

A statistical geoacoustic inversion scheme based on a modified radial basis functions neural network

George Tzagkarakis^{a)}

Department of Computer Science, University of Crete and Institute of Computer Science – FO.R.T.H.
FORTH-ICS, P.O. Box 1385, 711 10 Heraklion, Crete, Greece

Michael I. Taroudakis^{b)}

Department of Mathematics, University of Crete and Institute of Applied and Computational Mathematics –
FO.R.T.H. FORTH-IACM, P.O. Box 1385, 711 10 Heraklion, Crete, Greece

Panagiotis Tsakalides^{c)}

Department of Computer Science, University of Crete and Institute of Computer Science – FO.R.T.H.
FORTH-ICS, P.O. Box 1385, 711 10 Heraklion, Crete, Greece

(Received 14 March 2007; revised 25 July 2007; accepted 25 July 2007)

This paper addresses the task of recovering the geoacoustic parameters of a shallow-water environment using measurements of the acoustic field due to a known source and a neural network based inversion process. First, a novel efficient “observable” of the acoustic signal is proposed, which represents the signal in accordance with the recoverable parameters. Motivated by recent studies in non-Gaussian statistical theory, the observable is defined as a set of estimated model parameters of the alpha-stable distributions, which fit the marginal statistics of the wavelet subband coefficients, obtained after the transformation of the original signal via a one-dimensional wavelet decomposition. Following the modeling process to extract the observables as features, a radial basis functions neural network is employed to approximate the vector function that takes as input the observables and gives as output the corresponding set of environmental parameters. The performance of the proposed approach in recovering the sound speed and density in the substrate of a typical shallow-water environment is evaluated using a database of synthetic acoustic signals, generated by means of a normal-mode acoustic propagation algorithm.

© 2007 Acoustical Society of America. [DOI: 10.1121/1.2772232]

PACS number(s): 43.30.Pc, 43.60.Pt, 43.60.Lq [AIT]

Pages: 1959–1968

I. INTRODUCTION

Inverse problems in underwater acoustics are associated with measurements of the acoustic field performed in the frequency or in the time domain. In this framework, a set \mathbf{d} of observables is defined, which forms the input parameters of the inverse problem. The observables are related to the recoverable environmental parameters \mathbf{m} through a linear or nonlinear vector equation of the form $\mathcal{T}(\mathbf{d}, \mathbf{m})=0$.

The inversion procedure is based on the properties of the relationship between \mathbf{m} and \mathbf{d} , and it is considered to be more efficient if even small variations of the environmental parameters are associated with observables which can be clearly discriminated via the mapping $\mathcal{T}(\cdot, \cdot)$. Since the performance of a specific inversion procedure is directly related to the selected observables, defining observables which are easily identifiable and as sensitive as possible to changes of the environmental parameters, constitutes an important task.

Determining the sea-bed parameters from acoustic measurements obtained in the water column is among the most interesting inverse problems in underwater acoustics.^{1–4} It should be noted that most of the inversion procedures and

the associated observable identification (feature extraction) are based on deterministic approaches. In recent work,^{5,6} we proposed a novel observable for acoustic signals, based on a *symmetric alpha-stable* ($S\alpha S$) statistical modeling of the coefficients obtained after a transformation of the original signal using the one-dimensional (1D) dyadic wavelet transform (DWT). Then, a classification scheme was designed by combining the extracted features, that is, the estimated $S\alpha S$ parameters at each wavelet subband, with a closed-form expression of the Kullback-Leibler divergence (KLD) between $S\alpha S$ distributions. First results based on synthetic data showed that the proposed scheme provided a very accurate classification of a recorded acoustic signal in the true unknown environment, specified by several sets of parameters (e.g., sound speed profiles in the water and/or bottom domains, layer thicknesses and densities, source location, etc.).

Based on the above, in this paper we treat the inverse problem as a function approximation problem. In particular, we consider a nonlinear mapping \mathcal{T} with arguments the estimated $S\alpha S$ parameters and output the set of the corresponding environmental parameters. Our goal is to find an accurate approximation of \mathcal{T} . An efficient approximation of such a mapping between an acoustic field and its corresponding geoacoustic parameters is achieved using neural network-based approaches. In previous studies,^{7,8} the inversion process was carried out by employing multilayer feed-forward

^{a)}Electronic mail: gtzag@ics.forth.gr

^{b)}Electronic mail: taroud@iacm.forth.gr

^{c)}Electronic mail: tsakalid@ics.forth.gr

neural networks. In the present work, a neural network-based approach is employed using radial basis functions (RBF), since it is well known that RBF neural networks provide a suitable approximation to nonlinear functions in an efficient way. In particular, we propose a modified version of the standard RBF neural network, by replacing the Euclidean distance function, which measures the similarity between the input vectors and the centers of the hidden neurons, with the KLD between $S\alpha S$ distributions.

The paper is organized as follows: In Sec. II, the statistical modeling via $S\alpha S$ distributions is described, and the procedure for the design of the modified RBF network based on the KLD between $S\alpha S$ distributions is analyzed in detail. In Sec. III, the proposed RBF neural network is applied to a database of simulated acoustic signals generated in a shallow water environment to evaluate the inversion performance. Finally, we conclude in Sec. IV giving some future research directions.

II. THEORETICAL BACKGROUND

In the following, we introduce the main mathematical background of the proposed inversion scheme.

A. Statistical modeling via $S\alpha S$ distributions

The proposed geoacoustic inversion process is performed in two steps: First, a feature extraction procedure is applied, which represents the information content of a given acoustic signal with an appropriate set of features and second, the extracted features are used to build a RBF neural network.

During the feature extraction step, the acoustic signal is decomposed into several levels through a multiresolution analysis employing the 1-D DWT. This transform works as follows: Starting from the given signal $s(t)$, two sets of coefficients are computed at the first level of decomposition, (i) approximation coefficients A1 and (ii) detail coefficients D1. These vectors are obtained by convolving $s(t)$ with a low-pass filter for approximation and with a high-pass filter for detail, followed by dyadic decimation. At the second level of decomposition, the vector A1 of the approximation coefficients is decomposed in two sets of coefficients using the same approach replacing $s(t)$ by A1 and producing A2 and D2. This procedure continues in the same way, namely at the k th level of decomposition we filter the vector of the approximation coefficients computed at the $(k-1)$ th level. Thus, when $s(t)$ is decomposed in L levels, we obtain a set of L detail subbands (containing the high-frequency content) and one approximation subband (containing the low-frequency content).

For short-time, pulse shallow-water acoustic signals, the 1-D DWT seems to be a powerful modeling tool, providing a natural arrangement of the wavelet coefficients into multiple levels representing the frequency content of the signal in consecutive bands.⁹ Besides, it has been pointed out that the wavelet transforms of signals which present such a transient behavior tend to be sparse, resulting in a large number of coefficients with small magnitude and a small number of large magnitude coefficients.¹⁰ This property gives rise to

peaky and heavy-tailed *non-Gaussian* marginal distributions of the wavelet subband coefficients.¹⁰

Following the acoustic signal decomposition, an accurate fitting of the tails of the marginal distribution of the wavelet coefficients at each subband is achieved by modeling them as $S\alpha S$ random variables. A $S\alpha S$ distribution is best defined by its characteristic function:¹¹

$$\phi(t) = \exp(i\delta t - \gamma^\alpha |t|^\alpha), \quad (1)$$

where α is the *characteristic exponent*, taking values $0 < \alpha \leq 2$, $\delta (-\infty < \delta < \infty)$ is the *location parameter*, and $\gamma (\gamma > 0)$ is the *dispersion* of the distribution. The characteristic exponent is a shape parameter, which controls the “thickness” of the tails of the density function. The smaller the α , the heavier the tails of the $S\alpha S$ density function. The dispersion parameter determines the spread of the distribution around its location parameter, similar to the variance of the Gaussian.

In our previous work,⁵ the content of an underwater signal was accurately represented with a set of features, which is much smaller in size than the signal itself or any other representation in the frequency or wavelet domain. This set contains the maximum likelihood (ML) estimated parameters (α, γ) of the $S\alpha S$ distribution at each wavelet subband.

Thus, for a given acoustic signal S , decomposed in L levels, its feature vector \mathbf{d} is given by the following set of $L+1$ pairs:

$$S \mapsto \mathbf{d} = \{(\alpha_1, \gamma_1), (\alpha_2, \gamma_2), \dots, (\alpha_{L+1}, \gamma_{L+1})\}, \quad (2)$$

where (α_i, γ_i) are the estimated model parameters of the i th subband, using the consistent ML method described by Nolan,¹² which gives reliable estimates and provides the tightest confidence intervals. Note also that we follow the convention that $i=1$ corresponds to the detail subband at the first decomposition level, while $i=L+1$ corresponds to the approximation subband at the L th level.

B. Inversion using a RBF neural network

As we demonstrated in a recent work,⁵ the KLD¹³ is capable of distinguishing between two distinct acoustic signals, since the KLD between two signals generated in similar shallow-water environments is almost zero, whereas it increases when the signals are obtained from different environments. Thus, the function \mathcal{T} , which maps the estimated $S\alpha S$ parameters \mathbf{d} to the corresponding environmental parameters \mathbf{m} , is a well-defined nonlinear vector function. We also assume that \mathcal{T} defines a one-to-one correspondence between \mathbf{d} and \mathbf{m} .

This observation yields that there is a nonlinear vector function $\mathcal{T}: \mathcal{A} \mapsto \mathbb{R}^n$, where $\mathcal{A} \subseteq \mathbb{R}^{2(L+1)}$ contains the estimated parameters $\{(\alpha_1, \gamma_1), \dots, (\alpha_{L+1}, \gamma_{L+1})\}$ of each signal and n is the number of the environmental parameters we are interested in. In this paper, the shallow-water environment is modeled as a two-layered medium, with the first layer representing the water column and the second semi-infinite layer representing the substrate. Then, we focus on the recovery of the sound speed in the substrate, c_{sb} , and the substrate density ρ_{sb} . That is, the function \mathcal{T} maps the $S\alpha S$ parameters in

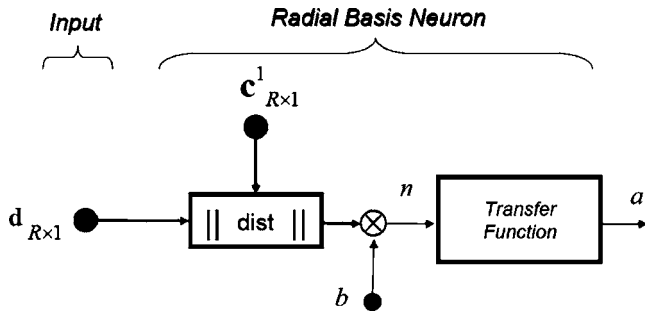


FIG. 1. The radial basis neuron model.

\mathbb{R}^2 , and particularly in vectors of the form $\mathbf{m}=[c_{sb}, \rho_{sb}]$. Note that a similar problem can be defined for each set of recoverable environmental parameters. In our previous work,⁵ it was shown that the KLD is capable of performing the identification of the parameters characterizing the sound speed profile in the water column as well.

Working in this framework, we proceed by converting the inverse problem into a function approximation problem. In particular, if we were able to approximate accurately the function \mathcal{T} , then we would be able to find a solution to the inverse problem, since the insertion of the estimated $S\alpha S$ parameters of a signal, recorded in an unknown environment, into \mathcal{T} , would result in the computation of the vector $\mathbf{m}=[c_{sb}, \rho_{sb}]$ with elements the substrate parameters of the unknown environment.

An efficient process to solve the above-mentioned function approximation problem is achieved by using a RBF neural network,¹⁴ as it is well known that such a network is suitable for the approximation of a nonlinear function. A RBF network consists of two layers: a hidden radial basis layer of N_1 neurons, and an output linear layer of N_2 neurons. In the test case described in Sec. II C, N_1 can be at most equal to the number of training samples (M) and $N_2=2$, since we are interested in recovering the two environmental parameters $[c_{sb}, \rho_{sb}]$.

Figure 1 shows the model of a single radial basis neuron in the hidden layer. In particular, the net input, n , to the transfer function is the vector distance between its center \mathbf{c} and the input vector \mathbf{d} , multiplied by the bias b . Thus, the output of the radial basis neuron, a , is equal to the value of the selected transfer function evaluated at $b \cdot \|\mathbf{d} - \mathbf{c}\|$.

Figure 2 shows the general architecture of a RBF network. Each radial basis neuron of the hidden layer is denoted by using its corresponding transfer function ϕ_i , $i=1, \dots, N_1$, while each neuron of the output linear layer is denoted by L_j , $j=1, \dots, N_2$. Following the notation of Fig. 1, the output of a single hidden radial basis neuron is given by $a_i = \phi_i(b_i \cdot \|\mathbf{d} - \mathbf{c}_i\|)$, where b_i and \mathbf{c}_i are the bias and the center of the i th radial basis neuron, respectively. Then, the output of a neuron in the linear layer is given by $\mathbf{m}_j = \sum_{i=1}^{N_1} w_{i,j} a_i + b_j$. In a common RBF network, the $\|\text{dist}\|$ box in Fig. 1 measures the Euclidean distance between the input vector \mathbf{d} and the center of the i th neuron, \mathbf{c}_i . The transfer function for a radial basis neuron is $\phi(n) = e^{-n^2}$, which has a maximum of 1 when its input is 0. As the distance between \mathbf{c}_i and \mathbf{d} decreases, the output increases. Thus a radial basis neuron acts as a detector

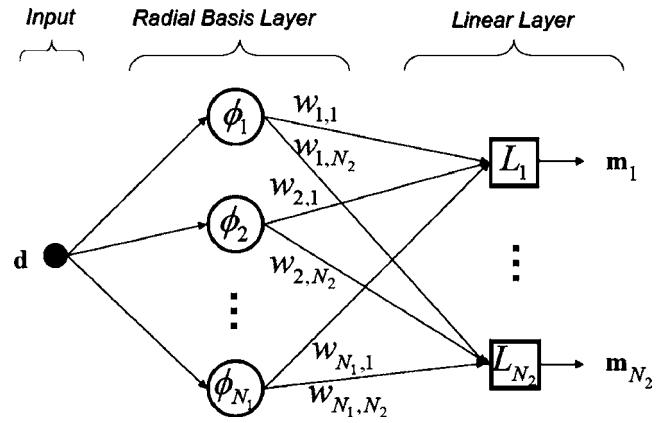


FIG. 2. The general architecture of a RBF network.

which produces 1 whenever the input \mathbf{d} is identical to its center \mathbf{c}_i . Besides, the bias b_i allows the sensitivity of the i th radial basis neuron to be adjusted. For example, if a neuron i had a bias of 0.1 its transfer function $\phi_i(n)$ would output 0.5 for any input vector \mathbf{d} at vector distance of 8.326 ($0.8326/b_i$) from its center \mathbf{c}_i .

Thus, the j th component, \mathbf{m}_j , of the vector function \mathcal{T} is approximated as a linear combination of the set of radial basis functions:

$$\tilde{\mathcal{T}}_j(\mathbf{d}) \equiv \tilde{\mathbf{m}}_j = \sum_{i=1}^{N_1} w_{i,j} \phi_i(\mathbf{d}) + b_j = \sum_{i=0}^{N_1} w_{i,j} \phi_i(\mathbf{d}), \quad (3)$$

where $w_{i,j}$ is equal to the weight of the edge connecting the i th radial basis neuron with the j th output of the network (see Fig. 2). In Eq. (3), the auxiliary radial basis function $\phi_0(\mathbf{d})=1$. The most common form of basis function used is the Gaussian

$$\phi_i(\mathbf{d}) \equiv \phi_i(b_i \cdot \|\mathbf{d} - \mathbf{c}_i\|) = \exp\left(-\frac{\|\mathbf{d} - \mathbf{c}_i\|^2}{2\sigma_i^2}\right),$$

where \mathbf{c}_i and σ_i are the mean and standard deviation of the basis function, respectively. This basis function is of the form e^{-n^2} with $n = b_i \cdot \|\mathbf{d} - \mathbf{c}_i\|$, where $b_i = 1/\sqrt{2}\sigma_i$. A hidden neuron is more sensitive to input vectors near its center. This sensitivity may be tuned by adjusting the widths (spread parameters) σ_i . For a given input vector, typically only a few hidden units will have significant activations. Besides, the spread parameters should be chosen large enough so that neurons respond strongly to overlapping regions of the input space, but they should not be too large so that each neuron would effectively respond in the same large area of the input space.

In the following, we study an inversion scheme based on two different kinds of RBF networks. The first one employs the standard RBF network architecture described so far, that is, the $\|\text{dist}\|$ in Fig. 1 is the common Euclidean distance and the transfer functions ϕ_i are Gaussians. The second novel scheme is based on a modification of the standard RBF architecture. In particular, we are interested in exploiting the results of our previous work,⁵ where it was illustrated that two distinct acoustic signals represented by their corresponding feature vectors can be discriminated very accurately by

employing a version of the KLD between two $S\alpha S$ distributions. Motivated by these results, we modify the standard RBF network by replacing the similarity measure between the input layer and the hidden radial basis layer, namely, by replacing the Euclidean distance with the KLD between $S\alpha S$ distributions.

C. Modified RBF network

In the modified RBF network the similarity measurement between an input vector \mathbf{d} and the i th radial basis neuron's center \mathbf{c}_i is carried out by employing the KLD between two $S\alpha S$ distributions. In our previous study,⁵ we showed that a chain rule can be applied in order to combine the KLDs from the multiple wavelet subbands. In particular, the overall distance between two acoustic signals S_1 and S_2 , which are represented by the feature vectors \mathbf{d}_1 and \mathbf{d}_2 given by Eq. (2), respectively, has the following expression:

$$D(S_1, S_2) \equiv D(\mathbf{d}_1, \mathbf{d}_2) = \sum_{k=1}^{L+1} D(\hat{q}_{S_1, k} \| \hat{q}_{S_2, k}), \quad (4)$$

where $D(\hat{q}_{S_1, k} \| \hat{q}_{S_2, k})$ is the KLD between the k th wavelet subbands of the two signals, which is evaluated using the estimated $S\alpha S$ parameters $(\alpha_{1, k}, \gamma_{1, k})$ and $(\alpha_{2, k}, \gamma_{2, k})$, respectively. This marginal KLD is given by

$$D(\hat{q}_{S_1, k} \| \hat{q}_{S_2, k}) = \ln\left(\frac{l_{2, k}}{l_{1, k}}\right) - \frac{1}{\alpha_{1, k}} + \left(\frac{\gamma_{2, k}}{\gamma_{1, k}}\right)^{\alpha_{2, k}} \cdot \frac{\Gamma\left(\frac{\alpha_{2, k} + 1}{\alpha_{1, k}}\right)}{\Gamma\left(\frac{1}{\alpha_{1, k}}\right)}, \quad (5)$$

with $l_{i, k}$, $i = 1, 2$, being a normalizing factor, which is equal to

$$l_{i, k} = \frac{2\Gamma\left(\frac{1}{\alpha_{i, k}}\right)}{\alpha_{i, k} \gamma_{i, k}}, \quad (6)$$

where $\Gamma(\cdot)$ is the gamma function.

We modify the standard radial basis neuron model, shown in Fig. 1, by replacing the Euclidean distance (computed by the $\|\text{dist}\|$ box) with the overall KLD between vectors containing estimated $S\alpha S$ parameters (4). Accordingly, the modified output of a single hidden radial basis neuron is given by

$$\phi_i^{\text{mod}} = \phi_i(b_i \cdot D(\mathbf{d}, \mathbf{c}_i)), \quad (7)$$

where $D(\mathbf{d}, \mathbf{c}_i)$ is the overall KLD between the input vector \mathbf{d} and the center \mathbf{c}_i of the i th radial basis neuron. The transfer functions $\{\phi_i(\cdot)\}_{i=1, \dots, N_1}$ have the form,

$$\phi_i(\mathbf{d}) \equiv \phi_i(b_i \cdot D(\mathbf{d}, \mathbf{c}_i)) = \exp(-b_i \cdot (D(\mathbf{d}, \mathbf{c}_i))^2),$$

where again the bias b_i is inversely proportional to the spread of the radial basis function, that is, $b_i \sim 1/\sigma_i$. Note that the KLD is always non-negative, $D(\mathbf{d}_1, \mathbf{d}_2) \geq 0$, with an equality if and only if the corresponding $S\alpha S$ parameter pairs are equal, $(\alpha_{1, k}, \gamma_{1, k}) = (\alpha_{2, k}, \gamma_{2, k})$, $k = 1, \dots, L+1$. Thus, the transfer function of the i th neuron has an output equal to 1 if and

only if its center \mathbf{c}_i is equal to the input vector \mathbf{d} .

Notice that the KLD between $S\alpha S$ distributions, given by Eq. (4), is not an actual distance function, since it does not satisfy the property of symmetry. However, we can define a symmetrized version of the KLD as follows:

$$D_{\text{sym}}(\mathbf{d}_1, \mathbf{d}_2) = D(\mathbf{d}_1, \mathbf{d}_2) + D(\mathbf{d}_2, \mathbf{d}_1). \quad (8)$$

It can be easily verified that $D_{\text{sym}}(\cdot, \cdot)$ satisfies all the properties of a distance function. In the subsequent illustrations, we will also test the performance of a RBF network based on the symmetrized KLD, that is, whose radial basis transfer functions have the form

$$\phi_i(\mathbf{d}) = \exp(-b_i \cdot (D_{\text{sym}}(\mathbf{d}, \mathbf{c}_i))^2).$$

D. Training process

In our proposed scheme, the (standard or modified) RBF network is trained using two different processes, namely, an exact and a more efficient one (let **P1** and **P2** denote the exact and the efficient process, respectively). The exact training process computes the weights $\{w_{i, j}\}_{i=1, \dots, N_1, j=1, \dots, N_2}$ and the biases $\{b_j\}_{j=1, \dots, N_2}$, such that the produced network achieves zero error on the training vectors. Besides, the exact process creates as many radial basis neurons as there are input vectors, where the i th input vector is used as the center \mathbf{c}_i of the corresponding neuron. The drawback of the exact process is that it produces a large network when many input vectors are needed to properly define a network.

On the other hand, the second and more efficient training process produces the RBF network iteratively, by adding one neuron at a time, starting with a single neuron. At each iteration, the input vector that will result in lowering the network error the most is used to create a radial basis neuron. The error of the new network is checked, and if it is low enough the training process terminates. Otherwise, the next neuron is added. Neurons are added to the network until the sum-squared error (SSE) falls below a specified error threshold or a maximum number of neurons is reached. Given a network with K radial basis neurons and input-output pairs $\{(\mathbf{d}_k, \mathbf{m}_k)\}_{k=1, \dots, K}$, the SSE is given by

$$\text{SSE} = \sum_{k=1}^K \|\tilde{\mathcal{T}}(\mathbf{d}_k) - \mathbf{m}_k\|^2, \quad (9)$$

where $\tilde{\mathcal{T}}(\mathbf{d}_k)$ is the approximation of the function \mathcal{T} at the input point \mathbf{d}_k , obtained at the output of the RBF network.

In both of the above-presented training processes it is important that the spread parameter of the transfer function of each radial basis neuron be large enough so that the neurons respond to overlapping regions of the input space. This also results in a better generalization for new input vectors occurring between input vectors used in the training process. However, the spread parameter should not be too large that all the neurons respond in essentially the same manner.

For the standard transfer function, $\phi(n) = e^{-n^2}$, the bias, b_i , of the i th radial basis neuron is related to the spread, σ_i , of the corresponding transfer function with the expression: $b_i = \sqrt{-\ln(0.5)}/\sigma_i$. This means that if the neuron's center \mathbf{c}_i is at

TABLE I. The shallow water environment.

Water Depth (H)	200 m
Range (R)	5 km
Central frequency (f_0)	100 Hz
Bandwidth (Δf)	40 Hz
Source/receiver depth	100 m
Sound speed profile in the water:	
$c_w(0)$	1500 m/s
$c_w(\min)$	1490 m/s
$c_w(H)$	1515 m/s
d [depth of min $c_w(z)$]	50 m
Semi-infinite substrate: (varying parameters)	
c_{sb}	[1550, 1650] m/s
ρ_{sb}	[1170, 1240] kg/m ³

a distance of σ_i from the input vector \mathbf{d} , then the output of the transfer function will be 0.5. In the subsequent illustrations, the influence of the exponent value of the transfer function is also studied. In particular, we design RBF networks whose transfer function has the general form $\phi(n) = e^{-n^r}$, where r is a positive integer, which is also known as a *sharpness* parameter. In this case, the bias is related to the spread parameter via the expression: $b_i = \sqrt[r]{-\ln(0.5)}/\sigma_i$.

A common heuristic for the selection of the spread parameter of a radial basis neuron is the following: Choose a spread constant larger than the distance between adjacent input vectors, so as to get good generalization, but smaller than the distance across the whole input space. Thus, in the standard RBF network it should be ensured that $d_{E,\min} \leq \sigma_i \leq d_{E,\max}$, where $d_{E,\min}$ and $d_{E,\max}$ are the minimum and maximum Euclidean distances between the training input vectors, respectively. Similarly, in the modified RBF network one should ensure that $d_{KLD,\min} \leq \sigma_i \leq d_{KLD,\max}$, where $d_{KLD,\min}$ and $d_{KLD,\max}$ are the minimum and maximum KLDs between the entraining input vectors, respectively, given by Eq. (4) or (8).

III. APPLICATION OF THE INVERSION SCHEME USING SYNTHETIC DATA

In this section, the efficiency of the proposed inversion scheme for shallow-water acoustic transmissions is evaluated using synthetic signals, based on the range independent and axially symmetric environment described in Table I. Figure 3 shows the sea environment of the experimental setup consisting of a shallow-water layer and a semi-infinite bottom (the substrate), which are considered fluid. The sound speed profile may vary with depth in the water layer, while it is constant in the substrate. Here, a linear sound speed profile in the water column is considered. For simplicity, the density of both layers is assumed to be constant.

As an attempt to simulate a geoacoustic inversion experiment, we consider a low-frequency sound source, with central frequency $f_0 = 100$ Hz and bandwidth $\Delta f = 40$ Hz. The source excitation function is modeled as a Gaussian, placed at a known depth of 100 m. A single receiver is placed at a distance of 5 km from the source and at the same depth. The

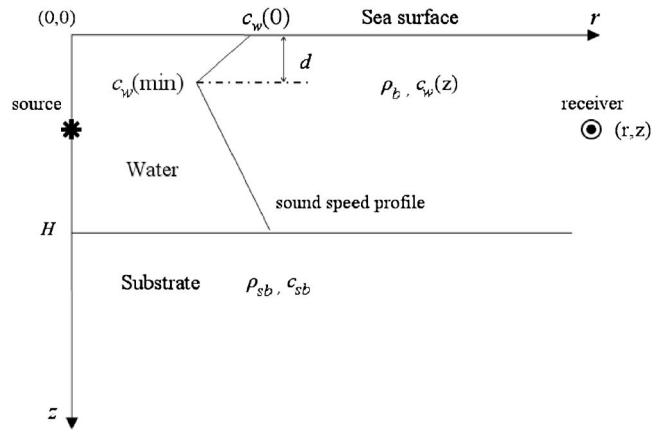


FIG. 3. The shallow water sea environment used in the present study.

experiment is aiming at the identification of the sea bed, characterized by the sound speed and density of the substrate.

We constructed a database by generating a set of synthetic signals according to the environmental parameters shown in Table I. The sound speed in the substrate varies in the interval [1550, 1650] m/s with a step size equal to 5 m/s and the substrate density varies in the interval [1170, 1240] kg/m³ with a step size equal to 1 kg/m³, resulting in a set with a total of 1491 synthetic acoustic signals. The signals are calculated using the Normal-Mode program MODE1 developed at FO.R.T.H-I.A.C.M. Then, the obtained data are provided as input to the inverse discrete Fourier transform to yield the signals in the time domain. Each of the time-domain signals is decomposed by implementing a three-level 1D DWT using the db4 wavelet function and, thus, each signal is represented by a vector \mathbf{d} with $2(L+1) = 8$ elements given by Eq. (2). The experimental results in our recent work⁵ showed that the best classification performance is obtained for the db4 wavelet function and using the estimated SaS parameters of the detail subbands only. Thus, in the subsequent illustrations each signal is represented by a vector \mathbf{d} with $2L=6$ parameters. In addition, the training set for the design of the RBF networks consists of 500 (out of the 1491) signals, obtained from distinct speed and density environmental parameters.

A. Inversion performance using the exact design process P1

In this section, we study the performance of a RBF network designed using the exact process **P1**. The RBF network is constructed using the estimated SaS parameters of a subset containing $M \in \{50:50:300\}$ signals chosen from the training set. Besides, $N_1 = M$ since the RBF network is designed using the exact process. As mentioned before, the selection of the spread parameters $\{\sigma_i\}_{i=1,\dots,N_1}$ is crucial for an improved performance of the inversion scheme. In the proposed method, we assume that all the radial basis neurons have equal spread parameter, that is, $\sigma_i = \sigma, \forall i$.

For the determination of σ , the heuristic described in Sec. II D is followed. In particular, in the case of the standard RBF network, an efficient design is ensured when $0 \leq \sigma$

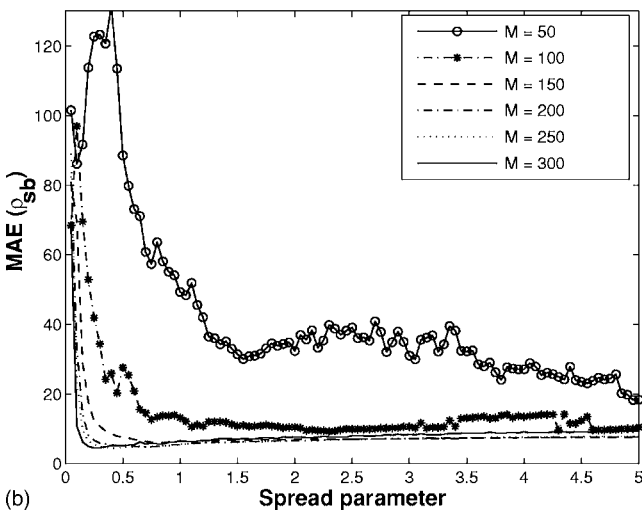
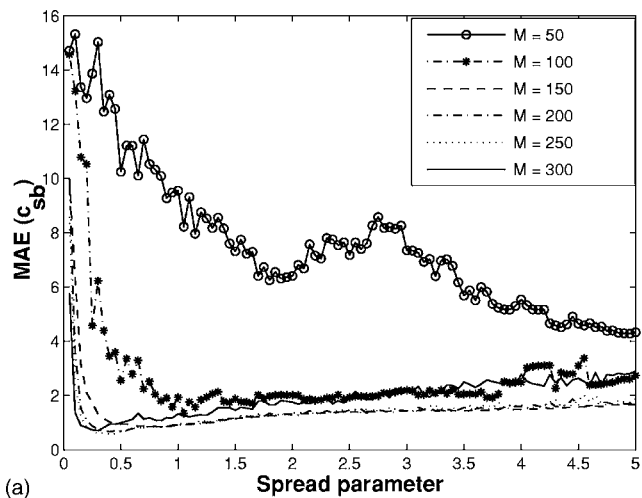


FIG. 4. Mean absolute error, given by the standard RBF network, as a function of the spread parameter σ and the number of training samples M for: (a) the sound speed c_{sb} and (b) the substrate density ρ_{sb} .

≤ 0.231 , where 0 and 0.231 are the minimum and maximum Euclidean distances, respectively, between the $S\alpha S$ parameter vectors of all the signals in the database. Similarly, in the case of the modified RBF network, the value of σ should be selected such that $2.2204 \times 10^{-16} \leq \sigma \leq 252.59$, where 2.2204×10^{-16} and 252.59 are the minimum and maximum KLDs, respectively, given by Eq. (4), between the $S\alpha S$ parameter vectors of all the signals in the database. Finally, when the modified RBF network employs the symmetrized KLD as a distance function, one should choose $8.8818 \times 10^{-16} \leq \sigma \leq 259.49$, where 8.8818×10^{-16} and 259.49 are the minimum and maximum values of the symmetrized KLDs, respectively, given by Eq. (8).

The inversion performance is evaluated using several values of σ , as well as of the sharpness parameter r of the general transfer function $\phi(n) = e^{-nr}$. In particular, in the case of the standard RBF network, the value of σ varies in the interval $[0.05, 5]$, while in the case of the modified RBF it varies in the interval $[0.001, 275]$. In both cases, r belongs to the set $\{1, \dots, 5\}$. For a given training size M and for fixed r and σ , we run 100 Monte Carlo iterations, where in each iteration a new RBF network is designed by randomly select-

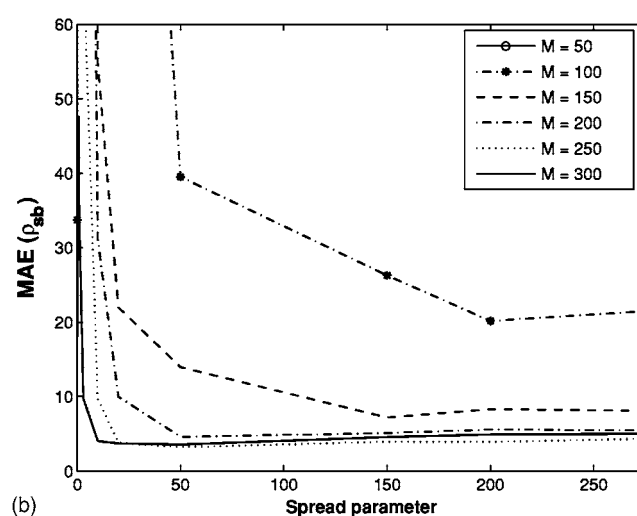
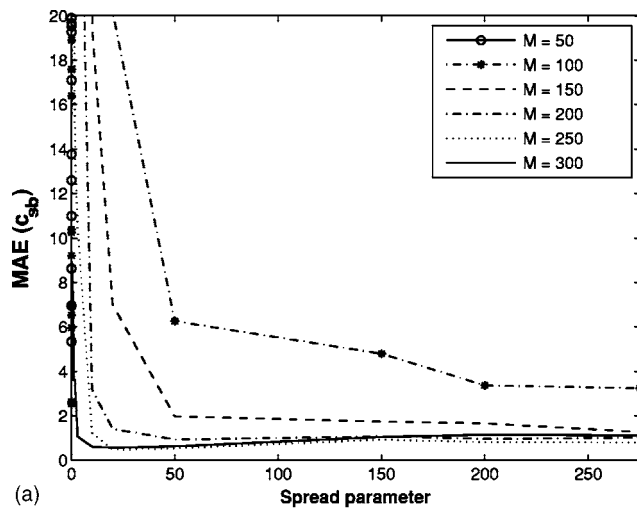
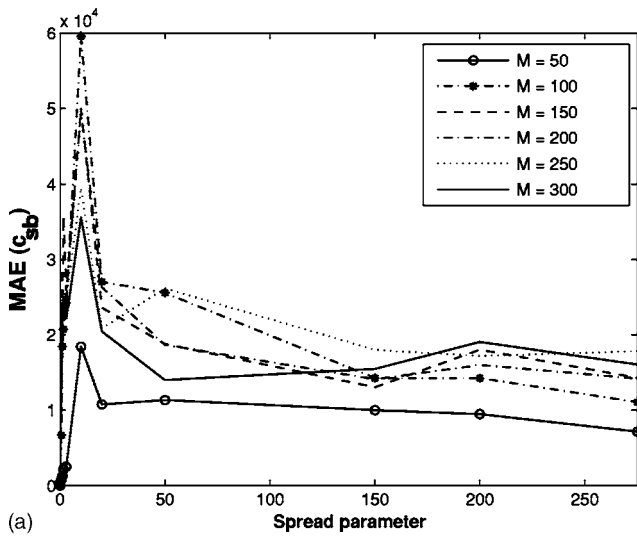


FIG. 5. Mean absolute error, given by the modified RBF network employing the standard KLD, as a function of the spread parameter σ and the number of training samples M for: (a) the sound speed c_{sb} and (b) the substrate density ρ_{sb} .

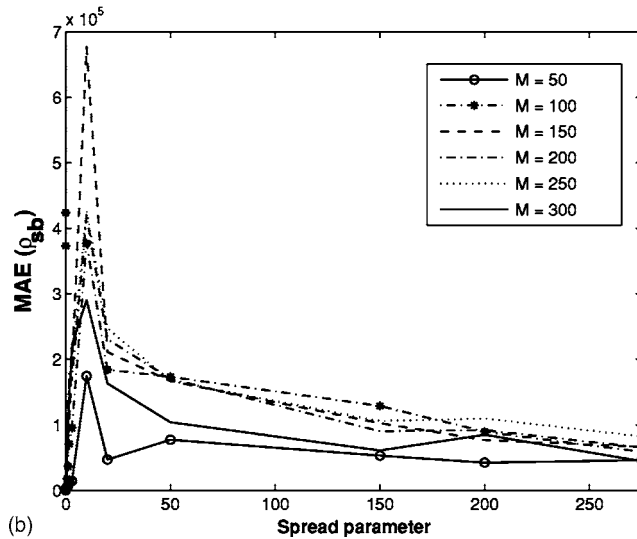
ing a training subset of size M from the training set. In the subsequent figures and tables, the results have been obtained using a sharpness parameter $r=2$, otherwise, the value of r will be mentioned explicitly.

Figures 4(a) and 4(b) show the total mean absolute error (MAE) of the estimated sound speed c_{sb} and substrate density ρ_{sb} values, respectively, over the whole test set, given by the standard RBF network, as a function of the spread parameter σ and the number of training samples M . First, it can be seen that in both cases, for a relatively large number of training samples, the minimum MAE is achieved for a value of σ satisfying the corresponding inequality $0 \leq \sigma \leq 0.231$. The second observation is that, as expected, in this region the performance is improved, that is, the MAE decreases, as M increases, while out of this region this rule is not valid, especially for the estimation accuracy of c_{sb} .

Figures 5(a) and 5(b)¹⁵ show the total MAE of the estimated c_{sb} and ρ_{sb} values, respectively, over the whole test set, given by the modified RBF network employing the standard KLD as a “distance” function, versus the spread parameter σ and for various training sample sizes M . It is clear that



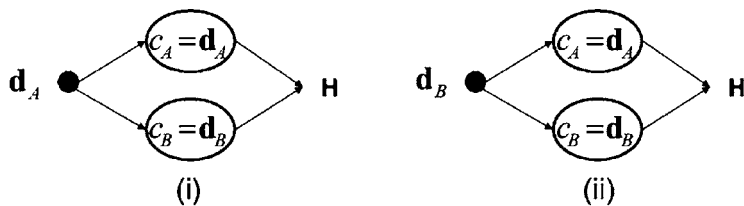
(a)



(b)

FIG. 6. Mean absolute error, given by the modified RBF network employing the symmetrized KLD, as a function of the spread parameter σ and the number of training samples M for: (a) the sound speed c_{sb} , and (b) the substrate density ρ_{sb} .

for both environmental parameters and a relatively large number of training samples, the minimum MAE is achieved for a value of σ satisfying the corresponding inequality



		Hidden Layer Output (H)	
		Standard KLD	Symmetrized KLD
(i)	\mathbf{d}_A	$H = D(\mathbf{d}_A, \mathbf{d}_A) + D(\mathbf{d}_A, \mathbf{d}_B)$ $= D(\mathbf{d}_A, \mathbf{d}_B)$	$H = D_{sym}(\mathbf{d}_A, \mathbf{d}_A) + D_{sym}(\mathbf{d}_A, \mathbf{d}_B)$ $= 0 + \{D(\mathbf{d}_A, \mathbf{d}_B) + D(\mathbf{d}_B, \mathbf{d}_A)\}$
(ii)	\mathbf{d}_B	$H = D(\mathbf{d}_B, \mathbf{d}_B) + D(\mathbf{d}_B, \mathbf{d}_A)$ $= D(\mathbf{d}_B, \mathbf{d}_A)$	$H = D_{sym}(\mathbf{d}_B, \mathbf{d}_A) + D_{sym}(\mathbf{d}_B, \mathbf{d}_B)$ $= \{D(\mathbf{d}_B, \mathbf{d}_A) + D(\mathbf{d}_A, \mathbf{d}_B)\} + 0$

FIG. 7. Comparison between the output of the hidden radial basis layer obtained by employing the standard and the symmetrized KLD.

$2.2204 \times 10^{-16} \leq \sigma \leq 252.59$. Besides, it can be seen that the MAE decreases as M increases, as expected. Also note that in both cases, c_{sb} and ρ_{sb} , and for a fixed M , the performance of the modified RBF network stabilizes for $\sigma > 50$. Finally, Figs. 5(a) and 5(b) show that as M increases, the corresponding value of σ , which minimizes the MAE, decreases. This should be expected, since as the number of training samples increases, the input space can be covered using radial basis functions with a smaller spread.

Figures 6(a) and 6(b) show the total MAE of the estimated c_{sb} and ρ_{sb} values, respectively, over the whole test set, given by the modified RBF network employing the symmetrized KLD as a distance function. The first observation is that for both environmental parameters, the MAE is too large, compared with the performance of the other two RBF networks. Besides, the behavior of this network does not follow the rule that the MAE decreases as the number of training samples increases. In particular, there is a region on the x axis ($[0, 50]$), where the MAE is minimized when the network is trained with a very small number of samples. On the other hand, similar to the case of the previous RBF network, which employs the standard KLD, the behavior of the MAE stabilizes for $\sigma > 50$, but the minimum MAE is still achieved for $M=50$.

An explanation for this unexpected behavior of the modified RBF network based on the symmetrized KLD is shown in Fig. 7. Consider the simple case of a RBF network consisting of two hidden units with centers $\mathbf{c}_A = \mathbf{d}_A$ and $\mathbf{c}_B = \mathbf{d}_B$, respectively, where the vectors \mathbf{d}_A and \mathbf{d}_B contain the estimated $S\alpha S$ parameters of two signals S_A, S_B recorded in two very distinct sea environments. Also assume that the same vectors, \mathbf{d}_A and \mathbf{d}_B , are given as inputs in the RBF network. As can be seen in Fig. 7, when the standard KLD is employed, the network outputs are different, $D(\mathbf{d}_A, \mathbf{d}_B) \neq D(\mathbf{d}_B, \mathbf{d}_A)$, since the standard KLD is not symmetric. On the other hand, if the symmetrized KLD is employed, we observe that the network gives the same output, although the input vectors correspond to two very distinct environments. Subsequently, the linear layer will result in an increased estimation error of the environmental parameters c_{sb}, ρ_{sb} .

Table II shows the average performance and the corresponding optimal design parameters for the three types of

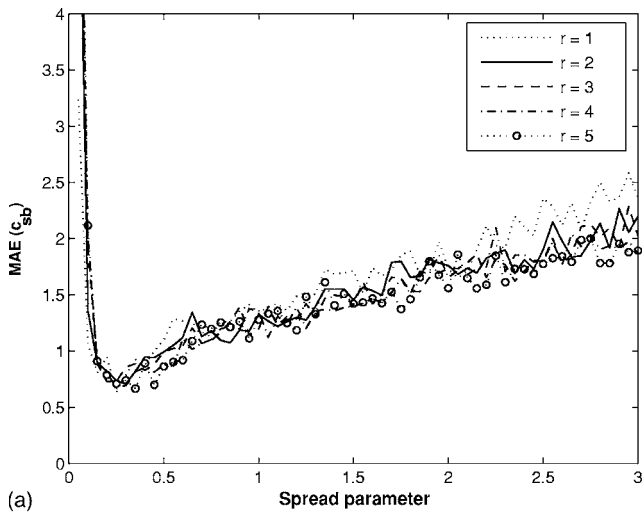
TABLE II. Average error performance and optimal design parameters for the three types of RBF networks ($r=2$).

RBF type	M	c_{sb} (m/s)		ρ_{sb} (kg/m ³)			
		σ	MAE	σ	MAE	Std of error	
Standard	250	0.4	0.5792	1.8933	0.45	4.4230	11.5974
Modified with $D(\cdot, \cdot)$	300	0.006	0.2481	0.7579	50	3.5687	2.5640
Modified with $D_{sym}(\cdot, \cdot)$	50	0.006	7.6462	3.4814	0.001	45.7241	25.4485

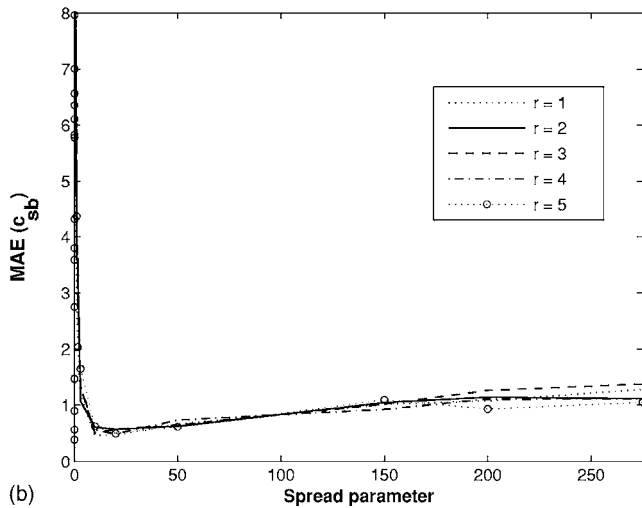
RBF networks. It is clear that the proposed modified RBF network, based on the standard KLD, outperforms the standard RBF network, based on the Euclidean norm, as well as the modified RBF network based on the symmetrized KLD. Besides, the increased estimation performance with respect to the substrate density becomes more evident, if we recall that the substrate density is one of the most difficult environmental parameters to be estimated by the majority of the previously developed inversion schemes. One should also note that distinct optimal σ parameters are necessary for estimating each environmental variable c_{sb} , or ρ_{sb} . Hence, for

optimal performance, two different single-output networks should be designed, one for each environmental parameter, using the corresponding optimal spread parameter. Of course, in this case there is a trade-off between the increased approximation accuracy and the computational complexity.

As mentioned before, we are also interested in studying the influence of the sharpness parameter (r) of a radial basis function, on the average performance of the proposed inversion scheme. Figure 8(a) shows the MAE of the estimated c_{sb} values using the standard RBF network, versus the spread σ ,

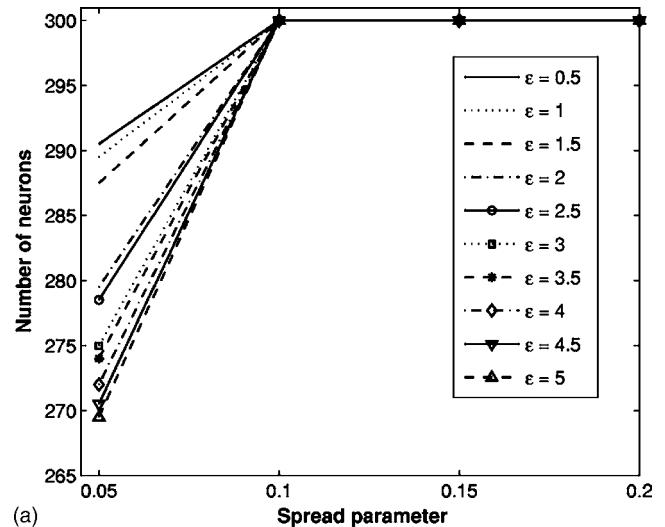


(a)

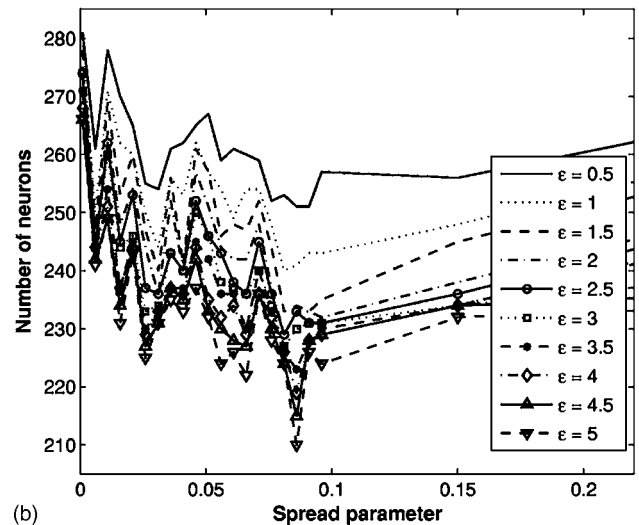


(b)

FIG. 8. Mean absolute error of the estimated c_{sb} values, as a function of the spread parameter σ and the sharpness parameter r , corresponding to: (a) the standard RBF network and (b) the modified RBF network employing the standard KLD.



(a)



(b)

FIG. 9. Number of neurons as a function of the spread parameter σ and for several error thresholds ϵ , corresponding to: (a) the standard RBF network and (b) the modified RBF network based on the standard KLD.

TABLE III. Comparison between the optimal number of neurons (N_1^{opt}) required to achieve a predetermined error threshold ε , for the standard and modified RBF networks ($r=2$).

ε	N_1^{opt}	
	Standard	Modified with $D(\cdot, \cdot)$
0.5	290	251
1	289	240
1.5	287	233
2	279	229
2.5	278	229
3	275	227
3.5	274	223
4	272	219
4.5	270	215
5	269	210

for various values of the sharpness parameter r . As can be seen, the performance is almost independent of the value of r . Figure 8(b) shows the average performance when the sound speed is estimated using the modified RBF network based on the standard KLD. It is clear that this network exhibits a higher degree of independence, with respect to the value of r , in comparison to the standard RBF network. Similar behavior was also found for both network types in the case of the substrate density estimation.

B. Inversion performance using the efficient design process P2

In this section, we study the performance of a network designed using the efficient process **P2**. The RBF network is constructed using the estimated $S\alpha S$ parameters of a subset containing $M=300$ signals chosen from the training set. Besides, let ε denote the error threshold, which we vary in the interval $[0.5, 5]$. In addition, the maximum number of neurons constituting the RBF network, if the error threshold is not achieved, is set equal to $M=300$. In the case of the standard RBF network, the value of σ varies in the interval $[0.05, 3]$, while in the case of the modified RBF based on the standard KLD, it varies in the interval $[0.001, 275]$. In both cases, the value of the sharpness parameter is set to $r=2$. For a given error threshold ε and for a fixed σ , we run 100 Monte Carlo iterations, where in each iteration a new RBF network is designed by randomly selecting a training subset of size $M=300$ from the training set.

Figure 9(a) shows the number of neurons in the standard RBF network, as a function of σ and for several values of the error threshold ε . We can see that for each value of ε , the maximum number of neurons is required to meet the corresponding error threshold, when the spread parameter is greater than 0.1. On the other hand, the number of neurons required to achieve the error threshold decreases as the value of ε increases, for $\sigma < 0.1$. This behavior should be expected, since we need more neurons in order to achieve an increased resolution of the input space, for a given value of σ .

Figure 9(b) shows the number of neurons in the modified RBF network employing the standard KLD, as a function of σ and for several values of the error threshold ε . We

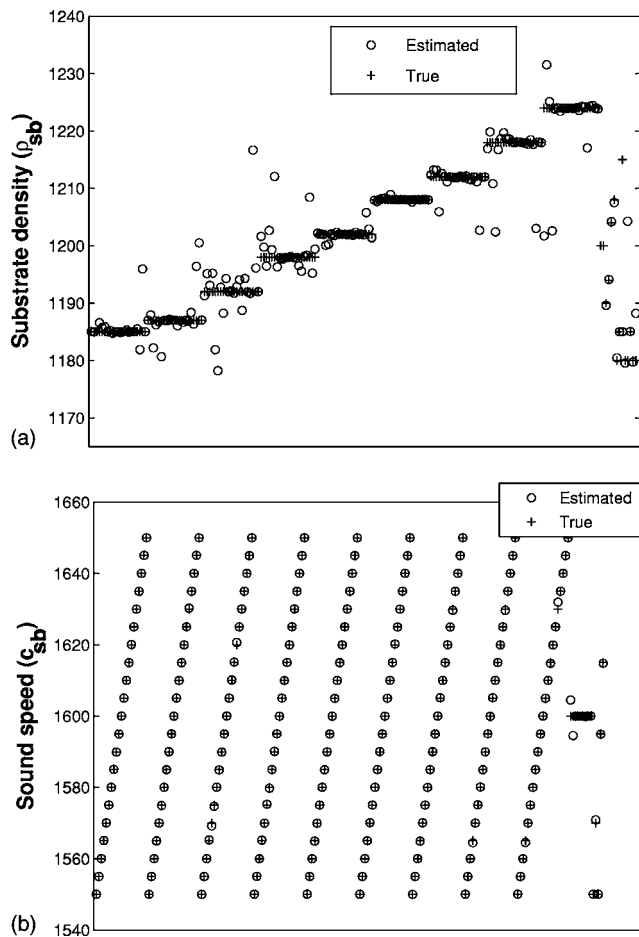


FIG. 10. True and estimated values of: (a) the substrate density and (b) the sound speed, obtained using the modified RBF network based on the standard KLD, with the optimal parameters of Table II.

can see that for each value of ε , the optimal number of neurons required to meet the corresponding error threshold does not follow the linear behavior as in the case of the standard RBF, but it strongly depends on the value of σ . However, the general rule that the number of neurons required to achieve the error threshold decreases as the value of ε increases, for a fixed σ , is still valid. Besides, by comparing Figs. 9(a) and 9(b), we can see that the proposed modified RBF network consists of less neurons than the standard RBF network, for the same error threshold ε . Table III SHOWS THE OPTIMAL number of neurons required to achieve the same error threshold, for both the standard and the modified RBF networks. It is clear that the proposed novel modified RBF network reduces significantly the number of hidden neurons, compared with its standard version. This reduction is important, since the number of hidden neurons affects the computational complexity of a neural network.

Finally, Figs. 10(a) and 10(b) show the true and estimated values for the substrate density and the sound speed, respectively, for a set of synthetic signals of our database (the horizontal axis is simply the signal index and, thus, omitted).¹⁶ The estimated values were obtained using the proposed modified RBF network based on the standard KLD, which is designed using the optimal parameters of Table II.

As was expected, the estimation accuracy of the sound speed is higher than the accuracy of the estimation of the substrate density.

IV. CONCLUSIONS

In this paper, an inversion scheme for acoustic signals recorded in shallow water was presented and evaluated, based on a $S\alpha S$ modeling of the coefficients of the 1D wavelet decomposition, followed by use of a novel RBF neural network. In particular, we demonstrated that the parameters of the $S\alpha S$ distributions constitute an effective set of features that can be employed for building a RBF neural network, which exploits the KLD between $S\alpha S$ distributions. This modified RBF network recovers efficiently the unknown environmental parameters of the recorded signal, achieving a decreased average error compared with the standard RBF network. Besides, our proposed RBF network requires less hidden neurons to achieve the same error threshold with its standard version. Our future work consists of testing the proposed scheme in real shallow-water environments, when the received signal is contaminated with noise. We also plan to study other neural network structures in order to achieve an even better performance of the inversion scheme.

¹*Inverse Problems in Underwater Acoustics*, edited by M. I. Taroudakis and G. N. Makrakis (Springer, Berlin, 2001).

²*Full Field Inversion Methods in Ocean and Seismic Acoustics*, edited by O. Diachok, A. Caiti, P. Gerstoft, and H. Schmidt (Kluwer Academic, Dordrecht, 1995).

³“Benchmarking geoacoustic inversion methods,” *J. Comput. Acoust.* (Special edition, K. Chapman and A. Tolstoy, eds.) **6**, 1–289 (1998).

⁴N. R. Chapman, S. Chin-Bing, D. King, and R. B. Evans, “Benchmarking

geoacoustic inversion methods for range-dependent waveguides,” *IEEE J. Ocean. Eng.* **28**, 320–330 (2003).

⁵M. Taroudakis, G. Tzagkarakis, and P. Tsakalides, “Classification of shallow-water acoustic signals via alpha-Stable modeling of the 1-D wavelet coefficients,” *J. Acoust. Soc. Am.* **119**, 1396–1405 (2006).

⁶M. Taroudakis, G. Tzagkarakis, and P. Tsakalides, “Characterization of an under-water acoustic signal using the statistics of the wavelet subband coefficients,” in *Theoretical and Computational Acoustics*, edited by A. Tolstoy, E.-C. Chang, and Y.-C. Teng (World Scientific, Singapore, 2006), pp. 167–174.

⁷Y. Stephan, X. Demoulin, and O. Sarzeaud, “Neural direct approaches for geoacoustic inversion,” *J. Comput. Acoust.* **6**, 151–166 (1998).

⁸J. Benson, N. R. Chapman, and A. Antoniou, “Geoacoustic model inversion using artificial neural networks,” *Inverse Probl.* **16**, 1627–1639 (2000).

⁹S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, New York, 1998).

¹⁰S. G. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation,” *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 674–692 (1989).

¹¹J. P. Nolan, “Parameterizations and modes of stable distributions,” *Stat. Probab. Lett.* **38**, 187–995 (1998).

¹²J. P. Nolan, “Numerical calculation of stable densities and distribution functions,” *Comput. Stat. Data Anal.* **13**, 759–774 (1997).

¹³S. Kullback, *Information Theory and Statistics* (Dover, New York, 1997).

¹⁴S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. (Prentice Hall, Englewood Cliffs, NJ, 1998).

¹⁵Notice that in both Figs. 5(a) and 5(b) the curve corresponding to $M = 50$ is not completely shown, since the y axis was scaled for a better visualization.

¹⁶Notice that the staircase-like view of Fig. 10(a) and the piecewise linear view of Fig. 10(b) is simply due to the way we grouped the signals in the data set. For convenience in the visual interpretation, we grouped the signals as follows: the first group contains signals with $c_{sb} \in [1550, 1650]$ m/s and the same density $\rho_{sb} = 1185$ kg/m³, the second group contains signals with $c_{sb} \in [1550, 1650]$ m/s and the same density $\rho_{sb} = 1187$ kg/m³, and so on.