Revista Brasileira de Engenharia Biomédica,
v. 23, n. 2, p. 143-151, abril 2007

ISSN 1517-3151

Artigo Original
Recebido em 27/02/2007, aceito em 05/06/2007

Generalization assessment of non-invasive data-driven temperature estimators from therapeutic ultrasound

Verificação da capacidade de generalização de estimadores de temperatura não invasivos "data-driven" em terapia ultra-sônica

César Alexandre Domingues Teixeira*
António Eduardo de Barros Ruano
Maria da Graça Cristo dos Santos Lopes Ruano
Centre for Intelligent Systems,
Faculty of Sciences and Technology,
University of Algarve
8005-139 Faro, Portugal
E-mail: {cateixeira, aruano, mruano}@ualg.pt

Wagner Coelho Albuquerque Pereira
Biomedical Engineering Program – COPPE/UFRJ
Caixa Postal 68510
21941-972 Rio de Janeiro, RJ
E-mail: wagner@peb.ufrj.br

*Corresponding author

Abstract
The objective of this work is the generalisation performance assessment, in terms of intensity, of non-invasive temperature models based on radial basis functions neural networks. The models were built considering data collected at three therapeutic ultrasound intensities, (among 0.5, 1.0, 1.5 and 2.0 W/cm^2) and then were validated in fresh data, which contain information from the trained intensities and form the untrained intensity. The models were built to estimate the temperature evolution (during 35 min) in a gel-based phantom, heated by physiotherapeutic ultrasound at four different intensities. It was found that the best models built without data from the intermediate intensities (0.5, 1.0 and 1.5 W/cm^2) perform well in validation at all the intensities. On the other hand, the models built without data from the extrapolated intensity (2,0 W/cm^2) presented unsatisfactory results in validation. This is because the models parameters were found considering a space bounded by the data used in their construction, and then the application of data outside this space resulted in poor performance. The models build without the intermediate data, for the three considered points, presented a maximum absolute error inferior to 0.5 ºC (which is accepted for therapeutic applications). The best models also presented a low computational complexity, as desired for real-time applications.
Keywords: Non-invasive temperature estimation, Data-driven models, Radial basis functions neural networks, Multi-objective genetic algorithms, Ultrasound, Physiotherapy.

Resumo
O objetivo deste artigo é a avaliação da capacidade de generalização, em termos de intensidade, de estimadores não invasivos de temperatura, baseados em redes neurais de funções de base radial. Os modelos foram treinados e a sua estrutura selecionada para estimar a evolução da temperatura em um phantom aquecido com ultra-som (tipicamente aplicado em fisioterapia) a quatro intensidades diferentes (0,5; 1,0; 1,5 e 2,0 W/cm^2). Neste trabalho, os modelos de temperatura foram construídos usando dados coletados em 3 das 4 intensidades e posteriormente validados em dados novos. Os dados de validação foram coletados nas quatro intensidades consideradas, ou seja, os modelos foram avaliados nas intensidades treinadas e numa intensidade não treinada, de forma a se poder realizar uma correta avaliação da capacidade de generalização. Os modelos construídos sem o uso dos dados coletados a intensidades intermédias (0,5; 1,0 e 1,5 W/cm^2), mostraram boa performance (erro máximo absoluto inferior a 0,5 ºC) em todas as intensidades. Por outro lado, quando utilizando os dados coletados na única intensidade extrema (2 W/cm^2), mostraram baixa performance em validação. Isto ocorre pelo fato da temperatura a ser estimada (coletada a 2 W/cm^2) estar fora do espaço delimitado pelos dados de treinamento, ou seja, o modelo estaria sendo usado para extrapolar valores de temperatura. Os melhores modelos apresentaram também uma baixa complexidade computacional, essencial para aplicações de tempo real.
Palavras-chave: *Estimação não invasiva de temperatura, Modelos "caixa preta", Redes neuronais de funções de base radial, Algoritmos genéticos multi-objetivo, Ultra-som, Fisioterapia.*

Introduction

A key factor for an accurate therapeutic device control system is the existence of precise temperature estimators, in time and space. For hyperthermia/diathermia applications, a spatial resolution of 1 cm^3 and a maximum absolute error of 0.5 ºC is desired (Arthur *et al*., 2005).

Many approaches for non-invasive temperature estimations were published, mainly based on electrical impedance tomography (Paulsen *et al*., 1996), microwave radiometry (Meaney *et al*., 1996), magnetic resonance imaging (MRI) (Hynynen *et al*., 1996), and backscattered ultrasound (BSU) (Arthur *et al*., 2005). It was reported that only the approaches based on MRI achieved the desired temperature resolution required for hyperthermia/diathermia. However, the cost of the MRI devices and the difficulty of handling it for some therapeutic modalities limit the development and success of temperature estimators. On the other hand, BSU needs much less expensive instrumentation, ultrasonic energy is non-ionizing, relatively simple signal processing tools are necessary to process data, and it is possible to use the same form of energy for heating and for temperature estimation. The application of BSU for non-invasive estimation was extensively reported in the past ten years. The most important features extracted from BSU that were used for estimation up to now are: temporal echo-shifts (*TS*) originated from changes on medium speed of sound, and medium expansion and contraction due to temperature changes (Simon *et al*., 1998); harmonic shifts, also due to temperature, induced expansion and contraction, which changes the mean scatterer spacing (MSE) of the medium (Seip and Ebbini, 1995); changes of the frequency dependent attenuation (Ueno *et al*., 1990); and changes on the backscattered energy form tissue inhomogeneities (Arthur *et al*., 2005). These published methods assume a linear relationship between features and temperature, need *a priori* determination of medium constants, and perform mathematical simplifications to reach linearity.

In this work, data were collected from a homogeneous gel-based phantom where three thermocouples were placed, representing the points where temperature is to be estimated. The medium was heated using therapeutic ultrasound at four different intensities (0.5, 1.0, 1.5 and 2.0 W/cm^2). The collected data were pre-processed and then fed to the entry of radial basis functions neural networks (RBFNNs) (Broomhead and Lowe, 1988). The multi-objective genetic algorithm (MOGA) (Fonseca and Fleming, 1993) was applied to find the best-fitted RBFNN structures. Well-fitted models for thermal therapeutic device control must have a good intensity generalisation capacity, in order to cope with intensity differences between the ones used in the therapeutic procedure and those used in the creation of the models, thus enabling a fine control. In this work four different types of MOGA runs were performed, corresponding to training and structure selection, always leaving one of the four intensities out. At the end of each MOGA run, the best attained models, selected according to a predefined parameterization (model errors, complexity and validity tests), were tested in fresh data (new data belonging to the intensities used in the models creation and data belonging to the intensity that was left out of the models creation), in order to assess the model's generalisation capacity.

Material and Methods

Experimental setup

The estimation studies presented in this work were based on backscattered ultrasound and temperature signals, collected from a homogeneous gel-based phantom. The phantom with dimensions 120 mm × 60 mm × 120 mm is composed of (in % weight): 86.5% of degassed water, 11% of glycerine, and 2.5% of agar-agar (Sato *et al*., 2003). The complete experimental setup is presented in Figure 1.

The phantom was immerged in a degassed-water tank to promote the coupling with the transducers, and to discard abrupt room temperature changes. The tank temperature was maintained at approximately 22 ºC, by using an aquarium heater (75 W) with an integrated thermostat. The phantom was heated by a therapeutic ultrasound (TUS) device (Sonopulse Generation 2000, Ibramed, Brazil), that emits a 1-MHz continuous wave. Four different TUS intensities were essayed (0.5, 1.0, 1.5 and 2.0 W/cm^2). The BSU signals were collected using an imaging ultrasound (IUS) transducer (V310SU, Panametrics-NDT, USA), working in pulse/echo mode at 5-MHz centre frequency. The IUS transducer is driven by a computer controlled pulser/receiver (5800PR, Panametrics-NDT, USA), which collects the radio frequency (RF) BSU signals and send them to an oscilloscope (TDS2024, Tektronix, USA) to be digitized at 50 MHz. Each RF signal is composed by 2,500 samples. The temperature was measured at the different points using k-type thermocouples, which were connected with a cold junction compensation (CJC) multiplexer (digital multimeter 2700/7700, Keithley, USA). In Figure 2, the position of the thermocouples is presented.

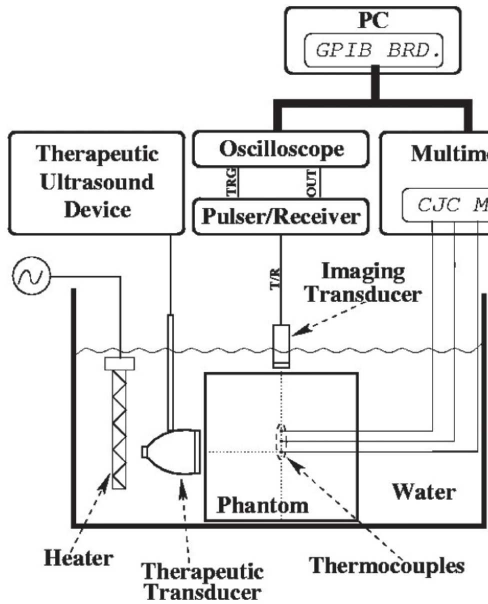

Figure 1. Experimental setup. The medium is heated by the therapeutic transducer, while an imaging transducer collect backscattered ultrasound signals from the medium. The temperature values were collected by three thermocouples placed at the points under study.

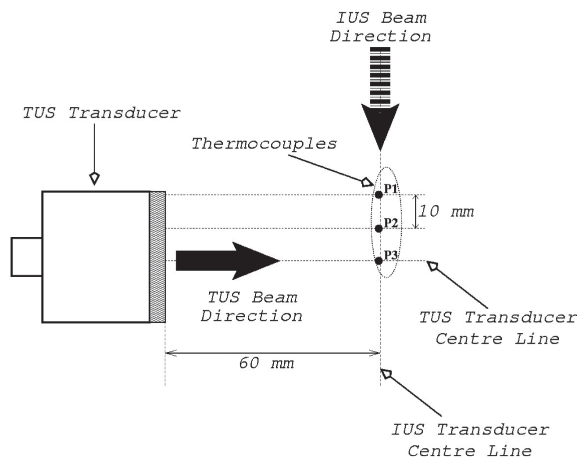

Figure 2. Thermocouples disposition in relation to the TUS and IUS beams.

The thermocouples were placed in points along the axial line of the IUS transducer, 10 mm spaced, and 60 mm apart from the TUS transducer face. In Figures 1 and 2 the relative position of the TUS and IUS transducers is presented. The transducers are perpendicular between each other, in order to reduce acoustic interference. Every 10 s the digitized BSU and temperature signals were transferred from the oscilloscope and from the multimeter respectively to a personal computer (PC) via a GPIB bus. On each experiment the temperature signal is composed by 5-min room-temperature baseline, 15-min heating curve and 15-min cooling curve.

Data pre-processing

In this work the temporal echo-shifts (*TS*) of the echoes coming from the thermocouples placed in the points under study, were extracted and assigned, in conjunction with the past temperature values, as RBFNN input variables. A method that directly evaluates continuous time-shifts from sampled data was applied. This method constructs a spline-based piecewise continuous representation of a sampled reference signal (in this case, the echoes of the first BSU signal in each experiment), then finds the minimum of the sum of squared errors between the reference and the delayed signals (in the case, the echoes of the other BSU signals throughout each experiment) to determine their relative time-shift (Viola and Walker, 2005). An example of the computed *TS* as compared to the measured temperature, for the three points (P1, P2 and P3), is presented in Figure 3.

Looking at Figure 3, it is possible to realise that for the phantom used, *TS* increases and decreases with temperature, meaning that the acoustic speed in the medium varies accordingly. However, there exists media where the acoustic speed decreases with the rise in temperature, e.g. fat tissue.

The computed *TS* and the measured temperature change signals were filtered and normalised to values between -0.5 and 0.5. It was applied a causal low-pass Butterworth filter (Proakis and Manolakis, 1995) with order 2 and cut-off frequency of 1/20 of the Nyquist frequency. These filter parameters were chosen to reduce noise while maintaining the signals fundamental behaviour. Normalization enables signals with different scales to be applied together as neural network inputs. To be noticed that in this work the temperature and the *TS* have very different scales, as it can be seen in Figure 3.

After normalisation and filtering, the temperature-change signals (ΔT) and the *TS* signals were arranged in separate files, according to the respective intensity and spatial point (P1, P2 and P3 in Figure 2), and applied to the RBFNN training, structure selection and validation. In the applied methodology, three different types of data sets were used: the training set, the test

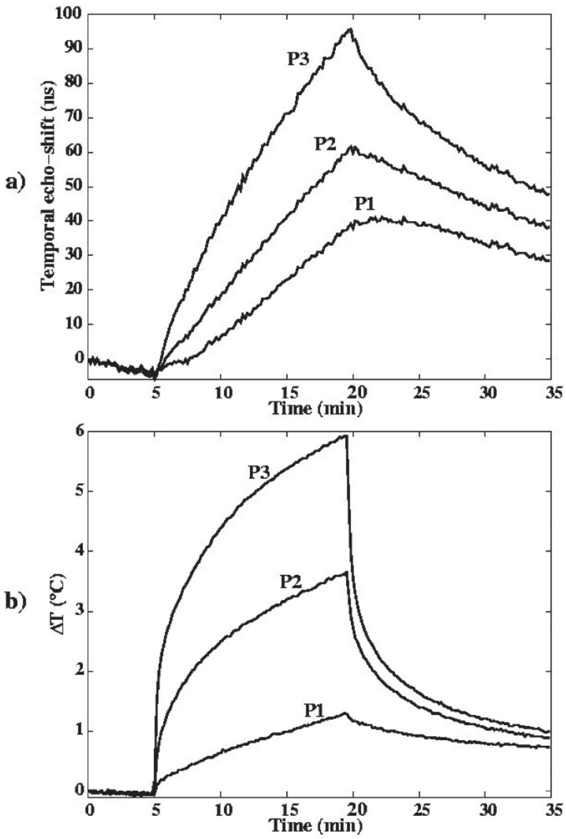

Figure 3. Example of the computed temporal echo-shifts as compared with the measured temperatures; a) temporal echo-shifts obtained for the different spatial locations, b) respective measured temperature change signals. P1, P2 and P3 refer to the phantom locations where temperature estimation is performed (see Figure 2).

set, and the validation set. The training set contains data used by the training algorithm to compute the RBFNN parameters. The test set is formed by data, different from those contained in the training set, which is used during the MOGA running to stop the training of each model, and to obtain generalisation performance descriptors, viewed as objectives to minimize. The validation set contains data, different from that present in both training and test sets, and was used to assess the models generalisation performance (in fresh data), after the MOGA structure selection. To be mentioned that both test and validation sets were subdivided into sub-sets. Each sub-set contains only data from one spatial point and from an intensity value. The test set is formed by nine sub-sets, which correspond to data from the three spatial points and from the three intensities applied. On the other hand, the validation set included 12 sub-sets, since four intensities were considered instead of three.

RBFNN model

In this work a non-linear autoregressive model with exogenous inputs (NARX) was applied. The model structure is presented in Figure 4.

This model is composed by a static RBFNN with external dynamics induced by their inputs. The model inputs considered were the past-lags of the computed *TS* and of the measured ΔT. A three layered feed-forward RBFNN was considered. The first layer is a set of inputs; the second layer performs a nonlinear transformation on the input data, and is composed by a set of processing elements (neurons). The third layer combines linearly the outputs of the hidden neurons to produce the overall output. The input/output relation is given by equation 1:

$$\Delta T(x_j) = b + \sum_{i=1}^{n} \alpha_i \varphi\left(\left\| x_j - c_i \right\|\right) \qquad (1)$$

where n is the number of neurons in the hidden layer, b is a bias term, $||.||$ represents the Euclidean norm, and $\{\varphi_i(.)\}_{i=1}^{n}$ is a set of radial basis functions centred

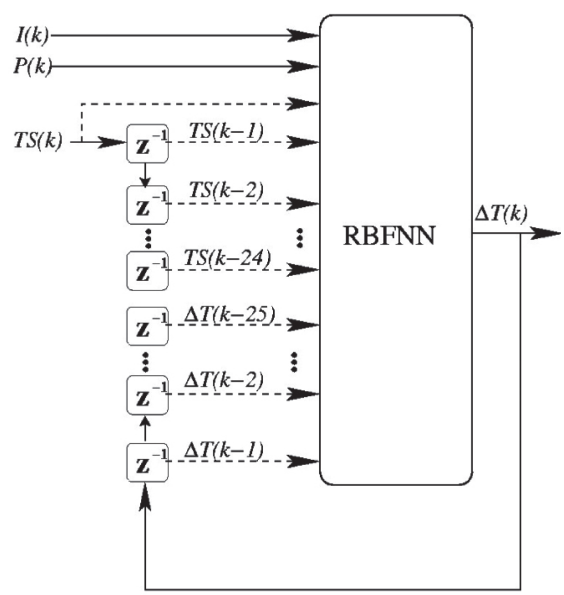

Figure 4. Model structure. The inputs represented by full line arrows are the fixed inputs, which were not selected by the MOGA and were always present. The inputs indicated by dashed line arrows represent inputs that can be selected by the MOGA. The symbol "$\mathbf{Z}^{-1}$" represents a pure time-delay of one sample.

at $\{c_i \in \Re^d\}_{i=1}^n$, being d the number of inputs. The basis functions are evaluated at points $x_j \in \Re^d$ and are weighted by $\{\alpha_i\}_{i=1}^n$ at the third layer. In this work it is applied the commonly used Gaussian basis functions (equation 2):

$$\varphi_i = e^{-\frac{1}{2\sigma_i^2}\left\|x_j - c_i\right\|^2} \quad (2)$$

where σ_i is the spread of the i^{th} function.

Model training and structure selection

The manual selection of the appropriate number of RBFNN neurons and the appropriate input set could be a high-time consuming task, due to the enormous number of possible solutions. The multi-objective genetic algorithm (MOGA) was used to evolve an initial and randomly generated population of models, in order to attain well-fitted non-invasive temperature estimations. As shown in Figure 4, the possible model inputs that MOGA can select were the lags of *TS* and *ΔT*. The maximum admissible lag for the input variables was considered as 25. This value was selected after studies on the temperature signals time-constants (Teixeira *et al*., 2006a). As the models were designed to estimate the temperature at different intensities and spatial positions, two additional inputs ($I(k)$ and $P(k)$ in Figure 4) were assigned to help the models to discriminate among intensities and spatial points. As these two inputs were always present, they were not selected by the MOGA. MOGA was allowed to evolve models that had a maximum of 20 inputs and a number of neurons in the interval [8,20]. This interval was experimentally selected after several runs considering wider values. In the MOGA loop, each individual must be trained in order to compute their parameters and consequently to obtain their performance descriptors, for the proposed estimation problem. Based on these descriptors, the population is ranked and the best models selected to create the new generation, by application of genetic operators. In this work the neural networks (NNs) were trained using a methodology that involves the Levenberg-Marquardt (LM) training algorithm and the minimization of an error criterion that exploits the separability of the parameters in non-linear ($\{c_i, \sigma_i\}_{i=1}^n$), and linear ($b$ and $\{\alpha_i\}_{i=1}^n$), improving the performance of the training (Ruano, 2005). The LM optimizes only the non-linear parameters, while the linear ones were found using the linear least-squares strategy.

As referred before, the search for appropriate model structures was formulated as a multi-objective optimization problem, where various model performance descriptors exist. These descriptors are viewed as objectives to minimize, in order to obtain well-fitted estimators. In this work, the performance descriptors were arranged in three groups: model errors, model complexity and model validity tests. The model errors considered were: the root mean square error in the training set (MSE_{TR}); the maximum mean square error in all the test sub-sets (MMSE_{TE}), obtained by feeding-back the measured temperature values; and the maximum absolute error obtained during estimation in all the test sub-sets (MAE_{TE}), obtained by feeding-back the estimated temperature values, as desired for a non-invasive temperature estimator. The model complexity descriptors under minimization were: the linear weights norm (LWN) and the total number of parameters (equation 3), which is taken as:

$$NP = NC \times NE + NS + NW \quad (3)$$

where NC is the number of centres, NE is the number of inputs, NS is the number of spreads, and NW is the number of linear weights. The choice of the LWN to minimize is because models with a high LWN usually are too specialised in the training data, and when considering other data sets tend to have an exacerbated error. The NP minimization is also important, because a good model should present a small error and also a small computational complexity, especially when real-time applications are desired. The model validity tests were proposed in Billings and Voon (1986) and used with success in similar modelling schemas (Teixeira *et al*., 2006b). These model validity tests involve the computation of first and higher order correlations between model inputs, outputs and residuals. In this work, as in Teixeira *et al*. (2006b), only the conditions involving the first order correlations were used, because the results obtained using the higher order correlations were not significantly better for the proposed estimation problem. The first-order correlations used were (equations 4 and 5):

$$R_{ee}(\tau) = \delta(\tau) \quad (4)$$

$$R_{ue}(\tau) = 0 \; \forall \tau \quad (5)$$

where $R_{ee}(.)$ is the auto-correlation of the error sequence, $R_{ue}(.)$ is the cross-correlation between inputs and the error, $\delta(.)$ is the Dirac's delta function, and τ is the time-shift or lag parameter associated with the correlation functions. In fact, $R_{ue}(.)$ will never be precisely zero for all lags, in this way the equality is considered true if the normalised value of $R_{ue}(.)$ lies within the 95% confidence limits defined as in equation 6:

$$CI = \frac{1.96}{\sqrt{N}} \tag{6}$$

where N is the number of training patterns. In the same way, the value of $R_{ee}(.)$ never equals the delta function, but the condition is considered true if the normalised value of the error auto-correlation enters the 95% confidence limit before lag one.

In order to obtain feasible solutions for a particular problem, goals and priorities should be defined for the previously referred objectives. The priorities enable the definition of relevance between the objectives. MOGA runs during a pre-defined number of generations evolving the initial randomly generated model population. At the end of a complete run, a set of best-fitted models, considering the goals and priorities defined, was obtained.

Results and Discussion

The MOGA was applied in four different run-types, each one corresponding to training and a structure selection without data from one of the intensities. The MOGA objectives, goals and priorities for the four run-types are presented in Table 1.

The error goals presented correspond to the normalised data. As previously said, the maximum accepted error for a desired estimator in hyperthermia/diathermia is 0.5 ºC, that is why the goal was arbitrary set for $MMSE_{TE}$ and MAE_{TE} as 0.43 ºC (a little bit bellow the threshold). If no goal is defined for MSE_{TR}, the MOGA tries to minimize towards zero, generating models too specialized in the training data, resulting in bad generalisation performance. It is possible to see that the objectives associated with the generalisation performance, i.e. the $MMSE_{TE}$ and MAE_{TE}, were defined with priority 2, having in mind the attainment of models with a good generalisation capacity. The LWN goal value was defined based on the maximum number of NN neurons defined, and in the data normalisation employed. The NP goal value was defined having in mind the search space defined.

The MOGA ran during 200 generation of 200 individuals each. At the end, the best fitted individuals were validated in data never applied in the training and structure selection. To note that this fresh data contains information from the trained and untrained intensities, in order to really assess the model's generalisation performance. In Table 2, the maximum absolute errors in the validation data (MAE_{VL}) presented by the best individual, in each run type are exposed.

Looking at this table it is possible to observe that for trainings and structure selections without the intensities 0.5, 1.0 and 1.5 W/cm^2, the applied methodology reaches maximum absolute errors inferior to 0.5 ºC (the gold-standard value for hyperthermia/diathermia) in the validation data. These intensities are intermediate (to note that the null intensity at the beginning and final part of each experiment curve is considered a intensity level, which makes 0.5 W/cm^2 an intermediate intensity), and the models can overcome the lack of their data in the training and structure selection by "seeing" data from both the correspondent lower and upper intensities. The 2.0 W/cm^2 is the upper intensity limit, which means that the heating and consequently the temporal echo are also upper limits. The lack of data from this intensity, during the training and structure selection, induces the applied methodology to obtain the model parameters with the inferior data limits (obtained with the inferior intensities). During the validation process, the application of data from 2.0 W/cm^2 results in bad performance (error greater than the gold-standard value) because it is out of the training and structure selection data domain.

The obtained values for the MOGA objectives for the best fitted-model in each of the four run types are presented in Table 3. Looking at this table it can be said that all the three error goals were fulfilled, in particular the high-priority ones, i.e. the $MMSE_{TE}$ and

Table 1. MOGA objectives, goals and priorities for the four run-types. The priorities enable the definition of relevance between the objectives.

Objective name	MSE_{TR}	$MMSE_{TE}$	MAE_{TE}	LWN	NP	$R_{ee}(.)$	$R_{ue}(.)$
Goal	3.0E-3 (0.043 °C)*	3.0E-2 (0.43 °C)*	3.0E-2 (0.43 °C)*	2.0	200	CI=2.4 E-2	CI=2.4 E-2
Priority	1	2	2	1	1	1	1

MSE_{TR} is the root mean square error in the training set; $MMSE_{TE}$ is the maximum mean square error in all the test sub-sets, obtained by feed backing the measured temperature values; MAE_{TE} is the maximum absolute error obtained in estimation for all the test sub-sets, obtained by feed backing the estimated temperature values; LWN is the linear weights norm; NP is total number of parameters; $R_{ee}(.)$ is the auto-correlation of the error sequence; $R_{ue}(.)$ is the cross-correlation between inputs and the error.

* Error goal values in the non-normalised data.

Table 2. Maximum absolute error in the validation data for the four run types applied.

Intensity excluded from training and structure selection (W/cm²)	**0.5**	**1.0**	**1.5**	**2.0**
MAE_{VL} (°C)	0.45	0.42	0.34	0.69

MAE_{TE}, as expected. The LWN goal was also fulfilled by all the four best individuals, showing that a small norm is desired for well-performing estimators. The goal defined for NP was fulfilled by three of the best-models, except by the best-model obtained in the run without data from 2.0 W/cm^2. The results obtained at the runs without the intermediate intensities show that it is possible to obtain low-complex models that present a MAE_{VL} inferior to 0.5 ºC, as desired in real-time applications. The goals defined for model-validity tests (R_{ue} and R_{ee}) were never fulfilled. In the case of R_{ue}, the obtained values are close to the defined goal. On the other hand, the values for R_{ee} are very far from the predefined goal, because the temperature waveforms are composed of two parts, an increasing and decreasing part, and the error among the different data sets, which compose the total training set, tends to have a common pattern (according to temperature increasing and decreasing) and the error autocorrelation function has high values after lag zero.

The structure parameters of the best individuals are presented in Table 4. Looking at this table, it can be said that the value defined for the maximum lag is sufficient for the attainment of good models, because it appears only two times in the best-models input set *TS*(k-24) in the run without data from 0.5 W/cm^2, and in ΔT(k-25) in the run without data from 2.0 W/cm^2). In terms of the number of inputs it can also be said that the chosen interval ([2,20]) is sufficient for the MOGA to achieve proper models, because the best models have inputs between 9 and 13, not converging to the maximum number. The bounds defined for the number of neurons were also well-chosen given that proper models do not have the maximum number of neurons. The best-model obtained in the run without the application of data from 2.0 W/cm^2 presents 18 neurons, showing that the smaller error obtained among the best models in this run-type demanded additional processing. On the other hand, an acceptable result (in terms of practical applications) was obtained, without increase in processing, in the runs without data from the intermediate intensities, where the models "know" the estimation space boundaries, and less processing requirements were necessary to achieve proper results for all the situations (three points and four intensities).

Future real applications encompass the attainment of the required error threshold in 1 cm^3 of tissue sample, as well as treatment duration of about 15-20 min, being the heating time per site of 1-2 min (Kitchen, 2003).

In Simon *et al*. (1998), high intensity focused ultrasound (HIFU) was applied to a rubber phantom, while BSU signals were collected. As in this work, temporal echo-shifts were then computed and applied for temperature estimation. The temperature was related with the derivatives of the temporal echo-shifts by means of a linear function. The accuracy of the method is assessed at three points (close to the two focuses of the HIFU), and using a unique intensity. The phantom was heated during 40 s and allowed to cool in the next 50 s; a maximum temperature increment of 4.22 ºC was reached. It is reported that a maximum error of 0.44 ºC was obtained. On the other hand, the proposed method in the present paper was validated at three different points and at four different intensities, which means that 12 operating points were considered (4 intensities × 3 points), instead of three (1 intensity × 3 points) in Simon *et al*. (1998). In terms of the maximum absolute error obtained, the proposed methodology reaches a value of 0.34 ºC for the best model built without data from 1.5 W/cm^2, which is 0.1 ºC less than the results pointed in Simon *et al*. (1998). This

Table 3. MOGA objectives presented by the best-individual, in each of the four run types, as compared with the a-priory defined goals.

		MOGA Objectives						
		MSE_{TR}	$MMSE_{TE}$	MAE_{TE}	LWN	NP	R_{ee}(.)	R_{ue}(.)
Intensity excluded (W/cm²)	0.5	2.1E-3	5.0E-3	2.7E-2	0.9	131	9.1E-1	8.5E-2
	1.0	1.6E-3	4.2E-3	2.8E-2	1.4	171	9.1E-1	5.3E-2
	1.5	6.4E-4	2.4E-3	2.9E-2	1.5	144	9.7E-1	3.3E-2
	2.0	2.9E-4	1.1E-3	2.8E-2	1.2	271	9.5E-1	3.7E-2
Goal		3.0E-3	3.0E-2	3.0E-2	2.0	200	*CI*=2.4E-2	*CI*=2.4E-2

Table 4. Structure parameters of the best-models obtained from the various run-types, as compared with the pre-defined possible values.

		Inputs (lags of *TS* and ΔT)			
		TS	ΔT	# of inputs	# of neurons
Intensity excluded (W/cm²)	0.5	0, 14, 20, 24	4, 5, 16, 20, 22	9	10
	1.0	1, 3, 6, 8, 16, 17, 18, 23	4, 5, 12, 18, 19	13	10
	1.5	4, 12, 13, 18, 20	2, 10, 14, 16	9	11
	2.0	2, 8, 15, 20	2, 3, 4, 7, 8, 16, 25	11	18
Pre-defined range values		[0, 24]	[1, 25]	[2, 20]	[8, 20]

accuracy was obtained for a total estimation time of 35 min (5-min baseline temperature + 15-min heating + 15-min cooling), and for a temperature increment of approximately 6 ºC.

Conclusion

The work hereby presented is on the assessment of the generalisation capacity of RBFNN, in non-invasive temperature estimation. The applied data-driven models were radial basis functions neural networks, with structure selected by the multi-objective genetic algorithm. The generalization studies performed were in function of the therapeutic ultrasound intensity, which means that the models were trained and their structure selected to perform well in data collected at three intensities, and then were validated in data collected at these three intensities and in data from one intensity not applied before. The applied methodology seems to work well for generalisations in data from the intermediate intensities, were maximum absolute errors inferior to the gold-standard value (0.5 ºC) accepted for hyperthermia/diathermia were obtained, values which are only competitive with those obtained with MRI temperature assessment as reported in literature. Excluding from training and structure selection the data from the extreme intensity, the maximum absolute error in validation is greater than 0.5 ºC, and the smaller error is obtained by increasing the processing complexity, which is reflected in an increase in computational demand. In runs without the intermediate intensities the models were trained and their structure selected knowing the estimation space boundaries, and inner estimations were obtained with success, in a natural way, without increase in processing requirements, even in situations not seen in the training and structure selection. From the above results it can be said that the maximum and minimum data limits must be provided during training and structure selection, and generalisations must be made inside the domain delimited by these data.

The merit of this work is the attainment of well-fitted estimators, which can overcome the lack of data from the intermediate intensities (0.5, 1.0 and 1.5 W/cm^2) in the training and structure selection, achieving maximum absolute errors inferior to 0.5 ºC in fresh data, collected at all the intensities. Another achievement is that both the medium parameters and structure were obtained from data, avoiding either mathematical simplifications or medium constant determination. In comparison with other temporal echo-shift based methods, such as the one published in Simon *et al*. (1998), a better accuracy was attained in a much more extensive estimation environment.

For the future it is planed to apply the estimation methodology to inhomogeneous media, and control the therapeutic instrumentation using the best fitted models. It is also planed to test the generalization performance of this kind of models, in terms of intensity and space, together.

Acknowledgements

The authors would like to thank: Fundação para a Ciência e a Tecnologia (grant SFRH/BD/14061/2003 and project POSC/EEA-SRI/61809/2004), Portugal; and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/CYTED/490.013/03-1), Brazil.

References

Arthur, R.M., Straube, W.L., Trobaugh, J.W., Moros, E.G. (2005), "Non-invasive estimation of hyperthermia temperatures with ultrasound", *International Journal of hyperthermia*, v. 21, n. 6, p. 589-600.

Billings, S., Voon, W. (1986), "Correlation based model validity tests for non-linear models", *International Journal of Control*, v. 44, p. 235-244.

Broomhead, D.S., Lowe D. (1988), "Multivariable functional interpolation and adaptive networks", *Complex Systems*, v. 2, n. 3, p. 321-355.

Fonseca, C.M., Fleming, P.J. (1993), "Genetic algorithms for multi-objective optimization: formulation, discussion and generalization", In: *Proceedings of the 5th International Conference on Genetic Algorithms*, Urbana-Champaign, p. 416-423, June.

Hynynen, K., Chung, A., Fjield, T., Buchanan, M., Daum, D., Colucci, V., Lopath, P., Jolesz, F. (1996), "Feasibility of using ultrasound phased arrays for MRI monitored non-invasive surgery", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, v. 43, n. 6, p. 1043-1053.

Kitchen, S. (2003), *Eletroterapia: Prática Baseada em Evidências*, 11th ed., São Paulo: Manole.

Meaney, P.M., Paulsen, K.D., Hartov, A., Crane, R.K. (1996), "Microwave imaging for tissue assessment: initial evaluation in multitarget tissue-equivalent phantoms", *IEEE Transactions on Biomedical Engineering*, v. 43, n. 9, p. 878-890.

Paulsen, K.D., Moskowitz, M.J., Ryan, T.P., Mitchell, S.E., Hoopes, P.J. (1996), "Initial *in vivo* experience with EIT as a thermal estimator during hyperthermia", *International Journal of Hyperthermia*, v. 12, n. 5, p. 573-591.

Proakis, J.G., Manolakis, D.G. (1995), *Digital Signal Processing: Principles, Algorithms, and Applications*, 3rd ed., New Jersey: Prentice-Hall International.

Ruano, A.E. (Ed.) (2005), *Intelligent Control Systems using Computational Intelligence Techniques*, London: IEE Press.

Sato, S.Y., Pereira, W.C.A., Vieira, C.R.S. (2003), "*Phantom* para medição da faixa dinâmica de equipamentos de ultra-som biomédicos", *Brazilian Journal of Biomedical Engineering*, v. 19, n. 3, p. 157-166.

Seip, R., Ebbini, E.S. (1995), "Noninvasive estimation of tissue temperature response to heating fields using diagnostic ultrasound", *IEEE Transactions on Biomedical Engineering*, v. 42, n. 8, p. 828-839.

Simon, C., VanBaren, P., Ebbini, E.S. (1998), "Two-dimensional temperature estimation using diagnostic ultrasound", *Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, v. 45, n. 4, p. 1088-1099.

Teixeira, C.A., Ruano, M.G., Ruano, A.E., Pereira, W.C.A., Negreira, C. (2006a), "Single black-box models for two-point non-invasive temperature prediction", In: *Proceedings of the 6th IFAC Symposium on Modeling and Control in Biomedical Systems* [*MCBMS'06*], Reims, p. 135-140, 20-22 Sep.

Teixeira, C.A., Ruano, A.E., Graça Ruano, M., Pereira, W.C.A., Negreira, C. (2006b), "Non-invasive temperature prediction of in vitro therapeutic ultrasound signals using neural networks", *Medical & Biological Engineering & Computing*, v. 44, n. 1-2, p. 111-116.

Ueno, S., Hashimoto, M., Fukukita, H., Yano, T. (1990), "Ultrasound thermometry in hyperthermia", In: *Proceedings of the IEEE Ultrasonic Symposium*, Honolulu, v. 3, p. 1645-1652, 04-07 Dec.

Viola, F., Walker, W.F. (2005), "A spline-based algorithm for continuous time-delay estimation using sampled data", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, v. 52, n. 1, p. 80-93.