

New Solutions to the Problem of Range Dependence in Bistatic STAP Radars

Fabian D. Lapierre (Research Fellow), Jacques G. Verly and Marc Van Droogenbroeck

University of Liège, Department of Electrical Engineering and Computer Science

Sart-Tilman, Building B28, B-4000 Liège, Belgium

{F. Lapierre, Jacques.Verly, M.VanDroogenbroeck}@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. However, in bistatic configurations, these snapshots are range-dependent. As a result, straight averaging results in poor performance. After reviewing existing methods for handling the range-dependence, we present new methods exploiting the precise shape of the bistatic direction-Doppler curves.

I. INTRODUCTION

Pulsed Doppler radars are used to detect moving targets and to measure their range and speed. They typically transmit a train of coherent pulses. Equipped with a linear array antenna, they are particularly well suited for detecting slow-moving targets in the presence of clutter and jammers. The technique of choice for dealing with this problem is space-time adaptive processing (STAP) [1], [2].

Radars typically operate in monostatic (MS) configuration, i.e., with the transmitter and the receiver colocated. While initial STAP research was focused on MS configurations, attention is now turning to bistatic (BS) configurations [3], [4], where the transmitter and the receiver are located on distinct, independently-moving platforms.

The data collected by STAP radars can be viewed as a sequence of 2D space-time arrays, called “snapshots”. Basic STAP methods compute a weighted linear combination of the snapshot elements. The calculation of the optimum weights generally involves the inversion of the covariance matrix of the snapshot. This covariance matrix must be estimated. This is typically done by averaging the snapshots at a series of neighboring ranges.

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

We begin by reviewing existing range-dependence compensation methods. Then, we present new methods based on the idea of estimating the covariance matrix at some reference range gate (indexed with l) by first applying transformations 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., to 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 . All these methods are thus based upon a mapping of the elements of the PSD at each range k .

The registration of the clutter ridges observed in experimental data is guided by analytical formulas describing the form of related “direction-Doppler (DD) curves”, which are mathematical curves fully determined by the configuration parameters and the range of interest. We distinguish between two types of range-dependence compensation methods: (a) “true-parameters (TP)” methods assume exact knowledge of the parameters and (b) “estimated-parameters (ET)” methods estimate the parameters from the data. Below, we describe these methods and discuss their performance.

II. BISTATIC (BS) RADAR-MEASUREMENT CONFIGURATION

Figure 1 shows a typical BS configuration with transmitter T and receiver R . The origin of the (x, y, z) coordinate system coincides with T . Its orientation is such that the x -axis points in the same way as the transmitter velocity vector \underline{v}_T and that the z -axis points vertically up. We assume that the receiver velocity vector \underline{v}_R is located in a horizontal plane. The angle between \underline{v}_R and \underline{v}_T is denoted by α_R . The antenna A is assumed to be linear and in a horizontal plane. Defining δ as the angle between the s -axis and \underline{v}_R , $\delta + \alpha_R$ fully describes the orientation of A . The sidelooking (SL) configuration corresponds to $\delta = 0$. The bistatic range R_b is the distance from T to S to R . Defining $v_R = |\underline{v}_R|$ and $v_T = |\underline{v}_T|$, any BS configuration is fully characterized by the vector of parameters

$$\underline{\theta} = (H, x_R, y_R, z_R, v_R, v_T, \alpha_R, \delta). \quad (1)$$

III. DIRECTION-DOPPLER CURVES AND SURFACES

A. Important parameters

The radars of interest are expected to determine at least three basic parameters for each scatterer S : the BS range R_b , the angular position ξ_s and the relative velocity v_r (with

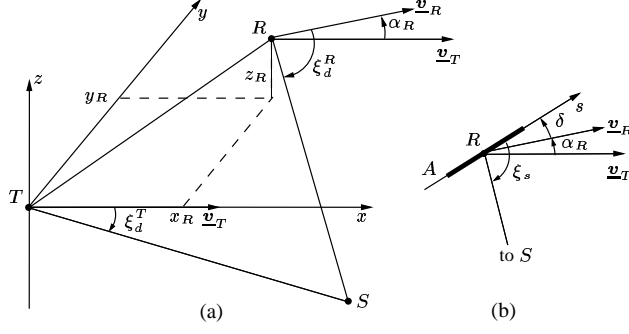


Fig. 1. BS configuration. (a) Transmitter (T) - Receiver (R) - Scatterer (S) geometry and related parameters. (b) Antenna A and related angles.

respect to R). The related parameters that are more directly measured from the radar returns are the roundtrip delay τ_{rt} , the spatial frequency f_s and the Doppler frequency f_d . For a stationary scatterer, they are respectively given by

$$\begin{aligned}\tau_{rt} &= R_b/c \\ f_s &= \lambda_c^{-1} \cos \xi_s \\ f_d &= \lambda_c^{-1} v_T \cos \xi_d^T + \lambda_c^{-1} v_R \cos \xi_d^R,\end{aligned}$$

where λ_c is the carrier wavelength and c is the speed of light. R_b , ξ_s and v_r are easily computed from τ_{rt} , f_s and f_d .

B. Isorange curves

All scatterers S characterized by the same range R_b are located on an isorange surface, which is an ellipsoid of revolution with foci at T and R . The intersection of this surface with the ground, modelled as a horizontal plane at $z = -H$, is called an isorange curve. It is in fact an ellipse, which we parameterize with the polar angle ψ .

C. Direction-Doppler (DD) curves

For any given configuration and range R_b , all stationary scatterers at this range map onto a curve showing the relation between f_s and f_d for any such scatterer. Each 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 $(\lambda_c/2(v_R + v_T))f_d$. Figure 2 shows that BS DD curves vary significantly with configuration and range. The fact that these curves vary with range for any particular BS configuration is the source of the range-dependence problem considered in this paper.

Deriving the equations of BS DD curves is a challenge. Our approach is as follows. First, we express ν_d as a function of ν_s . Since most BS DD curves have two possible values of ν_d for each value of ν_s , any BS DD curve is best described by two functions $\nu_d = f_1(\nu_s)$ and $\nu_d = f_2(\nu_s)$. Second, if we express ν_s and ν_d in terms of ψ , we find a parametric description of the DD curve, i.e.,

$$\nu_s = g_1(\psi) \quad \text{and} \quad \nu_d = g_2(\psi). \quad (2)$$

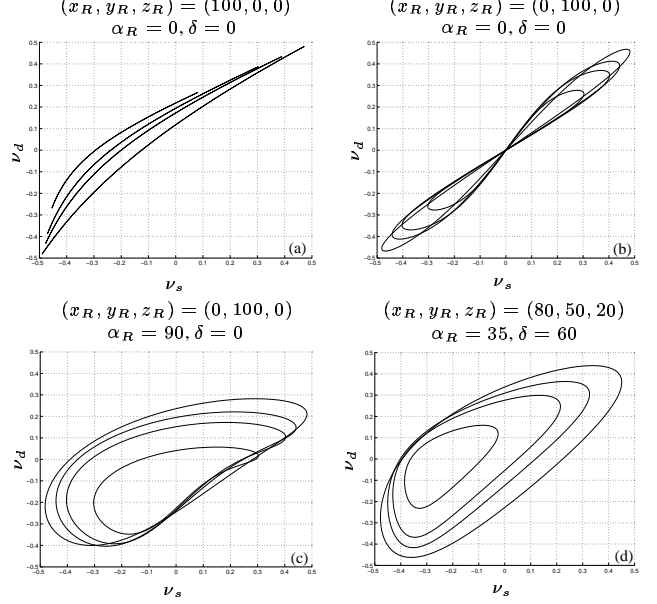


Fig. 2. Example BS DD curves. Units of x_R , y_R and z_R are km; units of α_R and δ are degrees. In all cases, $H = 50$ km, $v_R = 90$ m/s and $v_T = 90$ m/s. Ranges R_b are 170, 210, 250 and 400 km.

The derivation of the functions $f_1(\nu_s)$ and $f_2(\nu_s)$ or $g_1(\psi)$ and $g_2(\psi)$ is complicated and lengthy and is thus omitted.

D. Direction-Doppler (DD) surfaces

The surface obtained by stacking DD curves for various values of R_b is called a “direction-Doppler (DD)” surface. By construction, a slice of the DD surface at a given R_b produces the DD curve for that range.

E. Recovery of DD surface parameters

Consider the BS DD surface \mathcal{S} corresponding to an arbitrary BS configuration characterized by $\underline{\theta}_1 = (H, x_R, y_R, z_R, v_R, v_T, \alpha_R, \delta)$ and to all applicable values of R_b . One can show that the only other set of parameters that produces \mathcal{S} is

$$\underline{\theta}_2 = (H, x_R, -y_R, z_R, v_R, v_T, -\alpha_R, -\delta).$$

Thus, we know that the inverse problem of recovering $\underline{\theta}$ from \mathcal{S} does *not* have a unique solution, but that there are only *two* related solutions. From this, we can infer that the inverse problem of recovering $\underline{\theta}$ from a single slice of \mathcal{S} has *at least* two related solutions.

IV. 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 (called range gates) covering the range interval of interest. 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 specific ν_s , ν_d and R_b can be written as the $MN \times 1$ vector [2]

$$\underline{\mathbf{y}}(\nu_s, \nu_d) = \beta_r \underline{\mathbf{v}}(\nu_s, \nu_d) = \beta_r \underline{\mathbf{b}}(\nu_d) \otimes \underline{\mathbf{a}}(\nu_s), \quad (3)$$

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

$$\underline{\mathbf{a}}(\nu_s) = (1 \dots e^{j2\pi\nu_s n} \dots e^{j2\pi\nu_s(N-1)})^T \quad (4)$$

$$\underline{\mathbf{b}}(\nu_d) = (1 \dots e^{j2\pi\nu_d m} \dots e^{j2\pi\nu_d(M-1)})^T. \quad (5)$$

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

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

Since $\beta_c(\psi)$ is a random process, $\underline{\mathbf{y}}_c$ is a random vector. We assume it is wide-sense stationary. It is thus characterized by a constant covariance matrix $\underline{\mathbf{R}}_c = E\{\underline{\mathbf{y}}_c \underline{\mathbf{y}}_c^\dagger\}$. To find the power spectral density (PSD) associated with $\underline{\mathbf{y}}_c$, we use the minimum variance estimator [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 is [1]

$$\underline{\mathbf{w}}_o(\nu_s, \nu_d) = \underline{\mathbf{R}}^{-1} \underline{\mathbf{v}}(\nu_s, \nu_d), \quad (6)$$

where the covariance matrix $\underline{\mathbf{R}} = E\{\underline{\mathbf{y}} \underline{\mathbf{y}}^\dagger\}$ is the sum of the covariance matrices $\underline{\mathbf{R}}_c = E\{\underline{\mathbf{y}}_c \underline{\mathbf{y}}_c^\dagger\}$ for the clutter and $\underline{\mathbf{R}}_n = E\{\underline{\mathbf{y}}_n \underline{\mathbf{y}}_n^\dagger\}$ for the noise. We assume that the noise $\underline{\mathbf{y}}_n$ is spatially and temporally white so that $\underline{\mathbf{R}}_n = \underline{\mathbf{I}}$. In practice, $\underline{\mathbf{R}}$ is not known and must be estimated for each range.

The maximum-likelihood estimator $\hat{\underline{\mathbf{R}}}$ for range gate l is [2]

$$\hat{\underline{\mathbf{R}}}(l) = \frac{1}{N_l} \sum_{k \in S_l} \underline{\mathbf{R}}(k), \quad (7)$$

where N_l is the number of snapshots used for estimation, S_l is the set of snapshot indices k defined by $l - \frac{N_l-1}{2} < k < l + \frac{N_l-1}{2}$ and $\underline{\mathbf{R}}(k) = \underline{\mathbf{y}}(k) \underline{\mathbf{y}}^\dagger(k)$, where $\underline{\mathbf{y}}(k)$ is the snapshot at range (gate) k . Equation (7) provides an acceptable estimate for $\underline{\mathbf{R}}(l)$ only if the clutter ridge is range-independent. This never happens for BS configurations. Range-dependence compensation methods are thus needed.

The performance of a processor using weights $\underline{\mathbf{w}}$, whether optimal or not, is measured by the signal-to-interference-plus-noise (SINR) loss defined as [2]

$$\text{SINR}_L = \frac{\text{SINR}}{\text{SINR}_0} = \frac{|\underline{\mathbf{w}}^\dagger \underline{\mathbf{v}}|^2}{(\underline{\mathbf{w}}^\dagger \underline{\mathbf{R}} \underline{\mathbf{w}})(\underline{\mathbf{v}}^\dagger \underline{\mathbf{v}})}, \quad (8)$$

where SINR_0 is the SINR in the absence of clutter. Values of 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 for $\underline{\mathbf{w}} = \underline{\mathbf{w}}_o$. In practice, processor performance is degraded by estimation losses and by the range-dependence of the clutter ridge.

V. EXISTING METHODS BASED ON DOPPLER WARPING AND TAYLOR SERIES

A. Doppler Warping (DW) method

The Doppler Warping (DW) method was initially developed for nearly-SL MS configurations [5] and subsequently applied to BS configurations [4]. Keeping in mind that $\underline{\mathbf{y}}(k)$ is of the form given in Eq. (3), the principle of DW is to add to ν_d in Eq. (3) a Doppler shift $\Delta(k)$ that is chosen for each range k in such a way as to bring all clutter ridges in registration. Equation (5) shows that this can be achieved by premultiplying $\underline{\mathbf{y}}(\nu_s, \nu_d)$ by the matrix

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

Therefore, applying $\underline{\mathbf{T}}(k)$ to the snapshot $\underline{\mathbf{y}}(k)$ results in the Doppler-warped snapshot $\underline{\mathbf{y}}_{dw}(k) = \underline{\mathbf{T}}(k) \underline{\mathbf{y}}(k)$. $\underline{\mathbf{T}}(k)$ can provide perfect compensation at only one specific ν_s . Performance is poor for BS configurations [4]. The main advantage of the DW method is its simplicity of implementation. However, the configuration parameters must be known.

B. High-Order Doppler Warping (HODW) method

The High-Order Doppler Warping (HODW) method is a generalization of the Doppler Warping (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.

C. Derivative-based updating (DBU) method

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

$$\underline{\mathbf{w}}(k) = \underline{\mathbf{w}}(l) + (k-l) \dot{\underline{\mathbf{w}}}(l), \quad (9)$$

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

$$\begin{pmatrix} \underline{\mathbf{w}}_o(l) \\ \underline{\dot{\mathbf{w}}}_o(l) \end{pmatrix} = \underline{\tilde{\mathbf{R}}}^{-1} \begin{pmatrix} \underline{\mathbf{v}}(\nu_s, \nu_d) \\ \underline{\mathbf{0}} \end{pmatrix}, \quad (10)$$

where

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

The main advantage of the DBU method is that it does not require any knowledge of the configuration parameters. One disadvantage is that the performance of the method vary considerably from one configuration to another, since we assume that $\underline{\mathbf{w}}_o(k)$ varies linearly with range k . Another disadvantage is that the number of degrees of freedom is doubled: as a result, the number of samples required for estimating $\underline{\tilde{\mathbf{R}}}$ is doubled.

VI. METHODS BASED ON REGISTRATION OF DD CURVES

A. Conceptual transformation for $\underline{\mathbf{R}}(k)$

In all our range-dependence compensation methods,

$$\hat{\underline{\mathbf{R}}}(l) = \frac{1}{N_l} \sum_{k \in S_l} \underline{\mathbf{R}}'(k) = \frac{1}{N_l} \sum_{k \in S_l} T_{lk} [\underline{\mathbf{R}}(k)], \quad (11)$$

where $T_{lk}[\cdot]$ is the transformation bringing the clutter ridge of $\underline{\mathbf{R}}(k)$ into registration with that of $\underline{\mathbf{R}}(l)$. We will *not* provide an analytical expression for $T_{lk}[\cdot]$. Instead, we will provide algorithms implementing this transformation. Thus, $T_{lk}[\cdot]$ is primarily of conceptual interest. Since the manifestation of the range-dependence problem is in the spectral domain, range-dependence compensation methods are more naturally designed in the spectral domain.

As a result of stationarity, $\underline{\mathbf{R}} = \underline{\mathbf{R}}(k)$ is Toeplitz-block-Toeplitz [2]. Exploiting redundancy in $\underline{\mathbf{R}}$, we replace the $MN \times MN$ matrix $\underline{\mathbf{R}}$ by a $(2N-1) \times (2M-1)$ matrix $\underline{\Gamma} = \underline{\Gamma}(k)$ entirely equivalent to $\underline{\mathbf{R}}$. In contrast with $\underline{\mathbf{R}}$, $\underline{\Gamma}$ has one dimension devoted to space and the other devoted to time. Expressions equivalent to Eqs. (11) for $\underline{\mathbf{R}}$ can be written for $\underline{\Gamma}$.

B. Matching of DD curves

When designing our methods, we can either think in terms of the PSD clutter ridge or in terms of the corresponding DD curves. The DD curves are preferred, since we have analytical tools to deal with such curves. Ultimately though, the thinking must be translated into the PSD plane.

Consider a set of DD curves at various ranges for a specific configuration. The idea is to bring the DD curve at each 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 will be deformed into the destination curve, the terms *moving curve* (MC) and *fixed curve* (FC) are used.

C. Classes of mapping-based methods

We consider two classes of methods. In “true-parameters (TP)” methods, we assume that $\underline{\theta}$ is known and in “estimated-parameters (EP)” methods, we estimate $\underline{\theta}$ from the data. The general architecture of each class is shown in Fig. 3. For BS configurations, there are two preferred implementations of the “mapping-based compensation,” each relying on a particular geometrical transformation. Table I summarizes the various methods discussed below. A “mapping-based compensation” specifically designed for MS configurations is proposed in [7], where the TP or EP variants are also considered.

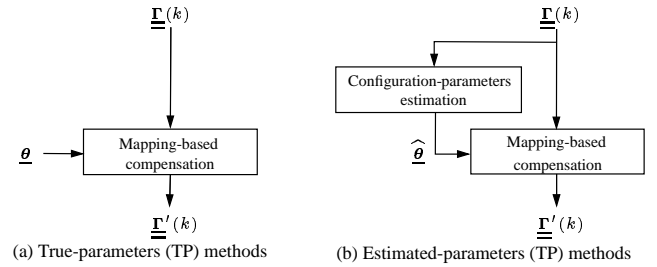


Fig. 3. Comparison of architecture of (a) true-parameters (TP) methods and (b) estimated-parameters (EP) methods.

TABLE I

	True-parameters meth.	Estimated-parameters meth.
Affine transformation	TP-AT-BS	EP-AT-BS
Warping transformation	TP-WT-BS	EP-WT-BS

VII. MAPPING-BASED COMPENSATION METHODS

A. Generic mapping-based compensation

Figure 4 shows 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.

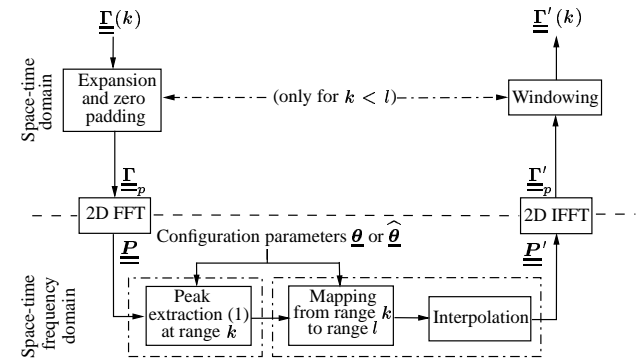


Fig. 4. Block diagram of processing steps for generic mapping-based compensation method.

Expansion and zero-padding (if required): For ranges $k < l$, the scaling of the PSD $\underline{\underline{P}}$ is a dilation, implying a contraction of $\underline{\underline{\Gamma}}$. The size $(2N - 1) \times (2M - 1)$ of $\underline{\underline{\Gamma}} = \underline{\underline{\Gamma}}(k)$ must thus be increased (by padding $\underline{\underline{\Gamma}}$ with zeros) in the ν_s and ν_d dimensions, respectively.

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

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

Mapping: The point $(\nu_s(k), \nu_d(k))$ on $\text{MC}(k)$ is mapped onto the point $(\nu'_s(l), \nu'_d(l))$ on $\text{FC}(l)$. The particular mapping of each method is described later.

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

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

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

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

In MS configurations, a simple scaling suffices since all MS DD curves are a scaled version of each other [7]. However, in BS configurations, a simple scaling does not lead to good compensation. A straightforward generalization of the scaling transformation is the affine transformation (AT). The coefficients of the AT are found as follows. First, we discretize ψ and record the related samples on $\text{MC}(k)$ and $\text{FC}(l)$. Second, we find the AT coefficients that are optimum in a LS error sense. The potential of the AT is illustrated in Fig. 5. 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.

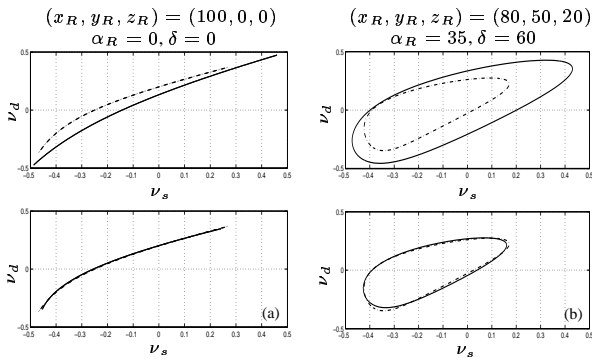


Fig. 5. Illustration of DD curve registration achieved with the AT. Dotted lines correspond to the FC. Solid lines correspond to the MC before (upper diagram) and after (lower diagram) AT. In all cases, $R_b(k)$ is 200 km and $R_b(l)$ is 350 km, $H = 50$ km and $v_R = v_T = 90$ m/s.

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

The relations in Eqs. (2) allow us to compute the “flow line” corresponding to each specific value of ψ . This is illustrated in Fig. 6(a). The various points on each flow line correspond to the various values of R_b . Typically, as R_b increases, we move along one direction along the flow line. The mapping is then simple : for any given ψ , we find the source point $A(\nu_s, \nu_d)$ on $\text{MC}(k)$ and map it into $A'(\nu'_s, \nu'_d)$ on $\text{FC}(l)$. Remember that mapping consists essentially in finding the “pixel” closest to A in the source PSD and assigning its intensity value to the “pixel” closest to A' in the destination PSD. The concept of mapping is illustrated in Fig. 6(b).

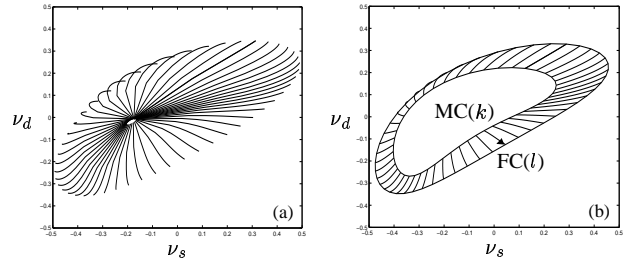


Fig. 6. (a) Illustration of flow lines. (b) Flow lines dictating the deformation of $\text{MC}(k)$ into $\text{FC}(l)$.

VIII. CONFIGURATION-PARAMETERS ESTIMATION METHODS

Figure 7 shows a block diagram of the processing steps in the generic configuration-parameters estimation method. The input is $\underline{\underline{\Gamma}}(k)$ and the output is the estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$. We assume that \bar{H} , v_R and v_T are known. We have developed estimation methods that work even when H , v_R and v_T are unknown. These methods are not described here.

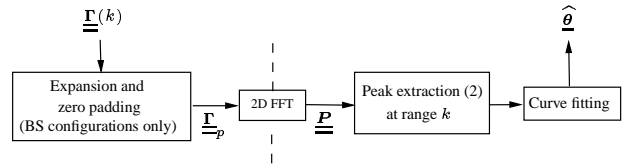


Fig. 7. Block diagram of processing steps for generic configuration-parameters estimation method.

Expansion and zero padding: The goal is to increase the size of $\underline{\underline{\Gamma}}(k)$ for curve fitting below. Experimentation shows that $\underline{\underline{\Gamma}}_p(k)$ should have a minimum size of 70×70 .

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

Peak extraction (2): Since $\underline{\theta}$ is not known, we cannot use the “peak extraction (1)” method. Instead, we use a watershed segmentation algorithm inspired from image processing [8].

Curve fitting: We have analytical equations for the DD curves. We also have the coordinates $(\nu_{s,j}, \nu_{d,j})$ of the peaks

just extracted. We can thus perform a LS fit of a parametric DD curve to the peaks. The result is the LS estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$. The statistics that is minimized is

$$E(\underline{\theta}) = \sum_{j=1}^{N_p} d^2((\nu_{s,j}, \nu_{d,j}), \mathcal{C}(\underline{\theta})), \quad (12)$$

where N_p is the number of detected peaks, $d(P_j, \mathcal{C})$ is the distance between point P_j and curve \mathcal{C} , and $\mathcal{C}(\underline{\theta})$ is the DD curve corresponding to $\underline{\theta}$. The parameters to be estimated are x_R, y_R, z_R, α_R and δ . Curve fitting is based on geometrical properties of the DD curves and is now described.

IX. CURVE FITTING METHOD

We first derive fundamental results leading to constraints on where R can be located in 3D space. Second, we use these constraints to develop a strategy for finding R .

A. Constraints on R

1) *Relationship between constant- R_b ellipsoid and ground:* Below, we carry most of our reasoning in terms of the radial cut of the 3D space by the vertical plane going through T and R . In this vertical cut, the ellipsoid $\mathcal{E}(R_b)$ appears as an ellipse $\mathcal{E}'(R_b)$.

The relationship between $\mathcal{E}'(R_b)$ and the ground is shown in Fig. 8. The major axis $2a$ of $\mathcal{E}'(R_b)$ must always be equal to R_b . However, the length $2b$ of the minor axis can be adjusted. Here, $\mathcal{E}'(R_b)$ is *tangent* to the ground at some point G . Since changing b amounts to changing the distance $2c$ between the foci T and R (and vice-versa), there are an infinite number of ellipses that (1) have T as one of their foci, (2) have $2a = R_b$ and (3) are tangent to the ground.

Via a geometrical argument, one can show that the ground, which is the tangent to $\mathcal{E}'(R_b)$ at G , is also the bisector of the angle $\widehat{TGT'}$, where T' is the mirror image of T with respect to the ground and is thus located at $z = -2H$ on the z -axis. Observe that $\overline{T'R} = R_b$.

First, assume that we are given T and some *fixed* point of tangency G on the ground. Thus, since $\overline{T'R} = R_b$, R is at the intersection (assumed to be above ground) of the line l through T' and G and the circle $c \equiv C(T', R_b)$ centered at T' and with radius R_b .

Second, assume that T is given but that G can move freely on the ground. Then, for each position G and for a given R_b , R is located at the intersection of l and c . Even though l changes with G , c remains invariant as G moves. We conclude that, for any ellipse $\mathcal{E}'(R_b)$ that corresponds to a range R_b and that is tangent to the ground, R is constrained to be on the circle $c(T', R_b)$. In 3D space, R is thus constrained to be on a sphere centered at T' and with radius R_b .

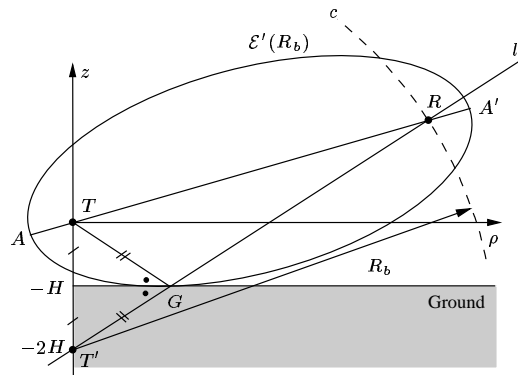


Fig. 8. (a) Key observation is that tangent to $\mathcal{E}'(R_b)$ at G is bisector of $\widehat{TGT'}$. (b) Circle c is the locus of all possible positions of R as we keep $T, \overline{AA'} = 2a = R_b$ and $\overline{TR} = 2c$ constant and $\mathcal{E}'(R_b)$ tangent to the ground.

2) *Maximum radial distance for R :* Figure 8 shows that, as G moves to the right, we reach a point where R and G coincide. This happens when $\overline{TG} + \overline{GR} = \overline{TR}$ or when $\overline{TR} = R_b$. Since R_b is also the length of the major axis of $\mathcal{E}'(R_b)$, $\mathcal{E}'(R_b)$ degenerates into the segment TR . Thus, the maximum radial distances of R and G are

$$\rho_R^{max} = \rho_G^{max} = \sqrt{R_b^2 - H^2}. \quad (13)$$

In other words, we have the constraint

$$\rho_R \leq \rho_R^{max}. \quad (14)$$

3) *Coarse region for R :* Consider a fixed combination of $T, \overline{AA'} = 2a = R_b$ and $\overline{TR} = 2c$. Starting from the tangent position, we preserve the intersection between $\mathcal{E}'(R_b)$ and the ground when we lower $\mathcal{E}'(R_b)$ and we loose it when we raise $\mathcal{E}'(R_b)$.

The above statement implies that, for a given combination of $T, \overline{AA'} = 2a = R_b$ and $\overline{TR} = 2c$, R is constrained to be on the arc segment $\widehat{RR_G}$ shown in Fig. 9(a), i.e., on the portion of the circle $C(T, 2c)$ limited by the circle $c = C(T', R_b)$ and the ground. Remember that $\widehat{RR_G}$ corresponds to a specific value of $2c$. If we change $2c$, we change the radius of $\widehat{RR_G}$. The locus of all possible positions of R for a given R_b (and H) is thus the dark-shaded region in Fig. 9(b). This 2D region is bounded by the circle $C(T', R_b)$, the circle $C(T', R_b - 2H)$ and the ground.

4) *Additional constraint from DD curve:* Now, we use our knowledge of the precise form of DD curves to further constrain the possible locations for R . Figure 10 shows a top view of the BS radar configuration of Fig. 1. P is the projection of R in the horizontal plane. The (x^ϵ, y^ϵ) -axes are defined so that the x^ϵ -axis points in the direction of the s -axis.

The important parameter is the abscissa x_P^ϵ of P in the

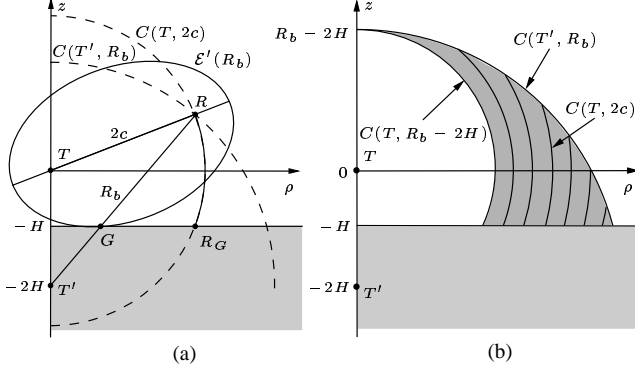


Fig. 9. Locus of all possible positions of R for a given R_b (and H) (a) for a fixed value of $2c$ and (b) for all possible values of $2c$.

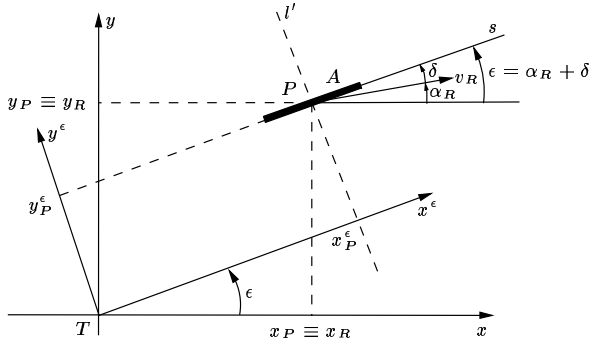


Fig. 10. Top view of BS configuration shown in Fig. 1.

(x^ϵ, y^ϵ) -axes. One can show that x_P^ϵ is given by

$$x_P^\epsilon = R_b \left(-\frac{4P_\nu + 1}{2S_\nu} + \text{sign}(S_\nu) \sqrt{\left(\frac{4P_\nu + 1}{2S_\nu} \right)^2 - 1} \right), \quad (15)$$

where $S_\nu = \nu_s^{\min} + \nu_s^{\max}$ and $P_\nu = \nu_s^{\min} \nu_s^{\max}$, ν_s^{\min} and ν_s^{\max} being the ν_s -coordinates of the extreme points of the DD curve along the ν_s -axis. Note that we can always compute x_P^ϵ without any knowledge of ϵ ! However, we should not conclude that we can locate the line l' shown in Fig. 10, which P must be located on. Indeed, we do not know ϵ ! However, the value of x_P^ϵ can be used to constrain the possible locations for P , and, thus, for R . Indeed, we can show that P must be located on the boundary or on the outside of a circle \mathcal{C}_{\min} of radius $\rho_P^{\min} = |x_P^\epsilon|$, i.e., $\rho_P \geq \rho_P^{\min}$. Since P is the projection of R in the (x, y) -plane, we also have

$$\rho_R \geq \rho_R^{\min}, \quad (16)$$

where

$$\rho_R^{\min} = |x_P^\epsilon|. \quad (17)$$

The net result of this constraint is to truncate the dark-shaded region of Fig. 9 at $\rho = \rho_R^{\min}$, as shown in Fig. 11.

Figure 11 shows that, for any radial distance ρ_R in $[\rho_R^{\min}, \rho_R^{\max}]$, there are constraints on the possible heights

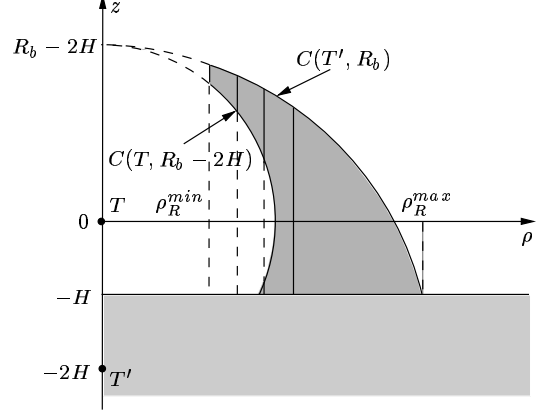


Fig. 11. Result of applying the constraint of Eq. (16) to the dark-shaded area of Fig. 9(b).

z_R . In practice, z_R can be expressed as a function of the point under consideration on the circle $C(T, \rho_R)$, S_ν , R_b and H . The derivation of this relation is complicated and is thus omitted.

B. Finding R

1) *Search-space annulus for P* : The constraints $\rho_R \geq \rho_R^{\min}$ and $\rho_R \leq \rho_R^{\max}$ indicate that P must be located in an annulus \mathcal{A} with radial extent $[\rho_R^{\min}, \rho_R^{\max}]$, as illustrated in Fig. 12. However, if $(x_R, y_R) \equiv (x_P, y_P)$ is the solution of the parameter estimation problem, we know that $(x_R, -y_R)$ is also a solution. Therefore, searching either the upper half or the lower half of the annulus will yield one solution. The second is automatically found by symmetry. To be more precise, if (y_R, α_R, δ) corresponds to one solution, $(-y_R, -\alpha_R, -\delta)$ corresponds to the other, all other parameters (x_R, z_R, v_R, v_T) remaining the same.

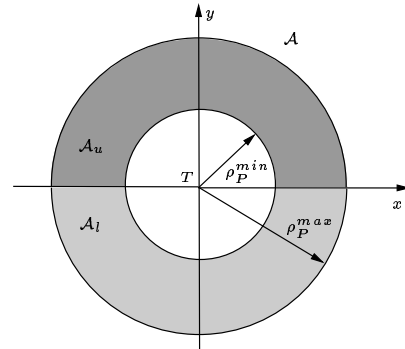


Fig. 12. The projection P of R is constrained to be located in annulus \mathcal{A} .

2) *Searching for P in annulus*: At this point, we have no choice but to test each point in, say, the upper search region \mathcal{A}_u . Given the annular shape of \mathcal{A}_u , \mathcal{A}_u is discretized using a polar grid. To find the location of P in \mathcal{A}_u , we need to explore all the points on the polar grid in \mathcal{A}_u . However,

to save computation time, we use *an adhoc* method that performs the search for the most likely value of $\theta_R = \theta_P$ at each radius ρ between ρ_R^{min} and ρ_R^{max} . The hope is that this value is close to the true ϵ . Note that there is no guarantee that this adhoc method will yield the right value of θ ! However, the method has worked perfectly in all cases tested. The algorithm for finding the estimates for ϵ and for the other unknown parameters is now described.

C. Curve fitting algorithm

First, using the positions $(\nu_{s,j}, \nu_{d,j})$ of the extracted peaks, we compute the estimates $\hat{S}_\nu = \hat{\nu}_s^{min} + \hat{\nu}_s^{max}$ and $\hat{P}_\nu = \hat{\nu}_s^{min} \hat{\nu}_s^{max}$, where $\hat{\nu}_s^{min} = \min_j \nu_{s,j}$ and $\hat{\nu}_s^{max} = \max_j \nu_{s,j}$.

Second, using the known value of R_b and the estimates \hat{S}_ν and \hat{P}_ν , we compute $\hat{\rho}_R^{min}$ according to Eqs. (17) and (15). Then, using the known values of R_b and H , we compute $\hat{\rho}_R^{max}$ according to Eq. (14). These values for $\hat{\rho}_R^{min}$ and $\hat{\rho}_R^{max}$ define the annulus \mathcal{A} .

Third, we proceed with one of the following approaches. The first approach is as follows. We discretize \mathcal{A}_u both angularly and radially. For each point $(x_P, y_P) \equiv (x_R, y_R)$, the estimate \hat{z}_R of z_R is found using the previously discussed relation linking (x_R, y_R) , \hat{S}_ν , R_b and H . We solve an estimation problem at each candidate point (x_R, y_R, \hat{z}_R) . The parameters to be estimated are α_R and δ . Parameter estimation is described below. This first approach is guaranteed to provide the correct solution, but it is slower than the second, which is as follows. We discretize \mathcal{A}_u only radially. We solve an ‘‘augmented’’ estimation problem at each candidate radius $\rho_P \equiv \rho_R$. Now, the unknown angular position $\theta_P \equiv \theta_R$ or, equivalently, ϵ , is included in the estimation problem. This approach is not guaranteed to provide the correct solution, but it is faster than the first. However, it has provided the correct solution in all cases tested.

Unknown parameters are estimated as follows. We express the fact that each of the data points $(\nu_{s,j}, \nu_{d,j})$ is on a BS DD curve parameterized with α_R , δ and, if applicable, θ_P . This is done repeatedly for each grid point or for each circle in \mathcal{A}_u . In each case, the parameters are found by LS estimation. Finally, we select the solution corresponding to the candidate (x_R, y_R, \hat{z}_R) with the smallest LS error.

X. PERFORMANCE COMPARISON

The performances of the methods of Table I are compared in terms of the SINRL. For reference, we also show the performances of the OP and of the straight-averaging processor (SAP). We assume omnidirectional sensors. Similar results are obtained with directional sensors. Figure 13(a)-(b) illustrates the performance of the AT method. The plots confirm (a) that TP-AT-BS performs nearly as well as OP and much better than SAP and (b) that AT-BS methods perform nearly identically in both their TP and EP modes. Figure 13(c)-(d)

similarly illustrates the performance of the WT method. As expected, WT provides better performance than AT, both in TP and EP modes.

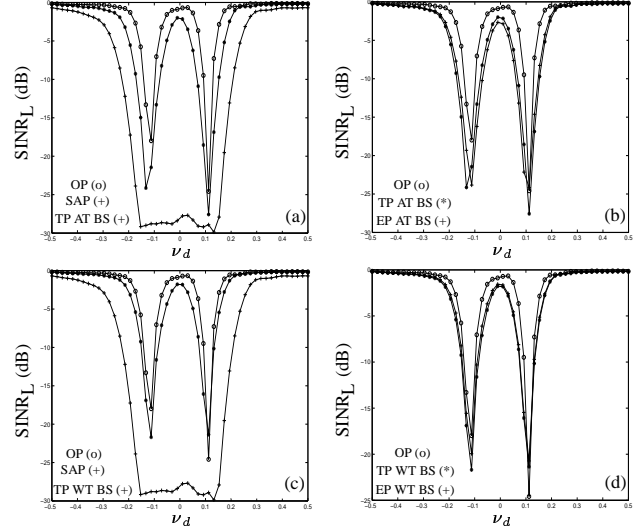


Fig. 13. Performance comparison of the methods of Table I in terms of $\text{SINRL}(\nu_d)$ at $\nu_s = 0$. Performances of the optimum processor (OP) and of the straight-averaging processor (SAP) are shown as references.

XI. CONCLUSIONS

We have examined the problem of range-dependence of the clutter ridge for BS STAP radars. New compensation methods based on the mathematical properties of DD curves have been proposed. Their performances were compared to those of the OP and SAP. Compensation is nearly perfect for all configurations, both in TP and EP modes. Compensation is achieved without increasing the number of degrees of freedom required for clutter rejection.

REFERENCES

- [1] R. Klemm, *Principles of space-time adaptive processing*, IEE Radar, Sonar, Navigation and Avionics, 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 Workshop, 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 Workshop, MIT Lincoln Laboratory, Lexington*, 13-14 March 2001.
- [7] F. Lapiere, 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. Meyer and S. Beucher, ‘‘Morphological segmentation,’’ *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 121–146, 1990.