RADIO SCIENCE, VOL. 43, RS6005, doi:10.1029/2008RS003842, 2008

A duct mapping method using least squares support vector machines

Rémi Douvenot,[1,2] Vincent Fabbro,[1] Peter Gerstoft,[3] Christophe Bourlier,[2] and Joseph Saillard[2]

Received 21 February 2008; revised 17 July 2008; accepted 24 July 2008; published 3 December 2008.

[1] This paper introduces a "refractivity from clutter" (RFC) approach with an inversion method based on a pregenerated database. The RFC method exploits the information contained in the radar sea clutter return to estimate the refractive index profile. Whereas initial efforts are based on algorithms giving a good accuracy involving high computational needs, the present method is based on a learning machine algorithm in order to obtain a real-time system. This paper shows the feasibility of a RFC technique based on the least squares support vector machine inversion method by comparing it to a genetic algorithm on simulated and noise-free data, at 1 and 5 GHz. These data are simulated in the presence of ideal trilinear surface-based ducts. The learning machine is based on a pregenerated database computed using Latin hypercube sampling to improve the efficiency of the learning. The results show that little accuracy is lost compared to a genetic algorithm approach. The computational time of a genetic algorithm is very high, whereas the learning machine approach is real time. The advantage of a real-time RFC system is that it could work on several azimuths in near real time.

Citation: Douvenot, R., V. Fabbro, P. Gerstoft, C. Bourlier, and J. Saillard (2008), A duct mapping method using least squares support vector machines, *Radio Sci.*, *43*, RS6005, doi:10.1029/2008RS003842.

1. Introduction

[2] In the low troposphere, anomalous propagation events occur owing to air movements and water evaporation. Survey radars usually work between 1 and 10 GHz in sea environment, and the detection range strongly depends on the atmospheric conditions at these frequencies. Anomalous propagation can cause "holes" in detection and either increase or decrease the radar detection range. This topic is especially critical for low-flying and floating targets for survey radar. Evaporation ducts, surface-based ducts, elevated ducts, and subrefractive layers are the identified anomalous propagation phenomena. These atmospheric structures are characterized by their vertical refractive index profile (Figure 1) [*Babin et al.*, 1997].

[3] The atmospheric conditions can be measured with radiosondes, rocket sondes, and buoys. Then the refractive index profile can be deduced from these measurements. Besides, refractometers can directly measure the refractive index. However, these techniques are difficult and expensive to implement. Atmospheric measurements also have a very low refreshment frequency and the delay between two measurements is too large for most situations. That is why "refractivity from clutter" (RFC) has been proposed to extract the refractivity directly from the radar clutter return [*Krolik and Tabrikian*, 1998; *Rogers et al.*, 2000; *Gerstoft et al.*, 2003a, 2003b; *Yardim et al.*, 2007]. Inferring the values of the modified refractivity profile from the sea clutter is a complex inverse problem because the relation between modified refractivity profile parameters and sea clutter is clearly nonlinear and ill-posed. Moreover, from a meteorological point of view, it implies several simplifying hypotheses to be solvable as described in section 2.

[4] The choice of the algorithm for the inversion is critical for RFC applications. Operational applications necessitate short computation time, less than ten minutes, to avoid error due to temporal evolution of refractivity [*Rogers*, 1996; *Douvenot et al.*, 2008]. Many fast nonlinear optimization methods exist. Neural networks,

[1] Département Électromagnétisme et Radar, Office National d'Etudes et de Recherches Aérospatiales, Toulouse, France.
[2] Institut de Recherche en Electrotechnique et Electronique de Nantes Atlantique, Radar Team, Ecole Polytechnique de l'Université de Nantes, Nantes, France.
[3] Marine Physics Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA.

0048-6604/08/2008RS003842$11.00

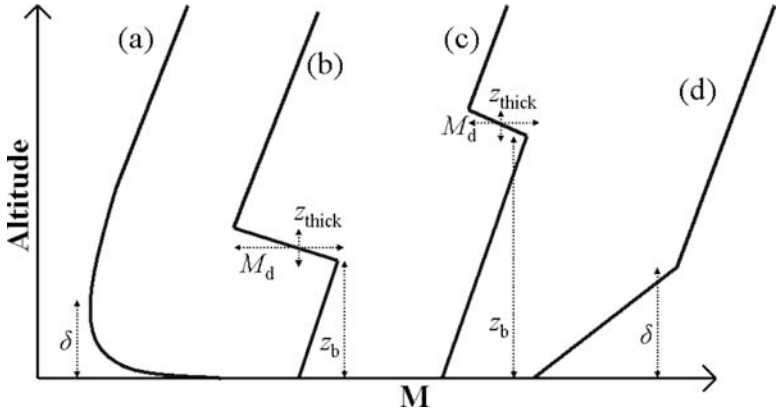

Figure 1. Modified refractivity profile and associated parameters for evaporation duct (curve a), surface-based duct (curve b), elevated duct (curve c), and subrefractive layer (curve d). For evaporation duct and subrefractive layer, δ is the duct height. For surface-based duct and elevated duct, z_b is the height of the duct base, M_d is the M deficit into the duct, and z_{thick} is the duct thickness.

Support Vector Machines (SVM), radial basis functions, and kriging are the most common [*Bartoli et al.*, 2006].

[5] In the present paper, a learning method using the least squares support vector machines algorithm (LS-SVM) [*Suykens et al.*, 2002] is proposed. The LS-SVM algorithm is based on a least squares optimization with linear constraints. This kind of problem is convex and avoids the local minima during the learning phase. Therefore, LS-SVM is preferred to neural networks and other optimization methods. The work presented here focuses on obtaining the modified refractivity profile from modeled propagation losses, without taking into account the value of the sea clutter radar cross section (RCS). This paper focuses on the feasibility of such a system, by comparing a LS-SVM method with a genetic algorithm (GA) [*Gerstoft et al.*, 2003a] in ideal conditions, with noiseless simulated propagation losses. RFC using LS-SVM method can provide a system giving a refractivity estimation in real time, which is important for operational conditions.

[6] First, the RFC is introduced and the parameterization of the refractivity profile is presented. Second, the GA and LS-SVM processes are briefly exposed. The two inversion methods for RFC applications are compared in terms of accuracy and speed, and finally, the results are discussed.

2. Definitions and Hypotheses

2.1. Atmospheric Ducts

[7] When modeling the atmospheric ducts, the modified refractivity M is preferred to the refractive index n. It is expressed as

$$M = (n - 1) \times 10^6 + 0.157z, \qquad (1)$$

where z is the altitude in meter. As the refractive index is very close to unity, the difference from 1 is highlighted in M. The second term corrects for the Earth curvature. A negative gradient of modified refractivity implies ducting conditions, subrefractive condition appears for a gradient superior to 0.127 M unit/m [*Yardim et al.*, 2007], and standard atmosphere corresponds to 0.118 M unit/m.

[8] Some meteorological conditions can cause abnormal propagation. The refractivity is modified by temperature and/or humidity gradients. There are four identified types of duct [*Turton et al.*, 1988] (Figure 1): curve a, evaporation duct; curve b, surface-based duct; curve c, elevated duct; and curve d, subrefractive layer. Elevated ducts are similar to surface-based ducts with a higher duct base. They are not taken into account in RFC because they have low influence on the power losses at the sea level. Actually, the waves trapped into an elevated duct are bent at high altitude and reach the sea surface beyond the maximum radar range considered for RFC. The subrefractive layer is a rare event and measurements are lacking to be accurately characterized, hence it is not taken into account in this study.

[9] The refractive conditions are considered constant with the distance for the entire propagation path. This hypothesis seems valid in open sea but could be a strong simplification in coastal environment, where the meteorological conditions may highly vary [*Kerr*, 1987]. A distance-varying refractive index could be introduced in further work. *Rogers* [1996] shows that the RMS error in propagation factor exceeds 6 dB or more after 30 min due to temporal decorrelation. Thus, a practical RFC system has to update environment parameter estimates much faster than 30 min. This calls for a simple range-independent model that can be estimated fast.

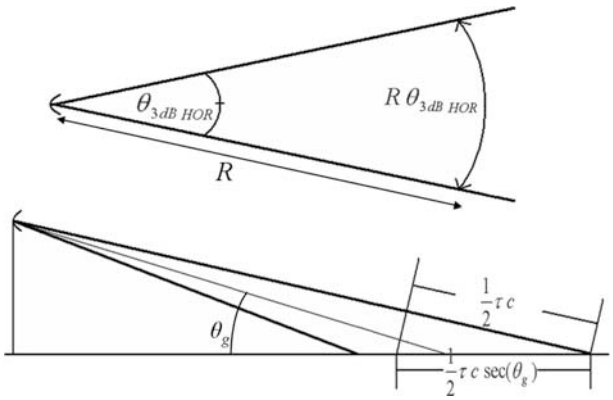

Figure 2. The area illuminated by the radar (top) from above and (bottom) from the side. The parameter τ is the radar pulse length, c is the velocity of wave propagation, R is the distance from the radar to the sea pixel, $\theta_{3dB\,HOR}$ is the horizontal antenna aperture, θ_g is the grazing angle, and sec is the secant function.

[10] The evaporation duct is a quasi-permanent event above the sea. It is due to a strong humidity gradient at the sea surface, which causes a refractive index gradient as well. The worldwide mean of the duct height δ, defined as the height above the sea mean level at which dM/dz = 0 M unit/m (Figure 1, curve a), is about 10 m. The duct height sometimes reaches 30 m and exceptionally 40 m [*Anderson*, 1989; *Patterson*, 1992]. If the value of δ is considered as constant with the distance, performing an inversion in the case of an evaporation duct is not a hard problem because δ is the only parameter to retrieve. A simple inversion method as gradient descent or quadratic regression could be used [*Rogers et al.*, 2000].

[11] On the contrary, the surface-based duct is less frequent, but its effect on propagation is much stronger. It highly increases the radar range or creates detection "holes." Its occurrence is about 15% of the time worldwide but until 50% of the time in the Persian Gulf [*Patterson*, 1992]. It is often due to an advection movement of a relatively warm and dry air, which forms a strong negative gradient of modified refractivity. The surface-based duct is typically modeled as a trilinear profile (Figure 1, curve b) [*Gossard and Strauch*, 1983], where z_b is the height of the duct base, M_d is the M deficit into the duct, and z_{thick} is the duct thickness. The following limits for the duct parameters are chosen from the work of *Gerstoft et al.* [2003a]: z_b varies from 0 to 300 m, M_d from 0 to 100 M unit, and z_{thick} from 0 to 100 m. The obtained profile is trilinear with slope discontinuities at the basis and at the top of the duct. A continuous arctangent-shaped surface-based duct model has been proposed by *Webster* [1983] to avoid discontinuities in the model. However, nor the trilinear model neither the arctangent-shaped model are based on a physical description and a model cannot be presented as better than the other. Moreover, these two types of ducts have very close effects on the wave propagation (see Appendix A). We finally choose the trilinear model, which is simpler to use. Actually, the duct strength can be linked to the parameters of the trilinear duct through the expression [*Turton et al.*, 1988]

$$\lambda_{max} = \frac{2C_{sbd} \times \sqrt{M_d} \times z_{thick}}{3}, \tag{2}$$

where C_{sbd} is a constant equal to 3.77×10^{-3} for surface-based ducts and λ_{max} is the maximum wavelength trapped into the duct in meter.

[12] Evaporation ducts are modeled using one or two parameters and are quite simple ducts to retrieve. The LS-SVM RFC method has been applied on real data in the presence of evaporation ducts [*Douvenot et al.*, 2007]. Therefore, surface based ducts are considered to test and validate the feasibility of an RFC system including the inversion method proposed in this paper. More complex and realistic profiles, as "double ducts" composed of an evaporation duct and a surface-based duct simultaneously, are to be studied in further works.

2.2. Sea Clutter Model

[13] In RFC technique, the propagation losses L are processed to determine the environmental parameters $\mathbf{M} = (z_b, M_d, z_{thick})$. In this paper, the refractive conditions are determined from simulated propagation losses. However, in operational conditions, only the power $P_R(\mathbf{M})$ observed by the radar is available. The received power has to be processed in order to isolate the propagation losses. The received power can be written (from the radar equation) [*Kerr*, 1987, p. 473]

$$P_R(\mathbf{M}) = P_E G^2 \frac{4\pi}{\lambda^2} \sigma L^2(\mathbf{M}), \tag{3}$$

where the emitted power P_E, the wavelength λ, and the antenna gain G are known. Only the radar cross section (RCS) of the sea clutter σ is needed to obtain the propagation losses L from the received power. Now, the RCS σ is the product of the illuminated area A_I by the Normalized RCS (NRCS) σ_0. The illuminated area can be expressed as (Figure 2)

$$A_I = \frac{1}{2} R \theta_{3dB\ HOR} \tau c \sec(\theta_g), \tag{4}$$

where τ is the radar pulse length, c is the velocity of wave propagation, R is the distance from the radar to the sea pixel, $\theta_{3dB\,HOR}$ is the horizontal antenna aperture, θ_g is the grazing angle, and sec denotes the secant function. The range resolution of the radar is

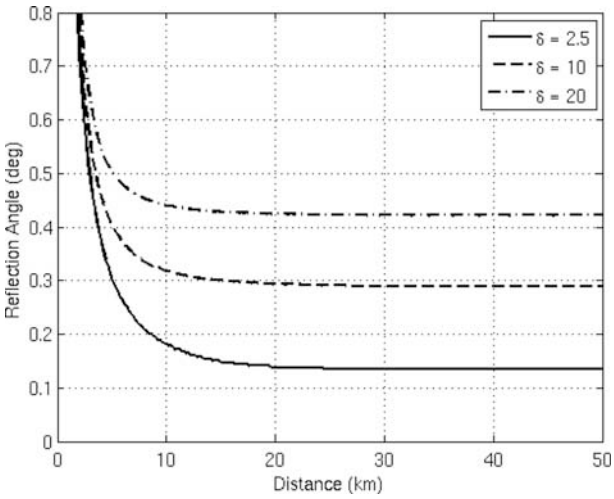

Figure 3. Reflection angle with respect to the distance in the presence of an evaporation duct for δ = 2.5, 10, and 20 m. Antenna is at a height of 25 m.

$(1/2)\tau c$ [*Kerr*, 1987]. Note that A_I is proportional to R if θ_g is constant.

[14] The problem is now to choose a model for the NRCS of the sea clutter σ_0. It depends on several meteorological parameters and on the grazing angle θ_g [*Feng et al.*, 2005]. However, the grazing angle is almost constant beyond the horizon in the presence of a range-independent evaporation duct [*Paulus*, 1990].

[15] For instance, with an antenna at a height of 25 m, the grazing angle hardly fluctuates beyond 10 km (Figure 3) for each refractivity condition. This result is obtained using ray tracing in heterogeneous and vertically stratified atmosphere [*Paulus*, 1990; *Rogers et al.*, 2000]. The incident angle of the electromagnetic wave θ_g is determined at sea level. Considering the sea roughness as uniform with the distance and the incident grazing angle of the trapped wave also as constant at the sea surface, σ_0 can be considered as constant beyond the horizon. Hence the model $\sigma = C^t R$, where C^t is a constant, could be relevant. Note that these results are obtained for range-independent evaporation ducts. Dependence with respect to the grazing angle [*Ward et al.*, 2006] could be introduced for the study of complex refractivity variations.

[16] Once the propagation losses L are known, an inversion algorithm is needed in order to retrieve the parameters of the modified refractivity profile.

3. Inversion Methods

[17] Retrieving the parameters of the ducts from the propagation losses is a complex problem. Two different inversion methods are presented in this paper. The first is a GA included in the SAGA code developed by *Gerstoft et al.* [2003a, 2003b]. The second, the LS-SVM, is a learning machine selected for its computational speed once trained. This latter method is based on a pregenerated and preprocessed database. As the main part of the computation is made prior to the operational use, the inversion using LS-SVM is real time.

3.1. Genetic Algorithm

[18] Genetic algorithms (GA) start with the selection of a population of q member models. Models consist of bit strings for each uncertain/unknown parameters (parameters are, thus, discretized in GA, in contrast to Simulated Annealing methods, which usually work with continuous variables).

[19] The ''fitness'' of each member is the value of the objective function for the particular model. On the basis of the fitness of the members, ''parents'' are selected and through a randomization, a set of ''children'' is produced. These children replace the least fit of the original population and the process iterates to develop an overall fitter population. The formation of child models is performed through the application of operators to the parents.

[20] Figure 4 shows the GA principle. Each child population P_{i+1} is processed in three steps: first, the more likely ''parents'' are selected in the parent population P_i. Then, crossover uses a part of the string corresponding to a parameter from one parent and supplements it with a part of the string for the same parameter from the other parent. The operation is applied individually to every parameter string (multipoint crossover), resulting in all-direction parameter perturbations.

[21] Mutation follows crossover and changes bit values in parameter strings in a random fashion. Bit changes occur with a low probability (usually 0.05). The small changes imposed on the new generation through the occasional bit changes assist the optimization process to escape from local minima. These three steps are applied on each population to obtain a final population containing an overall fitter population P_k. A more detailed description of genetic algorithms and their application to parameter estimation is given by *Gerstoft* [1994].

3.2. LS-SVM

[22] In the late 1990s, *Suykens and Vandewalle* [1999] developed LS-SVM from Vapnik's work on support vector machines (SVM) [*Vapnik and Lerner*, 1963; *Vapnik and Chervonenkis*, 1964; *Vapnik*, 1995]. This special case (quadratic loss function) simplifies the SVM theory formulation as exposed by *Smola and Schölkopf* [1998]. Moreover, the authors made great efforts to unify the theories of LS-SVM, neural net-

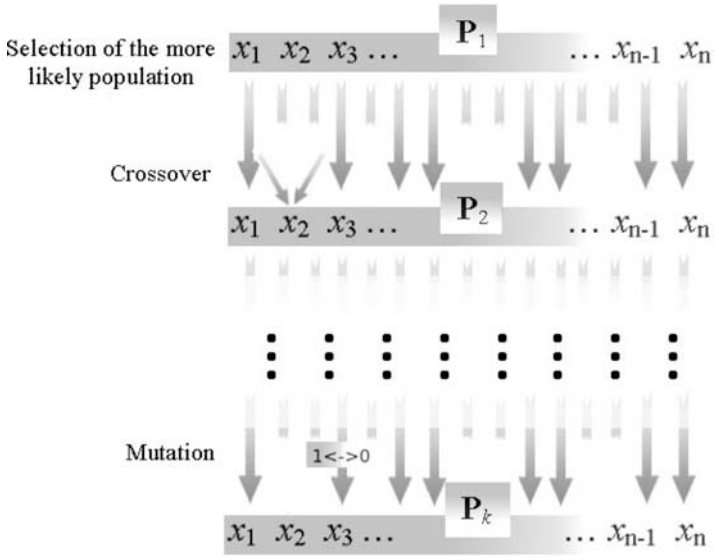

Figure 4. Scheme of the GA process. $\mathbf{P}_i$ are the populations, and $(x_1, \ldots, x_n)$ are the elements of each population.

works, optimization, and Gaussian processes. Following *Suykens et al.* [2002], the theory is outlined below.

[23] LS-SVM is a training process. The aim of the training process is to obtain an approximation of the nonlinear function f with respect to the vector of the propagation losses $\mathbf{L}$ at different ranges. The vector of the duct parameters $\mathbf{M}$ is the output: $\mathbf{M} = f(\mathbf{L})$.

[24] Figure 5 displays the process to generate the N element training database. First, N items of three-dimensional parameter sets $\mathbf{M}^{tr}$ describing N different surface-based ducts are drawn using a Latin hypercube sampling [*McKay et al.*, 1979]. Then, from these parameter values, N propagation losses vectors $\mathbf{L}^{tr}$ are obtained by carrying out N propagation simulations using the PWE solved by split-step Fourier (SSF) propagation method [*Barrios*, 1992]. This wave propagation method is accurate and takes into account the refractive index variations. Thereby, the training database is the set $(\mathbf{M}^{tr}, \mathbf{L}^{tr})$.

[25] Latin hypercube can generate a multidimensional training database more efficiently than using a regular sampling of each parameter z_b, M_d, and z_{thick} [*Loh*, 1996]. This method is extracted from the design of experiments theory [*Vivier*, 2002] and generalizes Leonard Euler's Latin square. Here is the principle of the generation of the N-sized three-dimensional database of the duct parameters as illustrated in Figure 6. In the three-dimensional space, the variations of the variables (z_b, M_d, z_{thick}) are represented from minimum to maximum value on each dimension. Each interval is divided into N equal sections. For the first draft, an interval is randomly drawn among the N intervals on the 3 dimen-

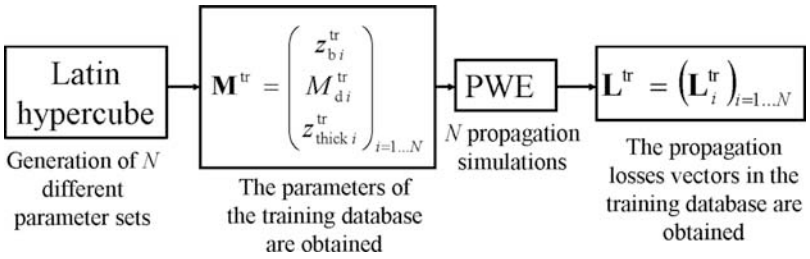

Figure 5. Process to generate the N element training database $(\mathbf{M}^{tr}, \mathbf{L}^{tr})$ by Latin hypercube sampling.

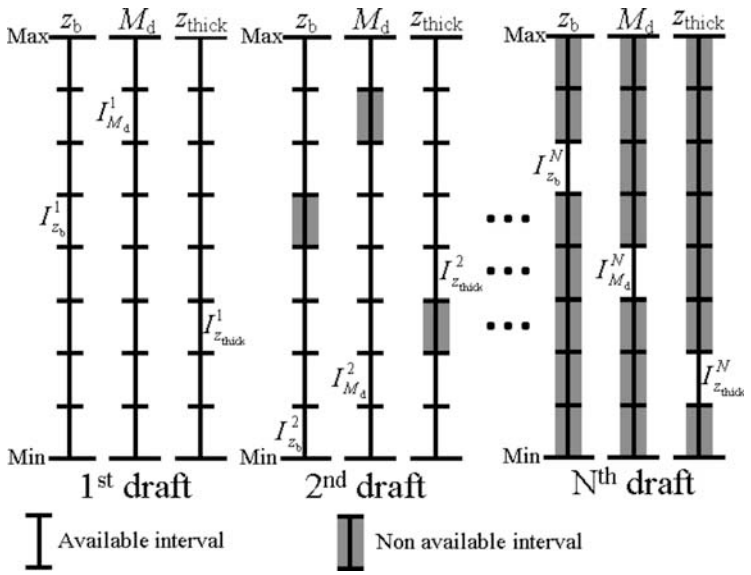

Figure 6. Illustration of the Latin hypercube sampling in the three-dimensional space of the trilinear surface-based duct parameters (z_b, M_d, z_{thick}) for $N = 8$.

sions. The intervals $I^1_{z_b}$, $I^1_{M_d}$, $I^1_{z_{thick}}$ are obtained. Then a vector value (z^1_b, M^1_d, z^1_{thick}) is randomly drawn into these intervals. For the second draft, the intervals $I^2_{z_b}$, $I^2_{M_d}$, $I^2_{z_{thick}}$ are drawn among the $N \times 3$ intervals deprived of the intervals I^1_{zb}, I^1_{Md}, I^1_{zthick}. The process is repeated N times until the last remaining intervals $I^N_{z_b}$, $I^N_{M_d}$, $I^N_{z_{thick}}$. All random distributions are uniform.

[26] Now that the database is generated, it has to be processed to obtain a nonlinear approximation of the aimed function. In learning theory, this step is the training of the system. Once the training database is generated, the system must be trained in order to obtain an approximation of the function f in the form

$$M_{\text{output}} = \sum_{j=1}^{N} \alpha_j \exp\left(-\frac{\|\mathbf{L}_{\text{input}} - \mathbf{L}_j^{\text{tr}}\|^2}{\sigma_K^2}\right) + B, \qquad (5)$$

where σ_K^2 is the width of the Gaussian function. It is a parameter of the inversion system. The support vectors $\alpha = (\alpha_j)_{j=1\ldots N}$ and the bias B are the values to optimize. The optimization is carried out during the training of the system, where the function is tested on the training database. M_{output} is scalar and represents z_b, M_d, or z_{thick}. So there is a function f for each parameter.

[27] Figure 7 shows the training process of LS-SVM. The system ''learns'' the best approximation of the function f by optimizing α and B on the training database itself. The Gram matrix $\Omega_{ij} = \exp(-\|\mathbf{L}_i^{\text{tr}} - \mathbf{L}_j^{\text{tr}}\|^2/\sigma_K^2)$, with $i, j \in \{1, \ldots, N\}$, is introduced. The optimized support vectors and bias are computed by solving the system

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & \mathbf{\Omega} + \gamma \mathbf{I_d} \end{bmatrix} \times \begin{bmatrix} B^{\text{opt}} \\ \boldsymbol{\alpha}^{\text{opt}} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{M}^{\text{tr}} \end{bmatrix}, \qquad (6)$$

where $\gamma \in]\, 0, +\infty\, [$ is the second parameter of the inversion algorithm: γ defines the trade-off between the accuracy on the training database and the ability of the function to find solutions outside the training database. It is called the regulation parameter; $\boldsymbol{\alpha}^{\text{opt}}$ are the support vectors, and B^{opt} is the bias of the function at the optimum.

[28] In theory, the propagation losses are mapped through a higher dimensional space using the Gaussian function in (5) and (6) in order to mimic the nonlinearity of the function f. Note that this Gaussian function can be replaced by another kernel function [*Mercer*, 1909]. The final system (6) is a Karush-Kuhn-Tucker system obtained by solving a ridge regression system as an optimization system under equality constraints. The ridge regression is carried out on the whole training data in order to obtain an approximation of the aimed function. For details, see *Suykens et al.* [2002].

[29] Determination of α^{opt} and B^{opt} in (5) is the training of the system. It requires inverting the $N \times N$ matrix $\mathbf{\Omega} + \gamma \mathbf{I_d}$. Once the system is trained, the inversion process is fast, less than 1 s. For an observed vector of propagation losses $\mathbf{L}_{\text{input}}$, the equation (5) to determine the duct parameter M_{output} is very fast. A nonlinear and

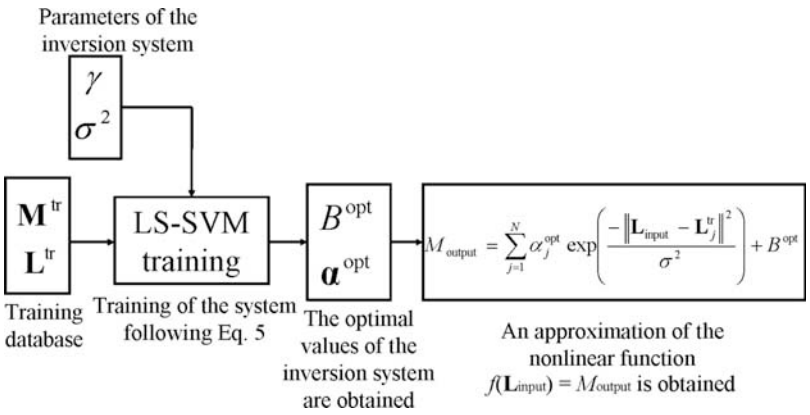

Figure 7. Scheme of the training process of LS-SVM algorithm.

real-time approximation of the aimed function $f(\mathbf{L}_{input}) = M_{output}$ is obtained.

4. Inversion Results

4.1. Comparison Layout

[30] The comparison of the inversion methods GA and LS-SVM for RFC is carried out on two simulated data sets. The first consists of propagation losses for 20 environments at 1 GHz. The antenna is set at a height of 50 m and its beam width is 30°. The second one is made on 10 environments at 5 GHz. The antenna is set at a height of 25 m, and its beam width is 4.6°. The values of the duct parameters are drawn randomly in the limits mentioned in section 2.1, i.e., $z_b \leq 300$ m, $M_d \leq 100$ M unit, $z_{thick} \leq 100$ m. The inversion is carried out on

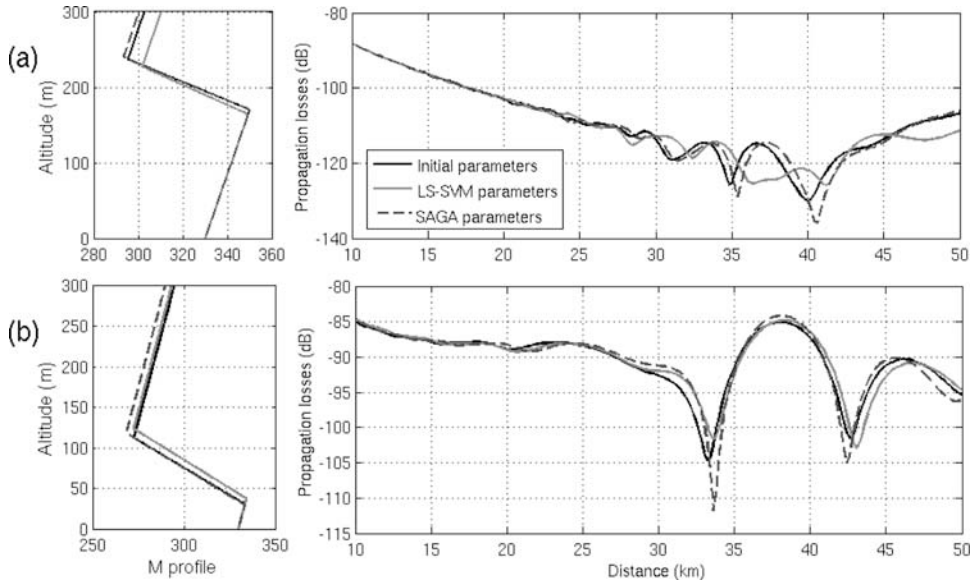

Figure 8. Two inversion cases at 1 GHz. (right) The propagation losses above the sea before inversion computed with the true parameters and after inversion computed with LS-SVM and SAGA are plotted. (left) Their associated profiles with respect to the altitude are depicted. True parameters are (a) $z_b = 171$ m, $M_d = 55$ M unit, and $z_{thick} = 67$ m and (b) $z_b = 31$ m, $M_d = 61$ M unit, and $z_{thick} = 82$ m.

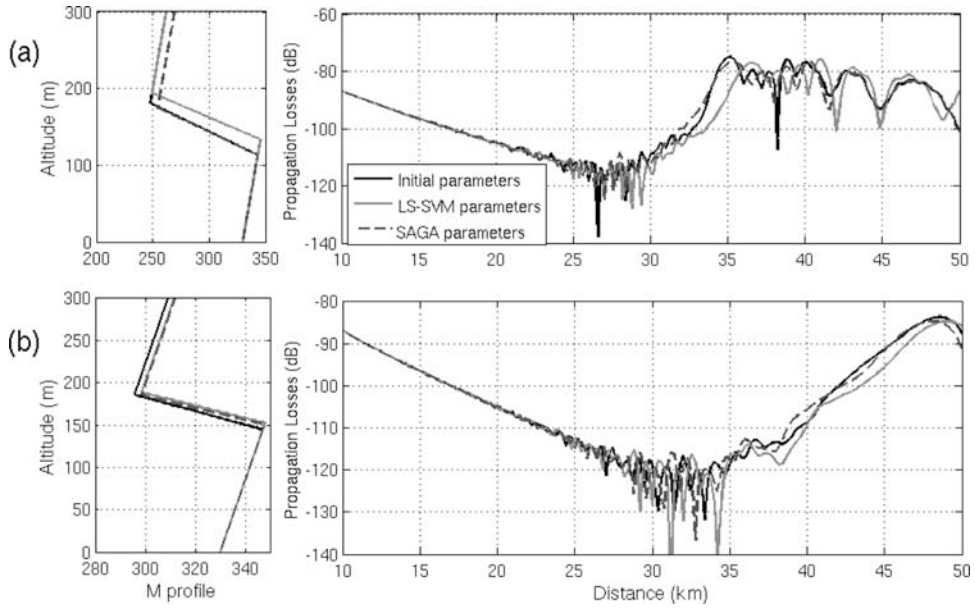

Figure 9. Two inversion cases at 5 GHz. (right) The propagation losses above the sea before inversion computed with the true parameters and after inversion computed with LS-SVM and SAGA are plotted. (left) Their associated M profiles with respect to the altitude are depicted. True parameters are (a) z_b = 114.05 m, M_d = 95.39 M unit, and z_{thick} = 67.72 m and (b) z_b = 145.86 m, M_d = 50.92 M unit, and z_{thick} = 40.92 m.

simulated propagation losses at the sea surface in the presence of constant surface-based ducts along the pathway in order to retrieve the duct parameters $\mathbf{M} = (z_b, M_d, z_{thick})$. For both GA and LS-SVM, the inversion is carried out between 10 and 50 km with 1 point each 200 m. Thus, each set of propagation losses contains 201 points.

[31] The GA software SAGA [*Gerstoft*, 2006] is delivered with the propagation code TPEM (terrain parabolic equation model) based on the PWE (parabolic wave equation) solved by SSF (split-step Fourier) [*Barrios*, 1992]. This method assumes a two-dimensional propagation, no backscattering, and perfectly conducting smooth sea surface, which is reasonable for very low grazing angles [*Fabbro et al.*, 2006]. Most of the default settings are kept for SAGA. Concerning the GA itself, 10 populations are launched with 2000 iterations each. Half of the population is kept at each iteration. The crossover rate is 50% and the mutation rate is 5%.

[32] For LS-SVM, a Gaussian kernel is chosen [*Smola and Schölkopf*, 1998], as described in (5) and (6). The training database is generated using a three-dimensional Latin hypercube [*McKay et al.*, 1979] of 10,000 environments for the parameters sets (z_b, M_d, z_{thick}). The corresponding propagation losses are modeled using a PWE propagation method solved by SSF.

4.2. Results

[33] Samples of the inversion results are shown in Figures 8 and 9. All the propagation losses have been computed by PWE, considering different sets of refractivity parameters. On Figures 8 and 9, "true parameters" indicates the propagation losses obtained with the true refractivity parameters. "LS-SVM parameters" indicates the propagation losses obtained with the refractivity parameters retrieved by inversion using the LS-SVM algorithm. "SAGA parameters" is the same with the parameters retrieved by GA. The associated profiles are also plotted.

[34] To quantify the comparison, the difference between the true propagation losses $\mathbf{L}^{true} = (L_i^{true})_{i=1\ldots201}$ and the propagation losses obtained considering inverted refractivity parameters has been computed every 200 m. Finally, the mean of the absolute difference in dB on the 201 points is calculated for each method:

$$\Delta\overline{\mathbf{L}}^{inv} = \frac{1}{201}\sum_{i=1}^{201}\left|L_i^{inv} - L_i^{true}\right|, \qquad (7)$$

where the superscript "inv" is the name of the inversion method (LS-SVM or GA). This result is finally averaged on all the tested cases to get a global error estimation $\Delta\overline{\mathbf{L}}^{\mathrm{inv}}_{\mathrm{mean}}$.

[35] The results of LS-SVM and GA inversions are summarized as follows: (1) at 1 GHz, $\Delta\overline{\mathbf{L}}^{\mathrm{LSSVM}}_{\mathrm{mean}}$ = 2.38 dB and $\Delta\overline{\mathbf{L}}^{\mathrm{GA}}_{\mathrm{mean}}$ (mean on 20 cases); and (2) at 5 GHz, $\Delta\overline{\mathbf{L}}^{\mathrm{LSSVM}}_{\mathrm{mean}}$ = 2.71 dB and $\Delta\overline{\mathbf{L}}^{\mathrm{GA}}_{\mathrm{mean}}$ = 1.54 dB (mean on 10 cases).

[36] The two algorithms are fundamentally different and this is observed in the results. GA is more accurate than LS-SVM, but GA uses significantly more CPU time. The mean of the computation time for the GA is $T^{\mathrm{GA}}_{\mathrm{mean}} \approx 90$ min at 1 GHz and $T^{\mathrm{GA}}_{\mathrm{mean}} \approx 1100$ min at 5 GHz. On the same computer, for the LS-SVM algorithm, $T^{\mathrm{LSSVM}}_{\mathrm{mean}} \leq 0.2$ s for both frequencies. Actually, as shown in section 3.2, LS-SVM only requires the calculus of the kernel function, whereas the GA requires a propagation computation, longer at 5 GHz than at 1 GHz, for each forward model. For LS-SVM approach, all the propagation computations have been preprocessed during the training step.

[37] Event though the GA gives more accurate results, the LS-SVM gives a good idea of the wave propagation and of the atmospheric duct as shown in Figure 8 (for 1 GHz) and Figure 9 (5 GHz). The one-way propagation losses above the sea are shown for the true surface-based duct parameters and for the parameters retrieved by inversion with the LS-SVM algorithm and with the GA. Then the associated refractive profiles are plotted. The two cases are shown at each frequency. The propagation losses obtained with the GA parameters fit the original ones. However, apart from small discrepancies, the LS-SVM curve follows the original one. Moreover, the ducts retrieved by LS-SVM are close to the original ones. Such accuracy might be sufficient to characterize the environment effect on propagation in order to predict trapping phenomenon and to detect shadow zones in the radar cover as the duct itself is well approximated by LS-SVM.

5. Summary and Conclusion

[38] In this paper, the feasibility of a RFC system using a learning machine has been shown on simulated data. The simplifying hypotheses have first been explained, then the method has been compared to a GA RFC system on noiseless simulated data. The results show a minor reduction in accuracy but a great improvement in computation time. Our system is especially interesting in higher frequencies, around 5 GHz, when the computation time of the GA-type RFC systems are very slow. The main interest of the LS-SVM learning machine is to make all the main calculations before the inversion is started. Then the obtained RFC system could work in real time.

[39] The LS-SVM inversion method is not as precise as the GA one. Yet, the system gives a good idea of the electromagnetic wave behavior and of the refractive index profile. Both methods are less precise at 5 GHz as the wave is more sensitive to refractive index variations.

[40] There are several ways to improve the RFC using learning machine: the first means is to improve the very inversion algorithm. An interesting way is to find a fitter kernel than the Gaussian kernel to the RFC problem. However, creating a kernel for a specific problem is a difficult task. Another possible improvement is the introduction of a correlation between the different dimensions of the duct during the inversion. The use of a multitask learning algorithm [*Argyriou et al.*, 2006] could improve the inversion by inverting the three duct parameters at once. To improve the inversion, this method could also be hybridized with a GA-type inversion: the LS-SVM inversion could give a first approximation of the duct parameters, and then the GA could refine the inversion. Therefore, the inversion could have the accuracy of GA with a lower computation time.

[41] Another way to improve RFC using a learning machine is to optimize the training database. Indeed, as LS-SVM is a learning machine, the generation of the training database is critical for the efficiency of the algorithm. To generate an efficient training database, one must describe at best all the possible cases. Thus, large and precise meteorological data could permit to refine the values of the parameters describing the duct in the training database. The more precise the database, the more likely the inversion is to succeed. Moreover, the choice of the sampling method is decisive. A sampling method better than Latin hypercube sampling could be applied for RFC.

[42] If the LS-SVM inversion is not precise enough to perfectly retrieve the refractivity profile, the advantage of this method is the real-time working. In operational conditions, the radar can observe several azimuths at several times. The obtained results could be smoothed over time and space in order to correlate the results. Having data at several frequencies or several elevation angles could also be useful.

Appendix A

[43] The debate of surface-based duct modeling is still open. One point of interest is the discontinuities in the vertical profile. To discuss the choice of a trilinear surface-based duct profile, which implies discontinuities, a comparison is carried out with a continuous arctangent-shaped profile, as advised by *Webster* [1983]. This

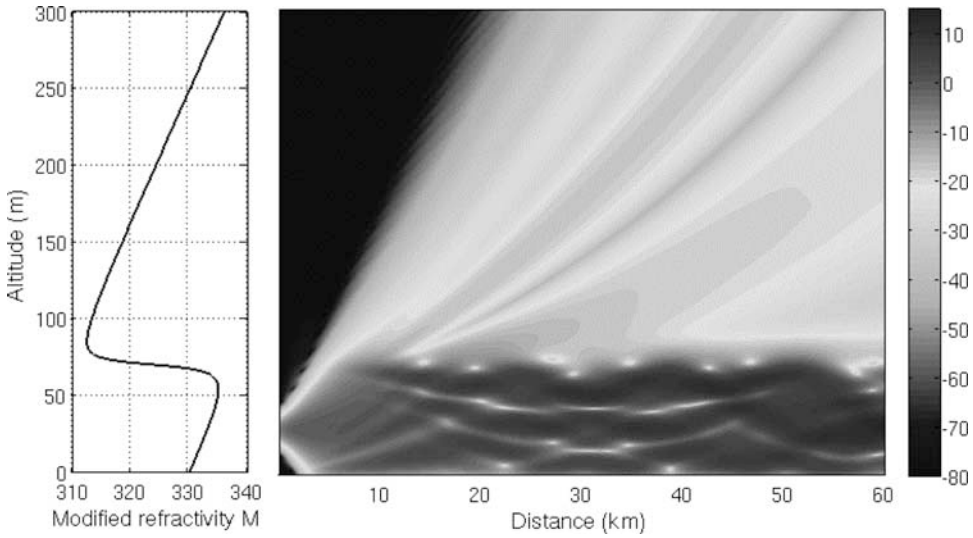

Figure A1. (left) The modified refractivity with respect to the altitude for an arctangent-shaped duct. (right) The 2-D associated propagation factor with respect to the distance and altitude.

comparison has been made at 2.84 GHz with an antenna at a height of 30.78 m.

[44] Figure A1 (left) represents the modified refractivity with respect to altitude for an arctangent-shaped duct, and Figure A1 (right) represents the 2-D map of the associated propagation factor with respect to the distance and altitude. Figure A2 is the same as the Figure A1 considering a trilinear duct instead of an arctangent-shaped duct. The trilinear duct is chosen very similar to the arctangent-shaped duct to highlight the effect of the discontinuities in the refractivity profile. Only small discrepancies can be observed between the two propagation factor maps. Globally, the discontinuities have no significant effect on the propagation factor.

[45] Figure A3 compares the propagation factor with respect to the distance at sea level obtained with the same

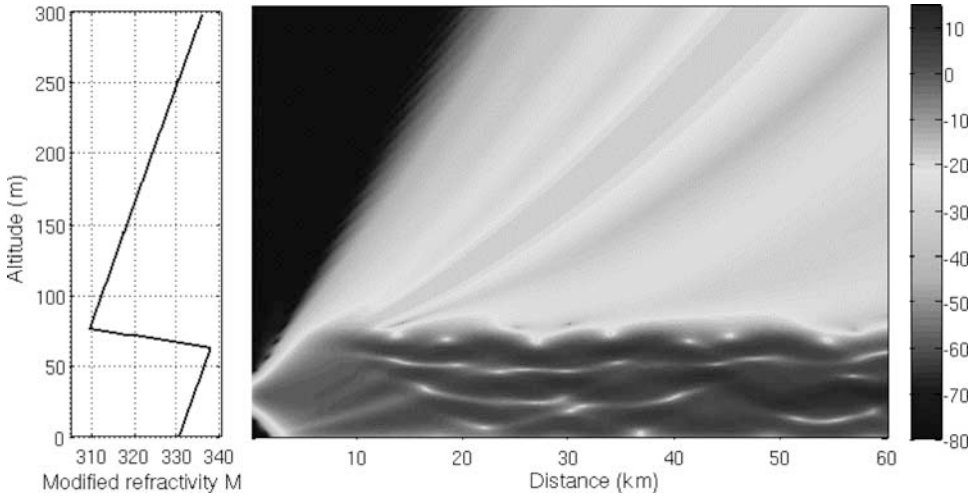

Figure A2. (left) The modified refractivity with respect to the altitude for a trilinear duct as close as possible to the arctangent-shaped duct from the Figure A1. (right) The 2-D associated propagation factor with respect to the distance and altitude.

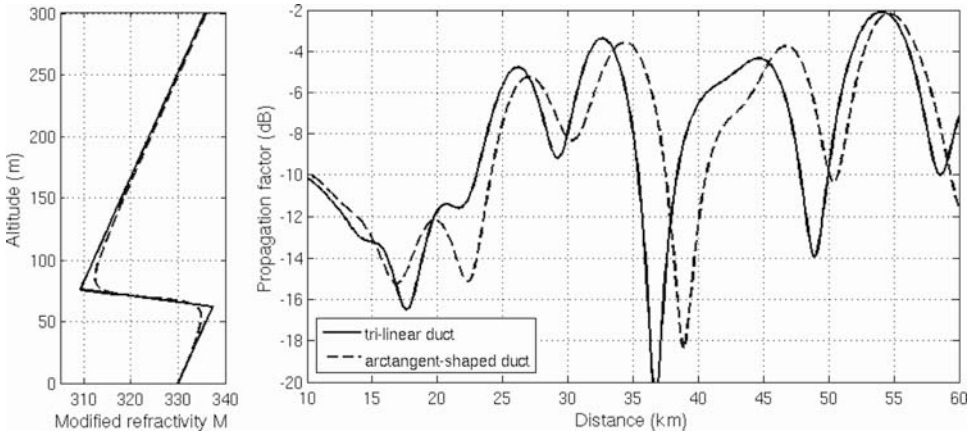

Figure A3. (left) The modified refractivity with respect to the altitude for the trilinear duct (continuous line) and for the arctangent-shaped duct (dotted line) from the Figures A1 and A2. (right) The associated propagation factors at sea level with respect to the distance.

arctangent-shaped duct (dotted line) and trilinear duct (continuous line). There are some discrepancies between the two propagation factors because the two refractivity profiles are not exactly the same. However, the dynamic of the propagation factors at sea level are very similar. Also, the discontinuities in the trilinear duct have not a great impact on the wave propagation.

[46] Note that the aim of RFC is not to give the exact refractivity profile, but to propose a "generic" model able to render an accurate approximation of the real atmospheric conditions.

[47] **Acknowledgments.** The authors would like to thank the French DGA-CNRS (Délégation Générale de l'Armement, Centre National de la Recherche Scientifique) for the funding.

References

Anderson, K. D. (1989), Radar measurements at 16.5 GHz in the oceanic evaporation duct, *IEEE Trans. Antennas Propag.*, *37*(1), 100–106, doi:10.1109/8.192171.

Argyriou, A., T. Evgeniou, and M. Pontil (2006), Multitask feature learning, in *Advances in Neural Information Systems Processing Systems*, *Proc. NIPS Conf.*, *19*, 41–48.

Babin, S. M., G. S. Young, and J. A. Carton (1997), A new model of the oceanic evaporation duct, *J. Appl. Meteorol.*, *36*, 193–204, doi:10.1175/1520-0450(1997)036<0193:ANMOTO>2.0.CO;2.

Barrios, A. E. (1992), Parabolic equation modeling in horizontally inhomogeneous environments, *IEEE Trans. Antennas Propag.*, *40*(7), 791–797, doi:10.1109/8.155744.

Bartoli, N., M. Samuelides, D. Bailly, and M. Marcelet (2006), *Revue de modèles réduits pour l'optimisation, Tech. Rep. RT 2/11576 DPRS/DTIM*, Off. Natl. d'Etudes et de Rech. Aérosp., Toulouse, France.

Douvenot, R., V. Fabbro, H.-H. Fuchs, H. Essen, C. Bourlier, and J. Saillard (2007), Retrieving evaporation duct heights from measured propagation factors, paper presented at Radar 2007, Inst. of Eng. and Technol., Edinburgh, U.K.

Douvenot, R., V. Fabbro, C. Bourlier, J. Saillard, H.-H. Fuchs, and H. Essen (2008), Refractivity from sea clutter applied on VAMPIRA and Wallops'98 data, paper presented at Radar 2008, Inst. of Eng. and Technol., Adelaide, Australia, 2–5 Sept.

Fabbro, V., C. Bourlier, and P. F. Combes (2006), Forward propagation modeling above Gaussian rough surfaces by the parabolic wave equation: Introduction of the shadowing effect, *Progress Electromagn. Res.*, *58*, 243–269, doi:10.2528/PIER05090101.

Feng, S., J. Chen, and X. Y. Tu (2005), Low angle reflectivity modeling of sea clutter using LS method, *IEEE Trans. Antennas Propag. Soc. Int. Symp*, *2B*, 187–190, doi:10.1109/APS.2005.1551969.

Gerstoft, P. (1994), Inversion of seismo-acoustic data using genetic algorithms and a posteriori probability distributions, *J. Acoust. Soc. Am.*, *95*, 770–782, doi:10.1121/1.408387.

Gerstoft, P. (2006), SAGA User Manual 5.3: An Inversion Software Package, an updated version of "SAGA Users Guide 2.0, an Inversion Software Package", manual, SACLANT, Undersea Res. Centre, La Spezia, Italy.

Gerstoft, P., L. T. Rogers, J. L. Krolik, and W. S. Hodgkiss (2003a), Inversion of refractivity parameters from radar sea clutter, *Radio Sci.*, *38*(3), 8053, doi:10.1029/2002RS002640.

Gerstoft, P., L. T. Rogers, W. S. Hodgkiss, and L. J. Wagner (2003b), Refractivity estimation using multiple elevation angles, *IEEE J. Oceanic Eng.*, *28*(3), 513–525, doi:10.1109/JOE.2003.816680.

Gossard, E. E., and R. G. Strauch (1983), *Radar Observation of Clear Air and Clouds*, Elsevier Sci., Amsterdam.

Kerr, D. E. (1987), *Propagation of Short Radio Waves*, *IEE Electromagn. Waves Ser.*, vol. 24, Peter Peregrinus, London.

Krolik, J. L., and J. Tabrikian (1998), Tropospheric refractivity estimation using radar clutter from the sea surface, in *Proceedings of the 1997 Battlespace Atmospherics Conference, SPAWAR Syst. Command Tech. Rep. 2989*, pp. 635–642, Space and Nav. Warfare Syst. Command Cent., San Diego, Calif.

Loh, W.-L. (1996), On Latin hypercube sampling, *Ann. Stat.*, *24*(5), 2058–2080, doi:10.1214/aos/1069362310.

McKay, M. D., W. J. Conover, and R. J. Beckman (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, *21*(2), 239–245.

Mercer, J. (1909), Functions of positive and negative type and their connection with the theory of integral equations, *Philos. Trans. R. Soc. London*, *209*, 415–446, doi:10.1098/rsta.1909.0016.

Patterson, W. (1992), Ducting climatology summary, report, Space and Nav. Warfare Syst. Cent., San Diego, Calif.

Paulus, R. (1990), Evaporation duct effects on sea clutter, *IEEE Trans. Antennas Propag.*, *38*(11), 1765–1771, doi:10.1109/8.102737.

Rogers, L. T. (1996), Effects of the variability of atmospheric refractivity on propagation estimates, *IEEE Trans. Antennas Propag.*, *44*(4), 460–465, doi:10.1109/8.489297.

Rogers, L. T., C. P. Hattan, and J. K. Stapleton (2000), Estimating evaporation duct heights from radar sea clutter, *Radio Sci.*, *35*(4), 955–966, doi:10.1029/1999RS002275.

Smola, A. J., and B. Schölkopf (1998), A tutorial on support vector regression, *NeuroCOLT Tech. Rep. NC-TR-98-030*, R. Holloway Coll., Univ. of London, London, U.K.

Suykens, J. A. K., and J. Vandewalle (1999), Least squares support vector machine classifier, *Neural Processes Lett.*, *9*(3), 293–300, doi:10.1023/A:1018628609742.

Suykens, J. A. K., Van T. Gestel, De J. Brabanter, B. De Moor, and J. Vandewalle (2002), *Least Squares Support Vector Machines*, World Sci., Singapore.

Turton, J. D., D. A. Bennetts, and S. F. G. Farmer (1988), An introduction to radio ducting, *Meteorol. Mag.*, *117*, 245–254.

Vapnik, V. (1995), *The Nature of Statistical Learning Theory*, Springer, New York.

Vapnik, V., and A. Chervonenkis (1964), A note on one class of perceptions, *Autom. Remote Control*, *25*, 821–837.

Vapnik, V., and A. Lerner (1963), Pattern recognition using generalized portrait method, *Autom. Remote Control*, *24*, 774–780.

Vivier, S. (2002), Stratégies d'optimisation par la méthode des plans d'expérience et application aux dispositifs électrotechniques modélisés par éléments finis, Ph.D. thesis, École Cent. Lille and Univ. des Sci. et Technol., Lille, France.

Ward, K. D., R. J. A. Tough, and S. Watts (2006), *Sea Clutter: Scattering, the K Distribution and Radar Performance*, The Institution of Engineering Technology, London, UK.

Webster, A. R. (1983), Angles-of-arrival and delay times on terrestrial line-of-sight microwave links, *IEEE Trans. Antennas Propag.*, *31*(1), 12–17, doi:10.1109/TAP.1983.1142976.

Yardim, C., P. Gerstoft, and W. S. Hodgkiss (2007), Statistical maritime radar duct estimation using hybrid genetic algorithm–Markov chain Monte Carlo method, *Radio Sci.*, *42*, RS3014, doi:10.1029/2006RS003561.

C. Bourlier and J. Saillard, Institut de Recherche en Electrotechnique et Electronique de Nantes Atlantique, Radar Team, Ecole Polytechnique de l'Université de Nantes, Rue Christian Pauc, La Chantrerie, BP 50609, 44306 Nantes, France.

R. Douvenot and V. Fabbro, Département Electromagnétisme et Radar, BP 4025, ONERA, 2 avenue Edouard Belin, F-31055 Toulouse CEDEX 4, France. (remi.douvenot@onera.fr)

P. Gerstoft, Marine Physics Laboratory, Scripps Institution of Oceanography, University of California, San Diego, MC0238, 9500 Gillman Drive, La Jolla, CA 92093-0238, USA.