Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online
CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-5 (2016) 297–302

# Non-invasive modelling of ultrasound-induced temperature in tissues: a b-splines neural network solution

**Ferreira, R.** * **Ruano, M.G.** ** **Ruano, A.E.** ***

* *Faculty of Science and Technology, University of Algarve, Faro, Portugal, (e-mail: raufer92@gmail.com).*
** *Faculty of Science and Technology, University of Algarve, Faro, Portugal and CISUC, University of Coimbra, Portugal (e-mail: mruano@ualg.pt )*
*** *Faculty of Science and Technology, University of Algarve, Faro, Portugal and IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal, (e-mail: aruano@ualg.pt )*

**Abstract:** Efficient hyperthermia therapy session requires knowledge of the exact amount of heating needed at a particular tissue location and how it propagates around the area. Until now, ultrasound heating treatments are being monitored by Magnetic Resonance Imaging (MRI) which, besides raising the treatment instrumental cost, requires the presence of a team of clinicians and limits the hyperthermia ultrasound treatment area due to the space restrictions of an MRI examination procedure. This paper introduces a novel non-invasive modelling approach of ultrasound-induced temperature in tissue. This comes as a cost effective alternative to MRI techniques, capable of achieving a maximum temperature resolution of 0.26 $^oC$, clearly inferior to the MRI gold standard resolution of 0.5 $^oC/cm^3$. Furthermore, we propose an innovative modelling methodology, where various similar models are built and are further combined through an optimization procedure, that we call neural ensemble optimization (NEO). This combination mechanism is shown to be superior to more simple schemes such as simple averages or evolutionary strategy based techniques.



*Keywords:* Ultrasound, Tissue Temperature Modelling, B-spline Neural Networks

## 1. INTRODUCTION

Efficient hyperthermia biomedical practice demands knowledge about the exact amount of heating required at a particular tissue location as well as information concerning the spatial heating distribution. Both processes are required to be accurately characterized. Until now ultrasound heating treatments are being monitored by Magnetic Resonance Imaging (MRI), recognized as being capable of achieving a 0.5 $^oC/cm^3$ temperature resolution (Wyatt et al. [2009]), the gold standard in this field. However, MRI-based techniques, besides raising the treatment instrumental cost, demand the presence of a team of clinicians and limits the hyperthermia ultrasound treatment area due to the space restrictions of an MRI examination procedure.

Different approaches have been proposed for non-invasive temperature estimation. These techniques are based on electrical impedance tomography (EIT) (Edd et al. [2005]), microwave thermometry (Arunachalam et al. [2009]), magnetic resonance imaging (de Senneville et al. [2005]), and backscattered ultrasound (BSU) (Straube et al. [2005]). Focusing on *ultrasound* based techniques to estimate the temperature, several methods have been reported, based on the extraction of temporal-echo shifts (Teixeira et al. [2008] and Teixeira et al. [2010]) and frequency shifts (Lei et al. [2013]), changes on the attenuation coefficient (Fukukita et al. [2002]).

Instead of relying on the physical properties of tissue, an intelligent computation model, based on measured data, is proposed in this work to model the temperature propagation inside tissues, when these are subject to an ultrasound heating session. In order to derive the model, as *in-vivo* studies are not ethical, a phantom was utilized, whose composition reflects the ultrasound reactions of human tissues. This way, although the data used to construct the model was captured in an invasive way, using the phantom, once the model is obtained, it can be applied in a non-invasive way, in human tissues.

The biological complexity inherent to human tissue results in a media with highly variable characteristics, giving shape to different dynamics involved in the temperature propagation process. One should expect slightly distinct behaviours in a response to changes with respect to different body parts, gender and age group. The temperature model to be designed should be robust enough to contemplate this variability, and capable of modeling temperature dynamics in terms of space and time dimensions. We propose to contaminate the data set, by

⋆ A. E. Ruano would like to acknowledge the support of the Portuguese Foundation for Science and Technology, through IDMEC, under LAETA, project UID/EMS/50022/2013.


10.1016/j.ifacol.2016.07.129

deliberately passing it through a Gaussian contamination process. Besides contemplating possible patient-dependent variabilities, this contaminated version can be seen as a simulated non-invasive data set, as if it was acquired by a non-invasive temperature monitoring method, whose maximum error can be approximated by the standard deviation that parametrizes the Gaussian distribution.

In this work, we propose additionally an innovative modelling methodology, where various similar models are built and further combined in a novel way, through an optimization approach denoted as *Neural Ensemble Optimization* (NEO).

The layout of the paper is: in Section 2, the material used and the experimental setup are described; Section 3 describes the models and the modelling approaches employed; Results are presented in Section 4, Conclusions and future work are drawn in Section 5.

## 2. MATERIALS AND EXPERIMENTAL SETUP

### *2.1 Materials*

In order to simulate human tissue, a phantom introduced in (Vieira et al. [2010]) was used. These mimicking solutions are named *phantoms* and they intend to exhibit similar characteristics to the ones found in human tissues. The basic composition of the solution employed is presented in Table 1.

**Table 1. Composition of the homogeneous solution used to mimify human tissue.**

| Material | % Composition |
|---|---|
| water | 86.5 |
| glycerol | 11.0 |
| agar | 02.5 |

For the localized heating of the phantoms, a therapeutic ultrasound device (US), *Sonopulse Generation 2000 Ibramed*, was used. Temperature at the spatial points (inside tissue) under study was measured using *type-K* thermocouples, connected to a compensation module (*80TK Fluke Everett, WA, USA*). This module is then connected to a digital multimeter (*2700/7700, Keithey*), which digitalises the temperature and makes it available to a general purpose PC. These temperature values were transferred to the PC via a GPIB bus (*GPIB-USB-B, National Instruments*).

### *2.2 Experimental Configurations*

The experiments were conducted following the experimental setup illustrated in Fig. 1.

A parallelepiped *phantom* holds five thermocouples, connected to a multimeter, invasively measuring the temperature inside the media. The therapeutic ultrasound device is placed at one exterior side of the *phantom*. Four beam intensities were considered: 0.5, 1.0, 1.5 and 1.8 $W/cm^2$. Each experiment trial has a 45 minute duration, divided in three phases: temperature stabilisation, heating and cooling. Temperature samples, measured by the sensors, are taken each 10 seconds. Therefore, we have available 4 beam intensities $\times$ 5 sensors $\times$45 minutes $\times$6 $(samples/min) = 5400$ temperature data points. Each

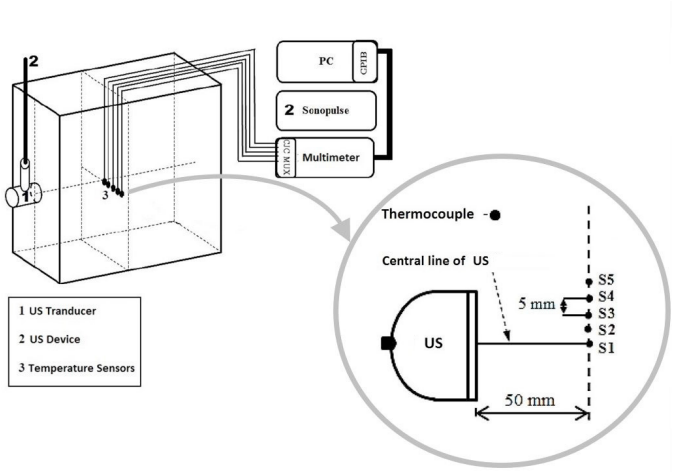


Fig. 1. Schematic diagram of the experimental configuration. Five thermocouples are placed inside the homogeneous *phantom*. These register the temperature variations induced by the ultrasound device (1).

*operating point* considers a single beam intensity, measured at a single *spatial* point, which results in $OP = 5$ sensors $\times$ 4 intensities = 20 operating points. Fig. 2.2 depicts the whole data set acquired, to be used in the construction of the models.

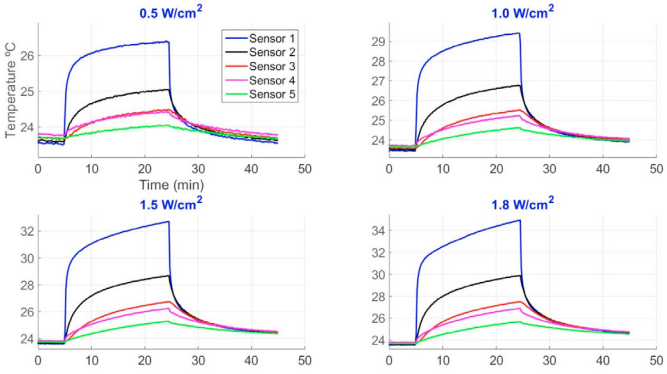


Fig. 2. Data acquired by all sensors considering four US beam intensities. Two distinct dynamics can be identified: one concerning the heating phase, and a second observed when the *phantom* naturally cools down.

Furthermore the data is interpolated using a cubic spline, from MATLAB's *Curve Fitting Toolbox*, resulting in curves with a temporal resolution of 1 second, as opposed to the 10 second resolution obtained in the experiment.

## 3. MODELING METHODOLOGY

During the modeling process, the complexity of the models is *gradually* increasing. We start by considering models for *single-point* and *single-intensity* estimation. We then gradually move the complexity of the models towards *multi-point* and *multi-intensity* estimation.

Fig. 3 summarises the key points involved in the modeling approach applied. B(asis)-splines functions (de Boor [1992]) are used to structure the networks, which results in structures known as *b-spline neural networks* (BSNN). B-spline functions allow for a local control of the curve shape,

since changing a control point has only a *local* effect, i.e. the change is not propagated through the entire shape. To model the unknown function of temperature propagation, a finite set of input-output observations is available. However, expanding the spline representation to a high dimensional space, by adding more input variables, implies increasing the number of parameters to impractical levels. Furthermore the identification of a proper model structure is of crucial importance. The structure should be general enough to capture the dynamics of the system, i.e. reduced bias, while preventing the occurrence of overfitting, i.e. high variance, which usually occurs when the number of free parameters is large, with respect to the true system. In order to overcome this issues, the ASMOD (Adaptive Spline Modelling of Observation Data) algorithm (Kavli [1993]) was employed. By making use of the observed data, coupled and decoupled dependencies are identified in the data set. Thus, the ASMOD output variable is modelled as a sum of several low dimensional submodels, having each submodel depend on a small subset of the input variables. This decomposition of higher dimensional input spaces results in more parsimonious models. The algorithm starts with a base simple model and iteratively *refines* it by adding more variables, coupling variables or changing the internal structure of the basis functions. Finally parsimonious models are forced in the *pruning* step, where variables are decoupled and the basis functions see their degree reduced and internal structure simplified. The refining/pruning steps continue until a stopping criterion is met. An ASMOD model $m$ is thus a linear combination of B-spline basis functions:

$$m(x) = \sum_{i=1}^{L} c_i b_i(x) = c^T b(x) \tag{1}$$

where $L$ is the number of basis functions, $c$ is a coefficient vector and $b(x)$ is a vector of B-splines. For each candidate model $m$ considered in the algorithm, the parameter $c$ must minimise the expected value of a chosen error function $f(.)$,

$$\hat{c} = \arg\min_{c \in A} E[f(y - c^T b(x)] \tag{2}$$

where $A$ is a set which can be used to introduced a priori knowledge about $\hat{c}$ domain.

However, notice that we only have available a finite set of observed data $D$. Additionally, the true underlying probability distribution is unknown, hence we cannot evaluate $E[f(y - c^T b(x)]$. An alternative is to use the maximum likelihood estimate, by making use of the available data. We are therefore interested to minimise an empirical cost function,

$$C_{emp} = \frac{1}{N} \sum_{i=1}^{N} (y_i - c^T b(x)_i)^2 \tag{3}$$

whose solution can be analytically found.

Once we have available an algorithm that provides us as output a accurate model with a suitable, parsimonious structure, is possible to go one step further and use this algorithm to obtain several of such models. By doing so, the accuracy of the predictions is enhanced by combining several individual predictions $y_{i \in N}$ into a final output $y$, with $N$ being the set of models considered.

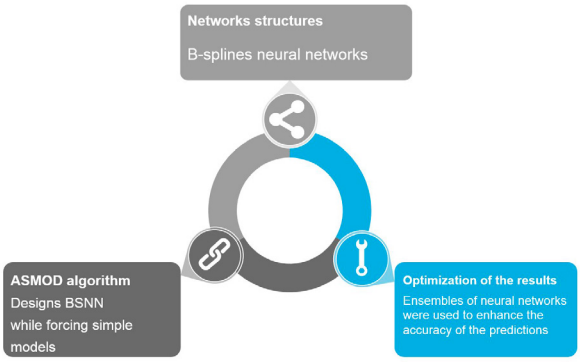


Fig. 3. A summary of the modeling approach applied, with the three main components identified: the model's structure, evolution process and optimization.

As explained, the contamination of the temperature samples registered by the sensors is central to assess our approach, since it simulates the use of non-invasive data to construct the models. We now characterize the *Gaussian contamination process*, crucial to the assessment of the results. By intentionally corrupting our original data set, we can account for the variability characteristic of the temperature propagation process. Here it is assumed that such method would have its error governed by a Gaussian distribution. The contamination is made on a *point-to-point* basis. This means that each temperature value in a curve is independently contaminated with a real number drawn from a Gaussian distribution, $e_i \sim \mathcal{N}(0, \sigma)$. Furthermore we use an additive contamination process, i.e. assuming $y_o$ is the original temperature sample, the corrupted version $y_c$ is:

$$y_c = y_o + e_i, \quad e_i \sim \mathcal{N}(0, \sigma) \tag{4}$$

The distribution is parametrised with a zero mean $\mu = 0$ and a standard deviation $\sigma = 0.2$,

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} = 1.99 e^{-12.5x^2} \tag{5}$$

Therefore, the probability of the error being lower than 0.6 $^oC$ is $P(error < 0.60) \approx 99.7\%$, i.e. a reliable estimator. The contamination methodology is applied as follows. Firstly, the original data set is contaminated and the result is used to train and validate the models. Afterwards the initial original set is used to test the model. The use of such methodology ensures the model is valid over the *true* process dynamics. Fig. 3 depicts the general BSNN structure used.

The structure uses $m$ past temperature values $T[k]$, with $m$ being the number of considered lags, which we constrained to 5, since it was observed that additional lags were only introducing more complexity in the system, without significant improvements in terms of performance. $\hat{T}[k+1]$ represents the one step ahead temperature value estimated by the network. $I[k]$ denotes the US device intensity, as measured in $W/cm^2$, providing information to the model about the current beam intensity. $P[k]$ provides information about the current spatial location of the input pattern. This input is numerically represented by the angle

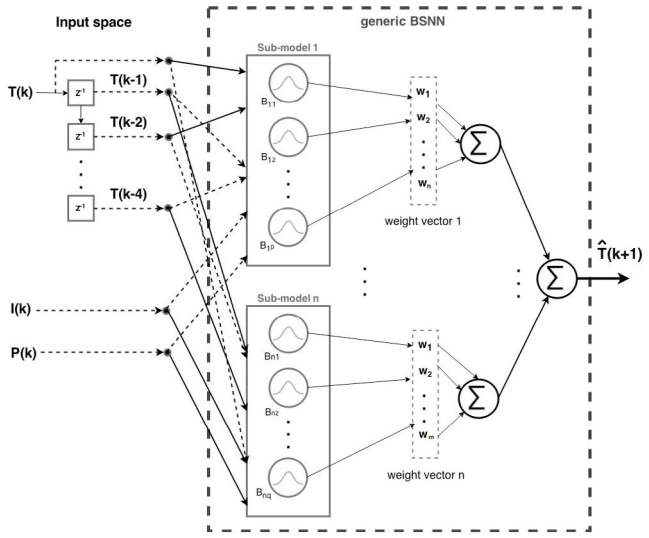


Fig. 4. B-spline neural network (BSNN) general structure considered in the most complex scenrario: multi-point multi intensity. The decomposition in *sub-models* is achived by the ASMOD algorithm, resulting in a simpler structures that are *additively* combined. Simpler environments don't make use of spatial information, $P[k]$, and/or intensity information $I[k]$.

that the sensor makes with the US device central line beam (Fig. 1). Each one of these inputs is mapped into a *single* sub-model, which can deal with one or more variable. The first layer in the structure performs a fixed *non-linear mapping* from the input space $\mathbf{X}$ to the output of the basis functions space $\mathbf{A}$, $\mathbf{X} \rightarrow \mathbf{A}$. The output space of the b-splines $\mathbf{A}$ is then *linealy mapped* to the network's output space $\mathbf{Y}$, $\mathbf{A} \rightarrow \mathbf{Y}$. The linearity of the last layer, with respect to the weights $\omega$, provides a mathematical convenient way of adapting the parameters, since we can analytically solve this minimization problem in one step, a solution know as *ordinary least squares* (OLS).

$$\hat{\mathbf{w}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y} \tag{6}$$

where $\mathbf{A}$ is a matrix of size ($m$ x $n$), whose $m^{th}$ row is composed of the transformed input vector for the $m^{th}$ pattern, assuming the network is built using $n$ basis functions in the second layer. $\mathbf{y}$ is the vector of desired outputs of length $m$ and $\hat{\mathbf{w}}$ is the optimal weight vector.

We started by considering models for *single-point and single-intensity* (SPSI) estimation. Then gradually, the complexity of the models is increased towards *multi-point and multi-intensity* (MPMI) estimation. Naturally, SPSI models only use temperature information in their input space ($T[k]$ to $T[k-4]$ in Fig. 4). On the other hand, MPMI models require additionally spatial and intensity information, $P[k]$ and $I[k]$ respectively. These inputs provide the model the ability to discriminate temperature variations according to the region and the US beam intensity being employed. Furthermore to improve the accuracy of the predictions, ensembles of neural networks are employed. Four models, with temperature lags varying from 2 to 5, are trained and combined into a final output. These techniques exploit improvements that one can obtain when combining a set of models that ideally are *negatively correlated* among

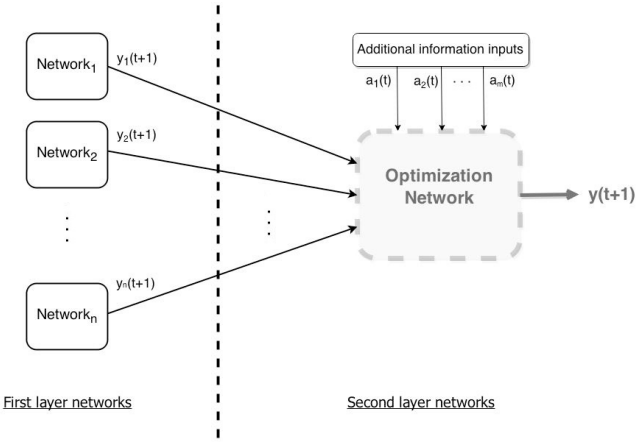


Fig. 5. Two layered network optimization aquitecture. The second layer combines all the individual outputs from the first layer network, by performing a non linear mapping $\mathbb{R}^{n+m} \rightarrow \mathbb{R}$. With $n$ being the number of individuals in the first layer, and $m$ is the number of additional inputs.

them. Assuming that the performance of the set of models $S$ models is comparable, and if their outputs are negatively correlated, simply averaging between them can result in major performance improvements, as elegantly introduced by [Krogh et al., 1995]. Instead of simply taking the average of the ensemble we propose an innovative approach to combine the individual networks. We make use of a *second* BSNN layer that acts as an optimizer agent, which we call *neural ensemble optimization* (NEO), combining the information from all the different sources. This concept is materialized in the archtecture illustrated in Fig. 5.

The final layer output is given by:

$$y[k+1] = f(y_1[k+1], \ldots, y_n[k+1], a_1[k], \ldots, a_m[k]) \tag{7}$$

where $y_i(t)$ represents the individual next-step prediction of network $i$, and $a_i(t)$ consists of additional information up to moment $k$, that may guide the network towards an optimal combination of $y_i(t)$. $f(\ldots)$ is the underlying function to the neural network, performing a non linear mapping $\mathbb{R}^{n+m} \rightarrow \mathbb{R}$. $n$ is the length of the ensemble, and $m$ is the number of (optional) additional inputs. In this application, the additional information is $I[k]$ and $P[k]$. The proposed ensemble combination approach is compared with two others combination schemes: a) simply averaging (SA) the individual outputs in the second layer; b) a weighted sum of all the individual outputs in the second layer, where each weight $\omega_i$ is obtained using an evolutionary strategy (ES) with uncorrelated mutations and $n$ step sizes [Eiben et al., 2003].

During the modeling process, we make the distinction between two phases observed during the evolution of the temperature in the *phantom*: a heating phase and a cooling phase. These two distinct phases are observed to have different dynamics (recall Fig. 2). This is intuitively true since the physical process that generates them is undoubtedly different and governed by different rules. Note that the heating is forced by the US device, whereas on the other hand, once the device is turned off, the phantom cools down in a natural way, without an external source doing work on the system. This observation suggests an

implementation that incorporates this distinction, which can be done by considering two models, one for each phase. The data is divided into two subsets, one for each cycle of the process. Each one of the subsets is used to construct the corresponding model.

## 4. RESULTS

The data depicted in Fig. 2 is corrupted by a Gaussian contamination process, resulting in the transformed data set presented in Fig 6.

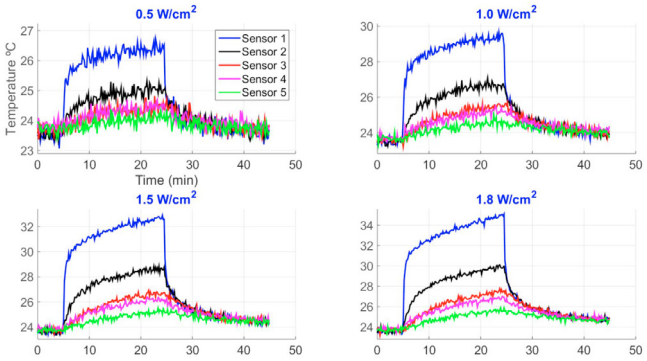


Fig. 6. Result of the Gaussian contamination process applied to the data of Fig. 2.2. This data was further divided and used to train and validate the models.

These temperature curves concern data collected by different sensors, placed at difference spatial locations (Fig. 1), using four different US beam intensities. Therefore, this data was used to construct a *multi-position multi intensity* (MPMI) model. Refering to the contaminated set, the validation set included the following curves: sensor 2: $0.5W/cm^2$, sensor 3: $1.0W/cm^2$, sensor 4: $1.5W/cm^2$ and sensor 5: $1.0W/cm^2$. The remaining curves were used to train the model. We considered one-step ahead predictions, with a prediction horizon $h = 60$ seconds. The models, once designed, were tested using the original uncorrupted data, Fig. 2.

The criteria used for comparing the different models are the$MSE$, $MSE_v$ and $MSE_t$ which represent the average (over the data considered for each set) mean square error (MSE) obtained in the training, validation and test set, respectively. Additionally, since this is a biomedical application, the average mean square error over the test set, $M_{ae}$, is a critical indicator, necessary to assure the patient safety. As a starting point, for comparison purposes, two additional ensemble mechanism were compared against the NEO approach: i) simple average (SA) and ii) an evolutionary strategy (ES) with uncorrelated mutations and n step sizes. The average MSE error obtained in the test set, i.e. the generalization error, is compared with the traditional keep-the-best model approach (KTB), considering 4 different BSSNs whose structures are depicted in Fig. 4, with temperature delays varying from 1 to 4.

Table 2 shows the improvement (in percentage) of performance of the three ensemble strategies, relatively to the KTB approach. The comparison was done over two model typologies: single-position single-intensity and single-position multi-intensity. NEO clearly achieved the best results in a consistent way, i.e. improving the performance in every case considered. The SA method is clearly a bad approach, followed by the ES option, whose action allowed, in some cases, considerable performance gains. NEO outperforms ES simply because the evolutionary approach, in the optimization phase, assigns the largest weight to the network that, on average, performs better, i.e. the KTB model. NEO explores a wider search space and has the flexibility of assigning dynamic weights as a function of the current operating point, thus retaining information about which model performs better at which region. Note that this second level information is lost in the ES approach.

**Table 2. Generalisation (test) error comparison between the various ensemble methods employed and the traditional KTB approach. The best values are denoted in bold. A negative value means a worst performance.**

| | | | Improvement (%) | | |
|---|---|---|---|---|---|
| Typ. | #S | I. ($W/cm^2$) | SA | ES | NEO |
| SPMI | 1 | 0.5 | -52.09 | 17.24 | **36.77** |
| | | 1.0 | -51.40 | -33.10 | **34.50** |
| | | 1.5 | -10.85 | -17.42 | **26.84** |
| | | 1.8 | -02.11 | -05.94 | 00.99 |
| | 2 | 0.5 | 57.04 | 85.24 | **95.33** |
| | | 1.0 | -166.7 | -22.60 | **33.23** |
| | | 1.5 | -129.7 | -52.12 | **17.00** |
| | | 1.8 | -78.54 | -175.9 | **17.87** |
| | 3 | 0.5 | -110.3 | **73.08** | 70.13 |
| | | 1.0 | -88.92 | 03.47 | **78.06** |
| | | 1.5 | -92.06 | 26.93 | **35.42** |
| | | 1.8 | -170.2 | 19.87 | **28.36** |
| | 4 | 0.5 | -149.4 | 45.17 | **83.53** |
| | | 1.0 | -117.4 | 40.10 | **49.99** |
| | | 1.5 | -100.0 | **38.88** | 21.15 |
| | | 1.8 | -107.7 | 37.42 | **64.08** |
| | 5 | 0.5 | -122.3 | 32.18 | **53.48** |
| | | 1.0 | -216.1 | 70.01 | **82.41** |
| | | 1.5 | -245.1 | **60.14** | 57.28 |
| | | 1.8 | -107.4 | 24.99 | **34.04** |
| SPMI | 1 | all | -00.35 | 16.26 | **73.69** |
| | 2 | all | -66.12 | 41.23 | **53.20** |
| | 3 | all | -101.4 | 69.83 | **70.98** |
| | 4 | all | -59.04 | **84.79** | 84.72 |
| | 5 | all | -104.6 | **80.33** | 78.48 |

Fig. 7 presents the results obtained in the test set, obtained by the NEO approach.

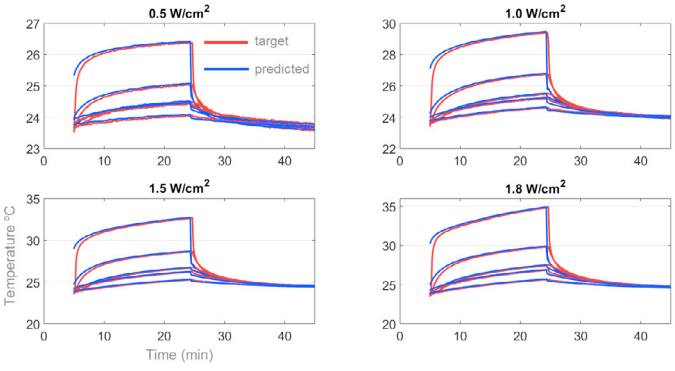


Fig. 7. Comparison of the desired behaviour (black line) and the values estimated by the network (grey line).

The proposed modelling approach was able to construct a reliable approximation of temperature propagation experienced by the phantom, throughout the two dimensions considered: space and US beam intensity. Table 3 reports the performance criteria obtained.

**Table 3. Performance criteria obtained for the MPMI model, trained with the data from Fig. 6 and tested with the data set from Fig. 2.**

| Criterion | Heating | Cooling |
|---|---|---|
| $MSE\ ^oC$ | 0.0324 | 0.0220 |
| $MSE_v\ ^oC$ | 0.0332 | 0.0354 |
| $MSE_t\ ^oC$ | 0.0606 | 0.0256 |
| $M_{ae}\ ^oC$ | 0.2629 | 0.2151 |

From the analysis of these criteria, the success of the modelling approach is clear. In particular, the $M_{ae}$ obtained is 0.2629, clearly below the gold standard of 0.5 $^oC/cm^3$.

The maximum absolute error $M_{ae}$ is a crucial indicator for a biomedical application. Table 4 exposes the variation of this criterion through all of the model typologies considered, confronting the results obtained using the uncorrupted data set, with the ones observed with a contaminated data set.

**Table 4. Maximum absolute error obtained for different modelling typologies, with different complexities.**

| | Maximum Absolute Error $^oC$ | | | |
|---|---|---|---|---|
| | uncorrupted | | corrupted | |
| **Typology** | **Average** | **Max** | **Average** | **Max** |
| SPSI | 0.1212 | 0.3230 | 0.0831 | 0.3039 |
| SPMI | 0.1379 | 0.3252 | 0.0564 | 0.1666 |
| MPMI | 0.0475 | 0.0510 | 0.2345 | 0.2629 |

The results show the maximum absolute errors were consistently kept under 0.24 $^oC$, providing a comfortable margin when confronted with the 0.5 $^oC$ MRI reference.

Is also interesting to note that the maximum absolute error did not scaled with the increasing model complexity, with all typologies registering comparable values. This is a very meritorious indicator that suggests the reliability of the system can scale along side with the modelling environment complexity.

## 5. CONCLUSIONS

In this paper a novel B-spline non-invasive approach for modelling ultrasound-induced temperature in tissues has been presented. Simulation results show that the use of BSNN predictive models, not requiring additional instrumentation, are able to supersede the MRI gold standard, 0.5 $^oC/cm^3$ temperature resolution.

As the structure of the model is a combination of BSNN models, which are piecewise polynomials, it means that these models can be inverted and used in a feedback control loop, therefore providing closed loop control of therapeutic ultrasound instruments, in contrast with the open loop use found nowadays. Work is being done to achieve this goal.

## REFERENCES


Arunachalam K., Stauffer P. R., Maccarini P.F., Jacobsen S., Sterzer F. *Characterization of a digital microwave radiometry system for noninvasive thermometry using temperature controlled homogeneous test load.* Physics in Medicine and Biology, 53(14):3883-3901, 2009.

Straube W. L., Arthur R. M., , Trobaugh J. W., Moros E. G. *Non-invasive estimation of hyperthermia temperatures with ultrasound.* Int. J. Hyperthermia, 21(6):589–600, 2005.

de Boor C. *The numerical evaluation of b-splines* J. Inst. Math. Appl, pages 134-1391, 1972.

Edd J.F., Horowitz L., Rubinsky B. *Temperature dependence of tissue impedivity in electrical impedance tomography of cryosurgery.* IEEE Trans. Biomed. Eng., 52(4):695-701, 2005.

Fukukita H., Ueno S., Hashimoto M., Yano T. *Attenuation coefficient and propagation speed estimates of rat and pig intercostal tissue as a function of temperature.* IEEE Trans Ultrason. Ferroelectr. Freq. Control, 49(10):1411-1420, 2002.

Kavli T. *Asmod: an algorithm for adaptive spline modeling of observation data.* International Journal of Control, 58(4):947-968, 1993.

Krogh A., Vedelsby J. *Neural network ensembles, cross validation, and active learning.* Advances in Neural Information Processing Systems, 7(38):231-238, 1995.

Lei S., Zhu-Huang Z., Wu Shui-Cai, Zeng Yi *Study frequency shift evaluation of ultrasound in heating fields of hyperthermia by AR model.* The 3rd International Conference on Biomedical Engineering and Technology, 7:138-144, 2013.

Wyatt C., Soher B., Maccarini P., Charles H. C., Stauffer P., MacFall J. *Hyperthermia MRI Temperature Measurement: Evaluation of Measurement Stabilization Strategies for Extremity and Breast Tumors.* Int. J. Hyperthermia, 25(6):422-433, 2009.

de Senneville D.B., Quesson B., Moonen C.T. *Magnetic resonance temperature imaging (review).* Int. J. Hyperthermia, 21(6):515-531, 2005.

Teixeira C. A., Ruano M. G., Ruano A. E., Pereira W. C. A. *A soft-computing methodology for non-invasive time-spatial temperature estimation.* IEEE Transactions on Biomedical Engineering, 55(2):572-580, 2008.

Teixeira C. A., Pereira W. C. A., Ruano A. E., Ruano M. G. *On the possibility of noninvasive multilayer temperature estimation using soft-computing methods.* Ultrasonics, pages 32-43, 2010.

Vieira C. R. S., Sato S. Y., Pereira W. C. A. *Phantom para medição da faixa dinâmica de equipamentos de ultra-som biomédicos.* Revista Brasileira de Engenharia Biomedica, 19(3):157-166, 2010.

Eiben A.E., Smith J.E. *Introduction to Evolutionary Computing.* Prentice Hall: Natural Computing Series Edition, pages 71-85, 2003.