR

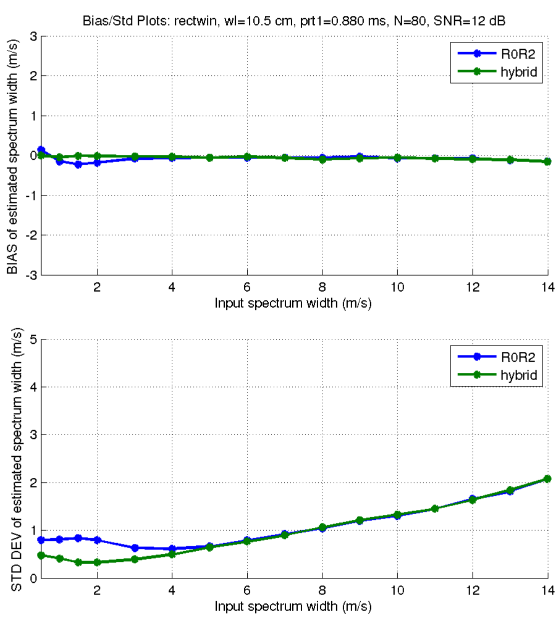
Figure 7: 2-D Histogram of true (x-axis) vs estimated spectrum widths (y-axis) for $T_1 = 1497\,\mu s$, $N = 46$. Shown on the left are the results from $R0R2$ and on the right are the results from the staggered PRT hybrid estimator. SNRs of 5 dB (top) and 12 dB (bottom) are shown.



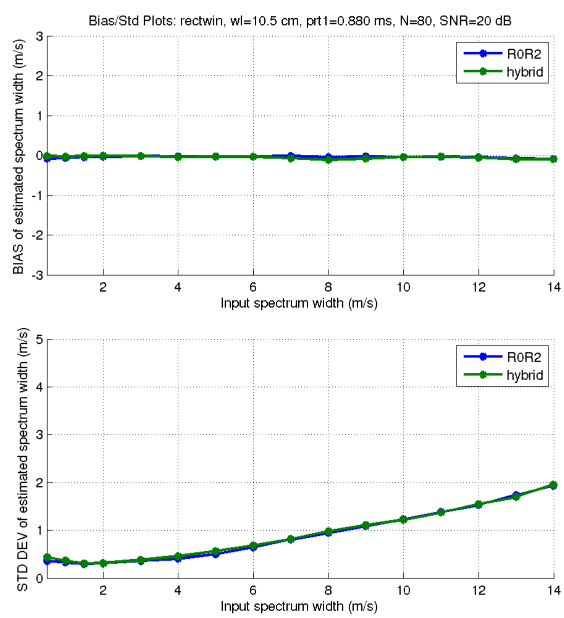
(a) 5 dB SNR



(b) 8 dB SNR

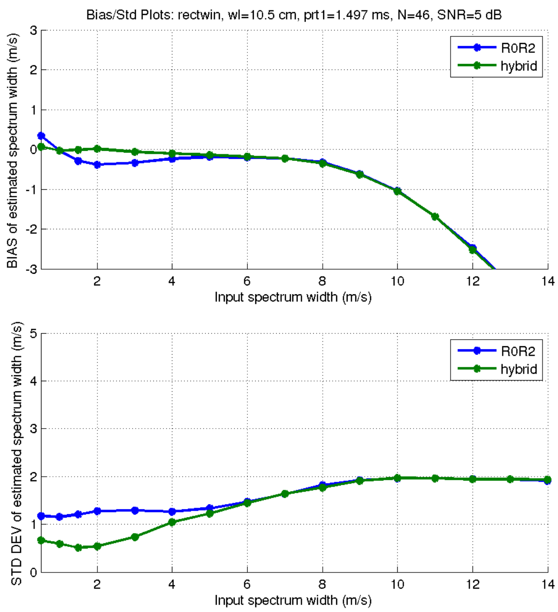


(c) 12 dB SNR

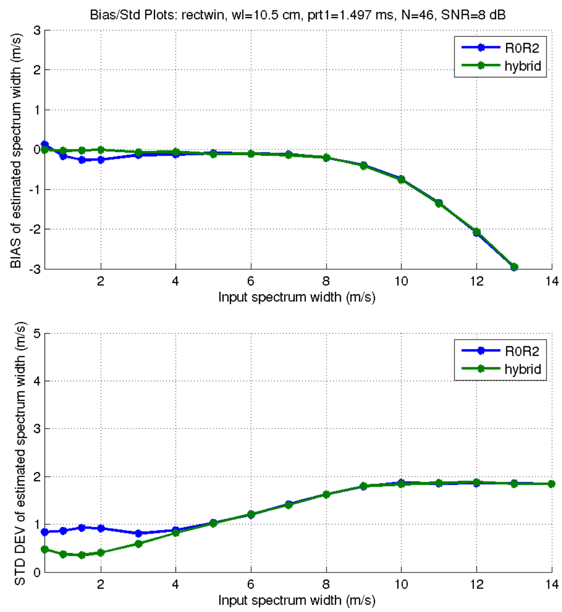


(d) 20 dB SNR

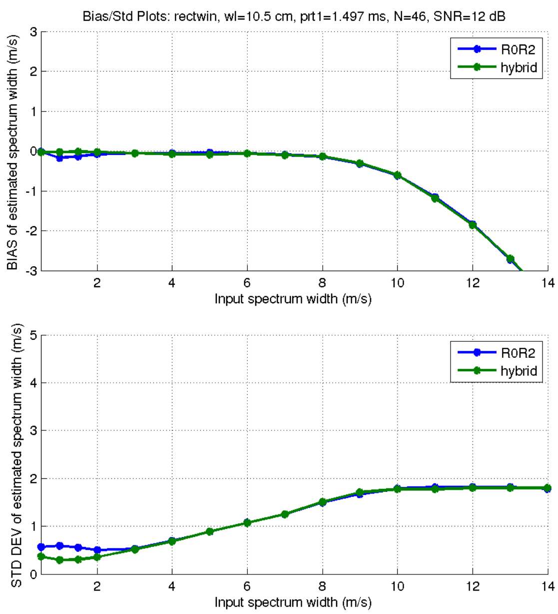
Figure 8: Bias and standard deviation for varying true (x-axis) spectrum widths for $R0R2$ and HSW. $T_1 = 880\,\mu s$, $N = 80$. SNRs shown are 5, 8, 12, and 20 dB.



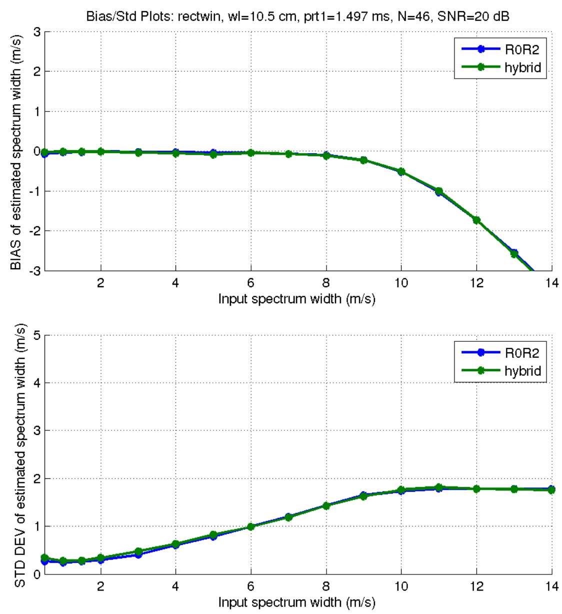
(a) 5 dB SNR



(b) 8 dB SNR



(c) 12 dB SNR



(d) 20 dB SNR

Figure 9: Bias and standard deviation for varying true (*x*-axis) spectrum widths for $R0R2$ and HSW. $T_1 = 1497\,\mu s$, $N = 46$. SNRs shown are 5, 8, 12, and 20 dB.