
Statistical Significance of Gabor Frames Expansions - Simple Filtering Principles for Radar Wind Profiler Data

G. Teschke^{1,2} and V. Lehmann²

¹ University of Applied Sciences Neubrandenburg, Institute for Computational Mathematics in Science and Technology, Brodaer Str. 2, 17033 Neubrandenburg, Germany

² Junior Konrad-Zuse-Fellow, Zuse Institute Berlin (ZIB), Takustr. 7, 14195 Berlin, Germany

³ Deutscher Wetterdienst, Meteorologisches Observatorium Lindenberg, D-15848 Lindenberg, Germany

Summary. A new signal processing method is presented for the suppression of intermittent clutter echoes in radar wind profilers. The technique presented makes use of a discrete Gabor frame expansion in combination with a statistical significance test. The rationale of this algorithm is outlined and an example using data obtained with an operational 482 MHz wind profiler is given.

1 Introduction

Radar wind profilers (RWP) were developed from MST-Radars and have meanwhile become standard instruments for measuring wind velocities in the atmosphere. Overviews of the technical and scientific aspects of RWP including its signal processing have been provided, among others, by e.g. [1]. Especially the routine application by weather services and the assimilation of the data in Numerical Weather Prediction Models is an indicator for the degree of maturation that this technology has achieved, see e.g. [6]. However, it is a matter of fact that sometimes large and unacceptable differences are observed between the profiler data and independent reference measurements. In many cases these differences are clearly attributable to either clutter echoes or Radio Frequency interference. Especially the problem of bird contamination has been well-known for more than a decade and it still is a research topic in RWP signal processing. There exist many attempts to reduce bird contamination, e.g. [5]. However, the disadvantage of all these methods is that the mitigation processing builds upon the Doppler spectra (either before or after spectral integration). Given the highly non-stationary characteristics of the intermittent clutter signal, it is necessary to deal with the problem before the Doppler

spectrum is estimated, because Fourier methods are generally inadequate for nonstationary signals. Further approaches that have tried to overcome these deficiencies by using wavelet representations were suggested by [2] and further by [4]. However, nonredundant wavelet filtering is in several cases also not best suited and causes undesired artifacts leading to erroneous filtering results. In this paper, we discuss a new signal-clutter separation method that circumvents these problems. It is based on a *Gabor frame decomposition* of the time series followed by the *statistical filtering approach* suggested by [5]. For an extensive description of the presented approach and a discussion in much greater detail we refer the interested reader to [3].

2 Classical signal model and its limitations

The classical RWP signal model assumption can be written as

$$\mathbf{S}[k] = \mathbf{I}[k]e^{i\omega k\Delta t} + \mathbf{N}[k], \quad (1)$$

where $\mathbf{I}[k] \sim N(0, \sigma_{\mathbf{I}}^2)$ and $\mathbf{N}[k] \sim N(0, \sigma_{\mathbf{N}}^2)$ are independent complex zero-mean Gaussian random vectors describing the atmospheric signal and the receiver noise, Δt is the sampling interval of the sequence and ω the mean Doppler frequency. Furthermore $\mathbf{I}[k]$ is narrowband compared to the receiver bandwidth and $|\omega| \leq \pi/\Delta t$ (Nyquist criterion). Because $\mathbf{S}[k]$ is the result of the demodulation of a real valued zero-mean and stationary Gaussian random process, the resulting Gaussian complex random process is also wide-sense stationary and zero-mean. Furthermore, the sequence has a vanishing pseudocovariance, that is we have $\mathbf{E}(\mathbf{S}[k]\mathbf{S}[l]) = 0$. Such a process is usually called proper, circular or phase-invariant. Therefore,

$$\begin{aligned} (\mathbf{R})_{k,l} &= \text{Cov}(\mathbf{S}[k], \mathbf{S}[l]) = \mathbf{E}(\mathbf{I}[k]\bar{\mathbf{I}}[l])e^{i\omega(k-l)\Delta t} + \mathbf{E}(\mathbf{N}[k]\bar{\mathbf{N}}[l]) \\ &= \sigma_{\mathbf{I}}^2 \varrho[k-l]e^{i\omega(k-l)\Delta t} + \sigma_{\mathbf{N}}^2 \delta_{k-l,0}, \end{aligned}$$

where ϱ is specified below. While this is a classical assumption in radar signal processing, it is unknown for which maximal time series length this assumption can be made safely. We found that bird clutter signals are significantly nonstationary over typically used dwell times of about 30 s to 60 s. The associated autocovariance function can be expressed as follows

$$\text{ACov}(k) = \sigma_{\mathbf{I}}^2 \varrho[k]e^{i\omega k\Delta t} + \sigma_{\mathbf{N}}^2 \delta_{k,0} = \sigma^2 \rho[k], \quad (2)$$

where we set $\sigma^2 := \sigma_{\mathbf{I}}^2 + \sigma_{\mathbf{N}}^2$ and $\rho[k] := \frac{\sigma_{\mathbf{I}}^2 \varrho[k]e^{i\omega k\Delta t} + \sigma_{\mathbf{N}}^2 \delta_{k,0}}{\sigma_{\mathbf{I}}^2 + \sigma_{\mathbf{N}}^2}$, while assuming $\varrho[k] = e^{-2\pi^2 w^2 k^2 \Delta t^2}$. In reality, however, there is sometimes a third component contributing to the signal, namely clutter [7], so that the signal model must be written as:

$$\mathbf{S}[k] = \mathbf{I}[k]e^{i\omega k\Delta t} + \mathbf{N}[k] + \mathbf{C}[k]. \quad (3)$$

Clutter is the totality of undesired echoes and interfering signals, therefore it is impossible to generalize the properties of $\mathbf{C}[k]$. In the case of RWP, clutter includes in particular echoes from airborne objects such as aircraft and birds as well as returns from the ground. Interfering signals may be caused by other radio transmitters operating in the RWP receiver band. In the remainder of the paper, we restrict ourselves to intermittent clutter signals and its removal from \mathbf{S} .

3 Gabor frame expansions for discretely sampled signals

Assume we are given some discrete and finite time (periodic) signal $\tilde{\mathbf{S}}$ with sampling points $n = 0, \dots, N - 1$, that is $\tilde{\mathbf{S}}[n] = \tilde{\mathbf{S}}[n + N]$. We therefore have to periodize the analysis and synthesis windows as well,

$$\tilde{\mathbf{h}}[n] = \sum_l \mathbf{h}[n + lN] \quad , \quad \tilde{\mathbf{g}}[n] = \sum_l \mathbf{g}[n + lN].$$

Slightly abusing the notation, we omit the tilde denoting periodic (finite) functions in the following. The signal \mathbf{S} can be discretely represented by

$$\mathbf{S}[n] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} a_{m,k} \mathbf{h}_{m,k}[n] \quad , \quad (4)$$

whereas the Gabor coefficients can be derived from

$$a_{m,k} = \sum_{n=0}^{N-1} \mathbf{S}[n] \bar{\mathbf{g}}_{m,k}[n] \quad . \quad (5)$$

Introducing integers ΔM and ΔK and the toral component $W_N = \exp[2\pi i/N]$, the discrete analysis and synthesis windows can be rewritten as

$$\begin{aligned} \mathbf{h}_{m,k}[n] &= \mathbf{h}[n - m\Delta M] W_N^{nk\Delta K} \quad , \\ \mathbf{g}_{m,k}[n] &= \mathbf{g}[n - m\Delta M] W_N^{nk\Delta K} \quad . \end{aligned}$$

As can be seen, ΔM denotes the time and ΔK the frequency step size. They correspond to T and Ω . In our setting they are constrained by $\Delta M \cdot M = \Delta K \cdot K = N$. The reconstruction formula becomes

$$\mathbf{S}[j] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} a_{m,k} \mathbf{h}_{m,k}[j] = \sum_{l=0}^{N-1} \mathbf{S}[l] \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} \bar{\mathbf{g}}_{m,k}[l] \mathbf{h}_{m,k}[j] \quad ,$$

where we have assumed that the following biorthogonality relation is fulfilled,

$$\sum_{m=0}^{M-1} \sum_{k=0}^{K-1} \bar{\mathbf{g}}_{m,k}[l] \mathbf{h}_{m,k}[j] = \delta_{l,j} \quad .$$

It can be shown that the biorthogonality relation is satisfied if

$$\sum_{j=0}^{N-1} \mathbf{h}[j + qK] W_N^{-jpM} \bar{\mathbf{g}}[j] = \frac{N}{MK} \delta_{p,0} \delta_{q,0} \quad (6)$$

for $0 \leq p \leq \Delta M - 1$ and $0 \leq q \leq \Delta K - 1$. System (6) can be rewritten in matrix form: Let $\mathbf{v} = (N/(MK), 0, \dots, 0)^T$ be a vector of length $\Delta M \Delta K$ and $\mathbf{g} = (\mathbf{g}[0], \dots, \mathbf{g}[N-1])$ the vector representing the discretely sampled dual frame, and let \mathbf{A} be the matrix of size $\Delta M \Delta K \times N$ with entries $\mathbf{A}_{(p,q),j} = \bar{\mathbf{h}}(j + qK) W_N^{jpM}$, then the dual frame atom \mathbf{g} is the solution of the linear system

$$\mathbf{A} \mathbf{g} = \mathbf{v} . \quad (7)$$

For oversampling $\Delta M \Delta K < N$, system (7) is under-determined, and the solution is no longer unique and therefore there is a variety of possible dual frame atoms \mathbf{g} . One suitable choice (beside optimal localizing window functions) is given by the minimum norm solution $\mathbf{g}_{min} = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1} \mathbf{v}$.

4 Statistical significance, filtering and a real example

First, we observe that

$$|a_\lambda|^2 = \sum_{n=0}^{N-1} \mathbf{S}[n] \mathbf{g}_\lambda[n] \sum_{l=0}^{N-1} \bar{\mathbf{S}}[l] \bar{\mathbf{g}}_\lambda[l].$$

With $\mathbf{E} \mathbf{S}[n] = 0$ and $\mathbf{E} \mathbf{S}[n] \bar{\mathbf{S}}[n+l] = \sigma^2 \rho[l]$ we obtain $\mathbf{E} |a_\lambda|^2 = \sigma^2 \langle \rho * \mathbf{g}_\lambda, \mathbf{g}_\lambda \rangle$ and $\text{Cov}(|a_\lambda|^2, |a_\eta|^2) = \sigma^4 |\langle \rho * \mathbf{g}_\lambda, \mathbf{g}_\eta \rangle|^2$. The ‘*’-symbol stands here for the discrete convolution. Therefore,

$$\text{Var} |a_\lambda|^2 = \sigma^4 |\langle \rho * \mathbf{g}_\lambda, \mathbf{g}_\lambda \rangle|^2 = (\mathbf{E} |a_\lambda|^2)^2 \text{ and thus } \frac{(\mathbf{E} |a_\lambda|^2)^2}{\text{Var} |a_\lambda|^2} = 1 \quad (8)$$

which holds true for independent as well as dependent samples $\mathbf{S}[n]$ that follow a distribution which is determined by its moments. In order to construct a statistical test that verifies property (8), we have to find optimal estimators for $\mathbf{E} |a_\lambda|^2$ and $\text{Var} |a_\lambda|^2$ that are based on a finite number of observations. To this end, we introduce an index subset $\Omega_\lambda \subset \Lambda$ containing λ and $L-1$ further different indices η , i.e. $|\Omega_\lambda| = L$. As an estimator for $\mathbf{E} |a_\lambda|^2 = \sigma^2 \langle \rho * \mathbf{g}_\lambda, \mathbf{g}_\lambda \rangle$, which is based on L neighboring observation variables, we define

$$\hat{E}(\Omega_\lambda) := \frac{1}{C_{\Omega_\lambda}} \sum_{\eta \in \Omega_\lambda} |a_\eta|^2 \text{ with } C_{\Omega_\lambda} = \sum_{\eta \in \Omega_\lambda} \frac{\langle \rho * \mathbf{g}_\eta, \mathbf{g}_\eta \rangle}{\langle \rho * \mathbf{g}_\lambda, \mathbf{g}_\lambda \rangle} > 1 . \quad (9)$$

Assuming there exists some small $\varepsilon > 0$ with $\sum_{\eta', \eta \in \Omega_\lambda} |\langle \rho * \mathbf{g}_{\eta'}, \mathbf{g}_\eta \rangle|^2 \leq C_{\Omega_\lambda}^{2-\varepsilon}$, estimator (9) consistent, for details see [3]. By the same reasoning, we define a consistent estimator for variance,

$$\hat{V}(\Omega_\lambda) := C \sum_{\eta \in \Omega_\lambda} (|a_\eta|^2 - \hat{E}(\Omega_\lambda))^2, \quad (10)$$

where the constant is defined by

$$C^{-1} := 2 \sum_{\eta \in \Omega_\lambda} \frac{c_\eta^2}{c_\lambda^2} + (L - 2C_{\Omega_\lambda}) \left(1 + \frac{1}{(\sum_\eta c_\eta)^2} \sum_{\xi, \alpha \in \Omega_\lambda} c_{\xi, \alpha}^2 \right). \quad (11)$$

To identify now intermittent clutter (the nonstationary signal component), we proceed as follows: In a first step, define the index set representing the k -th row, which we denote by $\Omega_k = \{(m, k) : m = 0, \dots, M-1\}$, and sort for each k the sequence $\{|a_{m,k}|^2\}_{(m,k) \in \Omega_k}$ in decreasing order. That is, we derive the order statistic of $\{|a_{m,k}|^2\}_{(m,k) \in \Omega_k}$ which we denote by $\{[a]_{m,k}\}_{(m,k) \in \Omega_k}$ ($[\cdot]$ stands for the order statistic map). Therefore, we have $[a]_{m,k} \geq [a]_{m+1,k}$ for all $(m, k) \in \Omega_k$. For $l = 0, \dots, M-1$, we define subsets $\Omega_k(l) = \{(m, k) : m = l, \dots, M-1\}$. The largest coefficients are stepwise discarded, which has the goal of eliminating the clutter signal component. Using the quantities $\hat{E}(\Omega_k(l))$ and $\hat{V}(\Omega_k(l))$ of the subset, the test statistics ϑ is computed for $l = 0, \dots, M-1$ as long as $\vartheta([a]_{l,k}) := \frac{(\hat{E}(\Omega_k(l)))^2}{\hat{V}(\Omega_k(l))} < 1$ holds. The largest coefficient of the first subset for which the test (positive for clutter) is not satisfied (a clutter-free subset) is then taken as a threshold for a frequency-dependent identification of the clutter component. All coefficients $|a_{m,k}|^2$ greater than the threshold are regarded as clutter. Based on this test, we introduce a clutter index set as $\Omega_k^c := \{(m, k) : \vartheta([a]_{m,k}) < 1, m = 0, \dots, M-1\}$. The coefficients $a_{m,k} \in \Omega_k^c$ are finally set to $t_k e^{i \arg a_{m,k}}$, where t_k is the average value of the remaining coefficients, $t_k = \frac{1}{|\Omega_k \setminus \Omega_k^c|} \sum_{(m,k) \in \Omega_k \setminus \Omega_k^c} |a_{m,k}|$. Consequently, the filtered signal \mathbf{S} is given by

$$\Phi(\mathbf{S})[n] = \sum_{k=0}^{K-1} \left\{ \sum_{(m,k) \in \Omega_k \setminus \Omega_k^c} a_{m,k} \mathbf{h}_{m,k}[n] + \sum_{(m,k) \in \Omega_k^c} t_k e^{i \arg a_{m,k}} \mathbf{h}_{m,k}[n] \right\}. \quad (12)$$

To show the performance of the proposed algorithm, we process data that were obtained during routine operation of a 482 MHz wind profiler radar of the Deutscher Wetterdienst at Bayreuth, Germany in the fall of 2005. We consider data taken in the south beam of the radar wind profiler at range gate 9 (1.6 km height agl, dwell at 00:09:45 UTC). Figure 1 shows a time series in which strong intermittent clutter (bird echo) can be recognized. The results of the filtering procedure illustrate that the method completely eliminates the nonstationary signal component.

References

- [1] Richard J. Doviak and Dusan S. Zrnic. *Doppler Radar and Weather Observations*. Academic Press, 1993.

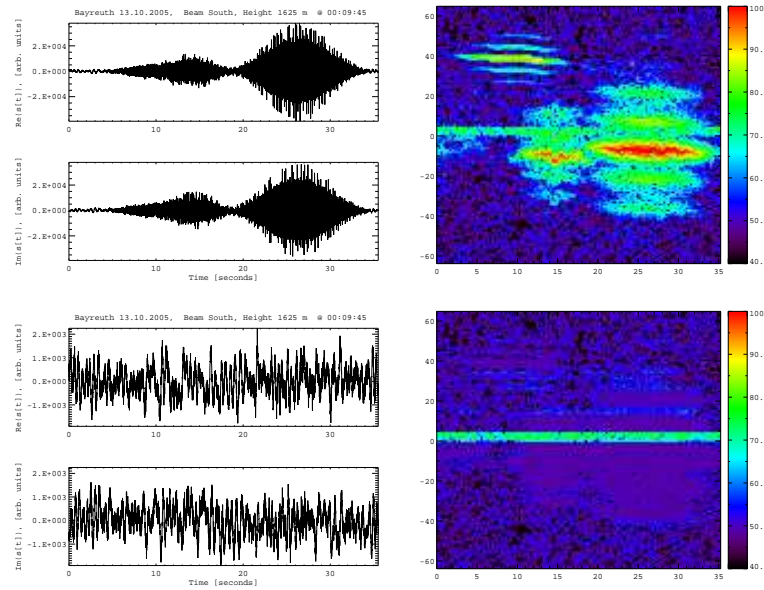


Fig. 1. Top left: real time series; top right: Gabor spectrum of this time series; bottom right: filtered Gabor spectrum; bottom left: reconstructed time series.

- [2] James R. Jordan, Richard J. Lataitis, and David A. Carter. Removing ground and intermittent clutter contamination from wind profiler signals using wavelet transforms. *J. Atmos. Ocean. Tech.*, 14:1280–1297, December 1997.
- [3] V. Lehmann and G. Teschke. Advanced intermittent clutter filtering for radar wind profiler: signal separation through a gabor frame expansion and its statistics. *Ann. Geophys.*, 26(4):759–783, 2008.
- [4] Volker Lehmann and Gerd Teschke. Wavelet based methods for improved wind profiler signal processing. *Ann. Geophys.*, 19:825–836, 2001.
- [5] Dave A. Merritt. A statistical averaging method for wind profiler Doppler spectra. *J. Atmos. Ocean. Tech.*, 12(5):985–995, 1995.
- [6] Wim A. Monna and Russell B. Chadwick. Remote-sensing of upper-air winds for weather forecasting: Wind-profiler radar. *B. World Meteorol. Organ.*, 47(2):124–132, April 1998.
- [7] Andreas Muschinski, Volker Lehmann, Lutz Justen, and Gerd Teschke. Advanced radar wind profiling. *Meteorol. Z.*, 14(5):609–626, 2005.