Application of model-based spectral analysis to wind-profiler radar observations
Eric Boyer, M. Petitdidier, W. Corneil, C. Adnet, Pascal Larzabal

To cite this version:
Eric Boyer, M. Petitdidier, W. Corneil, C. Adnet, Pascal Larzabal. Application of model-based spectral analysis to wind-profiler radar observations. Annales Geophysicae, European Geosciences Union, 2001, 19 (8), pp.815-824. <hal-00329177>

HAL Id: hal-00329177
https://hal.archives-ouvertes.fr/hal-00329177
Submitted on 1 Jan 2001

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Application of model-based spectral analysis to wind-profiler radar observations

E. Boyer\textsuperscript{1}, M. Petitdidier\textsuperscript{2}, W. Corneil\textsuperscript{2}, C. Adnet\textsuperscript{3}, and P. Larzabal\textsuperscript{1,4}

\textsuperscript{1}LESiR/ENS Cachan, UPRESA 8029, 61 avenue du président Wilson, 94235 Cachan cedex, France
\textsuperscript{2}CETP, 10-12 Avenue de l’Europe, 78140 Velizy, France
\textsuperscript{3}THALES Air Defense, 7-9 rue des Mathurins, Bagneux France
\textsuperscript{4}IUT de Cachan, CRIIP, Université Paris Sud, 9 avenue de la division Leclerc, 94 234 Cachan cedex, France

Received: 24 July 2000 – Revised: 13 February 2001 – Accepted: 5 March 2001

Abstract. A classical way to reduce a radar’s data is to compute the spectrum using FFT and then to identify the different peak contributions. But in case an overlapping between the different echoes (atmospheric echo, clutter, hydrometeor echo...) exists, Fourier-like techniques provide poor frequency resolution and then sophisticated peak-identification may not be able to detect the different echoes. In order to improve the number of reduced data and their quality relative to Fourier spectrum analysis, three different methods are presented in this paper and applied to actual data. Their approach consists of predicting the main frequency-components, which avoids the development of very sophisticated peak-identification algorithms. The first method is based on cepstrum properties generally used to determine the shift between two close identical echoes. We will see in this paper that this method cannot provide a better estimate than Fourier-like techniques in an operational use. The second method consists of an autoregressive estimation of the spectrum. Since the tests were promising, this method was applied to reduce the radar data obtained during two thunderstorms. The autoregressive method, which is very simple to implement, improved the Doppler-frequency data reduction relative to the FFT spectrum analysis. The third method exploits a MUSIC algorithm, one of the numerous subspace-based methods, which is well adapted to estimate spectra composed of pure lines. A statistical study of performances of this method is presented, and points out the very good resolution of this estimator in comparison with Fourier-like techniques. Application to actual data confirms the good qualities of this estimator for reducing radar’s data.

Key words. Meteorology and atmospheric dynamics (tropical meteorology)- Radio science (signal processing)- General (techniques applicable in three or more fields)

Correspondence to: E. Boyer (eric.boyer@wanadoo.fr)

1 Introduction

Wind-profiler radars generally operate in the VHF and UHF bands, and are used to observe the scattering of radiowaves from clear-air refractive-index fluctuations. Although not necessarily optimised for detecting other types of scatterers, the signals that the radars receive, nevertheless, contain non-clear air echoes, such as hydrometeors at UHF, lightning emissions at VHF, clutter at UHF and VHF, etc. As these radars become an operational tool for short-term prediction and forecasting, as well as a research instrument for boundary-layer tropospheric and stratospheric studies, then robust algorithms have to be implemented in order to determine the different Doppler frequencies present in the signal, to classify and select the important ones, and to compute the wind. Primarily, the ordinary algorithm used contains the following steps: coherent integrations of the signal, application of a window, computation of the spectrum using FFT and incoherent integrations (Tsuda, 1989). Sato and Woodward (1982) computed the autocorrelation function instead of the FFT. Many approaches have been developed to search the echoes and to extract the corresponding Doppler frequencies (Yamamoto et al., 1988; Hocking, 1997) with the selection of the right frequency as the final step in order to compute the wind velocity. When the sought echo is unique and well separated from the clutter, and located at 0 Hz, any algorithm works, yet the fastest is the best. But over a whole vertical wind-profile (i.e. wind as a function of the altitude at a given time) and in all weather conditions, this is not the case. Several echoes which may overlap could be present in the spectrum. The classical case of overlapping is the one between clutter and atmospheric echo. Fourier-like techniques provide poor results for close echoes; the frequency resolution is not enough to separate the different contributions. The atmospheric echo may be approximated by a Gaussian shape, but the clutter does not correspond to any analytical model, except when it may be considered as a single spectral
SPECTRUM OF THE SIGNAL

CEPSTRUM OF THE SIGNAL: \[ CE(S^*(f)) \]

CEPSTRUM OF THE PERIODICAL TERM

\[ CE(\delta(f) + \delta(f - f_0)) \]

Fig. 1. Cepstrum results (a) FFT spectrum, (b) cepstrum of the signal, (c) cepstrum of one echo, (d) cepstrum of the periodic term.

point. As a consequence, even with a very sophisticated algorithm devoted to peak identification in a spectrum, it may be difficult to separate overlapped echoes, especially when their amplitudes are very different. As an example, in UHF, if overlapped echoes could not be separated, the identified Doppler frequency may shift from clear-air to a hydrometeo one, from altitude to altitude, from time to time, or from beam to beam. As a consequence, spurious data or no data, if using a consensus window, are obtained.

In this paper, we have investigated model-based spectral methods on simulated and actual data. These methods provide the different main frequency-components present in the signal; these methods replace the peak-identification algorithm. The aim of this work is to point out the possibility of separating overlapped echoes directly and then to improving the number of reduced data and their quality relative to Fourier spectrum analysis. The radar data came from the thunderstorm campaign that took place at the National Astronomy and Ionospheric Center (NAIC) in Arecibo, Puerto Rico in 1998 (Petitdidier et al., 2000); UHF and VHF time series are both used in our tests. In the first part, an estimation based on cepstrum properties is studied. The cepstrum algorithm has been used to determine the frequency shift between two echoes, and it is known to provide good results when the echoes are alike (Oppenheim and Shafer, 1989; Balluet et al., 1981). For example, this algorithm has already been applied in Particle Image Velocimetry (PIV) (Fournel et al., 1992). In this paper, the algorithm will be applied to the power spectrum of time series provided by a VHF ST radar with the aim of estimating the shift between two different echoes (for example, the shift between the ground clutter and the atmospheric echo). Applied to our data, this method does not provide a major improvement versus the spectrum analysis. Then, autoregressive methods are studied in the second part. The autoregressive (AR) spectral estimator is a standard tool in the field of spectral estimation and time series
analysis (Kay and Marple, 1981). This method is applied directly to time series. Subspace-based methods are presented in the third part. The key idea consists in a decomposition of the observation space in two subspaces: the signal subspace containing the echoes, and the noise subspace (Bienvenue and Kopp, 1979, Schmidt, 1986). This decomposition is realized by computing the covariance matrix of the signal, which is then decomposed into its eigenvectors. The significant eigenvalues correspond to signal subspace eigenvectors and the other eigenvalues correspond to the noise subspace.

2 Exploitation of the power cepstrum

2.1 The power cepstrum algorithm

In the cepstrum algorithm, the power spectrum \( S^*(f) \) is supposed to be the sum of two identical echoes, \( f_0 \) shifted (Fig. 1a):

\[
S^*(f) = S(f) + S(f - f_0) \tag{1}
\]

\[
s^*(t) = s(t)(1 + e^{2j\pi f_0 t}) \tag{2}
\]

where \( s(t) \) is the Fourier transform of \( S(f) \) and \( s^*(t) \) is the Fourier transform of \( S^*(f) \). Let us define the complex logarithm by

\[
\ln c(S) = \ln |S| + i \arg(S) \tag{3}
\]

where \( \arg(S) \) represents the unwrapped phase of \( S \) to ensure the uniqueness of the function. In the development

\[
\ln c(s^*(t)) = \ln c(s(t)) + \ln c(1 + e^{2j\pi f_0 t}) \tag{4}
\]

the term \( \ln c(1 + e^{2j\pi f_0 t}) \) is a \( f_0 \) periodical: a Fourier transform of Eq. (4) exhibits a peak at the expected frequency, \( f_0 \). Consequently, the power cepstrum of \( S^*(f) \) is defined by:

\[
CE(S^*(f)) = \left| FT\left( \ln c\left[ FT\left( S^*(f) \right) \right] \right) \right| \tag{5}
\]

where \( CE \) denotes the cepstrum operator and \( FT \) the Fourier Transform operator. (5) can be rewritten as

\[
CE(S^*(f)) = CE(S(f)) + CE(\delta(f) + \delta(f - f_0)) \tag{6}
\]

where \( \delta(f) \) denotes the Dirac distribution.

The cepstrum operator is generally applied directly on a time series, but in our case, the cepstrum operator is applied on spectra because we estimate the shift between two Gaussian spectra. Consequently, cepstra are represented with a “frequency” axis in Fig. 1 (b, c and d).

The second term

\[
CE(\delta(f) + \delta(f - f_0)) = FT\left( \ln c\left[ 1 + e^{2j\pi f_0 t} \right] \right)
\]

of Eq. (6), represented in Fig. 1d, exhibits a peak at the \( f_0 \) frequency and harmonics which are easily interpreted by the Taylor expansion of the logarithm function. In Fig. 1b, we can easily detect this peak but in some cases, the first term \( CE(S(f)) \) (represented in Fig. 1c) can hide that peak and the estimation won’t be possible (especially when the second Gaussian echo presents a significant attenuation in comparison with the first one).

2.2 Simulation results

For simulations, the spectrum of the signal has been computed as proposed by Papouli (1965):

\[
S^*(f) = \frac{1}{N_{inc}} \sum_{m=1}^{N_{inc}} (S_{th}(f) + \sigma_n^2)|y_m|^2 \tag{7}
\]

where \( S_{th}(f) \) is the theoretical spectrum, \( \sigma_n^2 \) is an additive noise power, \( N_{inc} \) is the number of incoherent integration and \( y_m \) is a Gaussian random variable with unit standard deviation. We took 256 points for the FFT algorithm.
Fig. 3. AR method: tests on the number of pole (a) FFT spectrum, (b) reconstructed spectrum with a pole number of 12; (c) reconstructed spectrum with a pole number of 5; (d) reconstructed spectrum with a pole number of 50. The vertical line drawn on each plot corresponds to the 0 Hz frequency.

In Fig. 2, two different Gaussian echoes have been computed: the first one at zero frequency, modeling the ground clutter with a SNR of 15 dB and a normalized standard deviation of 0.0015 and the second one, with a normalized frequency of −0.2, a SNR of 20 dB and a normalized standard deviation of 0.025. Even if the power cepstrum exhibits a peak at the expected frequency, the presence of numerous other peaks does not permit an immediate reliable estimation of the frequency shift.

2.3 Conclusion

As shown in the simulation, robust extraction of the shift frequency is not possible due to the numerous peaks. Consequently, this method has not been applied on actual data. In Sects. 3 and 4, a very different approach has been investigated which presents great improvements in relation to the FFT algorithm. It consists of two powerful parametric methods which rely on different models for the received signal.

3 Exploitation of autoregressive methods

3.1 Presentation of the method

The autoregressive (AR) spectral estimator is a standard tool in the field of spectral estimation and time series analysis (Kay and Marple, 1981). AR models are often used to model data which are not necessarily generated by AR equations. For instance, high-order models are commonly used to estimate peaked spectrum. This method is well-known for its good behaviour in resolving spectral peaks from noisy data. Given a time series $x_n$, $n = 0, 1, \cdots, N - 1$, which is the sum of the output of an AR process, $s_n$, and a white noise $x_n = s_n + w_n$.
and given
\[ \sum_{k=0}^{p} a_k s_{n-k} = -\xi_n \quad a_0 = -1 \]
where \( w_n \) and \( \xi_n \) are uncorrelated white noise processes, it is well-known that
\[ S_{AR}(\omega) = \frac{\sigma^2}{|A(\omega)|^2} \]
where \( \sigma^2 \) is the variance of \( \xi_n \) and
\[ A(\omega) = \sum_{k=0}^{p} a_k e^{-j\omega k} \quad a_0 = -1 \] .

In the application investigated in this paper, \( a_k \) will be estimated by solving the Yule-Walker equations using the well-known Levinson-Durbin algorithm.

3.2 Selection of the order

The first step is to determine the length of the predictor or pole number \( p \) for a correct representation of the spectrum. This delicate choice is important for the frequency estimation, as shown in Fig. 3, where three different values of \( p \) have been tested. In the present case (Fig. 3a), the atmospheric echo is stronger than the clutter one; in many cases, it is the reverse situation. \( p = 5 \) (Fig. 3c) leads to a fairly rich representation of the spectrum; in the case where an atmospheric echo is weaker than the clutter one, the atmospheric echo may not be present in the spectrum. \( p = 50 \) (Fig. 3d) leads to a fairly rich representation of the spectrum with parasite peaks: it is no longer possible to safely select the right Doppler frequency. The intermediate case of \( p = 12 \) (Fig. 3b) leads to a correct representation of the spectrum.

In order to safely select a correct pole number \( p \), typical VHF spectra have been simulated, corresponding to two different extreme cases (Fig. 4a and Fig. 5a). For each spectrum, a statistical study of the performance (bias and rms) of the estimator according to \( p \) has been investigated over 500 Monte Carlo noise realizations. Figure 4 (b and c) and Fig. 5 (b and c) show that for these two typical situations, the optimal order is around \( p = 12 \); in Fig. 4b and Fig. 4c, the value of \( p = 13 \) minimizes both the mean bias and mean
rms; in Fig. 5b and Fig. 5c, the value of \( p = 12 \) minimizes the mean bias, but not the rms. An alternative to this classical statistical study is to select \( p \) according to the predictive noise spectral flatness. Such a spectral flatness coefficient \( F_c \) can be defined as follow (Stoïca, 1997)

\[
0 \leq F_c = \frac{\exp \left[ -\frac{1}{2} \int \ln \left( S_n(f) \right) df \right]}{\frac{1}{2} \int S_n(f) df} \leq 1,
\]

where \( S_n(f) \) is the spectrum of the noise given by

\[
S_n(f) = \frac{S(f)}{S_{AR}(f)}
\]

with \( S(f) \) as the FFT spectrum and \( S_{AR}(f) \) as the AR spectrum. Since this noise must be white, the pole number \( p \) for the AR estimation is given by the corresponding spectrum \( S_n(f) \) which is the flattest. The flatness of \( S_n(f) \) is characterized by the spectral flatness coefficient previously introduced. As the flatness of \( S_n(f) \) increases with \( F_c \), the pole number \( p \) for the AR estimation is given by the value which maximizes \( F_c \). The results of simulations lead to a choice of \( p = 13 \) in the first case (Fig. 4d) and \( p = 10 \) in the second case (Fig. 5d). Regarding all of these results, the value of \( p = 12 \) seems to be a good choice for the various scenarios.

### 3.3 Experimental Results

The AR method has been applied to a time series of VHF and UHF radar data. Since the tests were promising, this method has been applied to reduce the radar data obtained during two thunderstorms. The reconstructed spectrum is computed for each sampled altitude and time. The adequacy between the FFT spectrum and the reconstructed spectrum confirms the choice of \( p = 12 \) for the AR model. The following step is to select the right peak frequency. In the presence of a broadband echo, two very close peaks are obtained regardless of whether the pole number is even or odd; an average value,
representative of the mean peak frequency, is computed. In some cases, although fewer than in the regular FFT spectrum method, the Doppler frequency cannot be detected because it is merged into a large echo. This case occurs when the atmospheric echo is very close to the clutter and relatively broad. This method, as it was implemented, cannot provide either the standard deviation or the backscattered power; using the reconstructed spectrum, they are underestimated especially when \( p \) is small. In the data reduction process, these parameters were obtained by means of the FFT spectrum after the identification of the mean frequency of the concerned echo.

3.4 Conclusion

The AR method, as it was implemented, improved the peak identification relative to the FFT spectrum analysis. It was an interesting method to obtain a first guess of the wind data; therefore, a peak search was no longer necessary. In our case, one of the limitations is the need to compute the spectrum in order to retrieve the other parameters, such as standard deviation and backscattered power. But further investigations will be carried out to avoid this constraint.

4 Introduction of a model-based approach for subspace parametric estimation

Different algorithms are exploiting subspace properties of covariance matrix methods (Bienvenue and Kopp, 1979, Schmidt, 1986). For simulation and experimental results, the MUSIC algorithm has been computed.

4.1 Application of the MUSIC algorithm to the times series

MUSIC (Multiple Signal Characterization) is one of the “high resolution subspace methods”, which is well adapted to estimate spectra composed of pure lines. This was first used as an array processing to estimate the direction-of-arrival of different point sources. Our approach is original since the MUSIC algorithm is applied to the autocorrelation function of the time series. On the hypothesis of a spectrum containing echoes with a relatively small standard deviation, the theoretical autocorrelation function is considered as a sum of sinusoidal signals with a slowly varying amplitude,

\[
r(t) = \sum_i \alpha_i(t)e^{j\omega_i t}.
\]

The frequency estimation is conducted here as if the envelope was constant (the terms \( \alpha_i(t) \) are assumed to be independent of \( t \)). This hypothesis corresponds to a spectrum composed of pure lines and it leads to the use of the MUSIC algorithm. Consider an \( N \times 1 \) complex vector \( r(t) \) corresponding to the concatenation of \( N \) autocorrelation lags. As in array signal processing problem formulation, in order to obtain multiple vectorial sub-observations of \( r(t) \), the vector \( r(t) \) must be reshaped into \( K \times 1 \) vectors \( r_k \), as explained in Fig. 6, where \( T_S \) is the pulse repetition time, \( \Delta T_S \) is an adjustable shift and \( K = \text{int}[(N - m)/\Delta + 1] \), where \text{int}() denotes the integer part operator. A classical eigenvalue decomposition of the covariance matrix

\[
R_r = \frac{1}{K} \sum_{k=1}^{K} r_k r_k^H
\]

is then performed to estimate the first spectral moments of Gaussian echoes and where \((\bullet)^H\) denotes the Hermitian transpose. The influence of the temporal variation of \( \alpha_i(t) \) on the quality of the MUSIC estimation in the case of a single echo can be found in Besson and Stoica (1996).

4.2 Simulation results

In order to exhibit the improvement obtained with subspace approaches versus FFT algorithms, we compute a simulation with two identical Gaussian echoes, with a fixed normalized standard deviation of \( \sigma = 0.0125 \) for 256 points. For each value of the SNR shown in Fig. 7, the limit of resolution, obtained by the smallest value of the shift between echoes, which provides two well separated peaks in the spectrum, is evaluated. As expected, the limit of resolution is better for the MUSIC algorithm and enables a detection of the two
Fig. 8. Simulation example with two identical Gaussian echoes $N_{inc} = 1$, $SNR = 10dB$, $Nfft = 1024$, $f_1 = 0.0391$, $f_2 = 0.0586$ and $\sigma = 0.01$.

Fig. 9. Simulation example with two different Gaussian echoes $N_{inc} = 1$, $SNR = 10dB$, $Nfft = 256$, $\sigma_1 = 0.005$, $f_1 = 0.0781$, $\sigma_2 = 0.02$ and $f_2 = 0.0813$.

echoes, where the Fourier-like techniques lack in resolution. Moreover, one can notice that for a SNR greater than $-5$ dB, the resolution of the MUSIC algorithm reaches the zero limit: in this case, it is always theoretically possible to separate two echoes. With actual data, the noise and statistical fluctuations of the signal limit the possibility to separate very close echoes. This improvement of the resolution is illustrated in Fig. 8 and Fig. 9 where a strong overlapping has been simulated. In Fig. 8, considering only the spectrum, it is difficult to know whether the spectrum is composed of one or two echoes, except if this was ascertained from time or/and altitude continuity. Bias and variance have been estimated with 500 Monte Carlo simulations. Clearly, an estimation of the mean frequencies is still possible in a domain where Fourier-like techniques fail.

4.3 Experimental results

Figures 10 and 11 present the results of a MUSIC computation of a VHF times series obtained during a thunderstorm, in
E. Boyer et al.: Application of model-based spectral analysis to wind-profiler radar observations

Fig. 10. Experimental results: first example.

Fig. 11. Experimental results: second example.

Arecibo, PR during September and October 1998. These two examples present cases with overlapped echoes. For these two cases, MUSIC provides an estimation of the mean velocity which is consistent with the Fourier spectrum. In Fig. 11, MUSIC exhibits three significant peaks where a Fourier estimation would have some difficulties in estimating the relative contribution of the ground clutter and the “negative” frequency peak. Of course, we need to pay attention to the interpretation of the results: when Gaussian echoes have a large standard deviation, MUSIC will exhibit several peaks inside that echo, thereby complicating the analysis of the results. Thus, an a-posteriori analysis is necessary to interpret the different peaks obtained with MUSIC and to select the one which is representative of the atmospheric contribution. As a matter of fact, several contributions may be present in the spectrum: clear air, clutter, hydrometeors, in case of strong rain, and spurious echoes. Moreover, it is important to mention that the MUSIC algorithm can be implemented in real time.

4.4 Conclusion

If the standard deviation of the echo is small, subspace-based methods provide good results. Application of subspace-based methods to large standard echoes leads to poor statistical performances due to the inadequacy between the signal and the pure line spectra model assumed by subspace-based methods. An alternative to this problem is to take into account the algorithm for both the Doppler frequency and standard deviation, conducting so to a 2D pseudo-spectrum. This is ongoing work (Boyer and al., 2001).

5 Conclusion

Several approaches were presented to overcome the lack of resolution of Fourier-like techniques. The Cepstrum technique does not bring immediate improvements as it is not well adapted to broadband echoes and therefore does not permit overlapped-echo identification. The AR technique, due to existing routines, is easy to implement. The determination of the pole number has to be adapted to the concerned case. The peak selection is easier than in the FFT spectrum due to some filtering, but some cases need a more sophisticated search than a maximum search. The standard deviation and the backscattered power could be retrieved from the “reconstructed” spectrum, but further investigation must be done. The MUSIC algorithm, applied to the autocorrela-
tion function, appears to be a good alternative to Fourier-like techniques when echoes are overlapped, with a possible real time implementation. A decisive step is to estimate the two other moments of the echoes (for a joint estimation of the three spectral moments of the echoes, see Boyer et al., 2001). The next step will be to implement this algorithm on a work station, to reduce actual data routinely and to compare with methods based on peak identification in the power spectrum.

Acknowledgement. The National Astronomy and Ionospheric (NAIC) Center is operated by Cornell University under contract with the National Science foundation.

Topical Editor D. Murtagh thanks C. E. Meek for his help in evaluating this paper.

References


Fournel, T., Daniere, J., Moine, M., Pigeon, J., and Courbon, M., Utilisation du cepstre d’énergie pour la vélométrie par images de particules, Traitement du signal vol.9-n°3, 1992


