REGISTRATION-BASED SOLUTIONS TO THE RANGE-DEPENDENCE PROBLEM IN STAP RADARS

Fabian D. Lapierre (Research Fellow) and Jacques G. Verly

University of Liège, Department of Electrical Engineering and Computer Science
Sart-Tilman, Building B28, B-4000 Liège, Belgium
F.Lapierre@ulg.ac.be, Jacques.Verly@ulg.ac.be

ABSTRACT

We address the problem of detecting slow-moving targets using a space-time adaptive processing (STAP) radar. The construction of optimum weights at each range implies the estimation of the clutter covariance matrix. This is typically done by straight averaging of neighboring data snapshots. The range-dependence of these snapshots generally results in poor performance. After reviewing existing methods for handling the range-dependence, we present new methods exploiting the geometry of the direction-Doppler curves.

1. INTRODUCTION

Pulsed Doppler radars are used to detect moving targets and to measure both their ranges and speeds. They typically transmit a train of coherent modulated pulses. Equipped with a linear array antenna, they are particularly well suited at detecting slow-moving targets in the presence of clutter and jammers. One of the most sophisticated techniques for dealing with this problem is space-time adaptive processing (STAP) [1, 2].

STAP research was initially focused on monostatic (MS) configurations, where the transmitter and the receiver are colocated. However, much of the STAP challenge now resides in bistatic (BS) configurations [3, 4], where the transmitter and the receiver are located on distinct, independently-moving plateforms.

The data collected by STAP radars can be viewed as a 2D discrete sequence of space and time variables, called a "snapshot". Basic STAP methods compute a weighted linear combination of the snapshot elements. The calculation of the optimum weights generally involves the inversion of a covariance matrix of the data snapshot. This covariance matrix must be estimated. This is typically done by straight averaging (SA) of the snapshots at a series of neighboring ranges. However, this averaging works only for a limited number of configurations.

The range-dependence problem results from the fact that the clutter energy wanders in the power-spectral-density (PSD) domain as the range changes. We will see that this "range-walking" manifests itself by a deformation with range of an ubiquitous "clutter ridge". This results in a degradation of detection performance. The object of this paper is to compensate for this range-dependence of the clutter ridge with the goal of bringing performance as close as possible to its optimum level.

After reviewing exisiting range-dependence compensation methods, we present our new methods. The idea is to compute an estimate of the covariance matrix at some reference range gate (indexed with l) by first applying a transformation to the covariance matrices at a series of neighboring range gates (indexed with k) and then averaging the transformed matrices. However, the transformation is applied to the corresponding PSDs, i.e., in the spectral domain. The PSD corresponding to some range (gate) k is transformed to bring its clutter ridge into registration with that of the PSD at the reference range l. The methods proposed are thus based upon mappings of the elements of the PSDs at each range k.

The registration of the clutter ridges is equivalent to the registration of the "direction-Doppler (DD) curves," which are mathematical curves fully determined by the configuration parameters and the range.[1] We distinguish between two types of range-dependence compensation methods: (a) methods that assume exact knowledge of the configuration parameters are said to be "open-loop (OL)" and (b) methods that estimate the parameters from the data are said to be "data-adaptive (DA)". Below, we describe six of these methods and discuss their performances.

2. RADAR-MEASUREMENT CONFIGURATIONS

Figure 1 shows a typical MS configuration. The transmitter and the receiver are located on the same plateform $\mathcal{R}$. The scatterer S is either a target or a clutter patch. The origin of the (x, y, z) coordinate system is chosen to coincide with $\mathcal{R}$. Its orientation is such that the x-axis points in the same way as the radar velocity vector $\underline{\boldsymbol{v}}_{\mathcal{R}}$ and the z-axis points vertically up. We assume that the antenna A is linear and located in a horizontal plane. The orientation of A is then fully described by the angle δ between the antenna axis s and $\underline{\boldsymbol{v}}_{\mathcal{R}}$. The sidelooking (SL) MS configuration corresponds to $\delta = 0$. The MS range R_m is the one-way distance from R to S. Noting that $v_{\mathcal{R}} = |\underline{\boldsymbol{v}}_{\mathcal{R}}|$, any MS configuration is fully characterized by the vector of parameters

$$\underline{\boldsymbol{\theta}} = (\delta, v_{\mathcal{R}}, H). \qquad (1)$$

Figure 2 shows a typical BS configuration. The transmitter T and the receiver R are mounted on distinct, independent plateforms. The origin of the (x, y, z) coordinate system is chosen to coincide with T. Its orientation is such that the x-axis points in the same way as the *transmitter* velocity vector $\underline{\boldsymbol{v}}_T$ and that the z-axis points vertically up. We assume that the receiver velocity vector $\underline{\boldsymbol{v}}_R$ is located in a horizontal plane. The angle of $\underline{\boldsymbol{v}}_R$ with respect to $\underline{\boldsymbol{v}}_T$ is denoted by α_R. The antenna A is again assumed to be linear and in a horizontal plane. Defining δ as the angle between the s-axis and $\underline{\boldsymbol{v}}_R$, the angle $\delta + \alpha_R$ fully describes the orientation of A. The sidelooking (SL) BS configuration corresponds to

[1] We have developped a mathematical theory of DD curves but it is not discussed here.

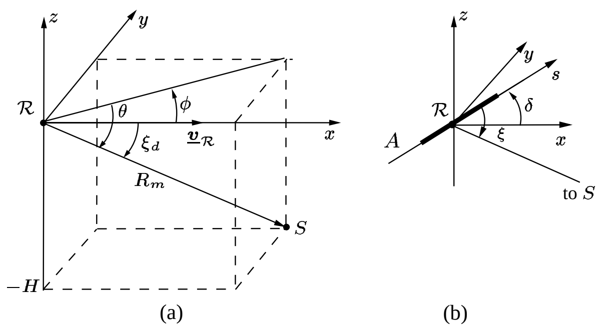


Figure 1: MS configuration. (a) Radar ($\mathcal{R}$)-scatterer (S) geometry. (b) Antenna (A) and related angles.

$\delta = 0$. The bistatic range R_b is the distance from T to S to R. Noting that $v_R = |\underline{\boldsymbol{v}}_R|$ and $v_T = |\underline{\boldsymbol{v}}_T|$, any BS configuration is fully characterized by the vector of parameters

$$\underline{\boldsymbol{\theta}} = (x_R, y_R, z_R, \alpha_R, \delta, v_R, v_T, H). \quad (2)$$

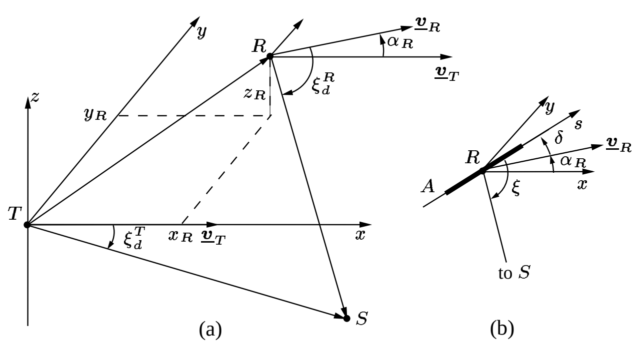


Figure 2: BS configuration. (a) Transmitter (T)-receiver (R)-scatterer (S) geometry. (b) Antenna (A) and related angles.

3. DIRECTION-DOPPLER CURVES

3.1. Radar and signal parameters

Radar systems are expected to determine at least three basic parameters for each scatterer S of interest: the angular position ξ, the range R_m or R_b and the relative velocity v_r. The related parameters that are more directly measured from the radar returns are the spatial frequency f_s [2], the roundtrip delay τ_{rt} and the Doppler frequency f_d. For a stationary scatterer, we have

$$\begin{aligned} f_s &= \lambda_c^{-1} \cos \xi \\ \tau_{rt} &= \begin{cases} 2R_m/c & \text{(MS)} \\ R_b/c & \text{(BS)} \end{cases} \\ f_d &= \begin{cases} \lambda_c^{-1} 2 v_R \cos \xi_d. & \text{(MS)} \\ \lambda_c^{-1} v_T \cos \xi_d^T + \lambda_c^{-1} v_R \cos \xi_d^R, & \text{(BS)} \end{cases} \end{aligned}$$

where λ_c is the carrier wavelength and c is the speed of the light. One can easily compute ξ, R_m or R_b, and v_r from f_s, τ_{rt} and f_d.

3.2. Direction-Doppler (DD) curves

For any given configuration and for any given range, all stationary scatterers at this range map onto a curve showing the relation between f_s and f_d for the scatterers. Any such curve is called a "direction-Doppler (DD)" curve. DD curves are typically represented in terms of the normalized spatial frequency ν_s equal to $(\lambda_c/2)f_s$ and the normalized Doppler frequency ν_d equal to $(\lambda_c/4v_{\mathcal{R}})f_d$ (MS) or to $(\lambda_c/2(v_R+v_T))f_d$ (BS). Figure 3 shows a variety of DD curves for different configurations and different ranges. Note that the DD curves vary significantly from one configuration to another and also from one range to another, especially for BS configurations. Figure 3(a) shows that, in the SL MS configuration, all DD curves are independent of range. In all other MS and in all BS configurations, these curves vary with range.

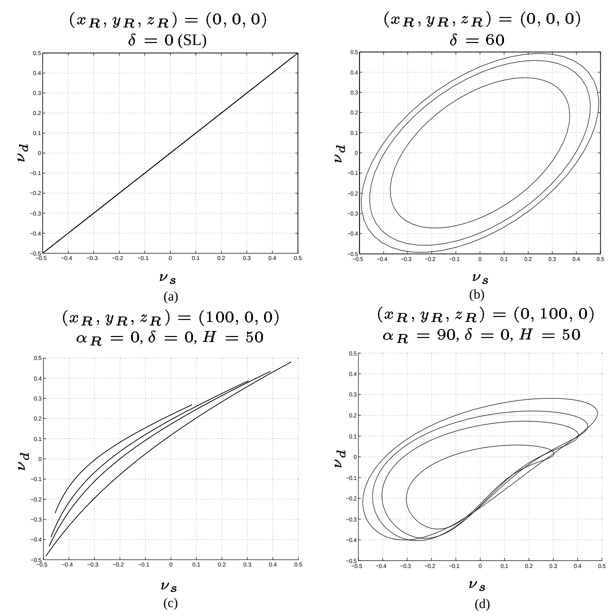


Figure 3: Example DD curves. (a)-(b): MS configurations. (c)-(d): BS configurations. MS ranges R_m are 75, 125 and 250 km and BS ranges R_b are 170, 210, 250 and 400 km. Units in text above graphs are degrees and kilometers.

Deriving the equations of these curves for MS configurations is straighforward [2, 1]. However, the case of BS configurations is much more challenging. Our approach is as follows. First, we express ν_d as a function of ν_s. Since most BS DD curves are closed curves, any BS DD curve is best described by two functions $\nu_d = f_1(\nu_s)$ and $\nu_d = f_2(\nu_s)$. One equation represents the upper part of the closed curve and the second the lower part. Second, if we express ν_s and ν_d as functions of the angle ψ parameterizing the isorange curve[2], we find a parametric description of the DD curve[3], i.e.,

$$\nu_s = g_1(\psi) \quad \text{and} \quad \nu_d = g_2(\psi). \quad (3)$$

[2] All scatterers S characterized by the same range are located on an isorange surface and its intersection with the ground is an isorange curve. If the ground is a horizontal plane, the isorange curve is a circle (MS) or an ellipse (BS), parameterizable with the polar angle ψ.

[3] The derivation of $f_1(\nu_s)$ and $f_2(\nu_s)$ or of $g_1(\psi)$ and $g_2(\psi)$ is quite complicated and long.

4. DATA SNAPSHOT AND OPTIMUM PROCESSOR

A train of M coherent pulses is transmitted, the returns are sensed at each of the N elements of a linear antenna array, and the sensed returns are sampled at a number of discrete ranges. We regard the data as a sequence of $M \times N$ data arrays at successive ranges. Each such array is called a "snapshot". The $M \times N$ snapshot corresponding to a single scatterer with normalized spatial frequency ν_s and normalized Doppler frequency ν_d and with range R_m or R_b can be written as the $MN \times 1$ vector [2]

$$\underline{\boldsymbol{y}}(\nu_s,\nu_d) = \beta_r \underline{\boldsymbol{v}}(\nu_s,\nu_d) = \beta_r \underline{\boldsymbol{b}}(\nu_d) \otimes \underline{\boldsymbol{a}}(\nu_s), \quad (4)$$

where β_r comes from the radar equation, $\underline{\boldsymbol{v}}(\nu_s,\nu_d)$ is the $MN \times 1$ steering vector, $\otimes$ the Kronecker product and $\underline{\boldsymbol{a}}(\nu_s)$ and $\underline{\boldsymbol{b}}(\nu_d)$ the $N \times 1$ spatial and $M \times 1$ temporal steering vectors

$$\begin{aligned} \underline{\boldsymbol{a}}(\nu_s) &= (1 \dots e^{j2\pi\nu_s n} \dots e^{j2\pi\nu_s(N-1)})^T & (5)\\ \underline{\boldsymbol{b}}(\nu_d) &= (1 \dots e^{j2\pi\nu_d m} \dots e^{j2\pi\nu_d(M-1)})^T. & (6) \end{aligned}$$

The clutter snapshot $\underline{\boldsymbol{y}}_c(\nu_s,\nu_d)$ is found by integrating $\underline{\boldsymbol{y}}(\nu_s,\nu_d)$ over the isorange curve parameterized by ψ, i.e.,

$$\underline{\boldsymbol{y}}_c(\nu_s,\nu_d) = \int_0^{2\pi} \beta_c(\psi)\, \underline{\boldsymbol{v}}(\nu_s(\psi),\nu_d(\psi))\, d\psi.$$

Since $\beta_c(\psi)$ is a random process, $\underline{\boldsymbol{y}}_c$ is a random vector. We assume it is stationary. It is thus characterized by a constant covariance matrix $\underline{\underline{\boldsymbol{R}}}_c = E\{\underline{\boldsymbol{y}}_c \underline{\boldsymbol{y}}_c^\dagger\}$. To find the power spectral density (PSD) associated with $\underline{\boldsymbol{y}}_c$, we use the minimum variance estimator (MVE) [1]. Clutter PSDs show a concentration of energy along a particular "curve" in the spectral plane. The support of this "clutter ridge" is in direct correspondence with the related DD curve.

The $MN \times 1$ weight vector providing optimum clutter rejection and thus defining the optimum processor (OP) is [1]

$$\underline{\boldsymbol{w}}_o(\nu_s,\nu_d) = \underline{\underline{\boldsymbol{R}}}^{-1} \underline{\boldsymbol{v}}(\nu_s,\nu_d), \quad (7)$$

where the covariance matrix $\underline{\underline{\boldsymbol{R}}} = E\{\underline{\boldsymbol{y}}\,\underline{\boldsymbol{y}}^\dagger\}$ is the sum of the covariance matrices $\underline{\underline{\boldsymbol{R}}}_c = E\{\underline{\boldsymbol{y}}_c \underline{\boldsymbol{y}}_c^\dagger\}$ for the clutter and $\underline{\underline{\boldsymbol{R}}}_n = E\{\underline{\boldsymbol{y}}_n \underline{\boldsymbol{y}}_n^\dagger\}$ for the noise. We assume that the noise $\underline{\boldsymbol{y}}_n$ is spatially and temporally white so that $\underline{\underline{\boldsymbol{R}}}_n = \underline{\underline{\boldsymbol{I}}}$. In practice, $\underline{\underline{\boldsymbol{R}}}$ is not known and must be estimated for each range. The maximum-likelihood estimator $\widehat{\underline{\underline{\boldsymbol{R}}}}$ for range gate l is [2]

$$\widehat{\underline{\underline{\boldsymbol{R}}}}(l) = \frac{1}{N_l} \sum_{k \in S_l} \underline{\underline{\boldsymbol{R}}}(k), \quad (8)$$

where N_l is the number of snapshots used for estimation, S_l is the set of surrounding snapshot indices k defined by $l - \frac{N_l-1}{2} < k < l + \frac{N_l-1}{2}$ and $\underline{\underline{\boldsymbol{R}}}(k) = \underline{\boldsymbol{y}}(k)\underline{\boldsymbol{y}}^\dagger(k)$ where $\underline{\boldsymbol{y}}(k)$ is the snapshot at range (gate) k. Equation (8) results in an unbiased estimator for $\underline{\underline{\boldsymbol{R}}}(l)$ only if the clutter ridge is range-independent. This happens only for SL MS configurations. In all other cases, range-dependence compensation methods must be used.

The performance of a processor using weights $\underline{\boldsymbol{w}}$ is measured by the signal-to-interference-plus-noise (SINR) loss defined as [2]

$$\mathrm{SINR_L} = \frac{\mathrm{SINR}}{\mathrm{SINR_0}} = \frac{\left|\underline{\boldsymbol{w}}^\dagger \underline{\boldsymbol{v}}\right|^2}{(\underline{\boldsymbol{w}}^\dagger \underline{\underline{\boldsymbol{R}}}\,\underline{\boldsymbol{w}})(\underline{\boldsymbol{v}}^\dagger \underline{\boldsymbol{v}})}, \quad (9)$$

where $\mathrm{SINR_0}$ is the SINR in the absence of clutter. Values of $\mathrm{SINR_L}$ range from a minimum equal to the noise-to-clutter ratio to a maximum of one, indicating that the processor performance is not degraded by clutter. Optimum performance is achieved with $\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_o$. In practice, processor performance is degraded by estimation losses and by the range-dependence of the clutter ridge.

5. EXISTING METHODS BASED ON DOPPLER WARPING AND TAYLOR SERIES

5.1. Doppler Warping (DW) method

The DW method was initially developped for nearly-SL MS radars [5] and later applied to BS configurations [4]. DW compensates the range-dependence of $\underline{\boldsymbol{y}}(\nu_s,\nu_d)$ in Eq. (4) by adding to ν_d a Doppler shift $\Delta(k)$ that is chosen for each range k in such a way as to bring all clutter ridges in registration at a specific ν_s. Equation (6) shows that this can be achieved by premultiplying $\underline{\boldsymbol{y}}(\nu_s,\nu_d)$ by the matrix

$$\underline{\underline{\boldsymbol{T}}}(k) = \left[1\, e^{j2\pi\Delta(k)} \,\dots\, e^{j2\pi(M-1)\Delta(k)}\right] \otimes \underline{\underline{\boldsymbol{I}}}.$$

Therefore, applying $\underline{\underline{\boldsymbol{T}}}(k)$ to the snapshot $\underline{\boldsymbol{y}}(k)$ results in the Doppler-warped snapshot $\underline{\boldsymbol{y}}_{dw}(k) = \underline{\underline{\boldsymbol{T}}}(k)\underline{\boldsymbol{y}}(k)$. Performance degrades as one moves away from the SL configuration in the MS case and is poor in the BS case [4]. The main advantage of the DW method is its simplicity of implementation. However, the configuration parameters must be known and the range-dependence compensation can be exact only for one specific ν_s.

5.2. High-Order Doppler Warping (HODW) method

The HODW method is a generalization of the DW method [6] and provides perfect compensation at more than a single ν_s. This method divides the Doppler frequency range into a finite number of Doppler bins. In each Doppler bin, a different range-dependent Doppler frequency shift $\Delta(k)$ is chosen and used in the same way as in the DW method. The main advantage of the HODW method is that the range-dependence compensation is nearly perfect. However, the configuration parameters must be known and the complexity of the Doppler filtering is significant.

5.3. Derivative-based updating (DBU) method

The DBU method was proposed in [4] to handle the range-dependence in BS configurations. The optimum weights $\underline{\boldsymbol{w}}_o(k)$ at range k are computed using a Taylor series expansion typically limited to first order,

$$\underline{\boldsymbol{w}}(k) = \underline{\boldsymbol{w}}(l) + (k-l)\dot{\underline{\boldsymbol{w}}}(l), \quad (10)$$

where l is the reference range, k the range of interest and $\dot{\underline{\boldsymbol{w}}}(\,.\,)$ the derivative of $\underline{\boldsymbol{w}}$ with respect to range. The values of $\underline{\boldsymbol{w}}_o(l)$ and $\dot{\underline{\boldsymbol{w}}}_o(l)$ are given by [4]

$$\widetilde{\underline{\boldsymbol{w}}}_o = \begin{pmatrix} \underline{\boldsymbol{w}}_o(l) \\ \dot{\underline{\boldsymbol{w}}}_o(l) \end{pmatrix} = \widetilde{\underline{\underline{\boldsymbol{R}}}}^{-1} \begin{pmatrix} \underline{\boldsymbol{v}}(\nu_s,\nu_d) \\ \underline{\boldsymbol{0}} \end{pmatrix}, \quad (11)$$

where

$$\widetilde{\underline{\underline{\boldsymbol{R}}}} = \frac{1}{N_l} \sum_{k \in S_l} \begin{pmatrix} \underline{\underline{\boldsymbol{R}}}(k) & (k-l)\underline{\underline{\boldsymbol{R}}}(k) \\ (k-l)\underline{\underline{\boldsymbol{R}}}(k) & (k-l)^2\underline{\underline{\boldsymbol{R}}}(k) \end{pmatrix}.$$

The main advantage of the DBU method is that it does not require any knowledge of the configuration parameters. However, performance vary considerably from one configuration to another, since we assume that $\underline{\boldsymbol{w}}_o(k)$ varies linearly with range k. In addition, the number of degrees of freedom is doubled, so that the number of samples required to estimate $\widetilde{\underline{\underline{\boldsymbol{R}}}}$ is doubled.

6. METHODS BASED ON REGISTRATION OF DD CURVES

6.1. Conceptual transformation for $\underline{\underline{\boldsymbol{R}}}(k)$

In all our range-dependence compensation methods, we have

$$\widehat{\underline{\underline{\boldsymbol{R}}}}(l) = \frac{1}{N_l} \sum_{k \in S_l} \underline{\underline{\boldsymbol{R}}}'(k) = \frac{1}{N_l} \sum_{k \in S_l} T_{lk} \left[\underline{\underline{\boldsymbol{R}}}(k) \right], \tag{12}$$

where $T_{lk}[\,.\,]$ is the transformation that brings the clutter ridge of $\underline{\underline{\boldsymbol{R}}}(k)$ into registration with the clutter ridge of $\underline{\underline{\boldsymbol{R}}}(l)$. Note that we do *not* attempt to provide an analytical expression for $T_{lk}[\,.\,]$. Instead, we provide algorithms that implement this transformation. Therefore, $T_{lk}[\,.\,]$ is primarily of conceptual interest. Since the manifestation of the range-dependence problem is in the spectral domain, the range-dependence compensation methods are more naturally designed in the spectral domain.

As a result of stationarity, $\underline{\underline{\boldsymbol{R}}} = \underline{\underline{\boldsymbol{R}}}(k)$ is Toeplitz-block-Toeplitz [2]. Exploiting redundancy in $\underline{\underline{\boldsymbol{R}}}$, we replace the $MN \times MN$ matrix $\underline{\underline{\boldsymbol{R}}}$ by a $(2N-1) \times (2M-1)$ matrix $\underline{\underline{\boldsymbol{\Gamma}}} = \underline{\underline{\boldsymbol{\Gamma}}}(k)$ entirely equivalent to $\underline{\underline{\boldsymbol{R}}}$. In contrast with $\underline{\underline{\boldsymbol{R}}}$, $\underline{\underline{\boldsymbol{\Gamma}}}$ has one dimension devoted to space and the other devoted to time. Clearly, we can write an equation similar to Eq. (12) for the $\underline{\underline{\boldsymbol{\Gamma}}}$ matrices.

6.2. Matching of DD curves

Here, one can think in terms of either the PSD clutter ridge or the corresponding DD curves. The DD curves are preferred since we have analytical tools to deal with them. However, ultimately, the thinking must be translated into the PSD plane.

Consider DD curves at various ranges for a specific configuration. The idea is to bring each DD curve at range $k \in S_l$ (the *source* range) into registration with the DD curve at reference range l (the *destination* range). Since the source curve is deformed into the destination curve, the terminologies of *moving curve* and *fixed curve* and the notations MC(k) and FC(l) are used.

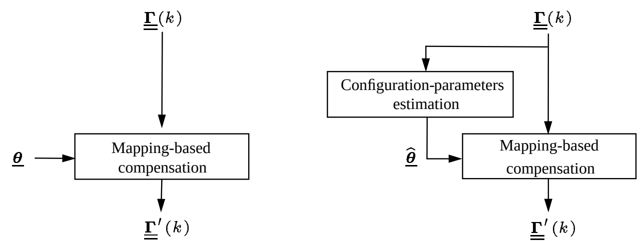
6.3. Classes of mapping-based methods

We consider two classes of methods. In "open-loop (OL)" methods, we assume that $\underline{\boldsymbol{\theta}}$ is known and in "data-adaptive (DA)" methods, we estimate $\underline{\boldsymbol{\theta}}$ from the data. The architecture of each class is shown in Fig. 4. There are three preferred implementations of the "mapping-based compensation" method, each relying on a different geometrical transformation. This results in the six different mapping-based methods listed in Table 1. They are discussed below.

7. METHODS FOR MAPPING-BASED COMPENSATION

7.1. Generic mapping-based compensation

Figure 5 shows a block diagram of the processing steps of all mapping-based compensation methods. The main step that differs from one particular method to the next is the "mapping" step.



(a) Open-loop (OL) methods (b) Data-adaptive (DA) methods

Figure 4: Comparison of architectures of (a) open-loop (OL) methods and (b) data-adaptive (DA) methods.

	Open-loop	Data-adaptive
Scaling transformation (MS)	OL-ST-MS	DA-ST-MS
Affine transformation (BS)	OL-AT-BS	DA-AT-BS
Warping transformation (BS)	OL-WT-BS	DA-WT-BS

Table 1: Classification names and abreviation for the range-dependence compensation methods.

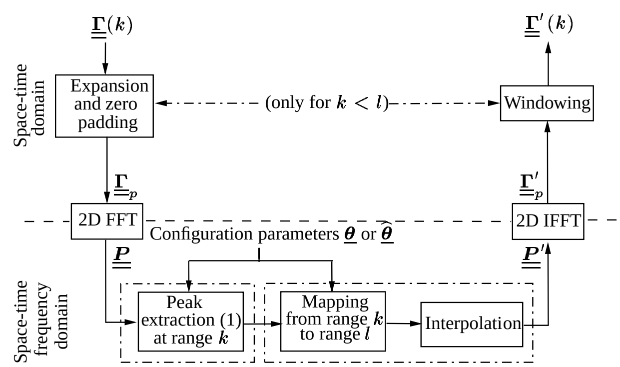


Figure 5: Block diagram of processing steps for generic mapping-based compensation method.

Expansion and zero-padding (if required): For ranges $k < l$, the scaling of the PSD $\underline{\underline{\boldsymbol{P}}}$ is a dilation, thus a contraction of $\underline{\underline{\boldsymbol{\Gamma}}}$. The size $(2N-1) \times (2M-1)$ of $\underline{\underline{\boldsymbol{\Gamma}}} = \underline{\underline{\boldsymbol{\Gamma}}}(k)$ is thus increased by mapping-specific factors along ν_s and ν_d and zero-padded.

Fourier transform: The 2D FFT of $\underline{\underline{\boldsymbol{\Gamma}}}_p$ gives the PSD $\underline{\underline{\boldsymbol{P}}}$.

Peak extraction (1): In order to dilate or contract the clutter ridge in $\underline{\underline{\boldsymbol{P}}}$, we find the position of the significant peaks in $\underline{\underline{\boldsymbol{P}}}$ by tracking these peaks down along the theoretical DD curve using $\underline{\boldsymbol{\theta}}$ or $\widehat{\underline{\boldsymbol{\theta}}}$.

Mapping: $(\nu_s(k), \nu_d(k))$ on MC(k) is mapped onto $(\nu'_s(l), \nu'_d(l))$ on FC(l). The three mappings are described later.

Interpolation: When $k < l$, linear interpolation is performed to ensure the "continuity" of the dilated ridge in $\underline{\underline{\boldsymbol{P}}}'$.

Inverse Fourier transform: The 2D IFFT of $\underline{\underline{\boldsymbol{P}}}'$ gives $\underline{\underline{\boldsymbol{\Gamma}}}'_p$.

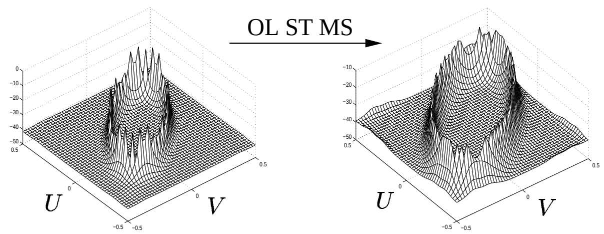
Windowing: If $k < l$, $\underline{\underline{\boldsymbol{\Gamma}}}'_p$ must be windowed to recover the desired $(2N-1) \times (2M-1)\ \underline{\underline{\mathbf{\Gamma}}}'$.

7.2. Scaling transformation for MS configurations (ST-MS)

This method applies only to MS configurations. It relies on the fact that all DD curves for MS configurations are scaled versions of each others. The mapping from $(\nu_s(k), \nu_d(k))$ on MC(k) to $(\nu_s'(l), \nu_d'(l))$ on FC(l) is given by the scaling transformation (ST)

$$\begin{pmatrix} \nu_s'(l) \\ \nu_d'(l) \end{pmatrix} = \begin{pmatrix} s_s(l,k) & 0 \\ 0 & s_d(l,k) \end{pmatrix} \begin{pmatrix} \nu_s(k) \\ \nu_d(k) \end{pmatrix}, \quad (13)$$

where $s_s(l,k) = s_d(l,k)$ are simple functions of H and of the ranges $R_m(l)$ and $R_m(k)$ corresponding to range gates l and k. Figure 6 illustrates MVE-based PSD before and after compensation.



(a) PSD before compensation (b) PSD after compensation

Figure 6: Illustration of the OL-ST-MS method for a MS configuration with $v_R = 90$ m/s, $\delta = 45$ deg, $R_m(k) = 50$ km and $R_m(l) = 125$ km. The MVE of $\underline{\underline{\mathbf{\Gamma}}}(k)$ (or $\underline{\underline{\boldsymbol{R}}}(k)$) is shown in (a). The MVE of $\underline{\underline{\mathbf{\Gamma}}}'(l)$ (or $\underline{\underline{\boldsymbol{R}}}'(l)$) is shown in (b).

7.3. Affine transformation for BS configurations (AT-BS)

In BS configurations, a simple scaling does not lead to good compensation. A straightforward generalization of scaling is the affine transformation (AT). The coefficients of the AT are found by first discretizing ψ and recording the related samples on MC(k) and FC(l). Then, an overdetermined MMSE problem is solved to find the optimum coefficients. The potential of the AT is illustrated in Fig. 7. The transformation works well when the source and destination curves have similar shapes. Given its limited number of degrees of freedom, the AT cannot be expected to bring into perfect registration two curves that have quite different shapes.

7.4. Warping transformation for BS configurations (WT-BS)

Using the pair of equations in Eq. (3), we can compute the "flow line" corresponding to each discrete value of ψ, as shown in Fig. 8(a). As R_b increases, we move along any particular flow line. The mapping is then simple : for any given ψ, we find the source point $A(\nu_s, \nu_d)$ on MC(k) and map it into $A'(\nu_s', \nu_d')$ on FC(l). Remember that mapping consists essentially in determining the intensity associated with A in the source PSD and in mapping this value into the pixel closest to A' in the destination PSD. An example of mapping is shown in Fig. 8(b).

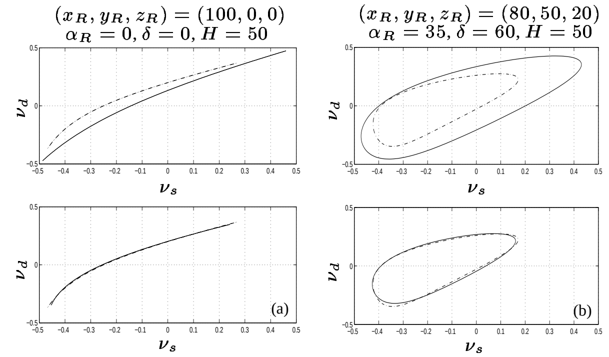


Figure 7: Illustration of DD curve registration achieved with the AT. The dotted lines correspond to the FCs. The solid curves are the MCs before (upper diagram) and after (lower diagram) transformation. In all cases, $R_b(k)$ is 200 km and $R_b(l)$ is 350 km.

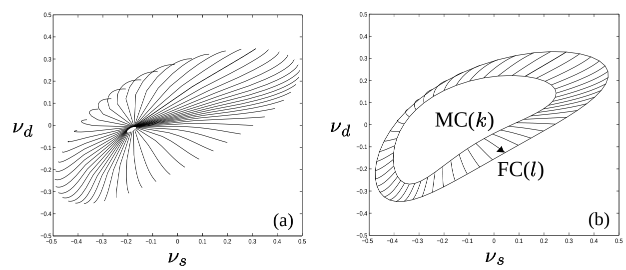


Figure 8: (a) Illustration of flow lines. (b) Flow lines dictating the deformation of MC(k) into FC(l).

8. ESTIMATION OF CONFIGURATION PARAMETERS

Fig. 9 shows a block diagram of the processing steps in the generic configuration-parameters estimation method. The input is $\underline{\underline{\mathbf{\Gamma}}}(k)$ and the output is the estimate $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$.

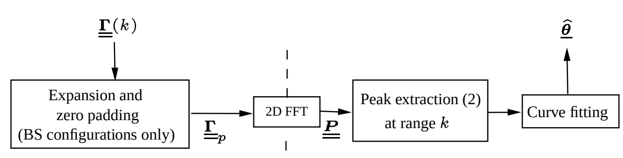


Figure 9: Block diagram of processing steps for generic configuration-parameters estimation method.

The main steps that differ for MS and BS configurations are the "peak extraction" and the "curve fitting" steps. In both case, we assume that the height H is known. In the MS case, we assume that $v_{\mathcal{R}}$ is known. This is a reasonable assumption, since the processing is typically done onboard. In the BS case, we similarly assume that both v_T and v_R are known. We have developped estimation method that work even in the case where H, $v_{\mathcal{R}}$, v_T and v_R are unknown. These methods are not described here.

Expansion and zero padding: The goal is to increase the size of the array $\underline{\underline{\boldsymbol{\Gamma}}}(k)$ for subsequent curve fitting below. This is done only for BS configurations: more data points are required given that there are more parameters to estimate.

Fourier transform: The 2D FFT of $\underline{\underline{\boldsymbol{\Gamma}}}_p$ gives $\underline{\underline{\boldsymbol{P}}}$.

Peak extraction (2): Since $\underline{\boldsymbol{\theta}}$ is not known, we cannot use the "peak extraction (1)" step. For MS configurations, we use a thresholding algorithm that uses the histogram of the peak amplitudes to find the optimum threshold [7]. For BS configurations, the MS thresholding algorithm performs poorly due to the larger number of parameters to estimate. Instead, we use a watershed segmentation algorithm inspired from image processing [8].

Curve fitting: We have analytical equations for the MS and BS DD curves. We also have the coordinates $(\nu_{s,j}, \nu_{d,j})$ of the peaks just extracted. We can thus perform a MMSE fit of the parametric curve to the peaks. The result is the MMSE estimate $\widehat{\underline{\boldsymbol{\theta}}}$ of the parameter vector $\underline{\boldsymbol{\theta}}$. The statistics that is minimized is

$$E(\underline{\boldsymbol{\theta}}) = \sum_{j=1}^{N_p} d^2\left((\nu_{s,j}, \nu_{d,j}), \mathcal{C}(\underline{\boldsymbol{\theta}})\right), \qquad (14)$$

where N_p is the number of detected peaks, $d(P_j, \mathcal{C})$ the distance between point P_j and curve $\mathcal{C}$, and $\mathcal{C}(\underline{\boldsymbol{\theta}})$ the DD curve corresponding to $\underline{\boldsymbol{\theta}}$. For MS configurations, the only unknown is δ. Its value is found by solving an overdertermined MMSE problem [7]. For BS configurations, the unknowns are x_R, y_R, z_R, α_R and δ. Their estimation is more complex and is described in [8].

9. PERFORMANCE COMPARISON

The $\mathrm{SINR_L}$ of Eq. (9) is used to compare the performances of the methods listed in Table 1 and those of the OP and of the SA. (OP uses the exact value of $\underline{\underline{\boldsymbol{R}}}(l)$ and SA the value of $\widehat{\underline{\underline{\boldsymbol{R}}}}(l)$ obtained by straight averaging of the $\underline{\underline{\boldsymbol{R}}}(k)$'s.) Here, we assume omnidirectional sensors. Similar results are obtained with directional sensors. Figure 10 compares the performances of the various methods. Its organization matches that of Table 1. Rows corresponds to ST, AT and WT. The graphs on the left compare the OL methods to OP and SA. Those on the right compare the OL and DA methods. Figures 10(a)-(b) deal with ST and confirm (1) that OL-ST-MS performs nearly as well as OP and much better than SA and (2) that OL and DA provide nearly identical performance. The remaining subfigures lead to similar conclusions for AT and WT. As expected, WT provides better performance than AT, in both OL and DA modes.

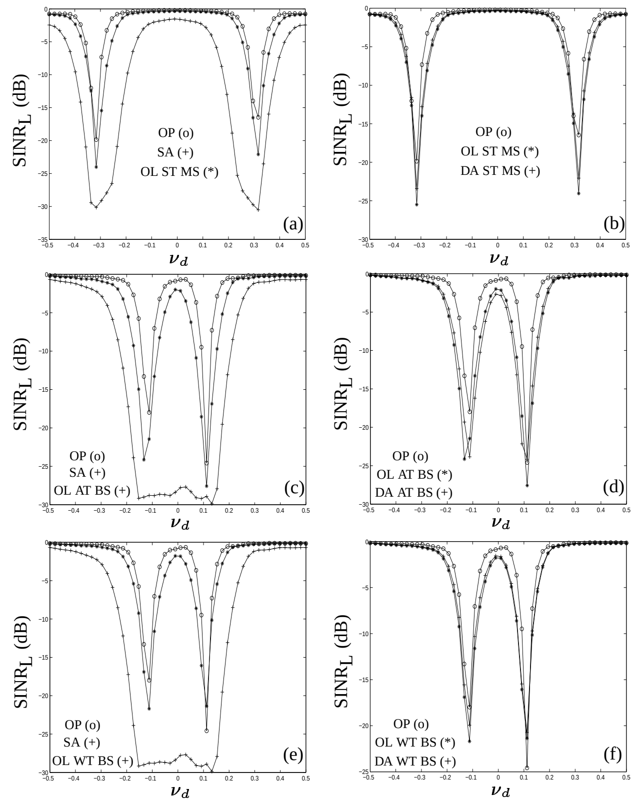


Figure 10: Performance comparaison of the methods of Table 1. Figures show graphs of $\mathrm{SINR_L}(\nu_d)$ at $\nu_s = 0$.

10. CONCLUSIONS

The range-dependence of the PSD clutter ridge is a serious problem for non-SL MS and BS STAP radars. This paper discusses six new compensation methods based on mappings exploiting the geometrical properties of DD curves. Their performance was discussed. The range-dependence compensation achieved is nearly perfect for all configurations, including in the DA mode. This is achieved without any increase in the number of degrees of freedom required for clutter rejection.

11. REFERENCES

[1] R. Klemm, *Principles of space-time adaptive processing*, IEE Radar, Sonar, Navigation and Avionics 9, 2002.

[2] J. Ward, "Space-time adaptive processing for airborne radar," *Technical Report 1015, MIT Lincoln Laboratory*, 1994.

[3] R.L. Fante, "Ground and airborne target detection with bistatic adaptive space-based radar," *IEEE Radar Conference*, pp. 7–11, 1999.

[4] S.M. Kogon and M.A. Zatman, "Bistatic STAP for airborne radar systems," *ASAP Conference, MIT Lincoln Laboratory, Lexington*, 13-14 March 2001.

[5] G.K. Borsari, "Mitigating effects on STAP processing caused by an inclined array," *IEEE National Radar Conference, Dallas*, pp. 135–140, 12-13 May 1998.

[6] F. Pearson and G. Borsari, "Simulation and analysis of adaptive interference suppression for bistatic surveillance radars," *ASAP Conference, MIT Lincoln Laboratory, Lexington*, 13-14 March 2001.

[7] F. Lapierre, M. Van Droogenbroeck, and J. Verly, "New methods for handling the range dependence of the clutter spectrum in non-sidelooking monostatic STAP radars," *ICASSP 2003*, April 2003.

[8] F. Lapierre, M. Van Droogenbroeck, and J. Verly, "New solutions to the problem of range dependence in bistatic STAP radars," *IEEE Radar Conference 2003*, May 2003.