Advanced Intermittent Clutter Filtering for Radar Wind Profiler: Signal Separation through a Gabor Frame Expansion and its Statistics

Volker Lehmann *
Deutscher Wetterdienst, Meteorologisches Observatorium Lindenberg, Germany

Gerd Teschke
Zuse-Institute Berlin
Takustr. 7, 14195 Berlin, Germany

Corresponding author address: Volker Lehmann, Meteorologisches Observatorium Lindenberg, Am Observatorium 12, D-15848 Tauche-Lindenberg, Germany.
E-mail: Volker.Lehmann@dwd.de

Abstract

A new method is presented for the suppression of intermittent clutter echoes in radar wind profilers. This clutter type is a significant problem during the seasonal bird migration and often results in large discrepancies between profiler wind measurements and independent reference data. The technique presented makes use of a discrete Gabor frame expansion of the coherently averaged time series data in combination with a statistical filtering approach to exploit the different signal characteristics between signal and clutter. The rationale of this algorithm is outlined and the mathematical methods used are presented in due detail. A first test using data obtained with an operational 482 MHz wind profiler indicates that the method outperforms the previously used clutter suppression algorithm.

Contents

1. Introduction

Radar wind profilers (RWP) were developed from MST-Radars (Van Zandt 2000) 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 Gage (1990); Röttger and Larsen (1990); Doviak and Zrnic (1993) and Muschinski (2004). 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. Monna and Chadwick (1998); Bouttier (2001); Benjamin et al. (2004b); St-James and Laroche (2005); Ishihara et al. (2006). 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. These spurious signals are often easily discernible in the Doppler spectrum by human experts, but not always adequately handled by the automatic processing. For that reason, research on improvements in wind profiler signal processing has remained a very active field over the last decade.

In this paper, we deal with so-called intermittent clutter and propose a new filtering algorithm for the detection and suppression of these clutter signals in the profiler raw data. Of particular importance are intermittent clutter echoes, which are caused by migrating birds in Spring and Fall. It is well known, that birds are effective targets for a wide range of radars from X-band to UHF (Vaughn 1985; Bruderer 1997a). In fact, most of the knowledge about migrating birds come from radar observations. That concerns in particular their flight behavior under the influence of environmental factors (Bruderer 1997b). Radar ornithology is meanwhile a mature field and it is therefore no surprise, that birds can also be detected by the the sensitive radar systems used for wind profiling. The susceptibility of wind profiler radar systems to bird echoes depends primarily on wavelength and antenna characteristics. It mostly affects L-Band and UHF-systems, that is Boundary Layer profilers and Tropospheric profilers, as discussed in Wilczak et al. (1995). Intermittent clutter is of course also an issue for the new generation of imaging radar systems, like the Turbulent Eddy Profiler (Cheong et al. 2006). We mention in passing that other remote sensing instruments used in Meteorology are also affected by migrating birds (Mastrantonio et al. 1999; Gauthreaux and Belser 1998; Gauthreaux et al. 1998; Zhang et al. 2005; Liu et al. 2005).

Intermittent clutter echoes caused by aircraft were already mentioned by Hogg et al. (1983), and a few years later it become obvious that especially echoes from migrating birds can be a serious issue in wind profiling (Ecklund et al. 1990; Barth et al. 1994). If present, such spurious signals can cause a significant deterioration of the quality of the derived winds. To give an example, the investigation of low-level jets using RWP data is hampered by bird migration clutter (Stensrud 1996). This makes it necessary to either use extensive quality control procedures to identify and skip contaminated data (Daniel et al. 1999; Song et al. 2005) or to limit the studies to periods where bird migration is negligible (Anderson and Arritt 2001). Many other investigations using RWP data have mentioned the bird contamination problem, e.g. Ralph et al. (1998); Locatelli et al. (1998); Parker and Johnson (2000); Lundquist (2003). While the need for an extensive manual data quality control and cleaning might be acceptable for research activities, it is surely not feasible in any operational setting. Nevertheless it is mandatory to avoid the assimilation of bird contaminated profiler wind data, as this can have significant effects on the quality of the forecasts (Semple 2005). Due to the nature of the problem, a bird migration check at the operational center itself is not the best approach (Benjamin et al. 2004a). While current

state-of-the art profilers nowadays run more or less sophisticated algorithms on site to reduce bird contamination (Merritt 1995; Jordan et al. 1997; Ishihara et al. 2006), practical experience supports the statement that the problem has not been fully resolved.

The problem of bird-contamination is well-known (Wilczak et al. 1995; Engelbart et al. 1998) and it has been a topic for research in RWP signal processing since. The first successful attempt to reduce bird contamination was made by Merritt (1995), who suggested a selective averaging method of the individual Doppler spectra based on a statistical criterion. The same method has also been applied off-line to averaged spectra, when data with higher resolution are not available (Pekour and Coulter 1999). A similar approach was taken by Weber (2005), who used neural networks for a classification of contaminated single spectra, followed by a selective averaging. Other proposals have concentrated on modified peak detection in the Doppler spectrum to address spurious flier returns, among other clutter types (Griesser and Richner 1998; Cornman et al. 1998; Morse et al. 2002; Weber et al. 2004). 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 bird problem at the earliest possible stage of RWP signal processing, that is before the Doppler spectrum is estimated. Fourier methods are generally inadequate for nonstationary signals, so it seems to be prudent to address the bird contamination problem before any Fourier transform is made. In other words, the necessary nonlinear filtering has to be performed in the time domain. This approach was first suggested by Jordan et al. (1997) and further by Lehmann and Teschke (2001), who suggested wavelet decomposition and wavelet coefficient thresholding, to remove the clutter part of the signal. However, the a-priori unclear choice of the mother wavelet and - at least for the dyadic wavelet transform - a suboptimal signal separation in the wavelet domain, especially near zero Doppler shift, makes an efficient separation of clutter and signal difficult.

Ideally one would like to have a intermittent clutter suppression algorithm that reduces the clutter part of the signal as best as possible, given the sampled data and that quantifies its *degree of contamination*, that is to provide some measure of clutter energy for quality control purposes. Furthermore the algorithm must not degrade both data quality and availability in the no-clutter case, but it should perform as well as the proven standard processing methods. This requirement is more stringent than it may appear at first glance. In this paper, we propose a new signal-clutter separation method that attempts to meet these objectives. It is based on a redundant frame decomposition of the time series followed by the statistical filtering approach suggested by Merritt (1995).

The paper is organized as follows. Section 2 gives an overview of RWP signal characteristics and signal processing and identifies shortcomings of the currently used methods when intermittent clutter signals are present. Section 3 reviews basic results of the mathematical theory of frames, which deals with linear discrete signal representations. The goal is here to find a signal representation, that achieves optimal separation between the atmospheric and the clutter part of the signal. This is achieved by the discrete Gabor representation, which is discussed next. Section 4 focuses on a statistical approach to objectively identify the atmospheric signal component, based on well-justified statistical assumptions. A comparison of the new algorithm with the previously used signal processing techniques is shown in section 5. The data used were obtained during routine operation of a 482 MHz wind profiler radar of the Deutscher Wetterdienst at Bayreuth, Germany in the fall of 2005. Finally, a summary and conclusions are given in Section 6.

2. RWP signal characteristics

a. General properties of the received signal

The relationship between the signal received by the radar and the scattering medium is the topic of radar instrument theory, which basically describes how atmospheric properties are mapped to the measurable function at the radar receiver output (Woodman 1991; Muschinski 2004). It is known that both models for the scattering processes and technical properties of the radar system need to be considered here. This task is formidable and requires simplifications. However, for the problem at hand it is not required to consider such theories in detail, because we are only interested in some rather general properties of the received signal, like statistical stationarity. For a pulsed RWP, the received signal at the antenna output has the following well-known properties:

1. Continuous real-valued random voltage signal: Every measurable physical quantity is real. The randomness is the result of the random nature of the scattering process.

2. Intrinsically nonstationary: This is due to the impulsive character of the transmitted signal and the inhomogeneous vertical structure of the atmosphere.

3. Multi-component: Beside the ubiquitous noise, there may be signal contributions from several independent scattering processes, like Bragg scattering at fluctuations of the refractive index, Rayleigh scattering at precipitation and scattering at various clutter targets.

4. Narrowband: The frequency spectrum is bandlimited, with a width much smaller than the transmit signals carrier frequency.

5. Large dynamic range: The signal varies easily over many orders of magnitude, which is typical for all radar systems.

Before the data are available for digital signal processing, the radar receiver performs the following pre-processing steps: Range-gate sampling, quadrature-demodulation and matched filtering. This is generic to all RWP receivers, both analog and digital implementations. The digital receiver output signal preserves the properties 3-5, provided processing is linear (e.g. no saturation effects due to hardware limitations). However, properties 1 and 2 are modified: Due to pulse repetition, the nonstationary continuous signal becomes quasi-periodic. Uniform sampling for N fixed heights at multiples of the radar inter-pulse period then generates N stationary sequences, provided the scattering medium at a fixed height does not change its properties significantly over the length of the time series (Woodman 1991). This is valid for atmospheric scattering, ground clutter and noise and one of the basic assumptions of signal processing for atmospheric radars (Keeler and Passarelli 1990). Furthermore, the quadrature demodulation step leads to a complex baseband representation of the narrowband signal, where the signal is described through the time series of its in-phase (I) and quadrature-phase (Q) components.

b. Classical signal model and its limitations

The classical RWP signal model assumption is that the demodulated voltage sequence at the receiver output can be written as

$$\boldsymbol{S}[k] = \boldsymbol{I}[k]e^{i\omega k\Delta t} + \boldsymbol{N}[k], \tag{1}$$

where $\boldsymbol{I}[k] \sim N(0, \sigma_{\boldsymbol{I}}^2)$ and $\boldsymbol{N}[k] \sim N(0, \sigma_{\boldsymbol{N}}^2)$ are independent complex zero-mean Gaussian random vectors describing the atmospheric signal and the receiver noise, respectively (Zrnić 1979), Δt is the sampling interval of the sequence and ω the mean Doppler frequency. Furthermore $\boldsymbol{I}[k]$ is narrowband compared to the receiver bandwidth and $|\omega| \leq \pi/\Delta t$ (Nyquist criterion). Because $\boldsymbol{S}[k]$ 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 pseudo-covariance, that is we have $\mathsf{E}(\boldsymbol{S}[k]\boldsymbol{S}[l]) = 0$. Such a process is usually called proper, circular or phase-invariant (Neeser and Massey 1993). We will use this property later in connection with a moments theorem for these processes (Reed 1962).

Because $\boldsymbol{S}[k]$ is Gaussian, it is completely characterized through its covariance matrix $\mathbf{R}$ with entries

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

Furthermore, stationarity is assumed over typical dwell-times of $\mathcal{O}(1 \text{ minute})$. Therefore we get the following expression for the autocovariance function

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

where we set $\sigma^2 = \sigma_{\boldsymbol{I}}^2 + \sigma_{\boldsymbol{N}}^2$. Finally, the autocorrelation function $\boldsymbol{\varrho}[k]$ is often assumed to follow a Gaussian correlation model, which corresponds to a Gaussian signal peak in the power spectrum. If the spectral width of the signal is w, then we have (Zrnić 1979; Frehlich and Yadlowsky 1994)

$$\boldsymbol{\varrho}[k] = e^{-2\pi^2 w^2 k^2 \Delta t^2} \,. \tag{3}$$

Note that this Gaussian correlation model must not be confused with the characterization of the random process as Gaussian, which covers a much wider class of signals. The assertions are normally very well justified and therefore successfully used in simulations of the radar signal (Zrnić 1975; Frehlich and Yadlowsky 1994; Muschinski et al. 1999).

In reality, however, there is often a third component contributing to the signal, namely clutter (Muschinski et al. 2005), so that the signal model must be written as:

$$\boldsymbol{S}[k] = \boldsymbol{I}[k] e^{i\omega k \Delta t} + \boldsymbol{N}[k] + \boldsymbol{C}[k] \,. \tag{4}$$

Clutter is the totality of undesired echoes and interfering signals, therefore it is impossible to generalize the properties of c_k. In the case of RWP, clutter includes in particular echoes from airborne objects such as aircraft and birds and returns from the ground. Interfering signals may be caused by other radio transmitters that operate in the RWP receiver band. In the remainder of the paper, we restrict ourselves to intermittent clutter signals.

While the properties of the intermittent clutter component have not been systematically investigated, it is instructive to take a look at a few examples. Such have been presented by various authors: Wilczak et al. (1995) described the distinct characteristic of bird contaminated I and Q data when seen in an A-scope display, but the shown time series taken with a 924 MHz RWP is only 0.5 s long, which is too short to see its essential characteristics. Jordan et al. (1997)

show an example of a 30 s long time series taken with a 915 MHz RWP during bird migration, which exhibits a variation in the envelope of the signal due to modulation of signal amplitude by the antenna beam pattern. Another example of intermittent clutter caused by airplanes and a simple theoretical model is given by Boisse et al. (1999). The most distinct feature here is the time-dependent amplitude of the signal. A 19 s time series of a 482 MHz RWP containing an airplane echo is discussed in Muschinski et al. (2005).

In the fall of 2005, time series data of the coherently integrated I/Q signal of the RWP at Bayreuth, Germany were saved in the wind low mode to get a unique dataset for the investigation of bird migration. For October 13, it was subjectively judged that the data showed a maximum of of bird echoes and we have therefore selected this day for further investigation. One particular dwell is shown in Figure 3. The time series has a length of about 35 s and its nonstationarity is striking.

When data containing intermittent clutter components are compared with uncontaminated clear air signals (and possibly ground clutter, as in the example shown in Muschinski et al. (2005)) it is very obvious, that the main difference is the transient character of the intermittent clutter signal component. Adopting the definition used by Friedlander and Porat (1989), we define a transient signal as a signal whose duration is short to the observation interval, in our case the dwell time. This reflects the clear nonstationarity of the underlying scattering process. It is not the sinusoidal signature that makes the difference, as a sufficiently strong clear air signal also exhibits a sinusoidal nature (see Fig. 1 and 2 in Muschinski et al. (2005)) - the most distinct property of intermittent clutter is the highly nonstationary character of the clutter component.

c. Consequences for signal processing

Signal processing is the art of extracting the maximum amount of information from a given measurement. This obviously means that the general properties of the signal determine the optimal mathematical processing methods. A stationary Gaussian stochastic process is without loss of information described by its time-independent second-order properties, that is the autocovariance function or, equivalently, the power spectrum. This assumption holds when equation (1) is valid, and the classical way to process RWP data is then based on a non-parametric estimation of the power spectrum using a discrete Fourier transform of the (usually coherently integrated) raw signal over the dwell-time. The power spectrum is usually called the Doppler spectrum. Its first three moments are estimated after the noise contribution to the spectrum has been subtracted, to describe the basic properties of the atmospheric signal (Woodman 1985). However, we have seen that the clutter contribution can be highly nonstationary. If the signal s_k contains nonstationary components, then the Doppler spectrum is no longer an adequate representation of the stochastic process because information regarding time dependency is already lost. So it cannot be expected, that a successful intermittent clutter filtering strategy can be developed based on the Doppler spectrum. Therefore it is tempting to try methods that were developed in the framework of nonstationary signal processing. A necessary condition is obviously a separation of $\boldsymbol{C}[k]$ from the stationary components $\boldsymbol{I}[k]e^{i\omega k\Delta t} + \boldsymbol{N}[k]$. To achieve this, we look for a representation of the signal in which we are able to discriminate between stationary and nonstationary signal components. This is the goal put forward in Wilczak et al. (1995): *Clearly, a superior technique would be one in which the bird signal and atmospheric signal could be differentiated from each other and processed independently.*

So far, we have considered either a pure time representation of the signal - namely the discrete time series or its complex Fourier transform as a pure frequency representation. Both

are not optimal for transient phenomena, although they are complete representations of the same information. Therefore we look for an intermediate representation that aims at the joint time-frequency structure of the signal, so it needs to depend both on time and frequency. This is the topic of the next section. If we are able to separate stationary and nonstationary signal components in such a representation, then we might be able to suppress the nonstationary clutter part while leaving the stationary signal component essentially intact.

3. Signal representation via Gabor frame expansions

a. The windowed Fourier transform and the time-frequency plane

Let us consider continuous signals first, although in practice we are always given a discretized signal. A quite natural way to analyze a continuous signal simultaneously in time and frequency is provided by the windowed Fourier transform (WFT), see Gabor (1946); Daubechies (1992); Kaiser (1994); Mallat (1999). It is essentially an extension of the well-known Fourier transform, where time localization is achieved by a pre-windowing of the signal with a normalized window function $h \in \mathbb{L}^2(\mathbf{R})$. For any given function $S \in \mathbb{L}^2(\mathbf{R})$, the WFT is defined as

$$V_h S(\tau, \omega) = \int_{-\infty}^{+\infty} S(t) h(t-\tau) e^{-i\omega t} dt \; . \tag{5}$$

The operator V_h maps isometrically between $\mathbb{L}^2(\mathbf{R})$ and $\mathbb{L}^2(\mathbf{R}^2)$, that is a one-dimensional function/signal is with no loss of energy transformed via the WFT into a two-dimensional function depending on both time τ and frequency ω. The (τ, ω)-plane is called the time-frequency (TF) plane or briefly the phase space. This representation was suggested by Gabor (1946) to illustrate that *both time and frequency are legitimate references for describing a signal*. The squared modulus of $V_h S$ is called the spectrogram, denoted by

$$P_h S(\tau, \omega) = |V_h S(\tau, \omega)|^2 \; , \tag{6}$$

and provides a measure for the energy of the signal in the time-frequency neighborhood of the point (τ, ω) and thus insight about the time-frequency structure of S around τ. However, due to Heisenberg's uncertainty relation, there is no arbitrary resolution in time and frequency simultaneously, i.e. a point-wise frequency description in time domain and a point-wise time description in frequency domain is impossible. Formally, one considers in the uncertainty context for some centralized signal h with $\|h\| = 1$, time and frequency variances

$$\sigma_t^2 = \int_{-\infty}^{+\infty} t^2 |h(t)|^2 dt \quad \sigma_\omega^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega^2 |\hat{h}(\omega)|^2 d\omega \tag{7}$$

for which the Heisenberg uncertainty relation yields

$$\sigma_t \cdot \sigma_\omega \geq \frac{1}{2} \; . \tag{8}$$

It can be shown, that equality in (8) is achieved when h is a translated, modulated or scaled version of the Gaussian function (equality means achieving optimal resolution in the time-frequency plane). Their time-frequency spread is visualized through a rectangle with widths σ_t and σ_ω in the TF-plane, this is called a Heisenberg box - see Figure 1. This optimality result

shall be used later on when elaborating a discrete version of (5). Since the WFT is an isometry, the inversion of V_h can be performed by its adjoint,

$$\langle S, S\rangle_{\mathbb{L}^2(\mathbf{R})} = \|S\|^2_{\mathbb{L}^2(\mathbf{R})} = \|V_h S\|^2_{\mathbb{L}^2(\mathbf{R}^2)} = \langle V_h S, V_h S\rangle_{\mathbb{L}^2(\mathbf{R}^2)} = \langle V_h^* V_h S, S\rangle_{\mathbb{L}^2(\mathbf{R})}$$

and therefore

$$S(t) = V_h^* V_h S(t) = \frac{1}{2\pi} \iint_{\mathbf{R}^2} V_h S(\tau, \omega) h(t-\tau) e^{i\omega t} d\omega d\tau \; . \tag{9}$$

Hence, in the continuous setting we still have signal analysis, transform (5), and signal synthesis, transform (9), in some straightforward way available and therefore time-frequency signal filtering can be performed in three simple steps (see e.g. Hlawatsch and Boudreaux-Bartels (1992)):

1. Analysis: Computation of the WFT using equation (5).
2. Modification of the WFT (e.g. time-dependent filtering).
3. Synthesis: Reconstruction of the modified signal using equation (9).

b. From windowed Fourier transform to Gabor frame expansions

For discrete signals, continuous transforms (5) and (9) are not suitable and would create very redundant representations of the signal. A first adjustment can be achieved when approximating (5) and (9) by discrete sums. Discretizing (9) means taking only values of the WFT at some discrete lattice in phase space. As it was pointed out, e.g. in Daubechies (1992), the sampling density in phase space plays a significant role for the existence and stability of a reconstruction formula, i.e. of a discrete version of (9).

Assume we are given some given discrete subset Λ (to be specified below) of the TF-plane, then a naive discrete version of the inversion formula (9) is given by

$$S(t) \stackrel{?}{\approx} \sum_{(m,k)\in\Lambda} V_h S(mT, k\Omega) h_{m,k}(t) \quad \text{with} \quad h_{m,k}(t) = h(t-mT) e^{ik\Omega t} \; , \tag{10}$$

where the parameter T controls the discrete linear shift mT along the time axis and Ω the sampling shift $k\Omega$ in the frequency domain. In order to verify whether (10) indeed exhibits a reconstruction formula, we first observe that for a family of elementary signals or so-called atoms $\{h_{m,k}\}_{(m,k)\in\Lambda}$ that is complete in $\mathbb{L}^2(\mathbf{R})$ any $S \in \mathbb{L}^2(\mathbf{R})$ can be represented by a linear expansion of the form

$$S(t) = \sum_{(m,k)\in\Lambda} a_{m,k} h_{m,k}(t) \; . \tag{11}$$

But only in very specific cases, e.g. when $\{h_{m,k}\}_{(m,k)\in\Lambda}$ forms a basis,

$$a_{m,k} = \langle S, h_{m,k}\rangle = V_h S(mT, k\Omega)$$

and therefore equality in (10) holds true,

$$S(t) = \sum_{(m,k)\in\Lambda} \langle S, h_{m,k}\rangle h_{m,k}(t) \; .$$

In general, this is not the case, i.e. we only have

$$S(t) \neq \sum_{(m,k)\in\Lambda} \langle S, h_{m,k}\rangle h_{m,k}(t) = F^*FS(t) \ ,$$

where the operator F^*F and its properties are briefly discussed in Appendix A. For a detailed analysis and discussion on this subject we refer the interested reader to, e.g., Daubechies (1992). To reconstruct S (i.e. to invert F^*F), special properties on Λ and on the analyzing atoms (the dual functions to h) are required. In what follows, we shall focus on the practically relevant biorthogonal case in which the construction of the analyzing atoms becomes simple and, moreover, numerically stable. To this end, suppose there is some auxiliary family $g_{m,k}(t) = g(t - mT)e^{ik\Omega t}$ (yet unknown) available that serves as a reservoir of analyzing atoms used to compute the Gabor coefficients $a_{m,k}$ via (5),

$$a_{m,k} = \langle S, g_{m,k}\rangle = V_g S(mT, k\Omega) = \int S(t')\bar{g}_{m,k}(t')dt' \ . \tag{12}$$

This approach was originally proposed by Bastiaans (1980). Inserting now (12) into (11) yields

$$S(t) = \sum_{(m,k)\in\Lambda} \int S(t')\bar{g}_{m,k}(t')dt' h_{m,k}(t) = \int S(t') \left(\sum_{(m,k)\in\Lambda} \bar{g}_{m,k}(t')h_{m,k}(t) \right) dt' \ .$$

Equality in the latter equation is assured as long as

$$\sum_{m,k} \bar{g}_{m,k}(t')h_{m,k}(t) = \delta(t - t') \ . \tag{13}$$

Condition (13) is called the *biorthogonality relation* and restricts the choice of g in dependence on the preassigned function h. The particular choice of the window function h (e.g. its variance σ_h), the time shift T and the frequency shift Ω directly controls the existence, uniqueness, convergence properties and the numerical stability of the Gabor expansion (11), which exists for arbitrary signals $S(t)$ only if $\Omega T \leq 2\pi$; this is a frame theoretical result, see (Daubechies 1990; Mallat 1999). The physical meaning of this inequality is nothing but the Nyquist sampling criterion and represents the sampling density. $\Omega T = 2\pi$ is called critical sampling. This was Gabor's original suggestion, as he was aiming at elementary signals *conveying exactly one datum or one 'quantum of information'*. In other words, there was no interest in any redundancy.

Again this can be visualized in the TF plane: The time-frequency concentration of the discrete elementary signals is represented by discrete rectangles with sides σ_t and σ_ω and area one-half, centered at the point $(mT, k\Omega)$. At critical sampling, the rectangles do not overlap, but fully cover the TF-plane. Gabor (1946) called this an *information diagram*. In his attempt to derive a theory of communication, each area represents one elementary quantum of information which Gabor proposed to call a *logon*. Although conceptually simple and appealing, the Gabor expansion at minimal sampling density in the TF-plane ($T\Omega = 2\pi$) has no nice mathematical structure. In particular, it does not form a basis with the basis functions localized in time and frequency. A relaxation of the equality $\Omega T = 2\pi$ is therefore required and generates a crucial degree of freedom in the Gabor expansion, this at the expense of oversampling and a possible non-uniqueness. For $\Omega T > 2\pi$ the stability of the expansion is lost.

c. Gabor frame expansions for discretely sampled signals

So far we have discretized (9) resulting in the Gabor frame expansion (11) for $S \in \mathbb{L}^2(\mathbf{R})$. But when it comes to real applications, only finitely many discretely sampled values of S are available; namely $\boldsymbol{S}[n] = S(m\Delta t)$. Therefore it becomes necessary to develop a fully discrete concept for evaluating the Gabor coefficients (12). Moreover, the discrete subset Λ in (11) is in general infinite and hence also not suitable for a numerical implementation: the sum needs to be appropriately truncated and, in addition, a discrete version of the dual function g needs to be derived.

In what follows, we illustrate how to proceed for discrete data $\boldsymbol{S}$. For greater detail we refer to the original paper by Wexler and Raz (1990) and Appendix B. Assume we are given some discrete and finite time (periodic) signal $\tilde{\boldsymbol{S}}$ with sampling points $n = 0, \ldots, N-1$, that is $\tilde{\boldsymbol{S}}[n] = \tilde{\boldsymbol{S}}[n+N]$. We therefore have to periodize the analysis and synthesis windows as well,

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

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

$$\boldsymbol{S}[n] = \sum_{m=0}^{M-1} \sum_{k=0}^{K-1} a_{m,k} \boldsymbol{h}_{m,k}[n] \ , \tag{14}$$

whereas the Gabor coefficients can be derived from

$$a_{m,k} = \sum_{n=0}^{N-1} \boldsymbol{S}[n] \bar{\boldsymbol{g}}_{m,k}[n] \ . \tag{15}$$

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

$$\boldsymbol{h}_{m,k}[n] = \boldsymbol{h}[n - m\Delta M] W_N^{nk\Delta K} \ , \ \boldsymbol{g}_{m,k}[n] = \boldsymbol{g}[n - m\Delta M] W_N^{nk\Delta K}.$$

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 $\Delta M \cdot M = \Delta K \cdot K = N$. From this it follows that $\Delta M \cdot \Delta K \leq N$ or $M \cdot K \geq N$. The reconstruction formula takes now the form

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

where we have assumed that the following discrete version of biorthogonality relation (13) for the sequences $\boldsymbol{h}$ and $\boldsymbol{g}$ is fulfilled,

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

It can be shown (for a proof see Appendix B) that the biorthogonality relation is satisfied if

$$\sum_{j=0}^{N-1} \boldsymbol{h}[j + qK] W_N^{-jpM} \bar{\boldsymbol{g}}[j] = \frac{N}{MK} \ \delta_{p,0} \delta_{q,0} \tag{16}$$

for $0 \leq p \leq \Delta M - 1$ and $0 \leq q \leq \Delta K - 1$. System (16) can be rewritten in matrix form: Let $\boldsymbol{v} = (N/(MK), 0, \ldots, 0)'$ be a vector of length $\Delta M \Delta K$ and $\boldsymbol{g} = (\boldsymbol{g}[0], \ldots, \boldsymbol{g}[N-1])$ the vector representing the discretely sampled dual frame, and let $\boldsymbol{A}$ be the matrix of size $\Delta M \Delta K \times N$ with entries $\boldsymbol{A}_{(p,q),j} = \bar{\boldsymbol{h}}(j + qK) W_N^{jpM}$, then the dual frame $\boldsymbol{g}$ is the solution of the linear system

$$\boldsymbol{A}\boldsymbol{g} = \boldsymbol{v} \; . \tag{17}$$

For critical sampling $\Delta M \Delta K = N$, $\boldsymbol{g}$ is unique if matrix $\boldsymbol{A}$ is nonsingular. For oversampling $\Delta M \Delta K < N$, system (17) is under-determined, and the solution is no longer unique and therefore there is a variety of possible dual frame functions $\boldsymbol{g}$.

d. On the choice of the analysis and synthesis atom and the TF-plane lattice

As we have seen, there is a high degree of freedom when constructing a frame representation of some signal $\boldsymbol{S}$. In particular,

i) the choice of the synthesis window $\boldsymbol{h}$

ii) the choice of the time-frequency sampling grid Λ, i.e. the choice of ΔM and ΔK that specifies the redundancy/non-redundancy and therewith the non-uniqueness/uniqueness of the Gabor frame expansion (14)

iii) the choice of $\boldsymbol{g}$ in case of $\Delta M \Delta K < N$, i.e. in the oversampling situation one may add further desirable constraints on the solution $\boldsymbol{g}$ of system (17), e.g. minimum energy-norm.

These three aspects shall now be discussed:

At *i)*: Any function $\boldsymbol{h}$ of finite energy (square integrable) is appropriate. However, as mentioned above, Heisenberg's uncertainty relation (8) requires for optimal time-frequency resolution a Gaussian function. Therefore, we chose

$$h(t) = \pi^{-1/4} \sigma_h^{-1/2} \, e^{-t^2/(2\sigma_h^2)}, \text{ such that } \|h\| = 1 \; , \tag{18}$$

where the scaling parameter σ_h (determined below) shall allow either a better resolution in time or frequency. As we shall in *iii)*, the time-frequency localization properties of synthesis function h carry over to analysis function g.

At *ii)*: The most important parameters that control the sampling density in the TF-plane are ΔK and ΔM. Together with the specification σ_h they fully determine (up to non-canonical choices of g) the discrete Gabor representation of some given function. In principle, the only requirement is $\Delta K \Delta M \leq N$. But because of Heisenberg's principle, too densely sampling (high redundancy) the TF-plane is not worth the trouble. More precisely, let Δt denote the sampling size of $\boldsymbol{S}$, i.e. $\boldsymbol{S}[n] = S(n\Delta t)$, with total period of $\boldsymbol{S}$ of $N\Delta t = T_d$ (often referred to as the dwell time). Then, in the classical FFT context, the frequencies are due to Nyquist's law automatically spaced with resolution $1/T$ within $[-1/2\Delta t, 1/2\Delta t]$. Through the flexibility of the Gabor representation, we may individually setup the time and frequency spacing. Let us consider to this end the Heisenberg box size, i.e. the time and frequency variances (7) which take for our particular h the form $\sigma_t^2 = \sigma_h^2/2$ and $\sigma_\omega^2 = (2\sigma_h^2)^{-1}$. If we restrict the spacing of the TF-plane to this box size (essentially smaller would produce an overlapping of the boxes),

i.e setting $\Delta\tau = \Delta M \Delta t = \sigma_t^2$ and $\Delta\omega = \Delta K / T = \sigma_\omega^2$, Heisenberg's uncertainty principle (8) and the solvability of (17) yields

$$N \geq \Delta M \Delta K \geq \frac{1}{4} N \; . \tag{19}$$

The right inequality in (19), represents an upper sampling bound that prevents an unnecessary Heisenberg box overlapping. If now an application requires a time resolution $\Delta\tau$ in the Gabor representation, we immediately obtain in the context of Heisenberg's uncertainty principle the optimal scaling factor for the synthesis (and therewith for the analysis) atom,

$$\sigma_h^2 = 2\Delta\tau \; ,$$

and a suggestion for the sampling density in time and frequency,

$$\Delta M = \lfloor \Delta\tau / \Delta t \rfloor \; , \qquad \Delta K \geq \frac{N}{4\Delta M} \; .$$

At *iii)*: In the oversampling situation ($\Delta M \Delta K < N$), the non-uniqueness can be used to add desirable constraints to the solution, for example minimum energy. This was discussed in greater detail in Qian and Chen (1993) and Qian et al. (1992): Since $\boldsymbol{A}$ is underdetermined we may rewrite (17) by applying the QR decomposition to its transposed form as

$$\left(\; \boldsymbol{R}^T \; \middle| \; 0 \; \right) \boldsymbol{Q}^T \boldsymbol{g} = \left(\; \boldsymbol{R}^T \; \middle| \; 0 \; \right) \begin{pmatrix} \boldsymbol{x} \\ \hline \boldsymbol{y} \end{pmatrix} = \boldsymbol{v}$$

and thus $\boldsymbol{x} = (\boldsymbol{R}^T)^{-1} \boldsymbol{v}$. Since $\boldsymbol{Q}\boldsymbol{Q}^T = \mathbb{1}$, it follows

$$\boldsymbol{g} = \boldsymbol{Q} \begin{pmatrix} \boldsymbol{x} \\ \hline \boldsymbol{y} \end{pmatrix} = \left(\; \boldsymbol{Q}_{\boldsymbol{x}} \; \middle| \; \boldsymbol{Q}_{\boldsymbol{y}} \; \right) \begin{pmatrix} \boldsymbol{x} \\ \hline \boldsymbol{y} \end{pmatrix} = \boldsymbol{Q}_{\boldsymbol{x}} \boldsymbol{x} + \boldsymbol{Q}_{\boldsymbol{y}} \boldsymbol{y} \; .$$

Since $\boldsymbol{h}$ is in the range($\boldsymbol{Q}_{\boldsymbol{x}}$) and because range($\boldsymbol{Q}_{\boldsymbol{x}}$)$\perp$range($\boldsymbol{Q}_{\boldsymbol{y}}$), one has $\boldsymbol{Q}_{\boldsymbol{y}}^T \boldsymbol{h} = \boldsymbol{0}$ (which is of interest below). Moreover, we observe that the analysis window $\boldsymbol{g}$ is the sum of two orthogonal vectors with $\|\boldsymbol{g}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$. Due to (17), $\boldsymbol{Q}_{\boldsymbol{x}} \boldsymbol{x} = \boldsymbol{Q}_{\boldsymbol{x}} (\boldsymbol{R}^T)^{-1} \boldsymbol{v}$, but $\boldsymbol{Q}_{\boldsymbol{y}} \boldsymbol{y}$ may depend on other constraints. When searching for the minimum norm solution, we simply set $\|\boldsymbol{Q}_{\boldsymbol{y}} \boldsymbol{y}\|^2 = \|\boldsymbol{y}\|^2 = 0$ and obtain

$$\boldsymbol{g} = \boldsymbol{Q}_{\boldsymbol{x}} \boldsymbol{x} = \boldsymbol{Q}_{\boldsymbol{x}} (\boldsymbol{R}^T)^{-1} \boldsymbol{v} = \boldsymbol{g}_{min}$$

which is nothing than $\boldsymbol{g}_{min} = \boldsymbol{A}^T (\boldsymbol{A}\boldsymbol{A}^T)^{-1} \boldsymbol{v}$. However, for a meaningful interpretation of the Gabor expansion, we would prefer an analysis window $\boldsymbol{g}$ which is locally concentrated in the TF-plane. The design of such a function $\boldsymbol{g}$ when the synthesis function $\boldsymbol{h}$ and ΔK and ΔM are given is a nontrivial problem and was also addressed in Qian and Chen (1993) and Qian et al. (1992). The problem can be formulated as follows: given an optimally concentrated function $\boldsymbol{h}$ (e.g. the preassigned synthesis function), find its biorthogonal function $\boldsymbol{g}$ whose shape best approximates time and frequency shifted versions of $\boldsymbol{h}$, i.e. minimize

$$E(\boldsymbol{g}, a, b) = \left\| \frac{\boldsymbol{g}}{\|\boldsymbol{g}\|} - \boldsymbol{h}_{a,b} \right\|^2 = 2 \left(1 - \frac{1}{\|\boldsymbol{g}\|} \Re \langle \boldsymbol{g}, \boldsymbol{h}_{a,b} \rangle \right) \text{ while } A\boldsymbol{g} = \boldsymbol{v}.$$

For fixed a and b, the optimal vector $\boldsymbol{y}$ in the representation for $\boldsymbol{g}$ ($\boldsymbol{x}$ is still fixed through the biorthogonality relation) is given by

$$\boldsymbol{y} = \frac{\|\boldsymbol{x}\|^2}{\Re\langle \boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{x}, \boldsymbol{h}_{a,b}\rangle}\boldsymbol{Q}_{\boldsymbol{y}}^T\boldsymbol{h}_{a,b} \ .$$

Choosing $\boldsymbol{h}_{a,b} = \boldsymbol{h}$ yields $\boldsymbol{Q}_{\boldsymbol{y}}^T\boldsymbol{h}_{a,b} = \boldsymbol{0}$ (see above) and thus $\boldsymbol{y} = \boldsymbol{0}$ and consequently, $\boldsymbol{g} = \boldsymbol{g}_{min}$, i.e. the shape of $\boldsymbol{g}_{min}$ best approximates the shape of $\boldsymbol{h}$. Therefore, the TF-plane localization properties of $\boldsymbol{h}$ carry over to $\boldsymbol{g}$ in this case. But note, that in principle any target function $\boldsymbol{h}_{a,b}$ is allowed and thus there is a large variety of possible analysis atoms $\boldsymbol{g}$.

e. Gabor representation of two examples

To illustrate the signal separation property of the discrete Gabor expansion for a single dwell, we consider two examples of simulated and measured RWP data. The method of Zrnić (1975) was used to simulate a signal in line with the classical signal model, which contains only noise and a stationary atmospheric component. In the frequency domain, the atmospheric signal peak is assumed to be a Gaussian centered at $f_d = \omega/2\pi = -10.9s^{-1}$ and with a spectral width of $w = 0.9s^{-1}$. The discrete spectrogram of this signal is shown in Figure 2. The atmospheric signal component is represented as a horizontal line (stationarity) centered at the prescribed Doppler frequency. Noise is spread over the complete TF plane.

Now lets take a look at real time series data containing an additional intermittent clutter component. This dataset is further discussed in section 5. The original I/Q data is shown in Figure 3. Clearly, this time series is not stationary but contains transient components due to migrating birds. Assuming that a time resolution of $\mathcal{O}(1s)$ is sufficient to resolve these transients, we select a time resolution of about 0.5 s for the Gabor expansion. This corresponds to a frequency resolution of about 2 Hz. An appropriate sampling density in the TF-plane is given with $\Delta M = 64$ and $\Delta K = 64$. Setting $M = 128$ and $K = 128$, we get an oversampling of factor 3.5; the optimal scaling is given by $\sigma_h^2 \approx 1$. In contrast to the simulated case, the spectrogram of the real signal shown in Figure 4 shows additional nonstationary signal components, which are a typical signature of contamination by intermittent clutter. Taking a look at the pure time representation of the signal it is difficult to identify the separate transients which show up as maxima of the envelope of the I/Q signal. However, Figure 4 shows the same signal, but this time its Gabor phase-space representation. This signal representation provides a far better picture of the signal transients, even if the spectrogram shows only the modulus of the Gabor coefficients, because the Gabor coefficients itself are complex. It becomes clear that the time series is contaminated by three transitory bird-events. Two of them overlap in time and can therefore not easily be distinguished in the time representation. All bird signals are much stronger in amplitude than the atmospheric signal of interest. The latter can be seen as a line of quasi-constant frequency centered at about a frequency of 3 Hz. By comparing Figure 2 with the real data shown in Figure 4, the goal of the filtering process becomes evident.

4. Filtering through the statistics of Gabor frame coefficients

a. Motivation for the statistical approach

With the tool of the Gabor representation at hand, the next step is to derive an appropriate filtering strategy for removal of the transient clutter signals. Our intention is to use the available

a-priori knowledge about the signal components (atmosphere, noise, clutter) to construct an objective decision process aiming at a proper signal component separation.

It is well-justified that both the atmospheric and the noise signal component are stationary Gaussian random processes. The atmospheric signal has a bounded spectral width much smaller than Nyquist interval, whereas noise is white and spread over the full TF plane. Not much is known in contrast about intermittent clutter, only the non-property that this signal component is nonstationary over typical dwell-times. We make use of this a-priori information to derive a filter that has a pass-characteristics for realizations of wide-sense stationary RP's and a stop-characteristics for all nonstationary processes. That is, signals looking like the simulated example shown in Fig. 2 should not be affected by the filtering process. The goal is thus to derive an objective procedure, that modifies the Gabor phase space representation of signals in such a way, that stationary Gaussian signal components are preserved.

One can imagine several strategies for implementing such a filter. For instance, this could be based on image processing techniques or a fuzzy-logic approach similar to the one used by Cornman et al. (1998). We follow a statistical approach, that has first been used by Merritt (1995) for the same problem and that is applied to the temporal sequence of Doppler spectral coefficients at fixed frequency bins. The goal is to construct a similar test, but this time in Gabor phase space. We therefore need to analyze the statistical properties of the Gabor coefficients with respect to the different signal components, in order to distinguish between clear air and clutter return. This immediately leads to the question of how the properties of Gaussian stationary processes are mapped to the Gabor coefficients $a_{m,k}$ or $|a_{m,k}|^2$. This problem is discussed in the next paragraph.

b. Mean and variance estimator for Gabor spectrogram coefficients

Since we aim to construct a statistical test (see the next section below) which is based on the expectation and the variance of the individual Gabor spectrogram coefficients $|a_{m,k}|^2$, we need to define adequate estimators for the expectation and the variance that are based on our observations (given through $\boldsymbol{S}$).

First, to simplify the notation, we introduce a_λ as a shorthand notation of $a_{m,k}$, i.e. in what follows we set $\lambda = (m,k)$. Then the Gabor spectrogram coefficients take the form

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

As mentioned in the previous section we may assume, the data sequence $\boldsymbol{S}$ satisfies for all $n = 0, \ldots, N-1$,

$$\mathsf{E}\boldsymbol{S}[n] = 0 \quad \text{and} \quad \mathsf{E}\boldsymbol{S}[n]\bar{\boldsymbol{S}}[n+l] = \sigma^2\boldsymbol{\rho}[l]\ .$$

With these two assumptions, the expectation and the covariance of the Gabor spectrogram coefficients are given by

$$\mathsf{E}|a_\lambda|^2 = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle\ ,$$
$$\mathsf{Cov}(|a_\lambda|^2, |a_\eta|^2) = \sigma^4|\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle|^2\ ,$$

which is shown in Appendix C (Lemma 3 and Lemma 4), where the '$*$'-operation stands here for the discrete convolution. The latter two formulas show the influence of the dependency of $\boldsymbol{S}$ and the redundancy of the Gabor frame expansion. In case, $\boldsymbol{S}$ would be i.i.d. (i.e. $\boldsymbol{\rho}[l] = \delta_{l,0}$), it follows

$$\mathsf{E}|a_\lambda|^2 = \sigma^2 \quad \text{and} \quad \mathsf{Cov}(|a_\lambda|^2, |a_\eta|^2) = \sigma^4|\langle \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle|^2.$$

If, moreover, $\{\boldsymbol{g}_\lambda\}_{\lambda\in\Lambda}$ forms an orthonormal system, the covariance matrix becomes diagonal; i.e. as long as we deal with a redundant frame, the Gabor spectrum is always correlated with a range of dependency described by the decay of the Gramian matrix of $\{\boldsymbol{g}_\lambda\}_{\lambda\in\Lambda}$ (up to the convolution with $\boldsymbol{\rho}$). The essential observation for our purpose is

$$\mathsf{Var}|a_\lambda|^2 = \mathsf{Cov}(|a_\lambda|^2, |a_\lambda|^2) = \sigma^4 |\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle|^2 = (\mathsf{E}|a_\lambda|^2)^2 \,.$$

Consequently,

$$\frac{(\mathsf{E}|a_\lambda|^2)^2}{\mathsf{Var}|a_\lambda|^2} = 1 \,, \tag{20}$$

which holds true for independent as well as dependent samples $\boldsymbol{S}[n]$ that follow a distribution which is determined by its moments. As property (20) constraints only the first two moments, it may hold true for a much richer class of distributions (in particular, it holds true for normal distributed random variables).

In order to construct a statistical test that verifies property (20), we have to find optimal estimators for $\mathsf{E}|a_\lambda|^2$ and $\mathsf{Var}|a_\lambda|^2$ that are based on a finite number of observations. To this end, we introduce a index subset $\Omega_\lambda \subset \Lambda$ containing λ and $L-1$ further different indices η, i.e. $|\Omega_\lambda| = L$. As an estimator for $\mathsf{E}|a_\lambda|^2 = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle$ which is based on L neighboring observation variables we define

$$\hat{E}(\Omega_\lambda) := \frac{1}{C_{\Omega_\lambda}} \sum_{\eta\in\Omega_\lambda} |a_\eta|^2 \,, \tag{21}$$

where the constant is given by

$$C_{\Omega_\lambda} = \sum_{\eta\in\Omega_\lambda} \frac{\langle \boldsymbol{\rho} * \boldsymbol{g}_\eta, \boldsymbol{g}_\eta\rangle}{\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle} > 1 \,.$$

For i.i.d. samples $\boldsymbol{S}[n]$, the correcting multiplier in estimator (21) reduces to $C_{\Omega_\lambda} = |\Omega_\lambda| = L$, and therefore (21) is then nothing than the well-known mean estimator,

$$\hat{E}(\Omega_\lambda) = \frac{1}{L} \sum_{\eta\in\Omega_\lambda} |a_\eta|^2 \,.$$

Assuming

$$\sum_{\eta',\eta\in\Omega_\lambda} |\langle \boldsymbol{\rho} * \boldsymbol{g}_{\eta'}, \boldsymbol{g}_\eta\rangle|^2 \le C_{\Omega_\lambda}^{2-\varepsilon} \,,$$

Lemmas 5 and 6 (see Appendix C) verify that (21) is a consistent estimator for $\mathsf{E}|a_\lambda|^2$, i.e.

$$\lim_{L\to\infty} \mathsf{E}|\hat{E}(\Omega_\lambda) - \mathsf{E}|a_\lambda|^2|^2 = \lim_{L\to\infty} (\mathsf{Var}(\hat{E}(\Omega_\lambda)) + (\mathsf{E}|a_\lambda|^2 - \mathsf{E}(\hat{E}(\Omega_\lambda))^2) = 0.$$

By the same reasoning, we define an estimator for variance,

$$\hat{V}(\Omega_\lambda) := C \sum_{\eta\in\Omega_\lambda} (|a_\eta|^2 - \hat{E}(\Omega_\lambda))^2 \,, \tag{22}$$

where the constant is defined by

$$C^{-1} := 2 \sum_{\eta\in\Omega_\lambda} \frac{c_\eta^2}{c_\lambda^2} + (L - 2C_{\Omega_\lambda}) \left(1 + \frac{1}{(\sum_\eta c_\eta)^2} \sum_{\xi,\alpha\in\Omega_\lambda} c_{\xi,\alpha}^2\right) . \tag{23}$$

Similar as before, it is shown (see Lemma 7 in Appendix C) that estimator (22) is unbiased (and certainly consistent but the proof is omitted). Switching to the i.i.d. case yields

$$C^{-1} = 2L + (L - 2L)\left(1 + \frac{1}{L^2} \sum_{\xi,\alpha \in \Omega_\lambda} c_{\xi,\alpha}^2\right) = L - \frac{1}{L} \sum_{\xi,\alpha \in \Omega_\lambda} |\langle \boldsymbol{g}_\xi, \boldsymbol{g}_\alpha \rangle|^2$$

and therefore (22) simplifies to

$$\hat{V}(\Omega_\lambda) = \frac{L}{L^2 - \sum_{\xi,\alpha \in \Omega_\lambda} |\langle \boldsymbol{g}_\xi, \boldsymbol{g}_\alpha \rangle|^2} \sum_{\eta \in \Omega_\lambda} (|a_\eta|^2 - \hat{E}(\Omega_\lambda))^2 \,,$$

which can be easily seen with the help of formula (C2). If, moreover, $\{\boldsymbol{g}_\lambda\}$ forms a basis, we end up with the classical variance estimator

$$\hat{V}(\Omega_\lambda) = \frac{1}{L-1} \sum_{\eta \in \Omega_\lambda} (|a_\eta|^2 - \hat{E}(\Omega_\lambda))^2 \,.$$

c. A statistical test performing signal identification

After having established estimators $\hat{E}(\Omega_\lambda)$ and $\hat{V}(\Omega_\lambda)$, we aim now to apply these quantities to the construction of a test that identifies Gabor coefficients that can be associated with clear air returns. Typically, an atmospheric return is stationary and assumed to follow a Gaussian distribution, i.e. a test on the first two moments of the signal will give us some indication if this is true.

The basic idea goes back to Merritt (1995), who statistically tested a sequence of single (non-averaged) Doppler spectra in order decide whether a particular Fourier power coefficient was caused by a Gaussian or non–Gaussian signal. For this, he used the classical test of Hildebrand and Sekhon (1974) in a modified way. Following this approach, we consider the squared modulus of the Gabor phase space coefficients, $|a_{m,k}|^2$. Because we are interested in stationary signal components, we consider the sequence $|a_{m,k}|^2$ for fixed frequency bins, i.e. we just pick individual rows and let only the time index change. For a fixed frequency index k, we define $\Omega_k := \{(m,k) : \ m = 0, \ldots, M-1\}$. If this sequence is stationary Gaussian, then it should hold that

$$\vartheta(|a_{m,k}|^2) := \frac{(\hat{E}(\Omega_k))^2}{\hat{V}(\Omega_k)} \geq 1 \tag{24}$$

This, however, is typically not observed if the sequence $|a_{m,k}|^2$ is affected by non-stationary intermittent clutter. It is a well known fact that intermittent clutter signals are almost always stronger than the (clear air) atmospheric return. In this case, we will get $\vartheta \leq 1$. In a next step, we therefore sort the sequence according to power. That is, we derive the order statistic of $\{|a_{m,k}|^2\}_{(m,k)\in\Omega_k}$ which is $\{|a_{\varphi(m),k}|^2\}_{(m,k)\in\Omega_k}$. φ stands for the permutation map. This ordered sequence has the property, that $|a_{\varphi(m),k}|^2 \leq |a_{\varphi(m)+1,k}|^2$ for all $(m,k) \in \Lambda_k$. For $l = 0, \ldots, M-1$, we define subsets $\Omega_k(l) = \{(\varphi(m), k) : \ \varphi(m) = l, \ldots, M-1\}$. The largest coefficients are discarded, which hopefully sorts out all coefficients affected by clutter. Using the quantities $\hat{E}(\Omega_k(l))$ and $\hat{V}(\Omega_k(l))$ of the subset, the statistics ϑ is again computed until

$$\vartheta(|a_{\varphi(m)=l,k}|^2) := \frac{(\hat{E}(\Omega_k(l)))^2}{\hat{V}(\Omega_k(l))} \geq 1$$

holds. The largest coefficient for which the test is satisfied is taken as a threshold for a frequency-dependent identification of the clutter component.

d. Signal separation through Gabor coefficient thresholding

The local threshold is computed as the average value of the remaining Gabor coefficients. All coefficients $|a_{m,k}|^2$ greater than this threshold are then regarded as clutter. One problem exists, if the subset $\Omega_k(l)$ becomes too small in this iterative process. Then the statistical estimate will become unstable and the estimation of a local threshold is no longer meaningful. This should not happen if the dwell time is sufficiently long, but it is not always known how long the dwell time must indeed be for various types of intermittent clutter. Further investigations are needed to clarify this question. However, it might nevertheless be attempted to clean data sets regardless of the dwell time used. In these cases, it can happen that the nonstationary components have sometimes a duration on the order of the dwell time. In this case, it can be useful to replace the local threshold with a global threshold if the criterion (24) is not reached after a certain number of iterations. Such a global threshold can be derived from stable estimates of local thresholds at other frequencies k. Such a global threshold should be constructed in such a way, that it reflects the noise level in the Gabor representation. For instance, such a global threshold could be estimated by averaging over a certain number of the smallest local thresholds.

Leaving this problem aside, we can formulate the filtering procedure as follows: A coefficient $|a_{\varphi(m),k}|^2$ with $\varphi(m) = l$ for which the (24) holds is associated with clear air return. Based on the test, we introduce a clutter index set as

$$\Omega_k^c := \{(m,k) : \ \vartheta(|a_{\varphi(m),k}|^2) \geq 1 \ , \ \varphi(m) = 1, \ldots, M-1\}$$

The average value of the remaining coefficients in the subset is taken as a local threshold t_k:

$$t_k = \frac{1}{|\Omega_k \setminus \Omega_k^c|} \sum_{(m,k)\in\Omega_k\setminus\Omega_k^c} |a_{m,k}| \ .$$

The *main result* of this paper - the nonlinear filtering - is now formulated in the following:

Let $\boldsymbol{S}$ be the given RWP signal. Based on our model assumptions, the filtered component is given by

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

Finally, we discuss a practical aspect of the filtering method. The evaluation of the clutter index set Ω_k^c requires the computation of the modified variance estimator. This can be done in a numerically thrifty way, assuming we have all the inner products $\langle \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta \rangle$ and $\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta \rangle$, respectively, at hand. Note that the efficiency can be additionally improved due to the fact that not all of them are required. This becomes clear because $\hat{V}(\Omega_k(l))$ is derived for fixed frequency indices k and thus the correction term relies only on inner products of the form

$$\begin{aligned} \langle \boldsymbol{g}_{m,k}, \boldsymbol{g}_{m',k} \rangle &= \langle \boldsymbol{g}(\cdot - m\Delta M) W_N^{\cdot k\Delta K}, \boldsymbol{g}(\cdot - m'\Delta M) W_N^{\cdot k\Delta K} \rangle \\ &= \langle \boldsymbol{g}(\cdot - m\Delta M), \boldsymbol{g}(\cdot - m'\Delta M) \rangle \ . \end{aligned}$$

Nevertheless, the computational overhead involved in computing the modified variance estimator is still greater than in the case of the classical variance estimator. Our experience has shown that the variance estimates obtained with the two methods do not differ much. It may therefore be appropriate to use the classical variance estimator if a saving of processing power is necessary for a real time implementation of the algorithm. However, a detailed consideration of this simplification is left for a future study.

5. A real example: Comparison with classical processing

a. Data set

Now let us illustrate the performance of the proposed filtering algorithm by applying it to real RWP data obtained with the 482 MHz wind profiler at Bayreuth, Germany on October 13, 2005. This radar is one of three operational systems that the Deutscher Wetterdienst uses in its operational aerological network. The technical characteristics are summarized in Table 1. More details and an overview of the standard signal processing steps of the radar system are given in Lehmann et al. (2003). For wind measurements, the system is running in a four-beam Doppler beam swinging configuration using two different pulse widths of 1700 ns (low mode) and 3300 ns (high mode). The averaging time is 26 minutes, the remaining 4 minutes are used for RASS measurements of the virtual temperature. For the investigation of bird migration we consider only low mode data. The sampling parameters for the low mode data are given in Table 2. Of interest below are the resolution of the time series $\Delta t = 0.007708$, the number of data samples $N = 4608$ and the total length or dwell time $T_d = N\Delta t = 35.518464s$.

During the bird migration season in October of 2005, full time series data of the coherently integrated demodulated receiver voltage signal were saved in the wind low mode. Both wind and spectral data were manually reviewed to identify a day with significant bird migration. It is well known, that a human expert can easily detect bird migration events by searching for typical patterns in the wind measurements (northeasterly directions in fall, discontinuities at sunrise and sunset), which are additionally accompanied by irregular and wide, sometimes multiple peaks in the Doppler spectra. In contrast to most clutter-free situations, those peaks exhibit a poor time and range gate continuity. Time-height plots of the estimated moments (power, radial velocity and spectral width) are helpful to get a quick overview of potentially interesting cases, and a closer look into the time series data then typically confirms the conjecture of bird migration. Particulary significant bird migration was noted on October 13 and we therefore selected that day as a test case for the new bird mitigation algorithm. A significant fraction of wind data were contaminated by bird returns, the effect is best seen in Figure 10. The winds shown have been computed without any intermittent clutter removal algorithm. The consensus method is normally not able to suppress the effect of the bird echoes because of their frequent occurrence. The operationally used intermittent clutter removal algorithm (ICRA), a particular implementation of the statistical averaging method proposed by Merritt (1995), could only alleviate the problem, see Figure 11. Also, the operational quality control (Weber-Wuertz continuity check, not shown) was only able to flag a small percentage of the contaminated data, because the erroneous wind data exhibited the typical intrinsic consistency.

b. Processing details and results

A software was developed to allow reading and writing of the profiler time series data using the proprietary binary format of the profiler vendor. This made it easy to process the data using the Gabor filter and to save them again in the original file format. The reprocessed data could therefore be seamlessly integrated in the off-line version of the operational wind profiler software, to compare the performance of the different algorithms.

As an example, we consider again the measurement taken in the south beam of the profiler at range gate 9 (1.6 km height agl) with a start time of the dwell at 00:09:45 UTC. This was already discussed in section 3. As described in section 4, local (constant frequency) thresholds were estimated to separate the clutter part of the signal from the stationary components atmosphere and noise. During processing of the complete day it was revealed that the dwell time of the data of about 35 seconds was apparently rather short to guarantee that every observed intermittent clutter signal exhibits a clear transient behavior. Sometimes the duration of the clutter signal component was on the order of the dwell time instead. If this is the case, then the estimation of the local threshold may become unstable and signal separation can partly fail with the result that clutter energy leaks through the filter. One way to remedy this problem is to replace local thresholds with a global threshold as described above. For the example data, this was done if more than 30 percent of the Gabor coefficients of a particular frequency were classified as clutter. The global threshold was computed as the median over 15 percent of the smallest local thresholds. Note that the global threshold is an estimate for the noise level. Another way to handle this situation would be to either flag this range gate as suspect or to set the whole spectrum to that of random white noise, so any signal would be suppressed. Further research is needed to learn more about typical intermittent clutter characteristics and to optimize both the data sampling and the performance of the filter. The method described in this paper should be a useful tool for such investigations.

Application of the filtering strategy yields a filtered Gabor phase-space representation, which is shown in Figure 5. Here, the moduli of the coefficients $a_{m,k}$ representing the transient (bird) contributions have been replaced by an estimation of the stationary signal component at that frequency (either noise or atmospheric signal). The reconstructed I/Q series which is obtained through back-transformation into the time domain is presented in Figure 6. The nonstationary signal components have been suppressed and also the amplitude has been significantly reduced. It is easy to measure the reduction of total power of the data by computing the difference in variance between the unfiltered and the filtered data, to get an information about how much clutter energy was removed by the filter.

Gabor filtering was performed for the complete dataset and the resulting bird-cleaned time series data were used for reprocessing of the whole day. This was compared with two additional processing methods: Method 1 used no intermittent clutter filtering algorithm whereas method 2 used the routinely employed Intermittent Clutter Algorithm (ICRA), an implementation of the Statistical Averaging Method (SAM) originally proposed by Merritt (1995). The results for all range gates for the dwell taken at 00:09:45 UTC (stacked Doppler spectra) are shown in Figures 7 (no filtering), 8 (ICRA filtering) and 9 (Gabor filtering). Without filtering, the lowest 17 range gates show spurious peaks and also large spectral widths due to the transient bird echoes. Note especially the discontinuity in height of the location of the estimated signal peak (derived Doppler velocity). With ICRA processing, the effect of the birds has been drastically reduced, however, there are still range gates which show spurious peaks. This indicates that ICRA was unable to reduce the clutter energy completely. Figure 9 shows the processing results

of the newly suggested filtering algorithm. The spurious remnants of the bird clutter are almost completely gone, although range gates 15 and 16 (2.49 and 1.64 km height agl) show apparently some bird clutter energy leaking through. This is also reflected in the somewhat larger spectral width at these heights. However, the spectral peak is now continuous across all heights and the spectral width estimates are mostly unaffected by the clutter.

Finally, the horizontal wind vector data derived through the three different processing methods using the same data are shown in Figures 10 (no clutter filtering), 11 (ICRA processing) and 12 (Gabor filtering), respectively . The color coding is due to the wind speed (magnitude of the horizontal wind vector). Obviously, the clutter contamination has been drastically reduced by the new algorithm.

6. Summary and conclusions

We have dealt with wind profiler signals that were obtained during bird migration and shown, how the signals can be decomposed into a time-frequency representation. Apparently, a Gabor frame representation achieves optimal time-frequency resolution and thus provides good signal-clutter separation. Previous attempts for intermittent clutter filtering have made use of the wavelet transform (WT) and its discrete versions (Jordan et al. 1997; Boisse et al. 1999; Lehmann and Teschke 2001), so it is interesting to briefly discuss the difference between the wavelet and the Gabor approach and to point out why we favor the Gabor approach. The WT is another way of analyzing nonstationary signals. The difference between both approaches lies in the tiling of the TF-plane by the elementary signals (or time-frequency atoms). In the Gabor (WFT) approach, the tiling is uniform with fixed resolution. This is in contrast to the wavelet approach, where the tiling is generally variable. For example, a wavelet orthonormal basis decomposes the frequency axis in dyadic intervals whose sizes grow exponentially. In other words, the frequency resolution gets worse the better the time resolution becomes. This is wanted if the signals under investigation have high-frequency components of short duration embedded within low-frequency components of slow temporal variation. For the RWP signals however, we found no evidence for such a behavior. The intermittent clutter components occur at nearly all frequencies within the typical Nyquist range with no obvious difference in temporal characteristics. In particular, they can occur close to zero frequency where the temporal resolution of the WT is the worst. Especially in this case, the WT seems not ideal to resolve the transient nature of the intermittent clutter. Examples of clutter signals in both representations shown by Justen and Lehmann (2003) illustrate this quite clearly. So the argument which is often used against the WFT, namely the constant time-frequency resolution of this representation turns out to be an advantage in our application. Additionally, the Gabor expansion using a Gaussian window achieves the best possible time-frequency resolution by reaching the lower limit of the Heisenberg uncertainty constraint.

To identify the clutter contribution of the signal, we make use of the a-priori information that the atmospheric signal component of interest can be adequately modelled as a stationary, proper complex Gaussian random process. Using this assumption, a test statistic is constructed to serve as a criterion for the discrimination between stationary and non-stationary signal components. This follows the approach first suggested by Merritt (1995). However, in case of the redundant Gabor transform it turns out, that the variance estimator has to be modified to guarantee its unbiasedness and consistency. Proofs for the necessary modifications are given in detail.

Finally, the algorithm has been applied to a real dataset that was obtained with a 482 MHz

wind profiler during bird migration season. It could be demonstrated that the performance of the new algorithm was superior to the performance of the operationally used intermittent clutter reduction algorithm, without obvious negative side effects. Application of the algorithm has shown, that sampling settings of the wind profiler apparently play an important role in the clutter mitigation capabilities of the algorithm. This is not unexpected, since both the sampling period and the dwell time determine the resolution of the Doppler spectrum and obviously also the resolution of any time-frequency representation. Furthermore, longer dwell times may ease the identification of transient clutter signals and the stable estimation of the thresholds for noise and the stationary atmospheric component. This is especially important for cases where atmospheric and clutter signal overlap in frequency.

Future work is suggested for a better quantitative characterization of intermittent clutter signals during dense bird migration. This should allow to optimize both sampling and processing settings for operational wind profiler systems. A long-term evaluation of the new algorithm would be useful to determine its limits and to estimate the performance improvements of the new methods in comparison with previously used algorithms. This would be facilitated by an online-implementation of the method and a means to compare the profiler wind measurement with an independent data source, e.g. radiosonde measurements.

APPENDIX A

Frame theory

We briefly review some basic facts on frames using the abstract notation of functional analysis, but the reader is advised to consult the rich literature for details (Heil and Walnut 1989; Daubechies 1990; Carmona et al. 1998; Mallat 1999; Christensen 2001).

A frame is a family of functions, that allow to characterize a signal S from its inner products $\{\langle S, h_\lambda\rangle\}_{\lambda\in\Lambda}$. It generalizes the notion of a basis in Hilbert space. We can always safely assume that S(t) is element of the Hilbert space $\mathbb{L}^2$, because the received signal has finite energy. Let H be some Hilbert space, the pair of parenthesis $\langle\cdot,\cdot\rangle$ denote the associated inner product and $\|\cdot\|_H^2 = \langle\cdot,\cdot\rangle$ the associated norm. A frame $\{h_\lambda\}$ in H is a system of functions for which there exist constants $0 < A \leq B < \infty$ such that for all $s \in H$

$$A\|S\|_H^2 \leq \sum_{\lambda\in\Lambda} |\langle S, h_\lambda\rangle|^2 \leq B\|S\|_H^2 \; . \tag{A1}$$

The map, $F : H \to \ell_2$, defined via $F : f \mapsto \{\langle f, h_\lambda\rangle\}$ is usually referred to as the frame operator (analysis operator). So the signal is characterized by inner products with the frame. To answer the question of how f can again be synthesized from the inner products $\{\langle f, h_\lambda\rangle\}$, we consider the adjoint frame operator given by $F^*c = \sum_{\lambda\in\Lambda} c_\lambda h_\lambda$. This allows us to write

$$F^*Ff = \sum_{\lambda\in\Lambda} \langle f, h_\lambda\rangle h_\lambda \; . \tag{A2}$$

If F^*F equals the identity $\mathbb{1}$, F^* performs a perfect reconstruction. This is the case when $\{h_\lambda\}$ forms an orthonormal system. However, in general one has to apply $(F^*F)^{-1}$ to (A2). This is possible since the inverse exists and is bounded because of (A1),

$$A \cdot \mathbb{1} \leq F^*F \leq B \cdot \mathbb{1} \quad \text{and thus} \quad B^{-1} \cdot \mathbb{1} \leq (F^*F)^{-1} \leq A^{-1} \cdot \mathbb{1} \; .$$

Since $(F^*F)^{-1}$ is self-adjoint and denoting $(F^*F)^{-1}h_\lambda = g_\lambda$, one consequently has

$$\sum_{\lambda\in\Lambda} \langle S, g_\lambda\rangle h_\lambda = F^*F(F^*F)^{-1}S = S = (F^*F)^{-1}F^*FS = \sum_{\lambda\in\Lambda} \langle S, h_\lambda\rangle g_\lambda \; . \tag{A3}$$

In frame lore, g_λ is referred to as the canonical dual frame with respect to h_λ.

In general, $(F^*F)^{-1}$ cannot be explicitly computed but must approximated by an iterative approach. However, the situation can be essentially relaxed when assuming that the frames $\{h_\lambda\}$ and $\{g_\lambda\}$ form not a primal-dual but a bi-orthogonal frame pair, i.e. $\langle h_\lambda, g_\eta\rangle = \delta_{\lambda,\eta}$. If $\tilde{F}$ denotes the frame operator with respect to g_λ, then $\tilde{F} = F(F^*F)^{-1}$ and one may write $(\langle h_\lambda, g_\eta\rangle)_{\lambda,\eta\in\Lambda} = \tilde{F}F^*$ which is diagonal and therefore $\tilde{F}$ is an analysis and F^* a synthesis operator yielding perfect reconstruction (and vice versa, i.e. exchanging the roles of $\tilde{F}$ and F^*). If now the bi-orthogonality relation yields a way to derive g_λ, the inverse of F^*F needs not to be computed.

APPENDIX B

Biorthogonal discrete Gabor frame expansion

The following lemma can be retraced to its original form in Wexler and Raz (1990), it gives an explicit proof of the biorthogonality relation.

Lemma 1 *Assume the relation*

$$\sum_{j=0}^{N-1} \bar{g}(j)h(j+qK)W^{-jpM} = N/(MK)\ \delta_{p,0}\delta_{q,0} \tag{B1}$$

is fulfilled for $0 \le p \le \Delta M - 1$ *and* $0 \le q \le \Delta K - 1$. *Then the biorthogonality relation*

$$\sum_{m=0}^{M-1}\sum_{k=0}^{K-1} \bar{g}_{m,k}(l)h_{m,k}(j) = \delta_{l,j}$$

holds true.

Proof. This assertion can be directly shown. Let

$$f(l,j) := \sum_{m=0}^{M-1}\sum_{k=0}^{K-1} \bar{g}_{m,k}(l)h_{m,k}(j),$$

then

$$f(l,j) = \sum_{m=0}^{M-1} \bar{g}(l-m\Delta M)h(j-m\Delta M)\sum_{k=0}^{K-1} W^{k(j-l)\Delta K}\ .$$

We still have,

$$\sum_{k=0}^{K-1} W^{k(j-l)\Delta K} = \sum_{k=0}^{K-1} e^{i2\pi k(j-l)/K} = \begin{cases} K & , \text{ if } (j-l)/K \in \mathbb{Z} \\ 0 & , \text{ else} \end{cases}\ .$$

Since $(j-l)/K \in \mathbb{Z}$ means there exists some $q \in \mathbb{Z}$ such that $q = (j-l)/K$ or $j-l-qK = 0$, we may consequently write (by the Poisson Summation Formula and the made assumption)

$$\begin{aligned} f(l,j) &= \sum_{m=0}^{M-1} \bar{g}(l-m\Delta M)h(j-m\Delta M)K\sum_q \delta_{j-l-qK,0} \\ &= K\sum_q \delta_{j-l-qK,0}\sum_{m=0}^{M-1} \bar{g}(l-m\Delta M)h(l+qK-m\Delta M) \\ &= K\sum_q \delta_{j-l-qK,0}\Delta M^{-1}\sum_{p=0}^{\Delta M-1}\left(\sum_{j'=0}^{N-1}\bar{g}(j')h(j'+qK)W^{-j'pM}\right)W^{lpM} \\ &= K\sum_q \delta_{j-l-qK,0}M/N\sum_{p=0}^{\Delta M-1} N/(MK)\delta_{p,0}\delta_{q,0}W^{lpM} \\ &= \delta_{j,l}\ . \end{aligned}$$

□

APPENDIX C

Statistical properties of the Gabor coefficients

Lemma 2 *Let S be given and assume $\mathsf{E}\boldsymbol{S}[n] = 0$ for all $n = 0, \ldots, N-1$ and that a_λ is as defined in* (15)*. Then $\mathsf{E}a_\lambda = 0$.*

Proof. By definition, $a_\lambda = \sum_{n=0}^{N-1} \boldsymbol{S}[n]\boldsymbol{g}_\lambda[n]$. Therefore, $\mathsf{E}a_\lambda = \sum_{n=0}^{N-1} \mathsf{E}\boldsymbol{S}[n]\boldsymbol{g}_\lambda[n] = 0$. □

Lemma 3 *Let S be given and assume $\mathsf{E}\boldsymbol{S}[n] = 0$ for all $n = 0, \ldots, N-1$ and that a_λ is as defined in* (15)*. Moreover, assume a range of dependency of neighboring samples of S which is characterized by the auto-covariance function $\boldsymbol{\rho}$ of S, i.e. $\mathsf{E}(\boldsymbol{S}[n]\bar{\boldsymbol{S}}[n+l]) = \sigma^2\boldsymbol{\rho}(l)$. Then*

$$\mathsf{Cov}(a_\lambda, a_\eta) = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle,$$

where '$*$' denotes the discrete convolution.

The latter lemma states that the Gabor coefficients a_λ turn into dependent random variables (even when $\boldsymbol{\rho}$ is a delta sequence, i.e. for independent samples of S). The range of dependency is determined by sampling density in the time–frequency space and the range of dependency of S. In case S is a sequence of i.i.d. random variables, the dependency of a_λ is fully characterized by the reproducing kernel $\langle \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle$.

Proof. By Lemma 2, $\mathsf{Cov}(a_\lambda, a_\eta) = \mathsf{E}(a_\lambda \bar{a}_\eta)$. Therefore,

$$\begin{aligned}\mathsf{Cov}(a_\lambda, a_\eta) &= \mathsf{E}\Big(\sum_{n=0}^{N-1} \boldsymbol{S}[n]\boldsymbol{g}_\lambda[n], \sum_{l=0}^{N-1} \bar{\boldsymbol{S}}[l]\bar{\boldsymbol{g}}[l]\Big) = \sum_{n=0}^{N-1}\sum_{l=0}^{N-1} \mathsf{E}(\boldsymbol{S}[n]\bar{\boldsymbol{S}}[l])\boldsymbol{g}_\lambda[n]\bar{\boldsymbol{g}}_\eta[l] \\ &= \sigma^2 \sum_{n=0}^{N-1}\sum_{l=0}^{N-1} \boldsymbol{\rho}(l-n)\boldsymbol{g}_\lambda[n]\bar{\boldsymbol{g}}_\eta[l] = \sigma^2 \sum_{l=0}^{N-1} (\boldsymbol{\rho} * \boldsymbol{g}_\lambda)[l]\bar{\boldsymbol{g}}_\eta[l] = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle.\end{aligned}$$

□

A special case of Lemma 3 is $\mathsf{E}|a_\lambda|^2 = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle$.

Lemma 4 *Make the same assumptions as in Lemma 3. Then*

$$\mathsf{Cov}(|a_\lambda|^2, |a_\eta|^2) = \sigma^4 |\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle|^2 .$$

Proof. First, note that for proper Gaussian complex random variables $\boldsymbol{S}[k]$ with $\mathsf{E}\boldsymbol{S}[k] = 0$ and $\mathsf{Cov}(\boldsymbol{S}[k]\bar{S}[l]) = \mathsf{E}(\boldsymbol{S}[k]\bar{S}[l]) = \sigma^2\boldsymbol{\rho}(l-k)$ we have (Reed 1962)

$$\begin{aligned}\mathsf{E}(\boldsymbol{S}[k]\bar{\boldsymbol{S}}[l]\boldsymbol{S}[n]\bar{\boldsymbol{S}}[m]) &= \mathsf{E}(\boldsymbol{S}[k]\bar{\boldsymbol{S}}[l])\mathsf{E}(\boldsymbol{S}[n]\bar{\boldsymbol{S}}[m]) + \mathsf{E}(\boldsymbol{S}[k]\bar{\boldsymbol{S}}[m])\mathsf{E}(\bar{\boldsymbol{S}}[l]\boldsymbol{S}[n]) \\ &= \sigma^4(\boldsymbol{\rho}(l-k)\boldsymbol{\rho}(m-n) + \boldsymbol{\rho}(m-k)\bar{\boldsymbol{\rho}}(n-l)).\end{aligned}$$

With the help of Lemma 3 (special case),

$$\mathsf{Cov}(|a_\lambda|^2, |a_\eta|^2) = \mathsf{E}(|a_\lambda|^2|a_\eta|^2) - \sigma^4\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle\langle \boldsymbol{\rho} * \boldsymbol{g}_\eta, \boldsymbol{g}_\eta\rangle$$

and thus it remains to derive $\mathsf{E}(|a_\lambda|^2|a_\eta|^2)$. Using the moment theorem of Reed (1962), we have,

$$\begin{aligned}
\mathsf{E}(|a_\lambda|^2|a_\eta|^2) &= \left(\sum_{k=0}^{N-1} \boldsymbol{S}[k]\boldsymbol{g}_\lambda[k] \sum_{l=0}^{N-1} \bar{\boldsymbol{S}}[l]\bar{\boldsymbol{g}}_\lambda[l] \sum_{n=0}^{N-1} \boldsymbol{S}[n]\boldsymbol{g}_\eta[n] \sum_{m=0}^{N-1} \bar{\boldsymbol{S}}[m]\bar{\boldsymbol{g}}_\eta[m]\right) \\
&= \sum_{k,l,n,m=0}^{N-1} \mathsf{E}(\boldsymbol{S}[k]\bar{\boldsymbol{S}}[l]\boldsymbol{S}[n]\bar{\boldsymbol{S}}[m])\, \boldsymbol{g}_\lambda[k]\bar{\boldsymbol{g}}_\lambda[l]\boldsymbol{g}_\eta[n]\bar{\boldsymbol{g}}_\eta[m] \\
&= \sigma^4 \sum_{k,l,n,m=0}^{N-1} (\boldsymbol{\rho}(l-k)\boldsymbol{\rho}(m-n) + \boldsymbol{\rho}(m-k)\bar{\boldsymbol{\rho}}(n-l))\, \boldsymbol{g}_\lambda[k]\bar{\boldsymbol{g}}_\lambda[l]\boldsymbol{g}_\eta[n]\bar{\boldsymbol{g}}_\eta[m] \\
&= \sigma^4\Bigg(\sum_{l,m=0}^{N-1} \bar{\boldsymbol{g}}_\lambda[l]\bar{\boldsymbol{g}}_\eta[m]\Big\{\sum_{k=0}^{N-1} \boldsymbol{\rho}(l-k)\boldsymbol{g}_\lambda[k]\Big\}\Big\{\sum_{n=0}^{N-1} \boldsymbol{\rho}(m-n)\boldsymbol{g}_\eta[n]\Big\} + \\
&\qquad \sum_{m,n=0}^{N-1} \boldsymbol{g}_\eta[n]\bar{\boldsymbol{g}}_\eta[m]\Big\{\sum_{k=0}^{N-1} \boldsymbol{\rho}(m-k)\boldsymbol{g}_\lambda[k]\Big\}\Big\{\sum_{l=0}^{N-1} \bar{\boldsymbol{\rho}}(n-l)\bar{\boldsymbol{g}}_\lambda[l]\Big\}\Bigg) \\
&= \sigma^4\Bigg(\sum_{l,m=0}^{N-1} \bar{\boldsymbol{g}}_\lambda[l]\bar{\boldsymbol{g}}_\eta[m](\boldsymbol{\rho} * \boldsymbol{g}_\lambda)[l](\boldsymbol{\rho} * \boldsymbol{g}_\eta)[m] + \\
&\qquad \sum_{m,n=0}^{N-1} \boldsymbol{g}_\eta[n]\bar{\boldsymbol{g}}_\eta[m](\boldsymbol{\rho} * \boldsymbol{g}_\lambda)[m](\overline{\boldsymbol{\rho} * \boldsymbol{g}_\lambda})[n]\Bigg) \\
&= \sigma^4\Big(\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle\langle \boldsymbol{\rho} * \boldsymbol{g}_\eta, \boldsymbol{g}_\eta\rangle + |\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle|^2\Big)\,,
\end{aligned}$$

and consequently,

$$\mathsf{Cov}(|a_\lambda|^2, |a_\eta|^2) = \sigma^4 |\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle|^2\,.$$

□

After having verified the basic properties of the Gabor power coefficients, we prove that estimator (21) is consistent and that estimator (22) is unbiased (the proof of consistency is omitted because its proof requires the computation of $8th$-mixed moment).

Lemma 5 *The estimator $\hat{E}(\Omega_\lambda)$ unbiased, i.e. it holds $\mathsf{E}\hat{E}(\Omega_\lambda) = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle$.*

Proof. This follows by the definition of C_{Ω_λ} and Lemma 3,

$$\mathsf{E}\hat{E}(\Omega_\lambda) = \frac{1}{C_{\Omega_\lambda}} \sum_{\eta\in\Omega_\lambda} \mathsf{E}|a_\eta|^2 = \frac{1}{C_{\Omega_\lambda}} \sum_{\eta\in\Omega_\lambda} \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\eta, \boldsymbol{g}_\eta\rangle = \sigma^2\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle\,.$$

□

Lemma 6 *Assume, that for the given dual frame $\{\boldsymbol{g}_\lambda\}$ exists some $\varepsilon > 0$ such that the condition $\sum_{\eta',\eta\in\Omega_\lambda} |\langle \boldsymbol{\rho} * \boldsymbol{g}_{\eta'}, \boldsymbol{g}_\eta\rangle|^2 \le C_{\Omega_\lambda}^{2-\varepsilon}$ is fulfilled. Then the estimator $\hat{E}(\Omega_\lambda)$ satisfies*

$$\mathsf{Var}(\hat{E}(\Lambda)) \le \sigma^4 C_{\Omega_\lambda}^{-\varepsilon}$$

and is therefore consistent.

Proof. Similar as in the proof of Lemma 4 we directly obtain

$$\begin{aligned}
\mathsf{Var}(\hat{E}(\Omega_\lambda)) &= \mathsf{E}(\hat{E}(\Omega_\lambda))^2 - \sigma^4|\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle|^2 \\
&= \frac{1}{C_{\Omega_\lambda}^2} \sum_{\eta',\eta\in\Omega_\lambda} \mathsf{E}(|a_{\eta'}|^2|a_\eta|^2) - \sigma^4|\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle|^2 \\
&= \frac{\sigma^4}{C_{\Omega_\lambda}^2}\Bigg(\sum_{\eta',\eta\in\Omega_\lambda} \Big\{ \langle \boldsymbol{\rho} * \boldsymbol{g}_{\eta'}, \boldsymbol{g}_{\eta'}\rangle\langle \boldsymbol{\rho} * \boldsymbol{g}_\eta, \boldsymbol{g}_\eta\rangle + |\langle \boldsymbol{\rho} * \boldsymbol{g}_{\eta'}, \boldsymbol{g}_\eta\rangle|^2 \Big\} - \\
&\qquad\qquad C_{\Omega_\lambda}^2|\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle|^2 \Bigg) \\
&= \frac{\sigma^4}{C_{\Omega_\lambda}^2} \sum_{\eta',\eta\in\Omega_\lambda} |\langle \boldsymbol{\rho} * \boldsymbol{g}_{\eta'}, \boldsymbol{g}_\eta\rangle|^2 \le \sigma^4 C_{\Omega_\lambda}^{-\varepsilon} .
\end{aligned}$$

□

Lemma 7 *The estimator $\hat{V}(\Lambda)$ is unbiased, i.e. it holds $\mathsf{E}\hat{V}(\Omega_\lambda) = \sigma^4|\langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle|^2$.*

Proof. With similar arguments as in the proof of Lemma 4 and with the following shorthand notations

$$c_\lambda := \langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\lambda\rangle \;\text{ and }\; c_{\lambda,\eta} := \langle \boldsymbol{\rho} * \boldsymbol{g}_\lambda, \boldsymbol{g}_\eta\rangle$$

we have the following expressions

$$\begin{aligned}
\mathsf{E}|a_\eta|^4 &= \sigma^4(c_\eta^2 + c_\eta^2) = 2\sigma^4 c_\eta^2 , \\
\mathsf{E}(|a_\eta|^2\hat{E}(\Omega_\lambda)) &= \frac{\sigma^4}{C_{\Omega_\lambda}} \sum_{\xi\in\Omega_\lambda} (c_\eta c_\xi + c_{\eta,\xi}^2) , \\
\mathsf{E}(\hat{E}(\Omega_\lambda))^2 &= \frac{\sigma^4}{C_{\Omega_\lambda}^2} \sum_{\xi,\alpha\in\Omega_\lambda} (c_\alpha c_\xi + c_{\xi,\alpha}^2) .
\end{aligned}$$

Therefore with $L = |\Omega_\lambda|$ and the definition of C in (23),

$$\begin{aligned}
\mathsf{E}\hat{V}(\Omega_\lambda) &= C \sum_{\eta\in\Omega_\lambda} \mathsf{E}(|a_\eta|^2 - \hat{E}(\Omega_\lambda))^2 \\
&= C \sum_{\eta\in\Omega_\lambda} \Big\{ \mathsf{E}|a_\eta|^4 - 2\,\mathsf{E}(|a_\eta|^2\hat{E}(\Omega_\lambda)) + \mathsf{E}(\hat{E}(\Omega_\lambda))^2 \Big\} \\
&= \sigma^4 C \sum_{\eta\in\Omega_\lambda} \Big\{ 2c_\eta^2 - \frac{2}{C_{\Omega_\lambda}} \sum_{\xi\in\Omega_\lambda} (c_\eta c_\xi + c_{\eta,\xi}^2) + \frac{1}{C_{\Omega_\lambda}^2} \sum_{\xi,\alpha\in\Omega_\lambda} (c_\alpha c_\xi + c_{\xi,\alpha}^2) \Big\} \\
&= \sigma^4 C \sum_{\eta\in\Omega_\lambda} \Big\{ 2c_\eta(c_\eta - c_\lambda) - \frac{2}{C_{\Omega_\lambda}} \sum_{\xi\in\Omega_\lambda} c_{\eta,\xi}^2 + c_\lambda^2 + \frac{1}{C_{\Omega_\lambda}^2} \sum_{\xi,\alpha\in\Omega_\lambda} c_{\xi,\alpha}^2 \Big\} \\
&= \sigma^4 C \Big\{ 2 \sum_{\eta\in\Omega_\lambda} c_\eta(c_\eta - c_\lambda) + Lc_\lambda^2 + \frac{L - 2C_{\Omega_\lambda}}{C_{\Omega_\lambda}^2} \sum_{\xi,\alpha\in\Omega_\lambda} c_{\xi,\alpha}^2 \Big\} \qquad \text{(C1)} \\
&= \sigma^4 c_\lambda^2 C \Big\{ 2 \sum_{\eta\in\Omega_\lambda} \frac{c_\eta^2}{c_\lambda^2} + (L - 2C_{\Omega_\lambda}) \Big(1 + \frac{1}{(\sum_\eta c_\eta)^2} \sum_{\xi,\alpha\in\Omega_\lambda} c_{\xi,\alpha}^2 \Big) \Big\} = \sigma^4 c_\lambda^2 .
\end{aligned}$$

□

References

Anderson, C. J. and R. W. Arritt, 2001: Representation of summertime low-level jets in the Central United States by the NCEP-NCAR reanalysis. *J. Climate*, **14**, 234–247.

Barth, M., R. Chadwick, and D. van de Kamp, 1994: Data processing algorithms used by NOAA's wind profiler demonstration network. *Ann. Geophys.*, **12**, 518–528.

Bastiaans, M., 1980: Gabor's expansion of a signal into Gaussian elementary signals. *Proc. IEEE*, **68 (4)**, 538–539.

Benjamin, S. G., D. Devenyi, S. S. Weygandt, K. J. Brundage, J. M. Brown, G. A. Grell, D. Kim, B. E. Schwartz, T. G. Smirnova, T. L. Smith, and G. S. Manikin, 2004a: An hourly assimilation-forecast cycle: The RUC. *Mon.Weather Rev.*, **132**, 495–518.

Benjamin, S. G., B. E. Schwartz, E. J. Szoke, and S. E. Koch, 2004b: The value of wind profiler data in U.S. weather forecasting. *Bull. Amer. Meteor. Soc.*, 1871–1886.

Boisse, J.-C., V. Klaus, and J.-P. Aubagnac, 1999: A wavelet transform technique for removing airplane echos from ST radar signals. *J. Atmos. Oceanic Technol.*, **16**, 334–346.

Bouttier, F., 2001: The use of profiler data at ECMWF. *Meteor. Z.*, **10 (6)**, 497–510.

Bruderer, B., 1997a: The study of bird migration by radar part 1: The technical basis. *Naturwissenschaften*, **84**, 1–8.

———, 1997b: The study of bird migration by radar part 2: Major achievements. *Naturwissenschaften*, **84**, 45–54.

Carmona, R., W.-L. Hwang, and B. Torresani, 1998: *Practical Time-Frequency Analysis*. Academic Press.

Cheong, B., M. Hoffman, R. Palmer, S. J. Frasier, and F. Lopez-Dekker, 2006: Phased-array design for biological clutter rejection: Simulation and experimental validation. *J. Atmos. Oceanic Technol.*, **23**, 585–598.

Christensen, O., 2001: Frames, Riesz bases and discrete Gabor/wavelet expansions. *Bull. Amer. Math. Soc.*, **38 (3)**, 273–291.

Cornman, L. B., R. K. Goodrich, C. S. Morse, and W. L. Ecklund, 1998: A fuzzy logic method for improved moment estimation from Doppler spectra. *J. Atmos. Oceanic Technol.*, **15 (6)**, 1287–1305.

Daniel, C. J., R. W. Arritt, and C. J. Anderson, 1999: Accuracy of 404-MHz radar profilers for detection of low-level jets over the Central United States. *J. Appl. Meteor.*, **38**, 1391–1396.

Daubechies, I., 1990: The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, **36 (5)**, 961–1005.

———, 1992: *Ten Lectures on Wavelets*. SIAM, Philadelphia.

Doviak, R. J. and D. S. Zrnic, 1993: *Doppler Radar and Weather Observations*. Academic Press.

Ecklund, W., D. Carter, B. Balsley, P. Currier, J. Green, B. Weber, and K. Gage, 1990: Field tests of a lower tropospheric wind profiler. *Radio Sci.*, **25 (5)**, 899–906.

Engelbart, D., U. Görsdorf, and W. Ruhe, 1998: Effects and observation of migrating birds on a Boundary-Layer Windprofiler in Eastern Germany. *Meteorol. Z.*, **NF 7**, 280–287.

Frehlich, R. and M. Yadlowsky, 1994: Performance of mean-frequency estimators for doppler radar and lidar. *J. Atmos. Oceanic Technol.*, **11**, 1217–1230.

Friedlander, B. and B. Porat, 1989: Detection of transient signals by the Gabor representation. *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-37 (2)**, 169–180.

Gabor, D., 1946: Theory of communication. *J. IEE (London)*, **93 (3)**, 429–457.

Gage, K. S., 1990: Radar observations of the free atmosphere: structure and dynamics. *Radar in Meteorology*, Atlas, D., Ed., Amer. Meteor. Soc., Boston, 534–565.

Gauthreaux, S. A. and C. G. Belser, 1998: Displays of bird movements on the WSR-88D: Patterns and quantification. *Wea. Forecasting*, **13**, 453–464.

Gauthreaux, S. A., D. S. Mizrahi, and C. G. Belser, 1998: Bird migration and bias of WSR-88D wind estimates. *Wea. Forecasting*, **13**, 465–481.

Griesser, T. and H. Richner, 1998: Multiple peak processing algorithm for identification of atmospheric signals in Doppler radar wind profiler spectra. *Meteorol. Z.*, **N.F.7**, 292–302.

Heil, C. E. and D. F. Walnut, 1989: Contionous and discrete wavelet transforms. *SIAM Review*, **31 (4)**, 628–666.

Hildebrand, P. H. and R. Sekhon, 1974: Objective determination of the noise level in Doppler spectra. *J. Appl. Meteorol.*, **13**, 808–811.

Hlawatsch, F. and G. Boudreaux-Bartels, 1992: Linear and quadratic time-frequency signal representations. *IEEE Sig.Proc. Magazine*, 21–67.

Hogg, D., M. Decker, F. Guiraud, K. Earnshaw, D. Merritt, K. Moran, W. Sweezy, R. Strauch, E. Westwater, and C. Little, 1983: An automatic profiler of the temperature, wind and humidity in the troposphere. *J. Climate Appl. Meteorol.*, **22**, 807–831.

Ishihara, M., Y. Kato, T. Abo, K. Kobayashi, and Y. Izumikawa, 2006: Characteristics and performance of the operational wind profiler network of the Japan Meteorological Agency. *J. Meteor. Soc. Japan*, **84 (6)**, 1085–1096.

Jordan, J. R., R. J. Lataitis, and D. A. Carter, 1997: Removing ground and intermittent clutter contamination from wind profiler signals using wavelet transforms. *J. Atmos. Oceanic Technol.*, **14**, 1280–1297.

Justen, L. and V. Lehmann, 2003: Radar wind profiler signal processing using redundant windowed fourier and wavelet transforms. *6th International Conference on Tropospheric Profiling - Extended Abstracts*.

Kaiser, G., 1994: *A Friendly Guide to Wavelets*. Birkhäuser, Basel.

Keeler, R. J. and R. E. Passarelli, 1990: Signal processing for atmospheric radars. *Radar in Meteorology*, Atlas, D., Ed., American Meteorological Society, Boston, chap. 20a, 199–229.

Lehmann, V., J. Dibbern, U. Görsdorf, J. W. Neuschaefer, and H. Steinhagen, 2003: The new operational UHF wind profiler of the Deutscher Wetterdienst. *6th International Conference on Tropospheric Profiling - Extended Abstracts*.

Lehmann, V. and G. Teschke, 2001: Wavelet based methods for improved wind profiler signal processing. *Ann. Geophysicae*, **19**, 825–836.

Liu, S., Q. Xu, and P. Zhang, 2005: Identifying Doppler velocity contamination caused by migrating birds. Part II: Bayes identification and probability tests. *J. Atmos. Oceanic Technol.*, **22**, 1114–1121.

Locatelli, J. D., M. T. Stoelinga, P. V. Hobbs, and J. Johnson, 1998: Structure and evolution of an undular bore on the high plains and its effects on migrating birds. *Bull. Amer. Meteor. Soc.*, **79 (6)**, 1043–1060.

Lundquist, J. K., 2003: Intermittent and elliptical inertial oscillations in the Atmospheric Boundary Layer. *J. Atmos. Sci.*, **60**, 2661–2673.

Mallat, S., 1999: *A Wavelet Tour of Signal Processing*. Academic Press.

Mastrantonio, G., J. Naithani, P. Anderson, S. Argentini, and I. Petenko, 1999: Quantitative analysis and interpretation of dot echoes observed with a Doppler sodar. *J. Atmos. Oceanic Technol.*, **16**, 1928–1940.

Merritt, D. A., 1995: A statistical averaging methode for wind profiler Doppler spectra. *J. Atmos. Oceanic Technol.*, **12 (5)**, 985–995.

Monna, W. A. and R. B. Chadwick, 1998: Remote-sensing of upper-air winds for weather forecasting: Wind-profiler radar. *Bull. WMO*, **47 (2)**, 124–132.

Morse, C. S., R. K. Goodrich, and L. B. Cornman, 2002: The NIMA method for improved moment estimation from Doppler spectra. *J. Atmos. Oceanic Technol.*, **19**, 274–295.

Muschinski, A., 2004: Local and global statistics of clear-air Doppler radar signals. *Radio Sci.*, **39 (RS1008)**, doi:10.1029/2003RS002 908.

Muschinski, A., V. Lehmann, L. Justen, and G. Teschke, 2005: Advanced radar wind profiling. *Meteorol. Z.*, **14 (5)**, 609–626.

Muschinski, A., P. P. Sullivan, D. B. Wuertz, R. J. Hill, S. A. Cohn, D. H. Lenschow, and R. J. Doviak, 1999: First synthesis of wind-profiler signals on the basis of large-eddy simulation data. *Radio Science*, **34 (6)**, 1437–1459.

Neeser, F. D. and J. L. Massey, 1993: Proper complex random processes with applications to information theory. *IEEE Trans. Inf. Theory*, **39 (4)**, 1293–1302.

Parker, M. D. and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale convective systems. *Mon. Weather Rev.*, **128**, 3413–3436.

Pekour, M. and R. Coulter, 1999: A technique for removing the effect of migrating birds in 915-MHz wind profiler data. *J. Atmos. Oceanic. Technol.*, **16 (December)**, 1941–1948.

Qian, S. and D. Chen, 1993: Discrete Gabor transform. *IEEE Trans. Signal Processing*, **41 (7)**, 2429–2438.

Qian, S., K. Chen, and S. Li, 1992: Optimal biorthogonal functions for finite discrete-time Gabor expansion. *Signal Processing*, **27**, 177–185.

Ralph, F., L. Armi, J.M.Bane, C.Dorman, W.D.Neff, P.J.Neiman, W.Nuss, and P.O.G.Persson, 1998: Observations and analysis of the 10-11 June 1994 coastally trapped disturbance. *Mon. Weather Rev.*, **126**, 2435–2465.

Reed, I. S., 1962: On a moment theorem for complex Gaussian processes. *IRE Trans. Inf. Theory*, **IT-8**, 194–195.

Röttger, J. and M. Larsen, 1990: UHF/VHF radar techniques for atmospheric research and wind profiler applications. *Radar in Meteorology*, American Meteorological Society, Boston, chap. 21a, 235–281.

Semple, A., 2005: Forecast error investigation 12th October 2003: Assimilation of contaminated wind profiler data into the global model (Forecasting Research Technical Report No. 465). Tech. rep., UK Met Office.

Song, J., K. Liao, R. L. Coulter, and B. M. Lesht, 2005: Climatology of the Low-Level Jet at the Southern Great Plains Atmospheric Boundary Layer experiments site. *J. Appl. Meteor.*, **44**, 1593–1606.

St-James, J. S. and S. Laroche, 2005: Assimilation of wind profiler data in the Canadian Meteorological Centre's analysis system. *J. Atmos. Oceanic Technol.*, **22**, 1181–1194.

Stensrud, D. J., 1996: Importance of low-level jets to climate: A review. *J. Climate*, **9**, 1698–1711.

Van Zandt, T., 2000: A brief history of the development of wind-profiling or MST radars. *Ann. Geophysicae*, **18**, 740–749.

Vaughn, C. R., 1985: Birds and insects as radar targets: A review. *Proc. IEEE*, **73 (2)**, 205–227.

Weber, B., D. Welsh, D. Merritt, D. Wuertz, D. Wolfe, and T. Wilfong, 2004: A new paradigm for Doppler radar wind profiler signal processing. Tech. Rep. OAR ETL-306, NOAA - Environmental Technology Laboratory.

Weber, H. C., 2005: Classification of contaminated data from wind profiler measurements by Neural Networks. Ph.D. thesis, Swiss Federal Institute of Technology Zürich.

Wexler, J. and S. Raz, 1990: Discrete Gabor expansions. *Signal Processing*, **21**, 207–220.

Wilczak, J., R. Strauch, F. Ralph, B. Weber, D. Merritt, J. Jordan, D. Wolfe, L. Lewis, D. Wuertz, J. Gaynor, S. McLaughlin, R. Rogers, A. Riddle, and T. Dye, 1995: Contamination of wind profiler data by migrating birds: Characteristics of corrupted data and potential solutions. *J. Atmos. Oceanic Technol.*, **12 (3)**, 449–467.

Woodman, R. F., 1985: Spectral moment estimation in MST radars. *Radio Sci.*, **20**, 1185–1195.

———, 1991: A general statistical instrument theory of atmospheric and ionospheric radars. *J. Geophys. Res.*, **96 (A5)**, 7911–7928.

Zhang, P., S. Liu, and Q. Xu, 2005: Identifying Doppler velocity contamination caused by migrating birds. Part I: Feature extraction and quantification. *J. Atmos. Oceanic Technol.*, **22**, 1105–1113.

Zrnić, D. S., 1975: Simulation of weatherlike Doppler spectra and signals. *J. Appl. Meteor.*, **14**, 619–620.

———, 1979: Estimation of spectral moments for weather echoes. *IEEE Trans. Geosci. Electron.*, **GE-17 (4)**, 113–128.

List of Figures

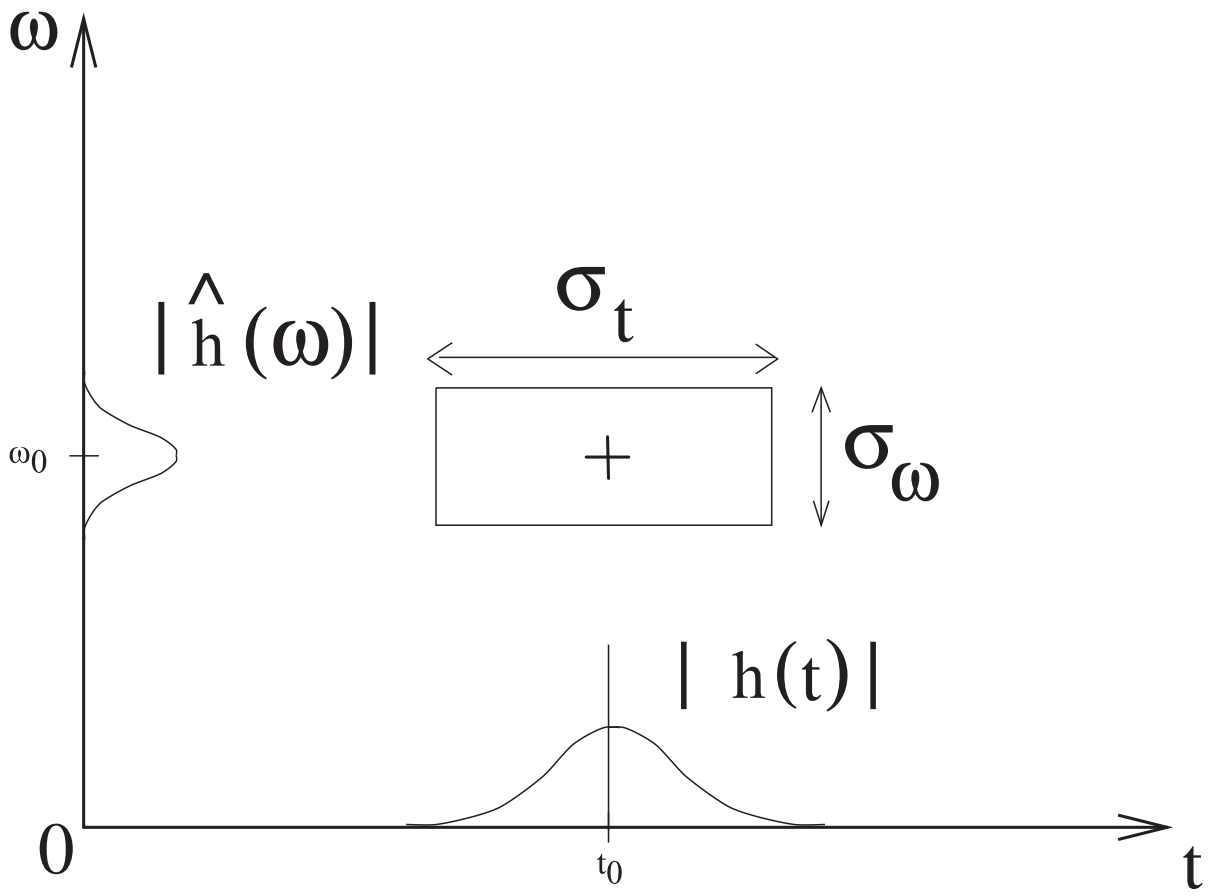

FIG. 1. Schematic representation of the Time-Frequency plane and the Heisenberg-box (resolution) of the window function $h_{\tau,\omega}(t)$, centered at time $\tau = t_0$ and frequency $\omega = \omega_0$.

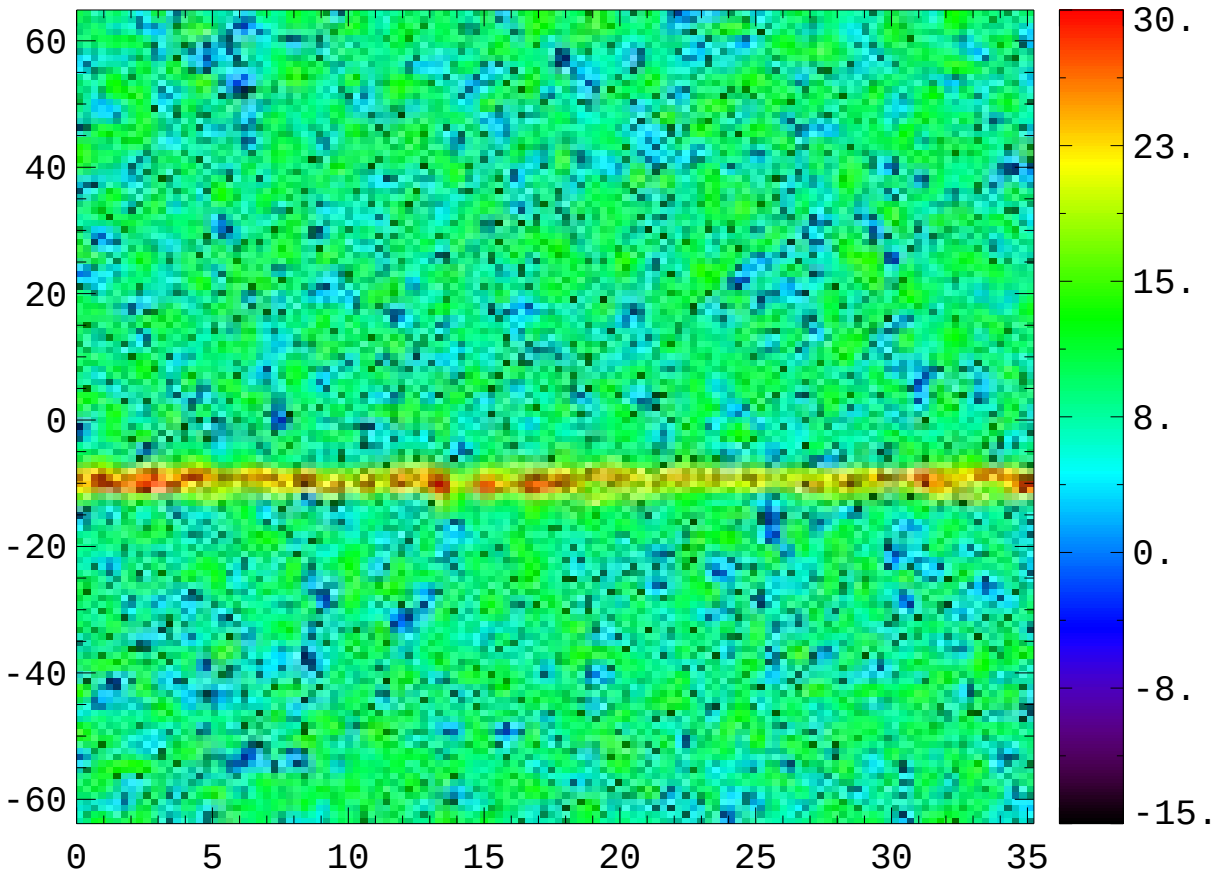

FIG. 2. Gabor phase space representation of a simulated RWP signal containing only noise and an atmospheric component. The x-axis shows time (in seconds) and the y-axis frequency (in Hz). Color contours (logarithmic scaling in dB) denote the power of the Gabor coefficients.

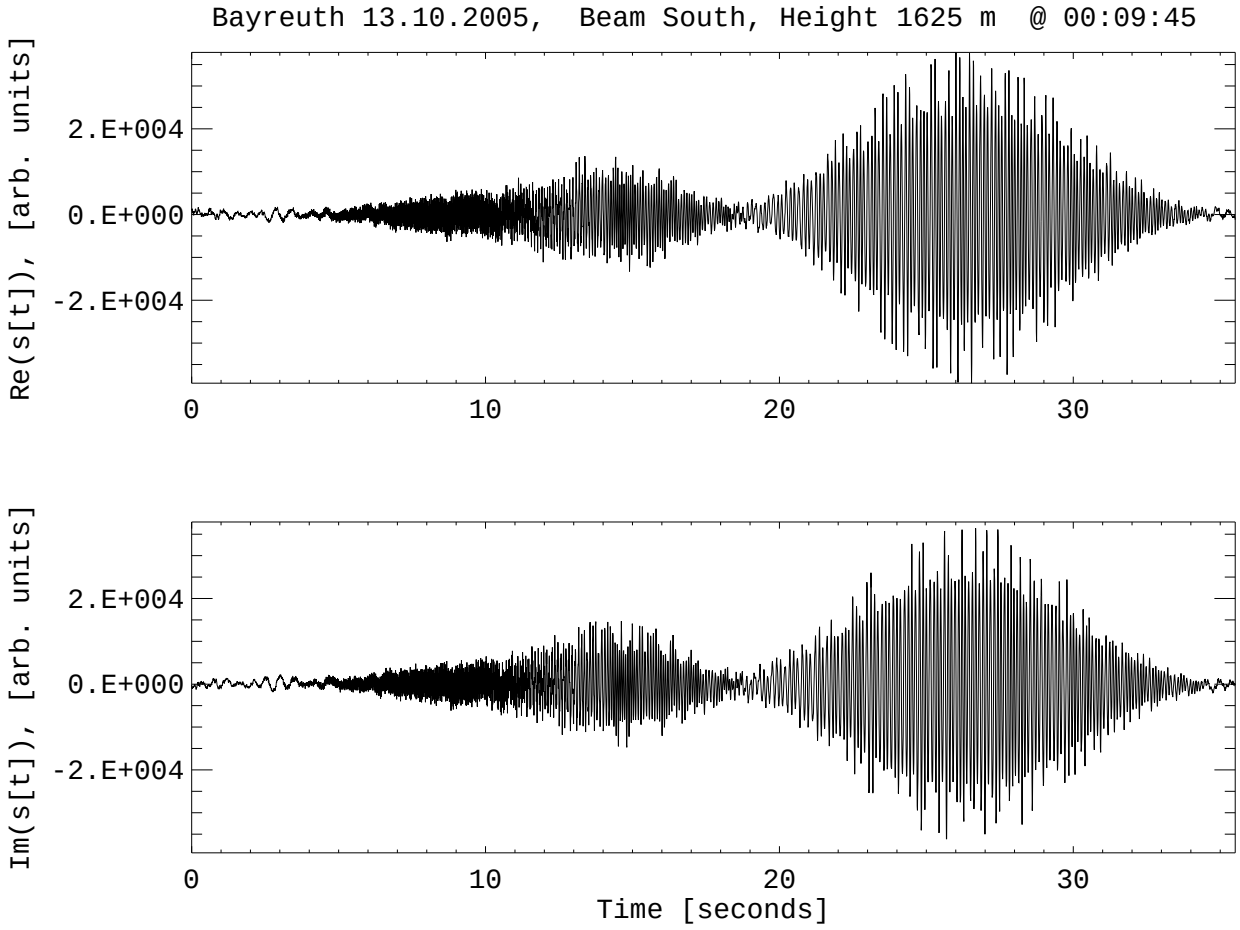

FIG. 3. Time series of the in-phase (upper plot) and quadrature (lower plot) component of the baseband signal measured at 00:09:45 UTC on October 13, 2005 (south beam, range gate 9) with the 482-MHz RWP at Bayreuth, Germany. The complex time series contains 4608 samples. Each sample is the coherent sum of 94 echoes from subsequent pulses.

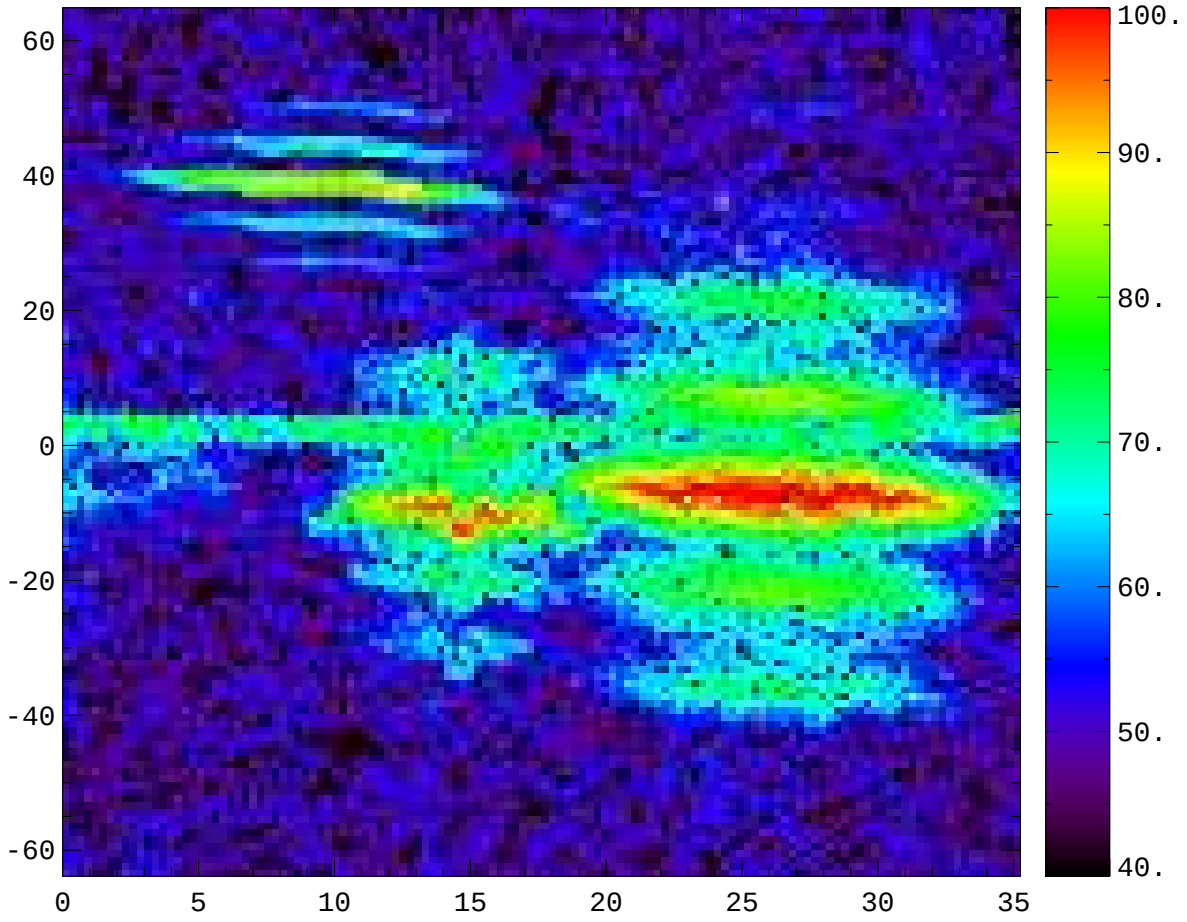

FIG. 4. Same representation as in Figure 2, but for the data shown in Figure3. The three transient signal components are clearly separated from the stationary atmospheric signal component.

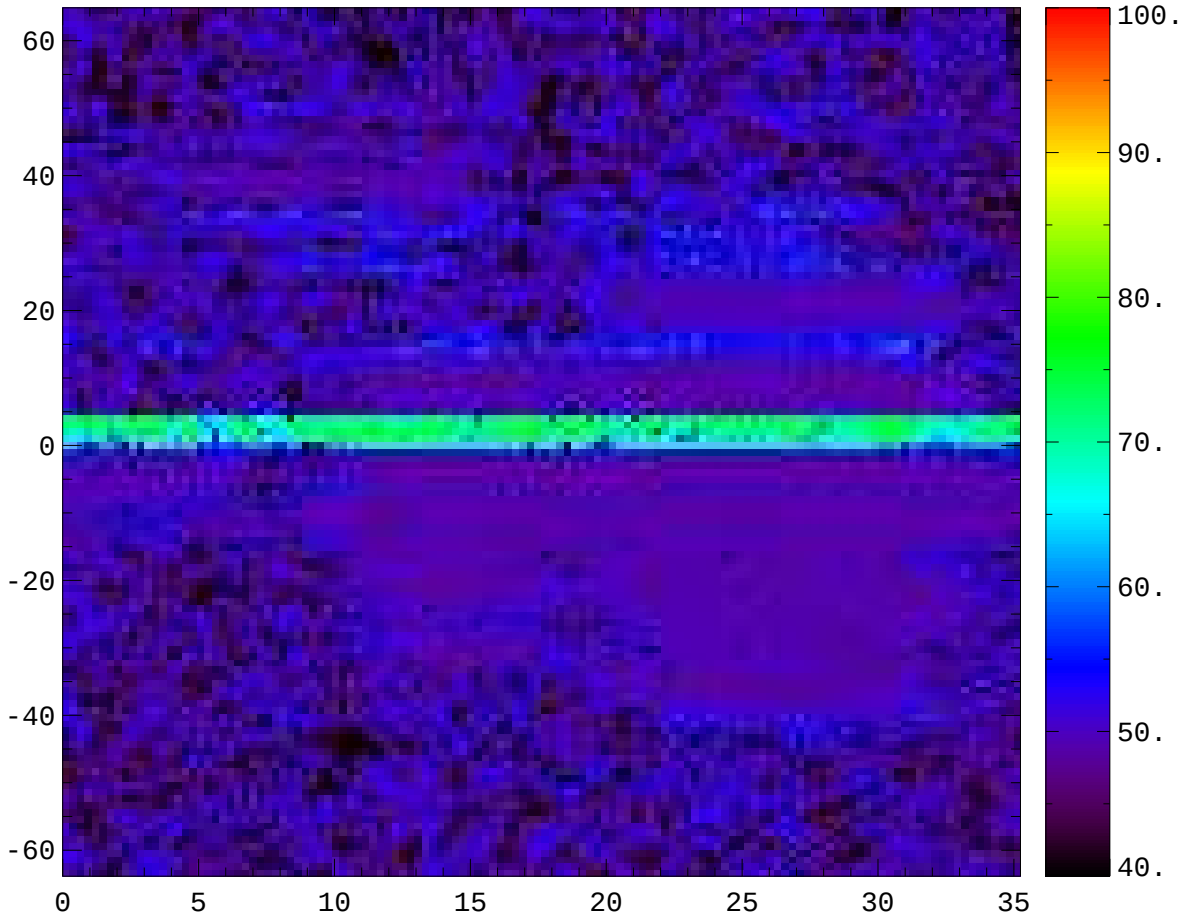

FIG. 5. Same as in Figure 4 after filtering. For the transient signal components, the Gabor coefficients were replaced by estimated thresholds for the stationary signal part at the given frequency.

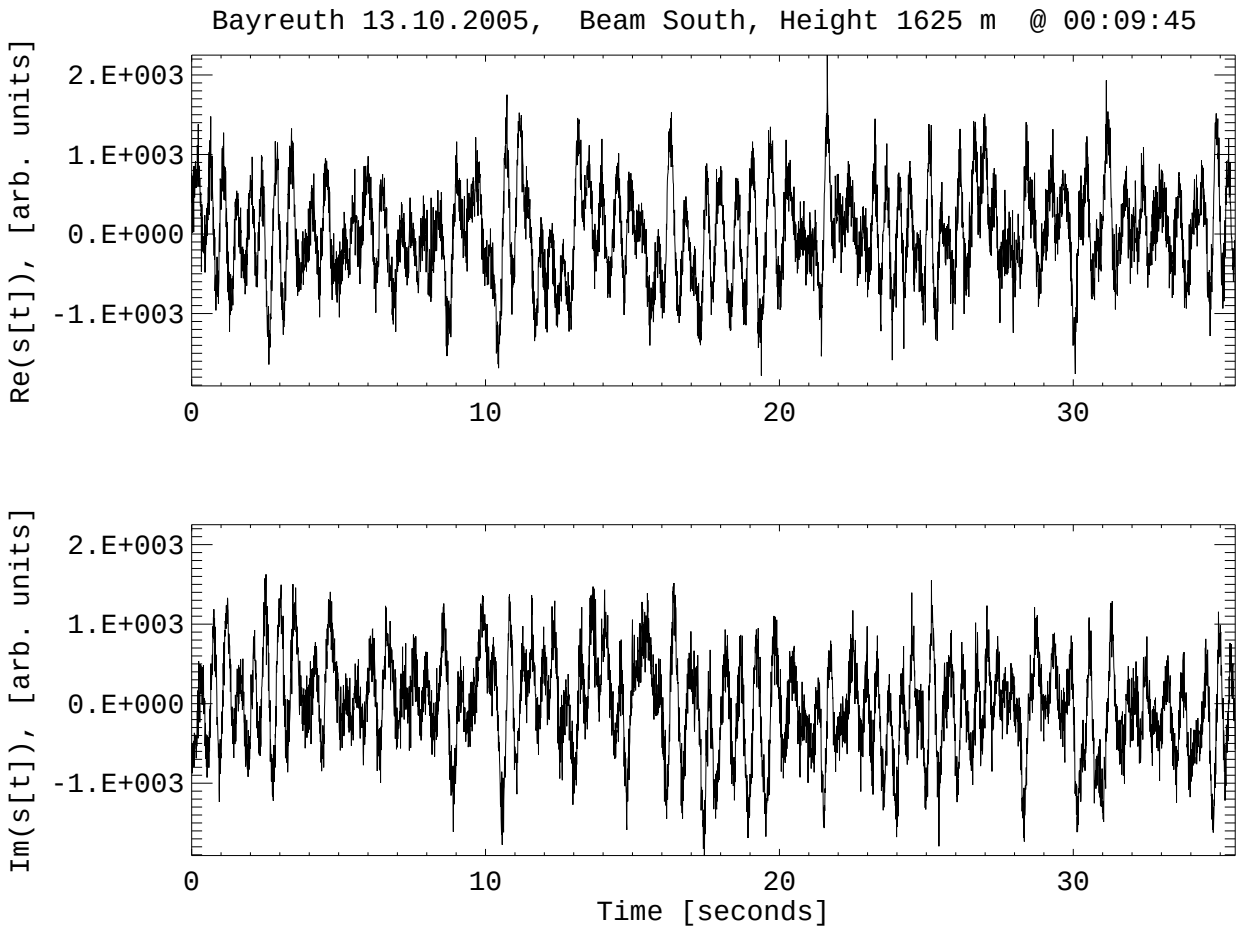

FIG. 6. Same as in Figure 3, but for the cleaned signal obtained from the filtered Gabor representation shown in Figure 5.

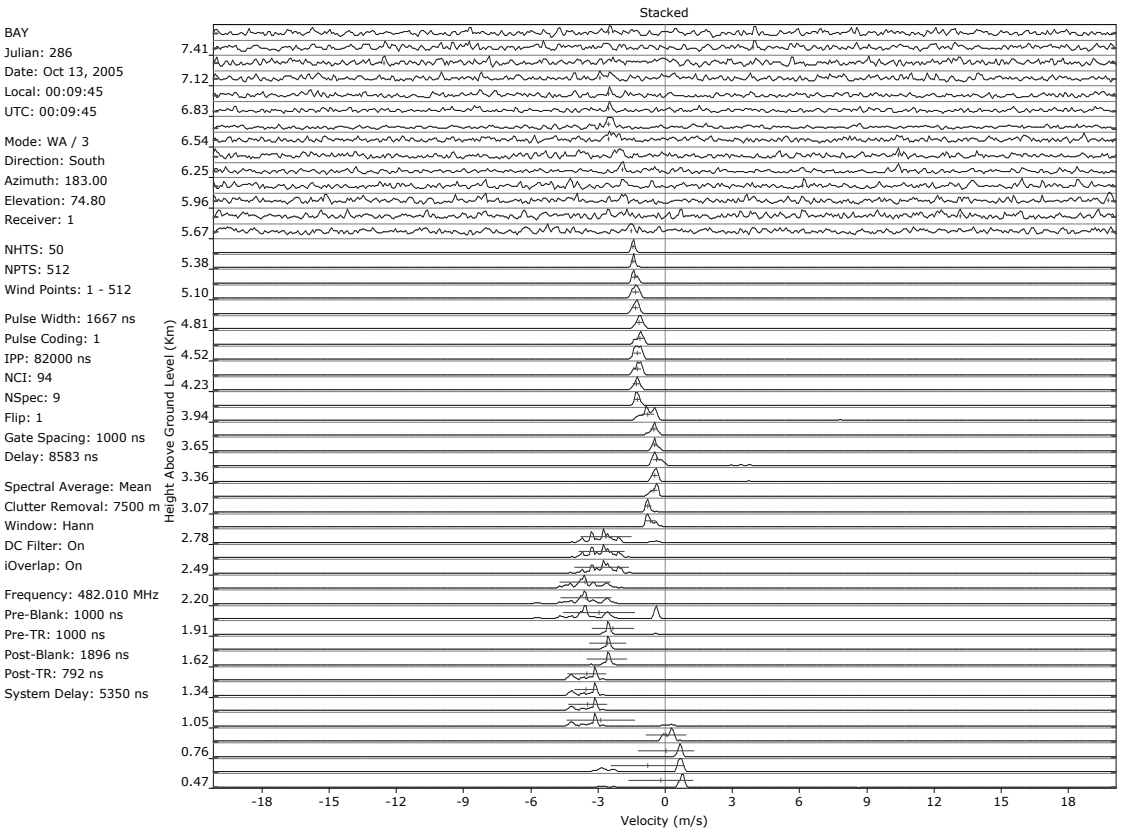

FIG. 7. Stacked plot of Doppler spectra for all low mode range gates, obtained through standard processing without any bird mitigation algorithm. Data were measured at 00:09:45 UTC on October 13, 2005 (south beam) with the 482-MHz RWP at Bayreuth, Germany. The estimated first and second moments are symbolized as a cross, where the vertical line shows the first moment (mean Doppler speed) and the horizontal line denotes spectral width. Massive bird contamination can bee seen in the range gates below 3.0 km height

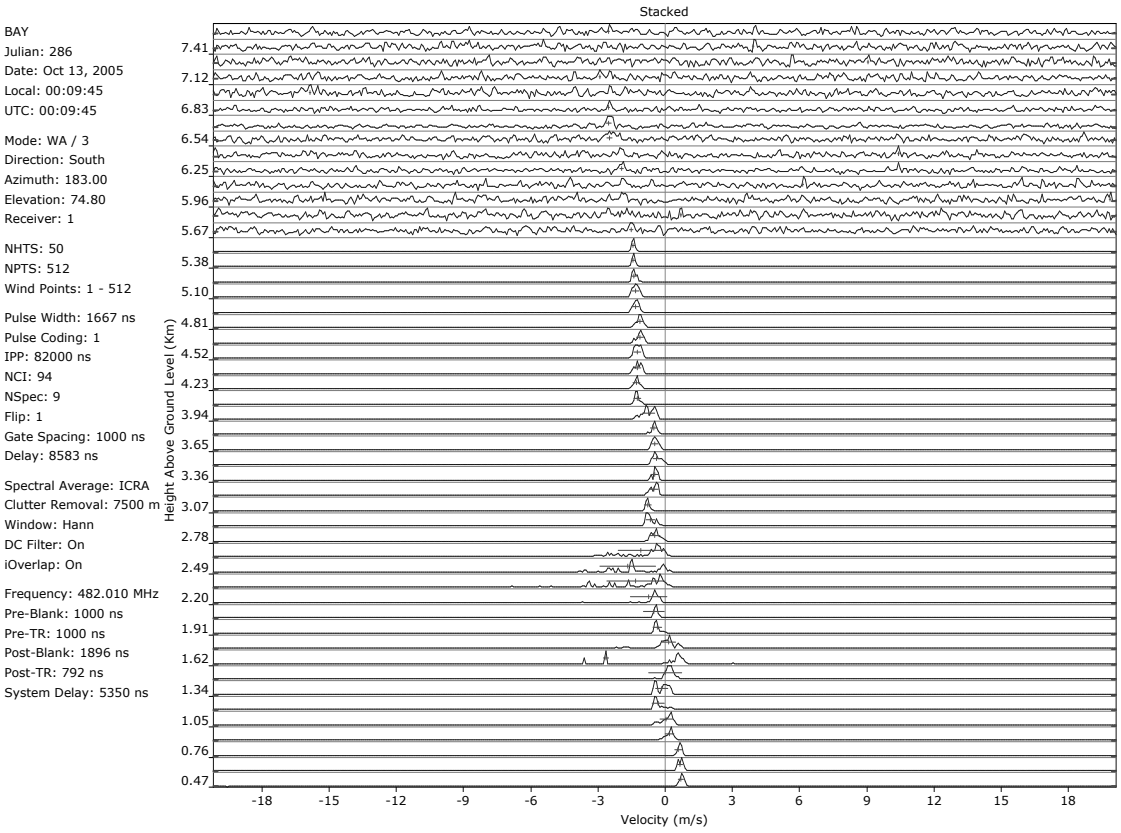

FIG. 8. Same as in Figure 7, but Doppler spectra were estimated using the operational bird-mitigation algorithm ICRA. Bird contamination below 3.0 km height is reduced compared to Figure 7, but still significant.

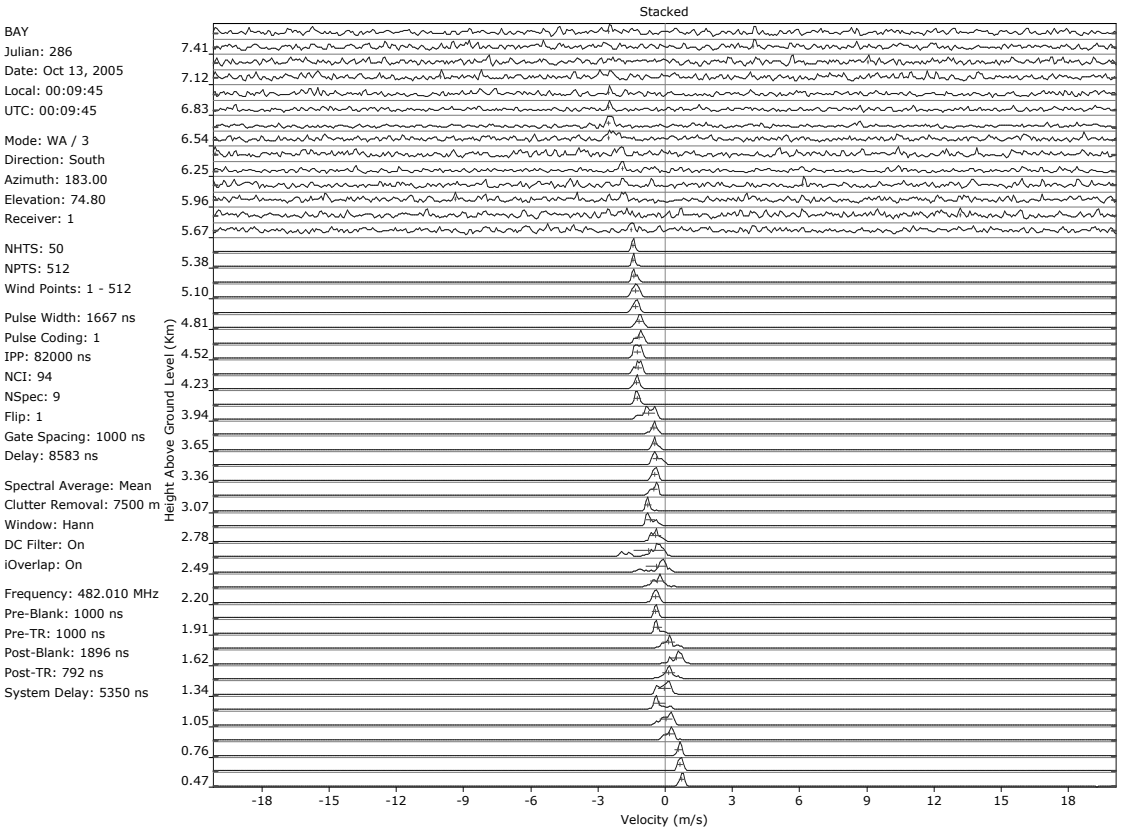

FIG. 9. Same as in Figure 7, but Doppler spectra were estimated after statistical Gabor filtering of the original time series. Only minor remnants of bird contamination can be seen in range gates 15 and 16 (at 2.5 and 2.6 km height).

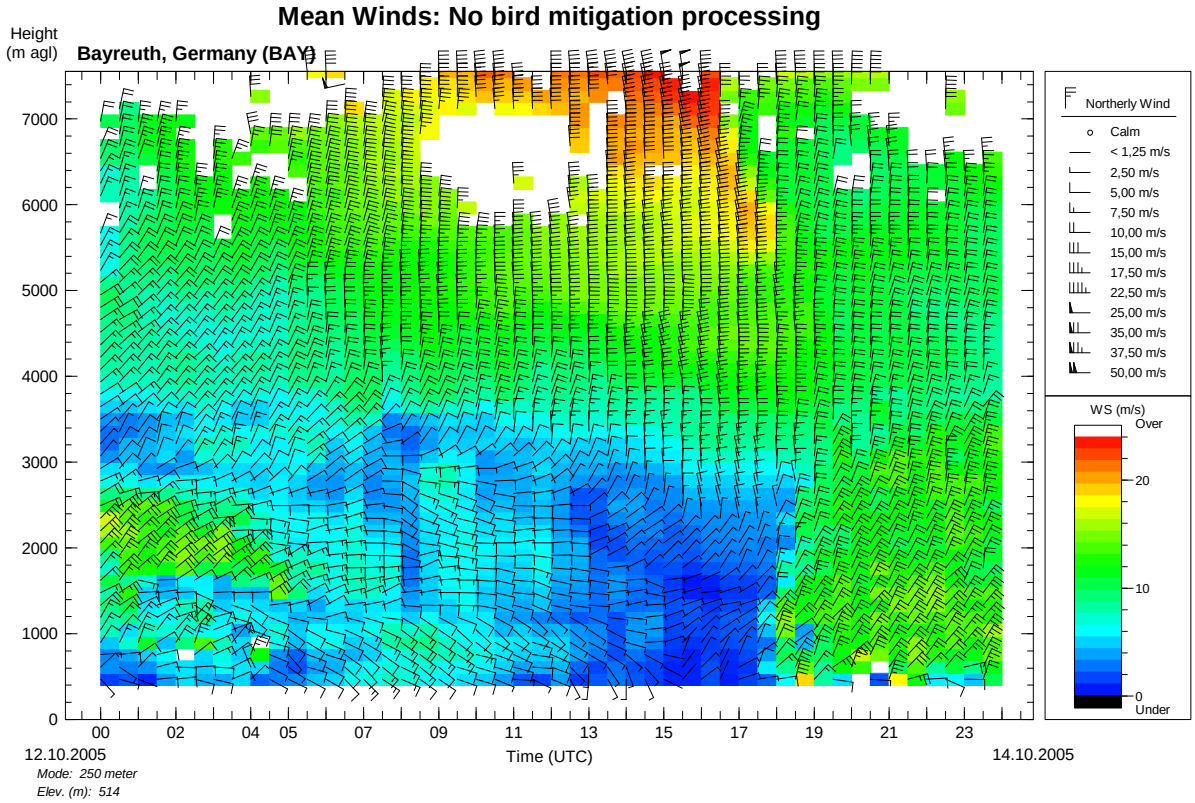

FIG. 10. Wind barb plot of horizontal winds measured in the low mode at Bayreuth on October 13, 2005. The x-axis shows time and the y-axis denotes height. Data have been color coded by wind speed. The signal processing was using no Bird mitigation algorithm. Relatively strong northeasterly winds below about 3.5 km indicate strong bird migration, this can be seen between 00 and 05 UTC at heights around 1000 m and above 1600 m and especially after 18 UTC from the lowest gate to about 3500 m.

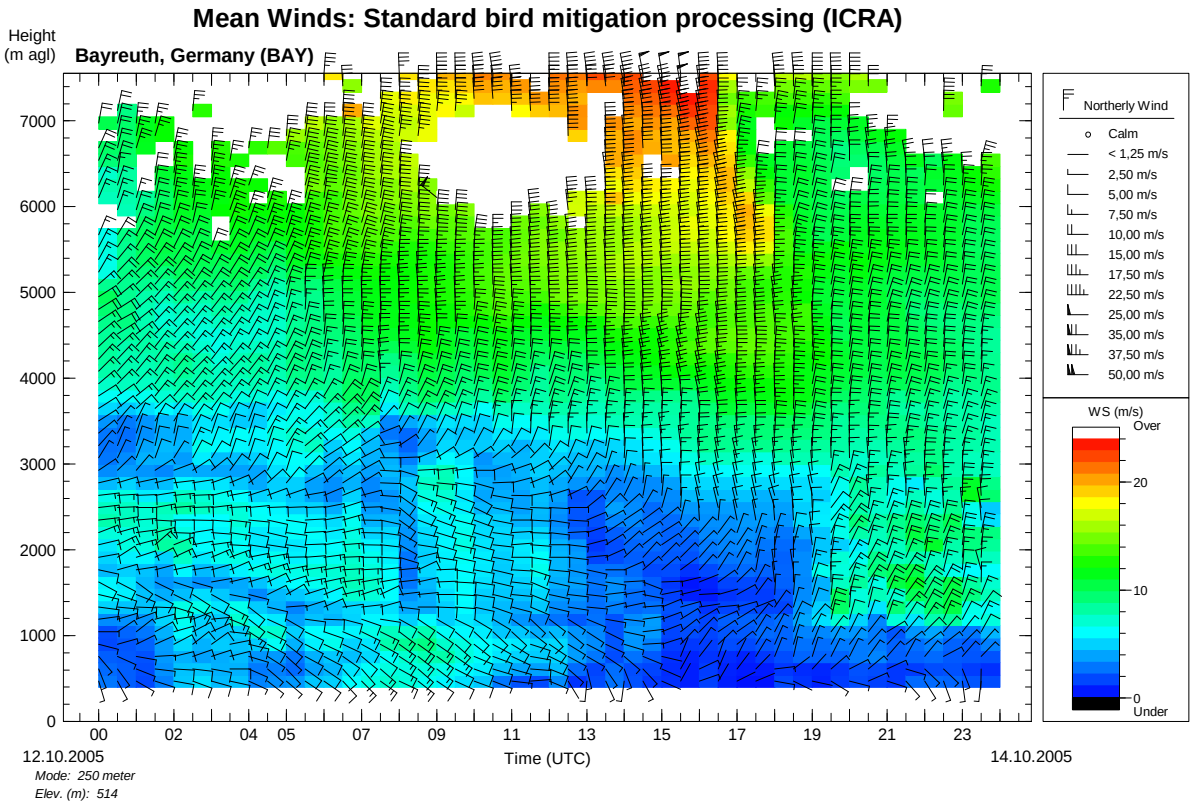

FIG. 11. Same as in Figure 10. The signal processing was using the standard ICRA algorithm. Bird contamination has been reduced compared to Figure 10, but is still significant after 19 UTC. A few other northeasterly wind barbs around 02 UTC are affected by intermittent clutter echoes.

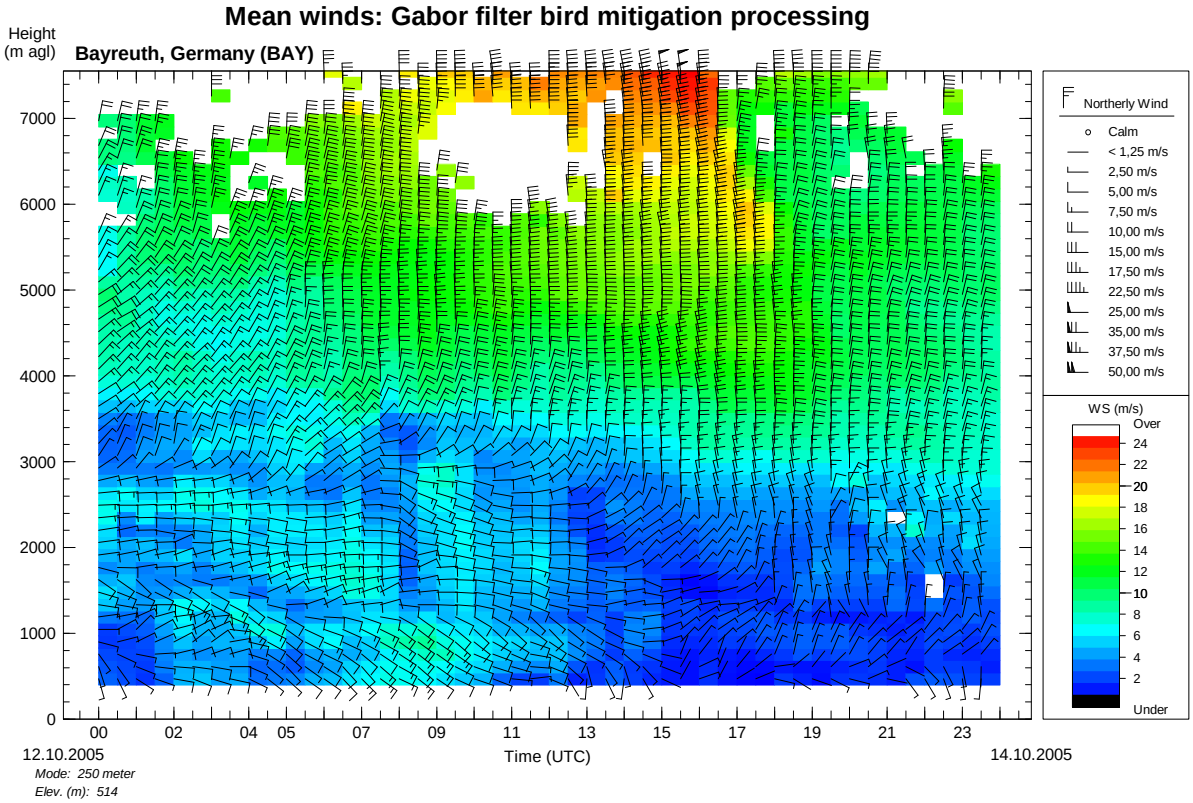

FIG. 12. Same as in Figure 10. The signal processing was using the new Gabor filter algorithm. Bird contamination has again been reduced compared to Figure 11. There are no indications of bird migration between 00 and 05 UTC, and only a few obvious outliers and missing data after 19 UTC.

List of Tables

Center frequency	482.0078 MHz
Peak (Average) RF envelope power (PEP)	16 (2.4) kW
Pulse modulation	Amplitude (B/W) Phase (pulse compression)
Pulse widths (vert. resolution)	1.7 μs (250 m) 2.2 μs (330 m) 3.3 μs (500 m) 4.4 μs (660 m)
Antenna type	Phased array of 180 CoCo antennas
Antenna aperture (area)	142 m^2 (12.4 $\times$ 11.5 m)
On-axis gain above isotropic	$\geqslant$ 34 dBi
One-way half power (3 dB) beamwidth	$\leqslant 3$
Oblique beam zenith distance	15.2
RX type	Heterodyne (IF 60 MHz), Digital IF
LNA noise figure	$\leqslant$ 0.6 dB
A/D conversion	14 bit (@ max 66 MHz)
Pulse compression	Bi-phase, complementary, max 32 bit
System sensitivity	$\leqslant$ -154 dBm
Vertical measuring range	16 km (wind), 4 km (virt. temp.)

TABLE 1. Technical parameters of the 482 MHz RWP/RASS at Bayreuth/Germany

	Wind Low-Mode
Inter Pulse Period	82 μs
Pulse Width	1.7 μs
Tx Duty	2.07 %
# of code bits	1 (phase flip)
Pulse Peak Power (PEP)	16 kW
Spacing (on RX)	1.0 μs
# of Gates	50
First Gate	8.6 μs

TABLE 2. TX and RX sampling parameters in routine operation