White paper on FITACF3

P. V. Ponomarenko[1], E. Bland[2], and K. Kotyk[1]

[1]University of Saskatchewan, Saskatoon, SK, Canada
[2]University Centre in Svalbard, Norway

1 Introduction and historic notes

The Super Dual Auroral Radar Network (SuperDARN) is a global network of over 30 high frequency (HF) radars designed for monitoring high-latitude plasma circulation in the northern and southern hemispheres (Greenwald et al., 1995; Chisham et al., 2007; Nishitani et al., 2019). The radars exploit the refractive property of the ionosphere for HF radiowaves to achieve perpendicularity with the Earth's magnetic field at high latitudes. This makes it possible to measure Doppler shift variations in 8–20 MHz coherent backscatter from field-aligned electron density irregularities in the ionosphere, which act as tracers for the plasma circulation.

The SuperDARN data are currently distributed to users as range-time arrays of complex-valued 'autocorrelation' (actually, autocovariance) functions (ACFs). These ACFs are then processed using the *Radar Software Toolkit (RST)*, which is the standard SuperDARN data analysis software, or with users' own packages. The RST contains all routines necessary for converting multi-radar ACF datasets into maps of high-latitude ionospheric electric potential (SuperDARN Data Analysis Working Group, 2021). The initial analysis stage involves processing the ACF data from each radar to estimate the main echo parameters and their respective measurement errors for each range gate and integration interval. These echo parameters are the Doppler velocity, spectral width, signal-to-noise ratio, and vertical angle of arrival. The standard method for estimating these parameters in the RST is a package called *FITACF*, which performs a series of least squares fits to the ACF amplitude and phase (Baker et al., 1988; Ponomarenko & Waters, 2006). Several other methods for processing ACFs have also been developed, including the *FITEX* algorithm, which is similar to FITACF but uses a different phase fitting method to estimate the Doppler velocity (Greenwald et al., 2008; Ribeiro et al., 2013). SuperDARN ACFs have also been analysed using the multiple signal classification (MUSIC) method in order to separate multiple echoes with different velocities and spectral widths contained within a single ACF (Barthes et al., 1998). Another method called *LMFIT* uses a Levenburg-Marquardt algorithm to perform a single model fit to the complex ACFs (Ribeiro et al., 2013; Reimer et al., 2018) while a two-component complex ACF fit was utilised to separate ground and ionospheric components in mixed-scatter echoes (Ponomarenko et al., 2008). The scope of this paper is restricted to the FITACF algorithm, which has remained essentially unchanged since its inception in the late 1980s (Baker et al., 1988).

FITACF processing is performed in two stages consisting of (I) preliminary data 'cleaning' (removal of 'bad' data) and (II) extracting physical information from the remaining 'good' data. In the original version of FITACF, the first step in the preliminary data cleaning is to remove ACFs with low signal-to-noise ratio (SNR). After that, several categories of 'bad' ACF lags are identified and removed from the analysis. They consist of (i) lags affected by blanking the samples coinciding with the pulses' emission times (Tx-overlap); (ii) lags whose level of cross-range interference (CRI, or self-clutter, when returns from different pulses in the sequence arrive to the receiver from different ranges at the same time) exceeds a certain threshold; (iii) lags whose

Corresponding author: Pavlo Ponomarenko, pasha.ponomarenko@usask.ca

power is 'buried' in the statistical fluctuations, and (iv) lags that do not conform to the single-component model ('bad' ACF power shape). The last criterion arises from the main physical assumption underlying the original FITACF algorithms: that any ACF is characterised by a single spectral maximum (a single average velocity) which implies a linear ACF phase change with time lag, τ, and decreasing ACF power with increasing τ.

At the second stage, the phase and power of a 'cleaned' ACF are fitted with model functions in order to estimate line-of-sight (LoS) Doppler velocity, spectral width, and SNR of the returns together with the related errors. In addition, similar fitting is applied to cross-covariance functions (XCF) between signals received by the main and auxiliary (interferometer) arrays in order to estimate the vertical angle of arrival (elevation angle) of the radar returns. A block diagram and a brief description of FITACF functionality are provided in Appendix A of (Ponomarenko & Waters, 2006). The obtained LoS velocity and spectral width estimates are used for distinguishing between the ionospheric scatter (IS) and ground scatter (GS) returns based on empirically derived criteria (e.g. introduction of Blanchard et al., 2009).

By the early 2000s it became apparent that FITACF estimates of spectral width for the IS component were too large to be attributed to any feasible physical mechanism. A detailed analysis of the FITACF code revealed an inaccurate treatment of the statistical fluctuation level that led to overestimating the spectral width. Furthermore, it was established that the CRI rejection threshold has been too liberal and had to be supplemented by additional procedures for removing 'bad' ACF lags. The necessary code corrections were proposed by Ponomarenko and Waters (2006) and eventually implemented in a revised package called FITACF2 (2006–2007).

Yet another revision of the FITACF algorithms was necessitated by the fact that the SuperDARN plasma drift velocities were generally lower and sometimes even showed an opposite sign as compared to those measured at the same location by other ground-based or in-situ instruments (e.g. Xu et al., 2001; Drayton et al., 2005). Another problem was that FITACF2 was producing unrealistic values of velocity fitting errors, which prevented accurate comparisons of SuperDARN velocity measurements with those from other instruments. The issue of underestimated SuperDARN velocities was resolved relatively quickly by taking into account the significant deviation of ionospheric refractive index from unity in the HF frequency range (Ponomarenko et al., 2009; Gillies et al., 2009, 2010). However, the velocity error problem required a much more extensive analysis of the FITACF code. By 2013 it was clear that there were several inadequacies with the implementation of the least-square fitting algorithms. These were related mainly to the intuitive rather than textbook treatment of backscatter signal statistics. These issues were reported in detail at the 2013 *SuperDARN Workshop* held in Moose Jaw, Canada, and stimulated creation of SuperDARN Data Analysis Working Group (DA-WG) charged with maintaining the RST software. By 2015 the next major revision of the FITACF software, FITACF3, had been developed within the framework of the DA-WG. As the data processing algorithms had undergone detailed revision and considerable modification, it became necessary to produce a comprehensive description of the updated package, including the differences between the previous versions and the statistical justification of the implemented changes.

2 Data acquisition and processing

2.1 'Raw' data

A typical SuperDARN radar collects information by scanning consecutively through sixteen azimuthal directions (beams) separated by $\simeq 3.5^\circ$ (total coverage of $\simeq 55^\circ$). Each beam direction is usually sampled in group range from 180 km to $\simeq$3500–4000 km

at 45-km steps forming 70–110 range cells (range gates). The integration time is 3.5–7 s/beam so that the full field of view is scanned within one or two minutes. In order to fulfill requirements to the maximum measured Doppler frequency shift and maximum group range, which are mutually exclusive in a mono-pulse emission mode, the SuperDARN radars every 100 ms emit sequences of 7 or 8 pulses in order to form a complex autocovariance function (ACF) $R(\tau)$ for each range-beam cell (e.g., Villain et al., 1996; Ponomarenko & Waters, 2006). The pulses in a sequence are unevenly separated in time in a way that allows to generate a nearly continuous set of unique time lags separated by a basic step of $\Delta\tau$=1.5 or 2.4 ms with the maximum time lag of τ_{max} ≃35-40 ms.

The SuperDARN echoes represent narrow-band signals, meaning that the effective frequency span is much smaller than the wave frequency f. In this case the signal $s(t)$ can be accurately approximated as a product of a slowly-changing envelope $E(t)$ and a harmonic oscillation at the carrier frequency ω:

$$s(t) \simeq E(t)e^{-j(\omega t+\phi_0)} \tag{1}$$

where $j = \sqrt{-1}$ and ϕ_0 is an initial phase at t=0. The analytic (complex) signal is built as

$$\bar{s}(t) = s(t) - jH\left(s\left(t\right)\right) \tag{2}$$

where $H(s(t))$ represents the Hilbert transform of the original signal. Real and imaginary parts of the ACF at each lag are obtained from averaging a cross-product of responses from a given pair of the pulses in a sequence (i.e., fixed time lag τ) over a number of pulse sequences, N_{ave} ≃20–70, (Villain et al., 1996; Ponomarenko & Waters, 2006),

$$R\left(\tau\right) = \langle \bar{s}(t)\bar{s}^*(t-\tau)\rangle = \langle E(t)E(t-\tau)\rangle\, e^{-j\omega\tau} = |R(\tau)|e^{j\phi(\tau)} \tag{3}$$

where * depicts the complex conjugate.

The ACF power (envelope) and phase are defined from the real and imaginary parts as

$$|R\left(\tau\right)| = \sqrt{\Re^2\left(\tau\right) + \Im^2\left(\tau\right)} \tag{4}$$

and

$$\phi\left(\tau\right) = \tan^{-1}\Im\left(\tau\right)/\Re\left(\tau\right) \tag{5}$$

respectively, where

$$R\left(\tau\right) = \Re\left(\tau\right) + j\Im\left(\tau\right) = |R\left(\tau\right)|e^{j\phi(\tau)} \tag{6}$$

As the radar returns represent narrow-band signals, their phase can be accurately approximated by a single harmonic oscillation with frequency ω, $s(t) = Ee^{j(\omega t+\phi_0)}$. In this case the ACF phase becomes a linear function of frequency

$$\phi\left(\tau\right) = \omega\tau. \tag{7}$$

Most of the SuperDARN radars are equipped with an auxiliary interferometer antenna array consisting of only four elements and located approximately 100 m in front of or behind the main 16-element array. The auxiliary array is used for measuring the elevation angle of the received signal. In order to achieve this, signal samples from two arrays are combined into cross-covariance functions $R_\times(\tau)$

$$R_\times\left(\tau\right) = \Re_\times\left(\tau\right) + j\Im_\times\left(\tau\right) = |R_\times\left(\tau\right)|e^{j\phi_\times(\tau)} \tag{8}$$

In contrast to $\phi\left(\tau\right)$, the cross-phase $\phi_\times\left(\tau\right)$ has an additional term (offset) ϕ_0 due to the spatial separation of the two antenna arrays

$$\phi_\times\left(\tau\right) = \omega\tau + \phi_0. \tag{9}$$

2.2 ACF models and estimated signal parameters

The main assumption made in FITACF about the radar returns is that they are generated by a single source (scatter volume) filled by plasma irregularities with spatially uniform characteristics, i.e., possessing the same statistical properties and drifting with the same average velocity.

The principal estimated parameters are

- signal-to-noise ratio, Q [dB]
- spectral width, W, [m/s]
- line-of-sight Doppler velocity, V_{LoS} [m/s]
- elevation angle, ϵ [deg]

2.2.1 Lag 0 power and signal-to-noise ratio

The signal-to-noise ratio is calculated from the lag 0 ACF power in decibels as

$$Q = 10 \log_{10} \frac{R(0) - R_n}{R_n} = 10 \log_{10} \left(\frac{R(0)}{R_n} - 1 \right), \tag{10}$$

where R_n is the noise power estimate calculated for each integration time (beam) from the ten lowest lag 0 ACF power values.

2.2.2 ACF power decay and spectral width

ACF power is assumed to decay with time lag due to either a finite lifetime of the irregularities or their random motion described by either an exponential or Gaussian decay model (Hanuise et al., 1993; Villain et al., 1996; Ponomarenko & Waters, 2006; Ponomarenko et al., 2007). The spectral width W is expressed in the velocity units (m/s) and corresponds to the half-power level of the model frequency spectrum $S(\omega)$. Through the Wiener-Khinchin theorem, it is inversely proportional to the decay time, but the exact relationship between these two parameters depends on the functional form of the power decay model [*FITACF tutorial* by K. Baker]:

- Exponential decay (Lorentzian spectrum)

$$|R(\tau)| = |R(0)| \, e^{-f_l \tau} \tag{11}$$

$$|S(\omega)| = \frac{f_l}{f_l^2 + (\omega_0 - \omega)^2} \tag{12}$$

$$W_l = \frac{1}{k} f_l \tag{13}$$

- Gaussian decay (Gaussian spectrum)

$$|R(\tau)| = |R(0)| \, e^{-f_g^2 \tau^2} \tag{14}$$

$$|S(\omega)| = \frac{\sqrt{\pi}}{f_g} \exp\left(-\frac{(\omega_0 - \omega)^2}{4 f_g^2} \right) \tag{15}$$

$$W_g = \frac{2\sqrt{\ln 2}}{k} f_g \tag{16}$$

Here, f_l and f_g are measured in Hz and represent the inverse of decorrelation times at the e-fold level for the exponential and the Gaussian models, respectively. $k = 2\pi/\lambda$ is radar wavenumber in vacuum where λ is the signal wavelength.

2.2.3 ACF phase and Doppler velocity

The signal frequency can be expressed as a combination of the emission frequency ω_0 and the Doppler frequency shift ω_D caused by dynamic processes in the ionosphere which produce variations in the phase path of the radar echoes (e.g., Ponomarenko et al., 2009; Scoular et al., 2013). Because for SuperDARN radars the signal frequency is downshifted to ω_0, the ACF phase slope is determined solely by the Doppler shift,

$$\phi(\tau) = \omega_D \tau. \tag{17}$$

In the SuperDARN data processing algorithms it is assumed that the radar echoes are produced by backscatter from electron density irregularities. In this case the Doppler shift is proportional to the line-of-sight component of the plasma drift velocity, V_{LoS}, and the signal wavenumber, k:

$$\omega_D = 2kV_{LoS} \tag{18}$$

so that

$$V_{LoS} = \frac{\omega_D}{2k} = \frac{\partial\phi}{\partial\tau}\frac{1}{2k}, \tag{19}$$

where $\frac{\partial\phi}{\partial\tau}$ is the ACF phase slope. By convention, a positive Doppler shift (phase slope) corresponds to plasma movement towards the radar, and a negative Doppler shift corresponds to plasma movement away from the radar.

2.2.4 Interferometer phase delay and elevation

The auxiliary (interferometer) antenna array makes it possible to estimate the vertical angle of arrival (elevation angle), ϵ, measured from the horizontal plane. The elevation angle is calculated from the phase offset between the signals received by the two antennas, $\Delta\phi$ (e.g., Milan et al., 1997):

$$\epsilon = \sin^{-1}\sqrt{\cos^2\Psi - \frac{\phi_0^2}{|k|^2 d^2}}, \tag{20}$$

where Ψ is the azimuth measured from the boresight direction at zero elevation ($\epsilon = 0$), and d is the interferometer base. These measurements are subject to the phase aliasing arising due to the interferometer base being larger than the radar wavelength, $d > \lambda$. As a result, additional assumptions have to be made in order to resolve this ambiguity (Milan et al., 1997).

2.3 Fitting model functions to observed data

The echo parameter estimates are all obtained through fitting model functions to $\phi(\tau)$, $\phi_\times(\tau)$ and $\ln|R(\tau)|$. Fitting is performed using the least square approach in which statistical uncertainty is present only in the dependent variable (phase, ACF power) and the independent variable (time lag) is known exactly (see Appendix for more detail).

2.3.1 Phase fitting

Phase fitting is based on the assumption of the linear phase dependence on the time lag (17), which corresponds to a uniform velocity field inside the scatter volume (single velocity component). Application of (A6) and (A7) to (17) provides the following expressions for the Doppler velocity estimate:

$$V_{LoS} = \frac{1}{2k}\frac{\sum_{i=1}^{N}\frac{\tau_i\phi_i}{\sigma_{\phi i}^2}}{\sum_{i=1}^{N}\frac{\tau_i^2}{\sigma_{\phi i}^2}} \tag{21}$$

and the respective errors in the phase slope and Doppler velocity are:

$$\delta_\omega = \frac{1}{\sum_{i=1}^{N} \frac{\tau_i^2}{\sigma_{\phi i}^2}} \longrightarrow \delta_V = \frac{1}{2k}\delta_\omega = \frac{1}{2k}\frac{1}{\sum_{i=1}^{N} \frac{\tau_i^2}{\sigma_{\phi i}^2}} \tag{22}$$

2.3.2 Cross-phase fitting

Fitting the cross-phase $\phi_\times(\tau)$ allows determination of the phase offset between the two arrays, $\phi_{0\times}$, which is later used for estimating the elevation angle of the echoes. While $\phi_{0\times}$ can be estimated directly from the $\phi_\times(\tau)$ at $\tau = 0$, utilising more XCF lags should provide a more accurate estimate. In contrast to fitting the ACF phase, the fitted linear model now contains an offset,

$$\phi_\times(\tau) = \phi_0 + \omega_D \tau.$$

Applying (A2)–(A5), we obtain the offset estimate and its error as

$$\phi_0 = \frac{S_{\tau\tau}^{\phi_\times} S_{\phi_\times} - S_\tau^{\phi_\times} S_{\tau\phi_\times}}{\Delta^{\phi_\times}}$$

and

$$\delta_{\phi 0} = \sqrt{\frac{S_{\tau\tau}^{\phi_\times}}{\Delta^{\phi_\times}}},$$

where

$$S^{\phi_\times} = \sum_{i=1}^{N} \frac{1}{\left(\sigma_i^{\phi_\times}\right)^2}, S_\tau^{\phi_\times} = \sum_{i=1}^{N} \frac{\tau_i}{\left(\sigma_i^{\phi_\times}\right)^2}, S_{\phi_\times} = \sum_{i=1}^{N} \frac{\phi_{\times i}}{\left(\sigma_i^{\phi_\times}\right)^2},$$

$$S_{\tau\tau}^{\phi_\times} = \sum_{i=1}^{N} \frac{\tau_i^2}{\left(\sigma_i^{\phi_\times}\right)^2}, S_{\tau Q} = \sum_{i=1}^{N} \frac{\tau_i \phi_{\times i}}{\left(\sigma_i^{\phi_\times}\right)^2}, \Delta^{\phi_\times} \equiv S^{\phi_\times} S_{\tau\tau}^{\phi_\times} - \left(S_\tau^{\phi_\times}\right)^2.$$

Elevation error dependence on the cross-phase fitting error is obtained from (20) by propagation of error:

$$\delta_\epsilon = \frac{\partial \epsilon}{\partial \phi}\delta_{\phi_0} = \frac{2\left(\cos^2\epsilon - \sin^2\Psi\right)^{\frac{1}{2}}}{kd\sin 2\epsilon}\delta_{\phi_0} \tag{23}$$

2.3.3 Power fitting

The lag 0 SNR, Q_0, and spectral width are estimated simultaneously by fitting either a linear or quadratic model to $L(\tau) = \ln(R(\tau))$:

- Exponential decay (linear $L(\tau)$ model)

$$L(\tau) = L_0 - f_l \tau \tag{24}$$

- Gaussian decay (quadratic $L(\tau)$ model)

$$L(\tau) = L_0 - f_g^2 \tau^2 \tag{25}$$

Application of (A2)–(A5) provides the following estimates for the offset and the slope of the model with exponential decay

$$L_0 = \frac{S_{\tau\tau}^L S_L - S_\tau^L S_{\tau L}}{\Delta^L} \tag{26}$$

$$\delta f_l = \frac{S^L S_{\tau L}^L - S_\tau^L S_L}{\Delta^L} \tag{27}$$

where

$$S^L = \sum_{i=1}^{N} \frac{1}{\left(\sigma_i^L\right)^2}, S_\tau^L = \sum_{i=1}^{N} \frac{\tau_i}{\left(\sigma_i^L\right)^2}, S_L = \sum_{i=1}^{N} \frac{L_i}{\left(\sigma_i^L\right)^2},$$
$$S_{\tau\tau}^L = \sum_{i=1}^{N} \frac{\tau_i^2}{\left(\sigma_i^L\right)^2}, S_{\tau L} = \sum_{i=1}^{N} \frac{\tau_i L_i}{\left(\sigma_i^L\right)^2}, \Delta^L \equiv S^L S_{\tau\tau}^L - \left(S_\tau^L\right)^2.$$

The respective fitting errors are

$$\delta_{L_0 l} = \sqrt{\frac{S_{\tau\tau}^L}{\Delta^L}} = \left[S_L - \left(S_\tau^L\right)^2 / S_{\tau\tau}^L\right]^{-1/2} \tag{28}$$

$$\delta_{fl} = \sqrt{\frac{S^L}{\Delta^L}} = \left[S_{\tau\tau}^L - \left(S_\tau^L\right)^2 / S_L\right]^{-1/2} \tag{29}$$

The expressions for the Gaussian model are identical except that τ_i in the above sums should be replaced with τ_i^2.

The lag 0 SNR estimate and its error are obtained by converting ln to $\log_{10}$. The conversion factors are the same for the Lorentzian and Gaussian models:

$$Q = \frac{L_0}{\ln 10} \tag{30}$$

$$\delta_Q = \frac{\delta_{L_0}}{\ln 10} \tag{31}$$

Spectral width errors are conventionally calculated in m/s using conversion factors between $f_{l,g}$ and $W_{l,g}$ from (13) and (16):

- Lorentzian

$$\delta_{Wl} = \frac{1}{k}\delta_{fl} \tag{32}$$

- Gaussian
  For Gaussian decay the fitting error characterises the square of the effective spectral width δf_g^2. Conversion to the actual spectral width error can be done through propagation of errors:

$$\delta_{fg} = \frac{\partial(f_g)}{\partial(f_g^2)}\delta_{f^2 g} = \frac{1}{2f_g}\delta_{f^2 g}$$

  The spectral width fitting error then is

$$\delta_{Wg} = \frac{2\sqrt{\ln 2}}{k}\delta_{fg} = \sqrt{\ln 2}\,\frac{\delta_{f^2 g}}{f_g k} \tag{33}$$

2.4 Statistical variations

In this subsection we will determine standard deviation (SD) values for ACF power and phase, σ_R and σ_ϕ, required for weighting the least-square sums. Statistical variations in ionospheric radar echo parameters are produced by a number of sources:

- radio wave scatter by random irregularities of ionospheric plasma
- internal electronic circuitry noise
- atmospheric noise
- radar self-clutter (multi-pulse sounding modes)

In our case the primary contributors are the signal's own statistical variability and the self-clutter (cross-range interference).

2.4.1 Power

To determine the level of statistical variability at each ACF lag, it is convenient to introduce a normalised ACF,

$$r(\tau) = \frac{R(\tau)}{R(0)}. \tag{34}$$

The exact expression for the relative statistical errors (standard deviation divided by the mean) in the ACF power is

$$\frac{\sigma_r(\tau)}{r(\tau)} = \sqrt{\frac{1 + r^2(\tau)}{2Nr^2(\tau)}}, \tag{35}$$

where τ is the time lag and N is the number of degrees of freedom (independent realisations of the random process) (Bendat & Piersol, 2010, eq. 8.109). Importantly, statistical errors in ACF power depend on the autocorrelation coefficient magnitude. We are interested in the absolute error, σ_r, so the respective expression is

$$\sigma_r(\tau) = \sqrt{\frac{1 + r^2(\tau)}{2N}} \tag{36}$$

Furthermore, for the ACF power which is the actual measured parameter, one would need to multiply the above expression by the lag 0 power, $R(0)$:

$$\sigma_R(\tau) = \sqrt{\frac{1 + r^2(\tau)}{2N}} R(0). \tag{37}$$

It is easy to see that σ_R changes between $R(0)/\sqrt{N}$ for $r = 1$ and $R(0)/\sqrt{2N}$ for $r = 0$ (top panel in Fig. 1).

Because FITACF fitting is applied to the natural logarithm of power, the log-power SD required for weighting can be obtained from linear power as

$$\sigma^L(\tau) = \frac{\partial L}{\partial R}\sigma_R(\tau) = \frac{\sigma_R(\tau)}{R(\tau)}$$

2.4.2 Phase

The phase variability is fundamentally different from that of the power by the virtue of being an arbitrary parameter. As a result, the phase SD value is independent of the phase value itself. While the phase SD is also determined by the normalised ACF power (i.e., correlation coefficient) and the number of averages, the functional dependence on these parameters is qualitatively very different (Bendat & Piersol, 2010, eq. 9.52):

$$\sigma_\phi(\tau) = \sqrt{\frac{1 - r^2(\tau)}{2Nr^2(\tau)}} = \sqrt{\frac{r^{-2}(\tau) - 1}{2N}} \text{ [rad]}, \tag{38}$$

In contrast to σ_r, the phase fluctuation magnitude ranges from 0 at $r = 1$ to infinity at $r = 0$ (bottom panel in Fig. 1).

Importantly, the original FITACF software utilises the same values of SD

$$\sigma(\tau) \propto r(\tau) \tag{39}$$

for weighting both power and phase fits.

2.5 Effects of cross-range interference

Cross-range interference (self-clutter) occurs when echoes from different pulses come to the receiver at the same time. Self-clutter acts as an external noise which increases

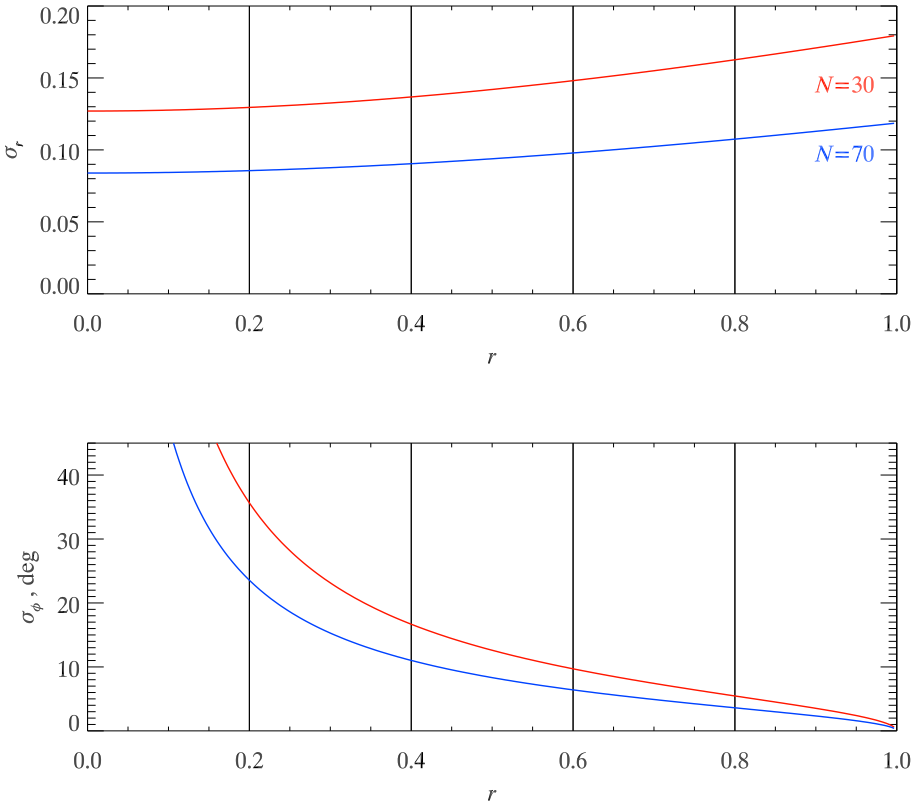

Figure 1. Dependence of standard deviations for power (σ_r) and phase (σ_{phi}) on correlation coefficient r for typical values of the number of averages N.

the magnitude of statistical fluctuations. Its effects on the power SD are described by the following expressions from Bendat and Piersol (2010):

$$\sigma_r(\tau) = \sqrt{\frac{1 + \alpha^{-2} r_S^{-2}(\tau)}{2N}}\, r_S(\tau). \tag{40}$$

Here α is a power correction factor,

$$\alpha^2 = \frac{P_S^2}{(P_S + P_{Nm})(P_S + P_{Nn})}, \tag{41}$$

where $P_S = R_S(0)$ is the power of the signal in the analysed range gate and $P_{Nm,n}$ are CRI power values affecting m^{th} and n^{th} pulses in the sequence, respectively.

The same correction can be applied to $r_S(\tau)$ in the expression for phase SD:

$$\sigma_\phi(\tau) = \sqrt{\frac{\alpha^{-2} r_S^{-2}(\tau) - 1}{2N}}\ [\text{rad}]. \tag{42}$$

3 Difference between FITACF1–2 and FITACF3 procedures

In FITACF3 the least square fitting technique is implemented in a textbook (optimal) way with respect to calculating phase and power SD, which is used for weighting in least-square sums while estimating fitted parameters. In contrast, in the original FITACF the fitting errors are obtained based on intuitive or predominantly qualitative

considerations rather than on accurate analytic treatment, and the same applies to data pre-selection procedures and background noise estimates.

- Weighting in least square fitting:
  - FITACF1–2: At any given lag, the same weighting coefficients are used for both power and phase. They are proportional to the square of ACF power at this lag.
  - FITACF3: At each lag, phase and power SD values are calculated separately based on theoretical formulas, which also account for CRI level so that different sets of weights are applied while fitting power and phase.
- Fitting errors
  - FITACF1–2: Due to the way the weighting was implemented, the fitting errors were estimated empirically based on the least square ACF power and phase residuals from the 'good' lags.
  - FITACF3: The fitting errors are calculated using the textbook formulas for least-square fitting based on SD estimates and number of averaged pulse sequences.
- 'Bad' lags
  - FITACF1–2: There are several kinds of 'bad' lags, most of which are determined using empiric criteria:
    - Tx overlap (accurate)
    - Low ACF power level (empiric threshold)
    - High CRI level (empiric threshold)
    - 'Bad shape' (empiric criteria for non-monotonic power decay)
  - FITACF3: There are only two types of 'bad' lags removed
    - Tx overlap (applied to both power and phase)
    - Low-power lags (applied to ACF power only)

    This approach allows more lags to be utilised, especially in phase fitting, where only Tx-overlap lags are abandoned. The low power lags are removed from power fitting because they produce a bias towards lower spectral width and lower lag 0 power due to the non-symmetric distribution of the power variations around the mean value. These lags are still used in the phase fitting because the phase variations are distributed symmetrically around the mean value.
- Mitigating cross-range interference effects
  - FITACF1–2: Lags with CRI level exceeding an empirically-determined threshold are completely excluded from the power and phase fitting.
  - FITACF3: A conventional (textbook) approach is applied: CRI effects are included in the power and phase SD estimates so that lags with larger interference levels make a proportionally smaller contribution to the least-square sums.
- Phase unwrapping
  - FITACF1–2: Conventional 2π phase unwrapping.
  - FITACF3: Phase deviations from the fitted linear models are compared for unwrapped and wrapped data in order to mitigate the 'Type 1 Echoes' artifact observed in low-power ground scatter echoes (presentation at SD2017).
- Noise determination

  - FITACF1–2: The background noise level for each integration time is determined as an average from 10 lowest lag 0 power values across all available range gates.
  - FITACF3: The original method leads to underestimation of the actual noise floor by 1-2 dB. In order to mitigate this, a correction techniques is implemented. It is based on expected statistical distribution of noise power which is defined by the number of averaged pulse sequences.

- Elevation angle estimate
  - FITACF1–2: Elevation angle is estimated as an intersect of a linear fit to the cross-phase. Fitting errors represented as lower and higher elevation limits based on the observed phase deviations from the fitted linear model (ELV_LOW and ELV_HIGH parameters in FIT structure)
  - FITACF3: Elevation angle is estimated from both (i) the linear fit to the cross-phase (used for estimating elevation angle errors) and (ii) a measured cross-phase offset at lag 0 of the XCF (elevation angle estimates). The latter provides significantly lower variability at the range gates affected by TX overlap as compared to the fitted values. A single value of the fitting error is estimated based on the textbook formulas.

- XCF power fitting
  - FITACF1–2: XCF power is fitted in the same way as that of the ACF, and the obtained estimates of lag 0 SNR and spectral width are stored in the standard output structure
  - FITACF3: No XCF power fitting is performed because the resulting estimates represent lower-quality replicas of the ACF power fits due to intrinsically lower SNR.

- Lag zero power data type
  - FITACF1–2: Raw lag zero power values are read by the software as integer data type, resulting in the rounding of lag zero power values down to the nearest integer.
  - FITACF3: Raw lag zero power values are read into the software as double precision floating point values.

4 Short description of FITACF3 procedures

1. Data preparation
   - Reading complex ACFs, XCFs and parameter structures from RAWACF files
   - Calculating ACF power and phase
   - Calculating the background noise level

     The noise level is estimated as the mean of the 10 lowest lag 0 power values (same as FITACF1–2). A correction factor is then applied based on a Gaussian distribution of the noise and assuming that we use the 10 lowest values of the distribution (low-power 'tail') out of 75 or 100.
   - Calculating SNR in dB
   - Rejecting whole ACFs with low SNR (< 0 dB)
   - Rejecting ACF lags affected by the transmitter emission (Tx overlap, same as in the old FITACF)
   - Calculating integral cross-range interference power for each receiver sample (i.e. sum of lag 0 powers from all interfering ranges)

- Calculating parameter α (eq. 41) for each ACF lag. This parameter is used for estimating levels of power and phase SD at different ACF lags, which are in turn used for calculating weighting coefficients in the least-square sums.

2. ACF power fitting
   - Removing a low-power 'tail' from the ACF based on analysing lags with low CRI levels ($1/\alpha < 2$): first determining if there is a lag with ($P < 2\sigma_R$), and if there is one, this and all larger lags are removed.

     Note: this is not applied to phase fitting!
   - Calculating linear power SD for each ACF lag based on α and number of averages
   - Converting linear power SD into log power SD through propagation of errors
   - Fitting linear function to log power using linear SD for weighting as it provides more accurate estimates of SNR and spectral width.
   - Repeating the above fitting but with log power SD as weights for correct characterisation of SNR and spectral width fitting errors.
   - Fitting quadratic function to log power using linear power SD as weights
   - Repeating the quadratic fitting with log power SD as weights
   - Calculating lag 0 SNR and spectral width for both models, exponential and Gaussian, together with the respective fitting errors
3. ACF phase fitting
   - Initial estimate of the phase-vs-lag slope as a mean for all available pairs of lags with phase difference $< \pi$
   - Subtracting the linear phase model from the real data and calculating the number of 2π 'wraps' for each lag
   - Initial 'unwrapping' of the data by adding the previously calculated number of 'wraps'.
   - Calculating theoretical estimates of the phase SD based on the fitted exponential power (correlation coefficient), CRI parameter α and the number of averages. The fitted value for the correlation coefficient is used because the observed values can exceed lag 0 power due to enhanced statistical fluctuations (e.g., high level of CRI) which will lead to imaginary values of SD in Eq. 38.
   - Fitting a linear function without offset to the phase using inverted SD for weighting the phase data at different lags. In contrast to the power, here we use all available lags except those affected by the Tx emission
   - Using the obtained slope for a second round of 'unwrapping' just in case we missed a jump or two.
   - Repeating the fitting procedure for the re-adjusted data
   - Fitting original ('wrapped') data.
   - Comparing the residual deviations for the original and 'unwrapped' phase and selecting the one with smaller deviations for further processing. In this way we eliminate the bulk of the false 'Type 1' echoes (for more detail see our 'Type 1' presentation at SD'17)
   - Calculating the line-of-sight velocity values and respective fitting errors.
4. XCF phase fitting
   - Calculating the power and phase for the interferometer data (XCFs)
   - Unwrapping the XCF phase in the same way as for the velocity except that the phase model is now having an offset (lag 0 XCF phase as an initial guess).

- Fitting a linear function with an offset to the 'unwrapped' data using the same weights as for the velocity (although, we need to use the normalised XCF power here)
- Repeating the 'unwrapping' using the fitted slope and offset values
- Second round of fitting a linear function with an offset to the adjusted data
- Calculating the elevation angle from the fitted lag 0 phase (offset) for comparison and for estimating the related errors

5. Recording the obtained estimates into FITACF files.

Appendix A Least-Square Fitting (textbook formulas)

Least-Square fitting is performed through minimising a sum of squared differences between the observed, $y(x)$, and modelled, $f(x)$, values where each term is weighted by the respective SD σ_y for a given value of x,

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - f(x_i)}{\sigma_y i} \right)^2 . \tag{A1}$$

Following Section 15.2 of Press et al. (1996), for a straight line with an offset,

$$y = a + bx,$$

minimisation of the least-square sum provides following expressions for the fitted parameters and the respective fitting errors are:

- offset

$$a = \frac{S_{xx} S_y - S_x S_{xy}}{\Delta(x, \sigma_y)}, \tag{A2}$$

$$\delta_a = \sqrt{\frac{S_{xx}}{\Delta(x, \sigma_y)}}, \tag{A3}$$

- slope

$$b = \frac{S S_{xy} - S_x S_y}{\Delta(x, \sigma_y)}, \tag{A4}$$

$$\delta_b = \sqrt{\frac{S}{\Delta(x, \sigma_y)}}, \tag{A5}$$

where

$$S = \sum_{i=1}^{N} \frac{1}{\sigma_{yi}^2}, S_x = \sum_{i=1}^{N} \frac{x_i}{\sigma_{yi}^2}, S_y = \sum_{i=1}^{N} \frac{y_i}{\sigma_{yi}^2},$$

$$S_{xx} = \sum_{i=1}^{N} \frac{x_i^2}{\sigma_{yi}^2}, S_{xy} = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_{yi}^2}, \Delta(x, \sigma_y) \equiv S S_{xx} - S_x^2.$$

For the simplest case of a straight line without an offset ($a = 0$),

$$y = bx,$$

one can easily find that

$$b = \frac{S_{xy}}{S_{xx}}, \tag{A6}$$

$$\delta_b = \frac{1}{\sqrt{S_{xx}}}. \tag{A7}$$

The above expressions are also applicable to the quadratic function with the offset,

$$y = a + bx^2,$$

one just needs to replace x_i with x_i^2.

References

Baker, K. B., Greenwald, R. A., Villain, J.-P., & Wing, S. (1988). *Spectral characteristics of high frequency (HF) backscatter for high latitude ionospheric irregularities: Preliminary analysis of statistical properties* (Tech. Rep.). John Hopkins University.

Barthes, L., André, R., Cerisier, J.-C., & Villain, J.-P. (1998). Separation of multiple echoes using a high-resolution spectral analysis for SuperDARN HF radars. *Radio Sci.*, *33*, 1005-1017. doi: 10.1029/98RS00714

Bendat, J. S., & Piersol, A. G. (2010). *Random data: Analysis and measurement procedures* (4th ed.). John Wiley & Sons, Inc.

Blanchard, G. T., Sundeen, S., & Baker, K. B. (2009). Probabilistic identification of high-frequency radar backscatter from the ground and ionosphere based on spectral characteristics. *Radio Sci.*, *44*, RS5012. doi: 10.1029/2009RS004141

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., ... Walker, A. D. M. (2007). A decade of the Super Dual Auroral Radar Network (SuperDARN): Scientific achievements, new techniques and future directions. *Surv. Geophys.*, *28*, 33–109. doi: 10.1007/s10712-007-9017-8

Drayton, R. A., Koustov, A. V., Hairston, M. R., & Villain, J.-P. (2005). Comparison of dmsp cross-track ion drifts and superdarn line-of-sight velocities. *Ann. Geophys.*, *23*, 2479–2486, SRef-ID: 1432-0576/ag/2005-23-2479. doi: 10.5194/angeo-23-2479-200

Gillies, R. G., Hussey, G. C., Sofko, G. J., McWilliams, K. A., Fiori, R. A. D., Ponomarenko, P., & St.-Maurice, J.-P. (2009). Improvement of SuperDARN velocity measurements by estimating the index of refraction in the scattering region using interferometry. *J. Geophys. Res.*, *114*, A07305. doi: 10.1029/2008JA013967

Gillies, R. G., Hussey, G. C., Sofko, G. J., Wright, D. M., & Davies, J. A. (2010). A comparison of EISCAT and SuperDARN F-region measurements with consideration of the refractive index in the scattering volume. *Journal of Geophysical Research: Space Physics*, *115*(A6). doi: https://doi.org/10.1029/2009JA014694

Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., ... Yamagishi, H. (1995). DARN/SuperDARN: A global view of the dynamics of high-latitude convection. *Space Sci. Rev.*, *71*, 761-795.

Greenwald, R. A., Oksavik, K., Barnes, R., Ruohoniemi, J. M., Baker, J., & Talaat, E. R. (2008). First radar measurements of ionospheric electric fields at sub-second temporal resolution. *Geophysical Research Letters*, *35*(3). doi: 10.1029/2007GL032164

Hanuise, C., Villain, J. P., Gresillon, D., Cabrit, B., Greenwald, R. A., & Baker, K. B. (1993). Interpretation of HF radar ionospheric Doppler spectra by collective wave scattering theory. *Ann. Geophys.*, *11*, 29–39.

Milan, S. E., Jones, T. B., Robinson, T. R., Thomas, E. C., & Yeoman, T. K. (1997). Interferometric evidence for the observation of ground backscatter originating behind the CUTLASS coherent HF radars. *Annales Geophysicae*, *15*(1), 29–39. doi: 10.1007/s00585-997-0029-y

Nishitani, N., Ruohoniemi, J. M., Lester, M., Baker, J. B. H., Koustov, A. V., Shepherd, S. G., ... others (2019). Review of the accomplishments of mid-latitude Super Dual Auroral Radar Network (SuperDARN) HF radars. *Progress in Earth and Planetary Science*, *6*(27). doi: 10.1186/s40645-019-0270-5

Ponomarenko, P. V., St-Maurice, J.-P., Waters, C. L., Gillies, R. G., & Koustov, A. V. (2009). Refractive index effects on the scatter volume location and doppler velocity estimates of ionospheric HF backscatter echoes. *Ann. Geophys.*, *27*(11), 4207–4219. doi: 10.5194/angeo-27-4207-2009

Ponomarenko, P. V., & Waters, C. L. (2006). Spectral width of superdarn echoes: Measurement, use and physical interpretation. *Ann. Geophys.*, *24*, 115–128. doi: 10.5194/angeo-25-675-2007

Ponomarenko, P. V., Waters, C. L., & Menk, F. W. (2007). Factors determining spectral width of HF echoes from high latitudes. *Ann. Geophys.*, *25*, 675-687. doi: 10.5194/angeo-25-675-2007

Ponomarenko, P. V., Waters, C. L., & Menk, F. W. (2008). Effects of mixed scatter on SuperDARN convection maps. *Annales Geophysicae*, *26*, 1517–1523. doi: 10.5194/angeo-26-1517-2008

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996). *Numerical recipes in Fortran 90 (2nd ed.): The art of parallel scientific computing.* New York, NY, USA: Cambridge University Press.

Reimer, A. S., Hussey, G. C., & McWilliams, K. A. (2018). Statistically self-consistent and accurate errors for SuperDARN data. *Radio Science*, *53*(1), 93–111. doi: 10.1002/2017RS006450

Ribeiro, A. J., Ruohoniemi, J. M., Ponomarenko, P. V., Clausen, L. B. N., Baker, J. B. H., Greenwald, R. A., ... de Larquier, S. (2013). A comparison of SuperDARN ACF fitting methods. *Radio Science*, *48*(3), 274–282. doi: 10.1002/rds.20031

Scoular, G., Ponomarenko, P. V., & St.-Maurice, J. (2013). A new type of Doppler velocity fluctuations in HF ground scatter from the polar cap. *Geophysical Research Letters*, *40*(19), 4992-4997. doi: 10.1002/grl.50960

SuperDARN Data Analysis Working Group. (2021). *SuperDARN Radar Software Toolkit (version 4.5).* (Participating members: Thomas, E.G., Schmidt, M.T., Bland, E.C., Burrell, A.G., Ponomarenko, P.V., Reimer, A.S., Sterne, K.T. and Walach, M.-T.) doi: 10.5281/zenodo.4435297

Villain, J.-P., André, R., Hanuise, C., & Gresillon, D. (1996). Observation of the high-latitude ionosphere by HF radars: Interpretation in terms of collective wave scattering and characterization of the turbulence. *J. Atmos. Sol. Terr. Phys.*, *58*, 943-958. doi: 10.1016/0021-9169(95)00125-5

Xu, L., Koustov, A. V., Thayer, J., & McCready, M. A. (2001). SuperDARN convection and Sondrestrom plasma drift. *Annales Geophysicae*, *19*(7), 749–759. doi: 10.5194/angeo-19-749-2001