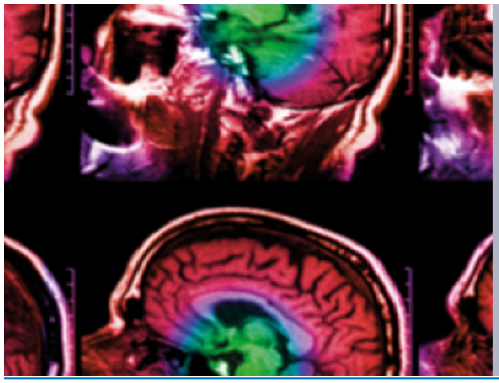
Biomedical Physics & Engineering Express

ENDORSED BY IPEM Institute of Physics and Engineering in Medicine

PAPER

A comparison between plausible models in layered turbid media with geometrical variations applying a Bayesian selection criterion



Biomedical Physics & Engineering Express

PAPER

A comparison between plausible models in layered turbid media with geometrical variations applying a Bayesian selection criterion

Demián A Vera[1], Guido R Baez[1,2], Héctor A García[1], Daniela I Iriarte[1] and Juan A Pomarico[1]

[1] Centro de Investigaciones en Física e Ingeniería del Centro de la Provincia de Buenos Aires (CIFICEN, UNCPBA—CICPBA - CONICET) Pinto 399, B7000GHG—Tandil, Buenos Aires, Argentina
[2] Plasmas Densos Magnetizados (PLADEMA). Pinto 399, B7000GHG—Tandil, Buenos Aires, Argentina

E-mail: dvera@alumnos.exa.unicen.edu.ar

RECEIVED
30 March 2020

REVISED
18 July 2020

ACCEPTED FOR PUBLICATION
11 August 2020

PUBLISHED
24 August 2020

Keywords: Turbid Media, Model Selection, Near Infrared Spectroscopy

Abstract

One possible application of Near Infrared techniques is to analyze human brain metabolic activity. Currently used models take into account the layered structure of the human head but, usually, they do not consider the non-planar surface of some of the boundaries, i.e. gray matter, which results in a much more complex structure, thus leading to more sophisticated models and longer calculation times. The main objective of this work is to determine if it is worth to replace a planar layered structure by a non-planar one. To this end we implement a Bayesian-based quantitative methodology for choosing between two competitive models describing light propagation in layered turbid media. Experiments of time-resolved diffuse reflectance measurements are performed in layered phantoms and complemented with numerical calculations. The resulting Distributions of Time of Flight of both models are compared using Bayesian model selection analysis. The non-planar interface was introduced in the simulations by a simple surface parametrization. Results suggest that, under certain conditions, a multilayer model with planar boundaries is good enough.

1. Introduction

In the last two decades, Near Infrared Spectroscopy (NIRS) techniques have been used to study light propagation in biological tissue, which can be considered as turbid media for NIR light [1–5]. In their way through turbid media, photons interact with different constituents of the matter, suffering absorption and scattering processes. In general, the propagation of light through turbid media is described by the Radiative Transfer equation (RTE), which under certain conditions can be replaced by its simplified approach, known as the Diffusion Approximation (DA) [2–4]. One possible application is to register the functionality of organs, and in the particular case of brain, NIRS techniques are a helpful tool to bedside diagnostic (or to follow the evolution) of stroke patients [6–8]. As early demonstrated by Jöbsis [1], variations in the way that NIR light behaves on its travel through biological tissue allows to determine, in a non-invasive way, oxygen level or general metabolic changes by monitoring possible increments or decrements of the absorption of radiation inside the studied medium. Because of this, it is of interest to have accurate models concerning light propagation in the human head, for which multiple layers need to be considered [9–12]. Many models, which describe light propagation in multi-layered turbid media, were proposed [9, 13–15]. In particular, Liemert and Kienle [10] developed a theoretical method to describe light propagation through multi-layered turbid cylinders of finite thickness. Most recently, García *et al* [11, 12] improved this model by assuming that the deepest layer is infinitely thick. In that way, diffusely reflected information of the turbid media can be obtained, and no diffusely transmitted radiation is reachable. Another work deals with rectangular rather than cylindrical geometries, also considering the deepest layer as infinitely thick [16]. At present, most of the literature related to light propagation in multi-layered turbid media treats the head model described above as if the boundaries between layers were plain [9–11, 13, 16, 17], i.e., ignoring the non-planar structure of the gray matter, which presents numerous convolutions, sulci and gyri. Concerning numerical exploration, the influence of a structured layer has been

studied by Okada *et al* [18]. They have investigated a sophisticated brain model with non-planar boundaries, that also considers the existence of a thin Cerebrospinal Fluid (CSF) layer that fills the gyri structure cavities. However, the criteria used to determine that variations introduced by the gyri and sulci structure is mainly empirical, and is subject to the more notorious influence of the CSF layer. Other authors, [19, 20] have studied numerical models containing very detailed structures. This kind of approach Requires some overhead due to segmentation and mesh generation processes and it is not always clear if that effort brings additional benefits with respect to some simpler approaches [11, 21]. In this work we study light propagation in a three-layered model considering a non-planar boundary between the two deepest layers, modeled by a simplified sinusoidal structure, and compare it with a model containing only planar boundaries. For either case, we compute time domain diffuse reflectances using numerical methods performed on MATLAB®. For comparison, experiments on time domain diffuse reflectance were considered as well. Moreover, dependency of the time-resolved diffuse reflectance is analyzed qualitatively for different orientations of the optodes relative to the structure of the non-planar boundary, by comparing the Distributions of Time of Flight (DTOF) of photons resulting for each case. To quantitatively asses the differences between reflectance results when changing optodes orientation a relative entropy criterion is applied. Finally, and as the main objective of this work, a *Bayesian Model Selection* (BMS) analysis is performed, in order to quantify the advantage of considering a more complex human head model or if the far more simpler planar boundary approach is enough for NIRS purposes. Although the main objective of this work is the comparison of models with different geometrical structures, and given the results of this first situation, we also perform an additional analysis in which the BMS approach is used for comparing models in the planar boundaries situation, but now considering the variation of the absorption coefficient of the two first layers and the thickness of the first one, which is also of interest for NIRS.

BMS is an extension of Bayesian inference methods [22] in which one does not only describe a parameter uncertainty through a prior distribution, but also the model uncertainty, obtaining posterior distributions for model parameter and for model themselves using Bayes' theorem, thus allowing for direct model selection. This approach has been applied in many areas of science such as Life Sciences and Medicine where it was used for example, for variable selection of factors associated with false positives in diagnostic kits [23], weight loss [24] and structure in Bayesian networks for patient-specific models [25]. In genetics it was used in the combination and selection of genetic information from metabolic pathways [26]. Other applications can be found in Humanities and Economy as well as in Physical Sciences and Engineering [27–30] and many others areas [31]. To the best of the authors knowledge, it has not been applied before to Light propagation through turbid media [32].

The paper is organized as follows: in section 2 we present the rationale as well as the description of the models, including the criterion used and the Bayesian Model Selection. Next, in section 3, the details concerning the experimental issues are described. The results and discussion of this work are given in section 4 and, finally, section 5 summarizes the main conclusions.

2. Rationale of the proposed approach

In general, when modeling a given physical situation it is necessary to define certain parameters that allow the model to better describe the desired phenomenon. However, the increasing number of parameters would need to be estimated *a priori*, or assumed, if there is lack or no information about them. Moreover, computational complexity can increase depending on the structure of the chosen model. A question arises naturally: considering a model, M_1, which uses less parameters, but describes the phenomenon in a poorer manner, and another model, M_2, that uses more parameters and describes the situation more precisely but having increased costs, which model should be used? Intuitively, if M_1 resembles enough M_2 (in the sense that similar results are obtained for similar situations), then there is no need to use M_2 considering its increased complexity, because the gain of using a more detailed model would be negligible. Conversely, if M_2 describes situations that are not reachable by M_1, then it is justified to use this more detailed model. We will consider two situations, namely: a semi-infinite three-layered medium with planar boundaries between its layers, and a semi-infinite three-layered medium with one non-planar boundary described by a sinusoid defined parametrically. They will be regarded as M_1 and M_2, respectively and both are to be considered as possible simplified representations of the human head. As it is necessary to define certain characteristics of the surface, the complexity increases for M_2 with respect to M_1.

2.1. Model description

As stated above, our simplified model M_1 describes light propagation through layered turbid media with planar boundaries. Let us consider an N layers medium, with the last layer extended to infinity. Light propagation through turbid media can be modelled using the Diffusion equation [33]. Considering the domain Ω with boundary $\partial\,\Omega$, the photon fluence is described by

$$\frac{\partial\Phi(x)}{\partial t} - \nabla(D(x)\cdot\nabla\Phi(x)) + \mu_a(x)\Phi(x) = 0,$$
$$\text{for } x \in \Omega$$
$$\Phi(x) + 2\tilde{A}D(x)\frac{\partial\Phi(x)}{\partial\hat{n}} = S(x, x_0),$$
$$\text{for } x \in \partial\Omega, \tag{1}$$

where $\mu_a(x)$ and $D(x) = [3(\mu_a(x) + \mu_s'(x))]^{-1}$ represent the absorption and diffusion coefficients, respectively, being $\mu_s'(x)$ the reduced scattering coefficient. S stands for the source term and $\tilde{A}$ is a factor that depends on the refractive index mismatch between the turbid medium and its surroundings [5]. In this work it will be assumed that the refraction index remains constant through the entire domain. This implies that there will be no internal reflections between layers. However, the surrounding non-scattering media will be considered as air, with $n = 1$. Let N represent the number of layers, $\mu_{a,i}$ and $\mu_{s,i}'$ the absorption and the reduced scattering coefficient of the i-th layer, respectively. Let L_j be the depth of the j-th boundary, measured from the interface between layer j and $j+1$ with j: 1, $\cdots$, $N-1$ and let $\mathcal{L} = (L_1, L_2, \cdots, L_{N-1})$ a vector with the layer depths. The light source is assumed to be located at r_s and a detector is placed at r_d. We can notate M_1 as a model in the following way:

$$\begin{aligned} M_1 &= M_1(\mu_a(\mathcal{L}),\ \mu_s'(\mathcal{L}),\ n,\ r_s,\ r_d) \\ &= M_1(\mathcal{L}), \end{aligned} \tag{2}$$

where the second equality means that the only variable used will be $\mathcal{L}$, while the rest will remain fixed. For the second model, M_2, light propagation will be described replacing one planar layer with a sinusoidal one, to emulate, in a simplified fashion, the sulci and gyri of the brain. Using the same notation as above, the second model will be described as

$$\begin{aligned} M_2 &= M_2(\mu_a(\mathcal{L}, A, \lambda),\ \mu_s'(\mathcal{L}, A, \lambda),\ n,\ r_s,\ r_d) \\ &= M_2(\mathcal{L}, A, \lambda), \end{aligned} \tag{3}$$

being A and λ the amplitude and wavelength of the structure of the sinusoidal interface, respectively. We will assume that the sine wave separates the last two layers and L_{N-1} is the depth at which the deepest layer begins. Moreover, in this work, we will restrict our study to the three layers case ($N = 3$), since the only structural alteration will be taken at the interface between the deepest layer and the one before it, thus being not relevant how many layers exist between the surface and this interface. In this configuration, the depth of layer 1 will remain fixed and the depth of the second layer will be variable. With this in mind, equations (2) and (3) can be expressed as:

$$M_1 = M_1(\mathcal{L}) = M_1(L_2) = M_1(L) \tag{4}$$
$$M_2 = M_2(L, A, \lambda), \tag{5}$$

where we remark that $L_2 = L$ in order to simplify the notation. A schematic representation of both models, is given in figure 1.

These two models can be seen as competitive when describing light propagation in the human head, where the complex structure formed by the circumvolutions, sulci and gyri is replaced by a sinusoidal interface, and the deepest layer represents gray matter tissue. Both models can be implemented computationally either using Monte Carlo simulations [34, 35] or numerical models of the Diffusion Approximation [36], depending on the optical properties used. Finally, they are relatively simple to reproduce experimentally.

2.2. Bayesian model selection

In this subsection we describe the main aspects of the Bayesian Model Selection (BMS). However, before proceeding, it is necessary to consider the effect of the optode orientation relative to the structure of the non-planar boundary, since in practical cases this orientation is unknown *a priori*. To this end, the Kullback-Leibler divergence (KL) will be used [37, 38]. Usually, this test is employed to quantify differences between two probability density functions (pdf), $p(x)$ and $r(x)$, defined as follows:

$$G \equiv \int p(x)\log_2\left[\frac{p(x)}{r(x)}\right]dx \equiv G[p, r], \tag{6}$$

where $G[p, r]$ is a semi-positive quantity [39, 40]. In our context, $p(x)$ and $r(x)$ represent DTOF measured from different orientations with $p(x)$ fixed. If the binary logarithm is employed, G can be interpreted as the amount of new information required to describe the new model $r(x)$ or configuration in terms of *bits*. Thus the KL divergence is a measure of how much information is gained or lost, for instance, when the optodes direction relative to the non-planar interface wavefront is changed. A low value of the KL divergence indicates that the replacement is worthless. Conversely, a high value of KL suggests that information may be lost if a given source-detector arrangement is used instead of another one. Note that if $p(x) = r(x) \Rightarrow G[p, r] = 0$, as expected. We will show that optode orientation produces no significant (or irrelevant or non measurable) variations in the DTOF and we can thus proceed to model comparison.

For the model selection criteria we will choose a model using BMS. Precisely, it uses the Bayes factor, which is defined as the ratio between the likelihood of M_1 given data Y and the likelihood of M_2 given data Y. In symbols:

$$K = \frac{p(M_1|Y)}{p(M_2|Y)} \tag{7}$$

Consider a set of R models $M = \{M_1, ..., M_R\}$ (in this work, $R = 2$) are under consideration for data Y (DTOF), and that under M_k, Y has density $p(Y|\theta_k, M_k)$, where θ_k is a vector of unknown parameters that indexes the members of M_k. The Bayesian approach proceeds by assigning a prior probability $p(M_k)$ to each model and a prior probability distribution $p(\theta_k|M_k)$ to the parameters of each model.

The model selection problem consists in finding which model in M can replace the data Y. The

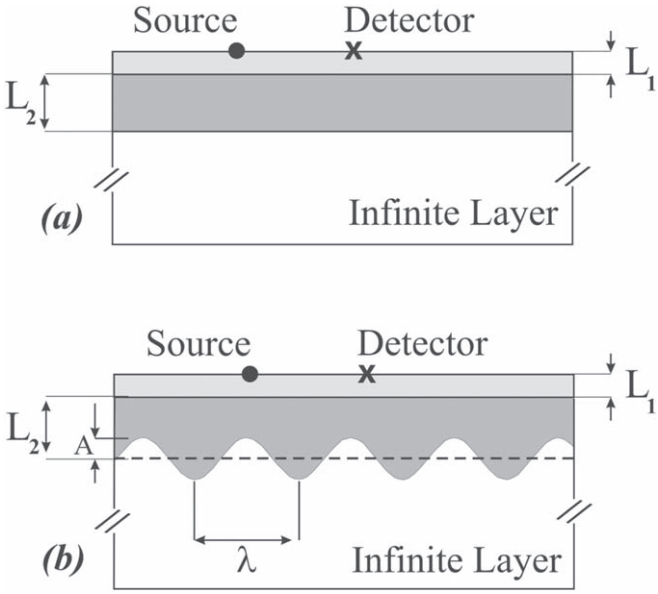


Figure 1. Schematic representation of the two types of three layers models: (a) the planar one, $M_1(L)$ and (b) the sinusoidal model, $M_2(L, A, \lambda)$.

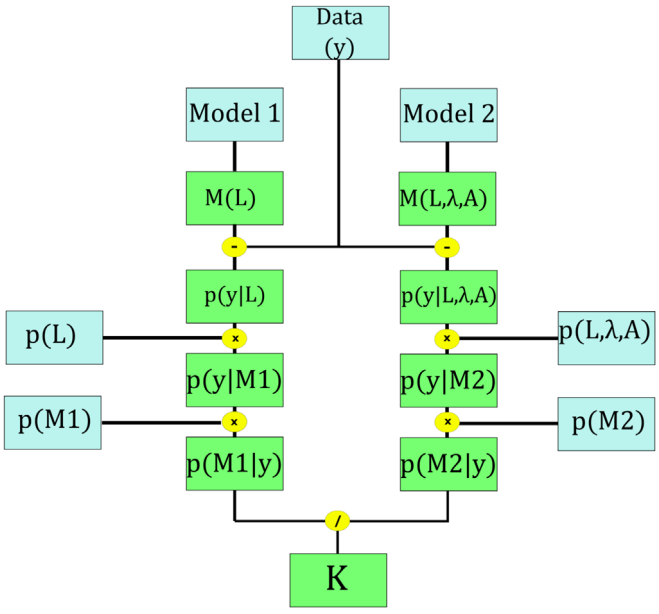


Figure 2. Diagram to illustrate the Bayesian Model Selection procedure. The light blue rectangles indicate the data that is known and provided, meanwhile the green rectangles show the inner computations performed with the given data.

probability that M_k was in fact this model, conditional on having observed Y, is the posterior model probability:

$$p(M_k|Y) = \frac{p(Y|M_k)p(M_k)}{\sum p(Y|M_k)p(M_k)}, \tag{8}$$

where

$$p(Y|M_k) = \int p(Y|\theta_k, M_k)p(\theta_k|M_k)d\theta_k \tag{9}$$

is the integrated likelihood of M_k. Based on this posterior probabilities, pairwise comparison of, say, two models M_1 (planar boundaries media) and M_2 (non-planar boundaries media) is summarized by the posterior odds:

$$K = \frac{p(M_1|Y)}{p(M_2|Y)} = \frac{p(Y|M_1)}{p(Y|M_2)} \times \frac{p(M_1)}{p(M_2)}. \tag{10}$$

This expression reveals how the data, through the Bayes factor K, equation (10), updates the prior odds $p(M1)/p(M2)$ to yield the posterior odds. Figure 2 shows a flowchart of the Bayesian Model Selection implementation.

The aim of the BMS is to discover a useful simple model from a large speculative class of models that allows to provide valuable scientific insights or less

Table 1. Optical properties of the three layers used in both, the FEM calculations and the experimental phantom.

		Optical Properties	
Layer	Thickness [cm]	μ_a[cm^{-1}]	μ_s'[cm^{-1}]
1	0.9	0.13 ± 0.01	10.90 ± 0.1
2	1.1	0.054 ± 0.01	9.11 ± 0.1
3	5	0.24 ± 0.01	17.80 ± 0.1

costly methods for prediction [41]. According to the value obtained for the Bayes factor it is possible to determine if it is convenient to replace M_1 by M_2 or not. The Bayes factor was computed for fixed optical parameters, $\mu_{a,i}$ and $\mu'_{s,i}$ (see table 1, in section 4) as well as keeping fixed the source and detector positions, r_s, and r_d, respectively. In this analysis only geometrical characteristics were taken as variables, i.e the parameter L in model $M_1(L)$, and L, λ and A in model $M_2(L, A, \lambda)$. When analysing experimental data, a 3D geometry must be considered. For that purpose, the FEM method used to describe light propagation was implemented. In general, the procedure concerning the BMS goes through the following steps:

- In each case, both models, M_1 and M_2, were weighted, first, by a uniform probability density function (pdf) and, second, by a Gaussian pdf.
- The limits of the Uniform pdf were set as the extremes of the range in which geometrical properties vary, i.e. L in M_1 and L, A and λ in M_2.
- The mean value and standard deviation of each Gaussian pdf were varied for different possible situations, i.e. considering either if one has good knowledge or not about the geometry of the medium of interest.

The first point deserves some explanation. The use of a Uniform or Gaussian pdf is based on a Bayesian approach of the user's prior knowledge with respect to the underlying situation. When the Uniform distribution is considered, there is no value with more plausibility than the others within the feasibility region. In this case, only the data likelihood interacts in the process of determining which model is more likely to be producing the data. For the Gaussian case the situation is different: there is a specific weight for each value and the highest weight is reserved for the mean value. The rest of the values will have higher or lower weight with respect to the mean depending on the value of the standard deviation and the distance to the mean. This situation is suitable for the case when the user has some knowledge about the values, for example, when some (or all) parameters can be inferred from literature or obtained from some other imaging modality. Strictly speaking, the Gaussian pdf used here is not a full Gaussian pdf in the sense that the entire domain is used, but a truncated version of it, i.e.

$$\pi_{\text{Gauss}}(x) = \frac{\pi_{\text{true Gauss}}(x) \times \pi_{\text{Uniform}}(x)}{\int_X \pi_{\text{true Gauss}}(x) \times \pi_{\text{Uniform}}(x)\,dx}. \quad (11)$$

It must be pointed out that it is possible to perform model selection using a Kullback-Leibler basis through the Akaike Information Criterion (AIC) [42, 43]. However, in this work we will use the more general approach which allows for arbitrary likelihood and prior distributions which models the measurement process and our prior knowledge appropriately. We are interested in studying the effects on equation (9) due to $p(\theta_k|M_k)$. We use this formalism because it allows us to explore certain characteristics which might be hidden in Akaike's method. Specifically, the resulting distributions may have multiple local minima due to the ill-posedness of the original problem. Akaike's method relies on knowledge of the maximum of the likelihood, but it can be hard to find due to the existence of many local minima. To avoid this problem, we also incorporate prior information on the parameter prior distributions. The formalism described above together with the corresponding results will be considered in section 4.

3. Experiments on phantoms

Time-resolved reflectance measurements were performed on a three-layered agarose phantom which represents a simplification of the human head. As it is shown in the following, the phantom was constructed in such a way as to include both models described above in section 2.1.

3.1. Construction and characterization of the phantoms

Phantoms were constructed inside a rectangular mould, using the following ingredients: distilled water (as carrier), milk (as scattering agent), ink (as absorption agent) and 2% in volume of agarose as solidifying agent. In a first step, agarose is added to the water and this mixture is heated up to about 98 °C while steering to achieve complete dissolution. Then, the milk containing the desired volume of ink, and pre-heated to 60 °C is added to the solution of water and agarose (preheating the milk/ink solution avoids the formation of solid lumps when getting in contact with the hot water/agarose mixture). Finally, after this mixture containing all four ingredients cools down to approximately 60 °C, it is ready to be poured into the mould. The desired structure of the (deep) interface, shown in figure 3, was previously made of plastic using a 3D printer, with lateral dimensions fitting exactly those of the mould. It contains both, a planar sector and a structured zone presenting a sinusoidal pattern. The zero of this pattern coincides with the height of the plane surface.

The structure in figure 3 was placed at the bottom of the mould, with the sinusoidal pattern facing up, and the solution with agarose was poured into it.

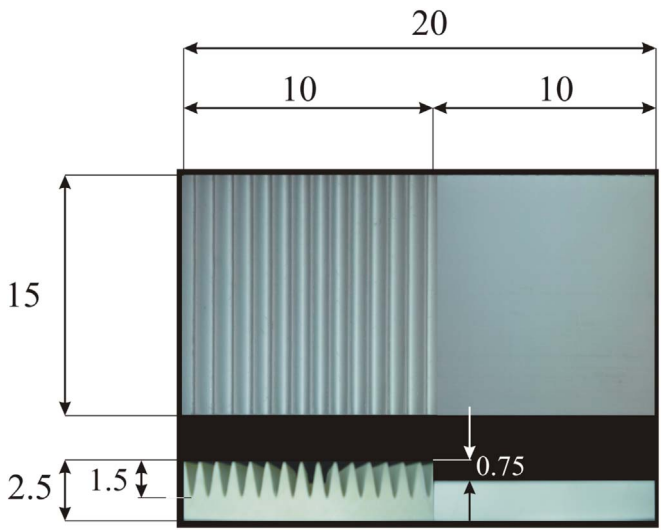


Figure 3. Photographs of the top and lateral views of the 3D-printed structure used to simultaneously elaborate the planar and non-planar interfaces of the agarose phantom. The flat surface is at a height equal to the zero line of the sinusoidal pattern. All dimensions in cm. Not to scale.

Figure 4. Photograph of a transverse cut of the deepest two layers of the phantom within before adding the third layer on top, illustrating the result of the elaboration procedure described in the text. In this picture, the absorption coefficient of the upper layer is lower than the other one.

Following this procedure, half of the phantom contains planar interfaces and the other half contains the corrugated interface; this ensures that both, the planar and non-planar layers result with exactly the same optical properties, thus disregarding any differences but those resulting from geometry.

The volume of agarose solution was calculated such as to obtain a thickness of about 5 cm when poured into the mould, emulating a semi-infinite medium. After this deep layer has completely solidified, it was carefully removed from the mould and placed once again inside of it, with the sinusoidal structure facing up. Finally, two layers were added on top of it. Each of these top layers was also made of a similar agarose solution as the one described above, but containing different amounts of ink, to achieve the desired absorption coefficient of the corresponding layers.

A photograph of the resulting structured interface in the agarose phantom and the layer on it is shown in figure 4, for the particular case in which the absorption coefficient of one layer is higher than that of the other. Note that both layers fit perfectly into each other, leaving no bubbles or void regions which could modify the temporal profile of the measured DTOF.

To characterize the phantoms, the optical parameters, μ_a and μ_s' of each layer were measured. To this end, homogeneous agarose slab phantoms were simultaneously prepared using the same mixture as that of the corresponding layer. Their lateral dimensions were 10 cm $\times$ 10 cm and their thickness was 3 cm. Using these slabs, time-resolved diffuse transmittance DTOF were measured independently on each homogeneous phantom, and used in a Levenberg-Marquardt fitting routine written in Python to recover the values of the optical properties [44]. For the fitting routine, the theoretical model from Contini *et al* [5] was used. The actual Instrument Response Function (IRF) of the measuring equipment was convolved with the theoretical DTOF produced by several sets of optical parameters (which were varied within proper intervals) and compared with the measured DTOF until the desired convergence was achieved. The values of these measured optical properties for the three layers are given in table 1.

3.2. Experimental setup

The experimental setup is sketched in figure 5, showing a schema of the transverse cut of a three layers phantom constructed using the procedure described in section 3.1. Note that, depending on the location of

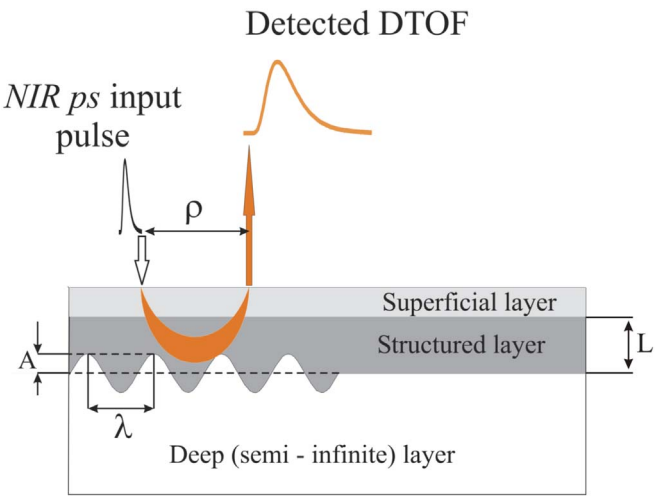


Figure 5. Experimental setup for measuring the diffusely reflected DTOF in a three layers phantom containing both, a structured layer and a planar one. A picosecond (*ps*) laser pulse is injected via an optical fiber on the top surface of the phantom and a fiber bundle collects the diffusely reflected photons at a distance ρ. The schema also shows the most probable photons path, also called *banana*, and the geometrical parameters of the phantom, namely L, the mean depth of the structured region and A and λ, the amplitude and the wavelength of the sinusoidal structure, respectively.

the input and output fibers, either the structured or the planar region is considered. For carrying on the experiments it was used a NIR pulsed laser source (Becker&Hickl, model BHLP 700) operating at $\lambda = 785$ nm and emitting 70 ps pulses at a repetition rate of 50 MHz. Its average power was set to 5 mW. The laser pulses, guided by a 600 μm core, $NA = 0.22$ optical fiber, impinged onto the upper layer of the phantom, labelled as the 'Superficial Layer' in figure 5. The collection of the diffusely reflected photons, after they have travelled inside the medium, was achieved by a fiber bundle ($\phi = 1/8''$, $NA = 0.33$) also in contact with this superficial layer but at a distance ρ from the input pulse. The photons collected by this bundle were detected by a Time Correlated Single Photon Counting system (TCSPC) consisting on a photomultiplier tube (Becker&Hickl, model PMC-100-20) and a TCSPC card (Becker&Hickl SPC-130), which builds up the desired DTOF. In figure 6 a typical DTOF obtained at a region of the phantom containing the structured layer and for $\rho = 4.5$ cm, together with the corresponding IRF of the measurement equipment is shown.

4. Results and discussion

In this section we will describe the results of implementing the BMS methodology over 2D and 3D data. For the 2D data we analyzed the following situations:

- A simulated DTOF obtained by convolution of a numerical DTOF with the actual IRF of our equipment.

And, in the 3D case:

- Using a DTOF obtained from a medium with a sinusoidal interface, we first applied the BMS approach to decide if it is worth to consider this kind of structure against the planar one.
- As a second illustration of the approach we used a DTOF resulting from a medium with three planar layers, in order to decide if variations in the absorption coefficient of the superficial layer and or in the scalp and skull thickness need to be taken into account.

In all cases the optical properties used are those given in table 1. Note that in this work it is not intended to emulate the actual properties of the brain, which are still matter of discussion [6, 13, 19, 20, 45, 46]. Thus, optical and geometrical properties used in 2D and 3D simulations were those obtained experimentally, using as a guide those given by Wu *et al* [19] but considering a higher value for the reduced scattering coefficient of the second layer, such that all calculations could be made within the DA. Geometrical parameters were taken as a guide from associated neuroscience literature [32, 47].

4.1. Considerations about the optodes orientation

To explore the influence of the optode orientation, the KL approach, described in section 2, was implemented. Calculations were carried out for a source located at $r_s = (0, 0, 1/\mu_s')$ [48], and a set of 45 detectors radially distributed, placed at a distance ρ from the source and on the top surface of the medium. In this way, while light propagates from the source to the detector through the medium, it encounters the sinusoidal boundary differently, depending on the orientation. This procedure was repeated by changing the optodes orientation in a range from 0 to $\pi/2$ radians, and for $\rho = 2.5$cm, $\rho = 3.5$ cm and $\rho = 4.5$ cm. For the numerical simulation three layered medium already described and with the optical

Table 2. KL entropy values ($G[p, r]$), for each source-detector distance, computed from experimental data, varying optodes orientation, for a **three layered** agarose medium. When $\theta = \pi/4$, the optodes are placed obliques to the interface wavefront, and for $\theta = \pi/2$, the optodes are placed parallel to it.

	Kullback-Leibler divergence		
	ρ		
Angle θ	2.5 cm	3.5 cm	4.5 cm
$\theta = \pi/4$	0.006 88	0.001 23	0.000 71
$\theta = \pi/2$	0.000 31	0.000 82	0.000 54

properties of table 1 was used. Results can be seen in figure 7 for $\rho = 2.5$ cm. For the other two cases considered, namely $\rho = 3.5$ cm and $\rho = 4.5$ cm, results show KL divergence values much lower than for $\rho = 2.5$ cm, and they are thus not shown in figure 7. Additionally, to verify these results, we measured the diffuse reflectance DTOF by using the constructed phantom, considering all source—detector distances and only three angles, at $\theta = 0, \pi/4$ and $\pi/2$ radians as shown in figure 8.

Results for these experiments can be seen in table 2. For the evaluation, the DTOF corresponding to the orientation perpendicular to the interface wavefront was left fixed ($p(x)$ in equation (6)) while the DTOF of the remaining orientations were left to vary ($r(x)$ in equation (6)).

The information gained by using the actual orientation instead of a fixed one is, at most, of order 10^{-3} of bit, suggesting that there is a negligible gain in considering the orientation. These results are consistent with the fact that diffusion makes the most probable photon paths *(banana)* to be spread in comparison to the structure of the interface.

In summary, the effect of the source-detector orientation relative to the interface wavefront is *almost* negligible according to our results. For this reason, for the rest of this work, the orientation considered will be the one that is perpendicular to the interface wavefront ($\theta = 0$, figure 8) and we shall make no further considerations on this subject.

4.2. 2D Simulations

Although it is an obvious statement that light propagation in tissue is a 3D phenomenon (as it is every physical process), given the complexity and vastness of the parameters space, and the huge computational cost of the 3D calculations, we have first implemented a set of 2D simulations. This allowed us to better visualize the geometrical aspects of the mesh selection and to test the whole algorithm. Of course results of this subsection are of relative value, but they are useful to gain some insight about the entire process. The reader can skip this section and proceed directly to section 4.3 if not interested. The simulated bidimensional experiment was performed in a space Ω which consisted of a rectangle of height $h = 4$ cm and width $w = 16$ cm. An isotropic source was placed at $r_s = [0, 1/\mu'_{s,1}]$ cm and the detector was placed at $r_d = [4.5, 0]$ cm. The unstructured mesh consisted of 19 831 nodes and 39 397 triangular elements. The optical properties were defined for each layer. Let the index $i = 1, 2, 3$ refer to layer i, $\mu_{a,i}$ and $\mu'_{s,i}$ stands for the absorption coefficient and reduced scattering coefficient for layer i, respectively. In figure 9 the used mesh is presented, whose statistics are the following:

The mesh statistics were the following: the minimum element quality was 0.1998, average element quality 0.8472, the mesh area was 64 cm^2. The maximum and minimum element size were 0.18 cm and 0.008 cm.

4.2.1. Bayes model selection in 2D medium

The complete simulation of 2D data is composed by a set of 16 000 DTOF for medium M_2, one for each different geometrical configuration considered. Every DTOF was simulated using a configuration $\{L_i, A_j, \lambda_k\}$, for $0.6 < L_i \leqslant 1.6$ cm in 40 steps; $0 < A_j \leqslant 1.2$ cm in 20 steps, and $0.3 \leqslant \lambda_k \leqslant 2$ cm in 20 steps that generated it after passing through the numerical model. For model M_1, 40 DTOF were considered for parameter L. The parameter space was $\{L_i\}_{i=1}^{40}$, where $0.6 < L \leqslant 1.6$ cm in 40 steps. The first case of study was to compute the Bayes factor for simulation data, K_{sim}, for 200 DTOF of the model M_2 that were selected randomly (taken as Y in equation (10)). In the second case of study, a randomly chosen noisy DTOF was obtained by performing a convolution between the simulation data and the experimental IRF. This procedure is repeated with the entire samples for both models (by convolution with another experimental IRF). In this way, the data is comparable and is possible to proceed onto K_{IRF} calculations.

There are three elements which must be defined in order to estimate K, say, $p(y|\theta_k, M_k)$ for the data likelihood, $p(\theta_k|M_k)$ for the prior parameter distribution, and $p(M_k)$ for the model prior distribution. For the likelihood, a normal distribution is used

$$p(y|\theta_k, M_k) \propto \exp\left(-\frac{1}{2}\,\|y - M_k(\theta_k)\|_R^2\right). \quad (12)$$

which is a reasonable choice given that the measurement noise can be approximated with a Gaussian distribution (using the Central Limit theorem after averaging the measurements). Note that the likelihood is written in terms of the norm of the discrepancy between the measurement and the model ($\|y - M_k(\theta_k)\|_R = \|x\|_R$). It states how likely is to replace the data with the model once the parameters are set. For the prior parameter distribution, which models the knowledge about the parameters of the respective model, two different cases were studied, namely, a Uniform prior distribution and a Gaussian prior distribution. The former is a case of a non-

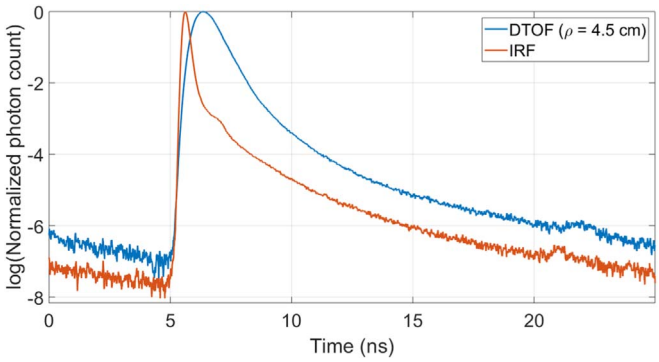


Figure 6. The actual IRF of the measurement equipment and a DTOF measured for $\rho = 4.5$ cm.

Table 3. K values for simulation data for three-layered medium; computed for a Uniform and Gaussian prior distribution, for different standard deviations.

Bayes factor K			
Uniform	Gaussian		
$\sigma \to \infty$	$\sigma = 0.8$ cm	$\sigma = 0.5$ cm	$\sigma = 0.2$ cm
[1.0025; 2.6716]	[1.0009; 2.9139]	[1.0141 ; 3.3027]	[1.0022; 10.0956]

informative distribution in the sense that there is no preference among the support of the distribution; the latter is a case of an informative one because values in the support, i.e. the region where the distribution is not zero, have different weight. In equation (12), $\|x\|_R^2 = x^T R^{-1} x$ where $R^{-1} = diag\,(1/\sqrt{y})$ because we are considering a Poisson process with uncorrelated noise [49]. Both cases represent different states of information that may be considered according to the situation. For example, in the presence of another imaging modality such as a MRI, a Gaussian prior is suitable given that it is possible to estimate the geometrical parameters from the image. Finally, for the model prior distribution, a Uniform distribution was chosen, i.e. $p(M_1) = p(M_2) = 1/2$.

The Bayes factor for the three-layered medium described by model M_2 was computed fixing $L_1 = 0.9$ cm. The geometrical variables were $L = L_2$ for model M_1 and L_2, λ and A, for model M_2. Results are shown in table 3, and the corresponding histograms in figure 10 for 200 random DTOF, where the orange lines corresponds to the kernel density estimation [50]. When we let the standard deviation increase, the possible values of K are less spread than when the standard deviation is small. The prior knowledge increases the certainty thus increasing the pragmability of M_2 as our knowledge is more precise. In table 3, results are presented in intervals which contains 200 random DTOF computed using a Uniform or a Gaussian prior pdf. For the latter, the standard deviations considered were $\sigma = 0.2$, 0.5, 0.8 cm. K values DTOF tend to increase for greater deviation values when the Gaussian prior is chosen. For $\sigma = 0.5$ cm and $\sigma = 0.8$ cm, K indicates that is barely worth mentioning the plausibility of a more complex model M_2. As expected, the Uniform prior acts as the limiting case for increasing deviations on the Gaussian prior. On the other hand, for $\sigma = 0.2$ cm the Bayes factor grows up indicating that having a good initial guess of the geometrical properties of the medium makes the likelihood of M_2 much higher, suggesting that this model is the preferred one.

4.3. 3D medium: experiments and simulations

4.3.1. Study of the structured interface

A three-layered phantom with layers structure as the one shown in the scheme of figure 6 was elaborated. The optical and geometrical properties obtained are shown in table 1. The simulations on the 3D medium were performed in similar fashion as the 2D case. The mesh statistics are the following: the minimum and mean element quality were 0.04 and 0.645, respectively, with a number of elements of 84 253. The minimum and maximum volume size were 0.003 cm^3 and 16.37 cm^3, respectively. The total number of nodes was 15 844. Due to its high computational cost, only $21 \times 15 \times 20 = 6300$ computations were performed. In table 4 the values taken by each parameter as well as the number of steps for all them are shown.

Experimental data was compared against simulations taking two different prior pdf for the parameters: Uniform and Gaussian. First, taking a Uniform distribution as prior means that every geometrical parameter, L, A and λ, has the same probability of occurring ($p(L) = c_1$ for the planar case, and $p(L, A, \lambda) = c_2$, where c_1 and c_2 are positive constants). As there are no preferences for any value in the domain, this is a good representation of ignorance [51] due to the bounded domain, as long as there is no need of a change of parameters. This situation can be interpreted as the user not knowing *a priori* whether there is a sinusoidal or planar boundary layer, nor which values describe the geometry appropriately. This is the fairest comparison between the models before adding some prior knowledge, because the prior knowledge (either it is correct or incorrect) adds a bias towards it. Results are shown in table 5. The value of K is an

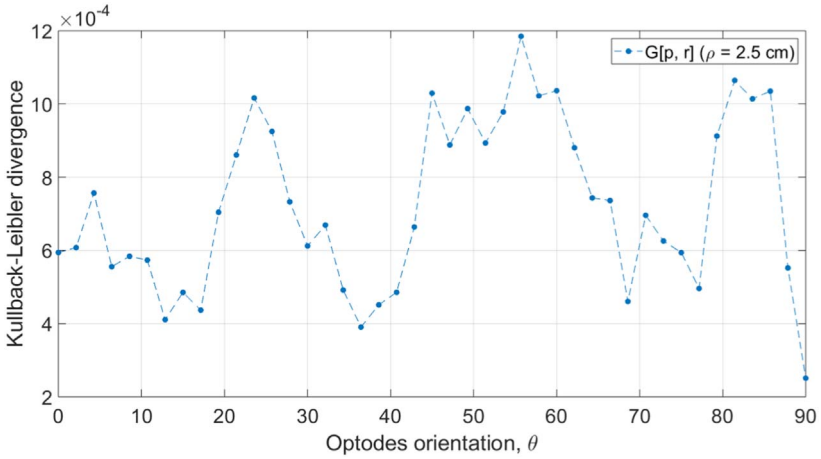


Figure 7. Kullback-Leibler divergence values in function of angle of orientation. The pdf taken as p corresponds to the DTOF (normalized by area) obtained considering the optodes orientation perpendicular to the interface wavefront, and r represents the 45 different orientations for increasing θ in the range from 0 to $\pi/2$ radians. Here, $\rho = 2.5$ cm.

Table 4. Set of parameters chosen for the 3D slab simulations including a sinusoidal interface.

Planar versus sinusoidal interface. **Bayes factor** K. Uniform prior			
Parameter	Values	Number of steps	Step size
L	$(0.6 \leqslant L \leqslant 1.6)$ cm	21	0.05 cm
A	$(0 \leqslant A \leqslant 1.2)$ cm	15	0.085 715 cm
λ	$(0.3 \leqslant \lambda \leqslant 2)$ cm	20	0.089 475 cm

Table 5. K value for Uniform prior. A value close to 1 means that is almost the same to consider Model 1 or Model 2, that is, is *barely worth mentioning* the difference between two plausible models.

Planar versus sinusoidal interface. **Bayes factor** K. Uniform prior			
Prior pdf	K	Selected model	Conclusion
Uniform	1.05	Model 2	Barely worth mentioning

indicator of which underlying model is the most plausible after have seen the data. To get the idea behind the indicator, let us recall the definition of K:

$$K = \frac{p(M_1|y)}{p(M_2|y)} = \frac{\int p(y|M_1,\, \theta_1)p(\theta_1|M_1)\,d\theta_1}{\int p(y|M_2,\, \theta_2)p(\theta_2|M_2)\,d\theta_2}. \quad (13)$$

By exploring the right hand side of equation (13) it is possible to note that the integrand evaluates mainly two components. Namely, the replace-ability of the data by the model M_i ($p(y|M_1,\, \theta_1)$) and the complexity of that model ($p(\theta_1|M_1)$). The term *complexity* is used here as the amount and quality of the model parameters. The amount is obtained from the dimension of θ_i, and the quality from the shape and uncertainty of the random variable θ_i. In the particular case shown here, the prior parameter distributions were:

- For model 1, $p(\theta_1|M_1) = \frac{1}{1.6 - 0.6}$.
- For model 2, $p(\theta_2|M_2) = \frac{1}{1.6 - 0.6}\frac{1}{1.2}\frac{1}{1.7} = p(\theta_1|M_1)\underbrace{\frac{1}{1.2}\frac{1}{1.7}}_{<1}$.

which corresponds to the Uniform distribution with bounds as stated in table 4. The uncertainty in the parameters of M_2 that are not part of M_1 decreases the plausibility of M_2. This means that M_2 should describe the data better in order to be chosen as the selected model. This will happen as well in the Gaussian analysis.

A value of K near 1 means that both models are almost equally likely, meaning that one can choose the model with lower complexity. That is, M_2 may describe the data better than M_1 but the increased complexity of considering three parameters instead of one gives M_1 a more pragmatical advantage. As the results suggest, M_1 compensates model description with reduced complexity and both are comparable, i.e M_2 is not as a good descriptor as it should be when considering this information.

The reason behind this lies in the fact that the geometrical estimation is an ill-posed problem because of its non-unicity. Any inverse solver (a technique that recovers geometrical parameters) that depends on gradient estimation will converge to the nearest local minimum, making the inverse problem hard to approach in this conditions. To tackle this ill-posedness, here we perform an analysis using Gaussian prior information.

In the second case, an experimental DTOF was compared to simulations considering a Gaussian pdf as prior information, in order to determine what model is convenient when one has knowledge about medium characteristics. To translate this knowledge to a Gaussian distribution, consider the expected (or most likely) value using any available source of information (i.e. bibliography, results of imaging techniques, etc.) and the uncertainty can be modeled using

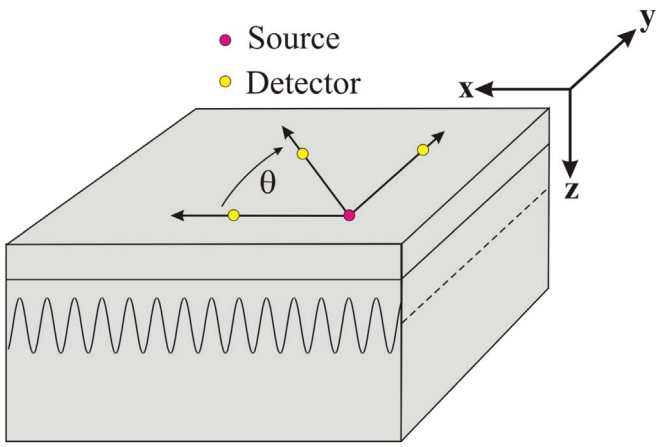


Figure 8. The three different source-detector orientations relative to the sinusoidal structure used to determine their influence on the measured pulse.

Table 6. K values taking a Gaussian prior for different geometrical guesses, i.e. varying the mean value.

Bayes factor *K*. Gaussian prior, variable mean. non-planar layers				
Guess	$\mu_L, \mu_A, \mu_\lambda$ [cm]	$\sigma_L, \sigma_A, \sigma_\lambda$ [cm]	K	Model
Experimental set up	1.1, 0.75, 1	0.1, 0.1, 0.1	2.70	Model 2
Incorrect 1	0.7, 0.085, 1.11	0.1, 0.1, 0.1	1.20	Model 1
Incorrect 2	0.6, 0.43, 0.3	0.1, 0.1, 0.1	1.73	Model 1
Incorrect 3	1.6, 0, 0.3	0.1, 0.1, 0.1	1.22	Model 1

Table 7. K values taking a Gaussian prior for the true geometrical parameters using varying standard deviation, but maintaining the initial guess fixed.

Planar versus sinusoidal interface. **Bayes factor *K***. Gaussian prior			
$(\sigma_L, \sigma_A, \sigma_\lambda)[cm]$	K	Model	Conclusion
(0.025, 0.025, 0.025)	4.03	Model 2	Substantial
(0.05, 0.05, 0.05)	3.22	Model 2	Substantial
(0.1, 0.1, 0.1)	2.70	Model 2	Barely worth mentioning
(0.2, 0.2, 0.2)	1.73	Model 2	Barely worth mentioning
(0.4, 0.4, 0.4)	1.22	Model 2	Barely worth mentioning
(0.8, 0.8, 0.8)	1.09	Model 2	Barely worth mentioning

the variance. Here, it is assumed that the parameters are mutually independent. A useful guideline to provide the standard deviation is to consider the 99% percent probability interval as $\mu \pm 3\sigma$ and derive σ from it. Once the user chooses this approach, a bias is added to the distribution which means that higher probability is assigned to regions nearby the chosen mean. The higher the variance the higher the uncertainty and, in the limit as σ tends to infinity the normal distribution tends to an improper Uniform distribution (i.e. its area is infinity). There is a special condition that must be observed, that is, the selection of the mean is critical. A good mean is one that approaches the real values and, conversely, an incorrect mean may lead to incorrect results. Here, both situations are addressed. Also, when considering the correct mean but increasing the standard deviation, it is possible to observe how the value of K approaches to that obtained when using Uniform prior information. Results of K using correct and incorrect means are shown in table 6. An incorrect information results in values of K close to unity and returns Model 1 as the suggested one. The effect of increasing standard deviation for the correct one can be seen in table 7, where it can be seen that for small deviation values (greater certainty), K tends to increase. For higher deviations (greater uncertainty) K approximates to the result obtained considering the Uniform prior.

To compare our results, we also performed the so-called Akaike information criterion (AIC) [43] which, using a KL divergence criterion, performs model selection based on the optimized log-likelihood (the maximum of the log-likelihood estimated) and the number of parameters. It is similar to the method used here -in fact, both come from the same formalism- but BMS allows for prior information in a sense that AIC does not. The explicit formula is:

$$AIC = -2\log(\nu) + 2k \tag{14}$$

where ν is the maximum of the log-likelihood according to equation (12) and k is the number of parameters. This formula is applied to each model against experimental data resulting in $AIC_1 = 3.86$ for model M1

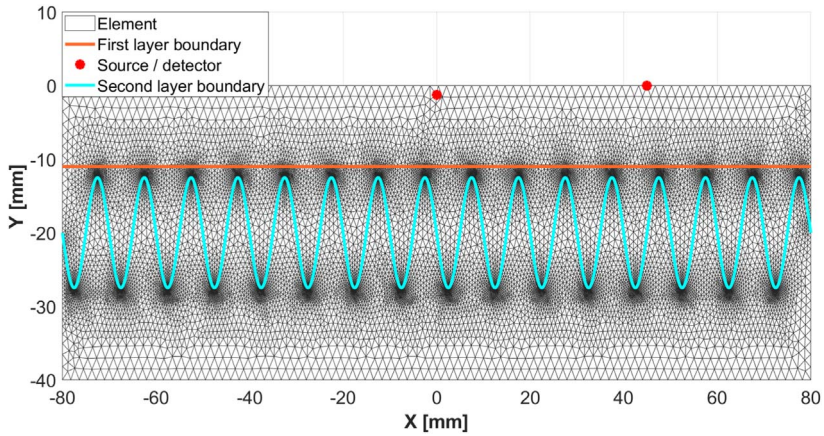


Figure 9. Mesh used for the 2D simulations. The red point inside the medium corresponds to the source and the superficial one corresponds to the detector. Layer boundaries are drawn for better mesh visualization.

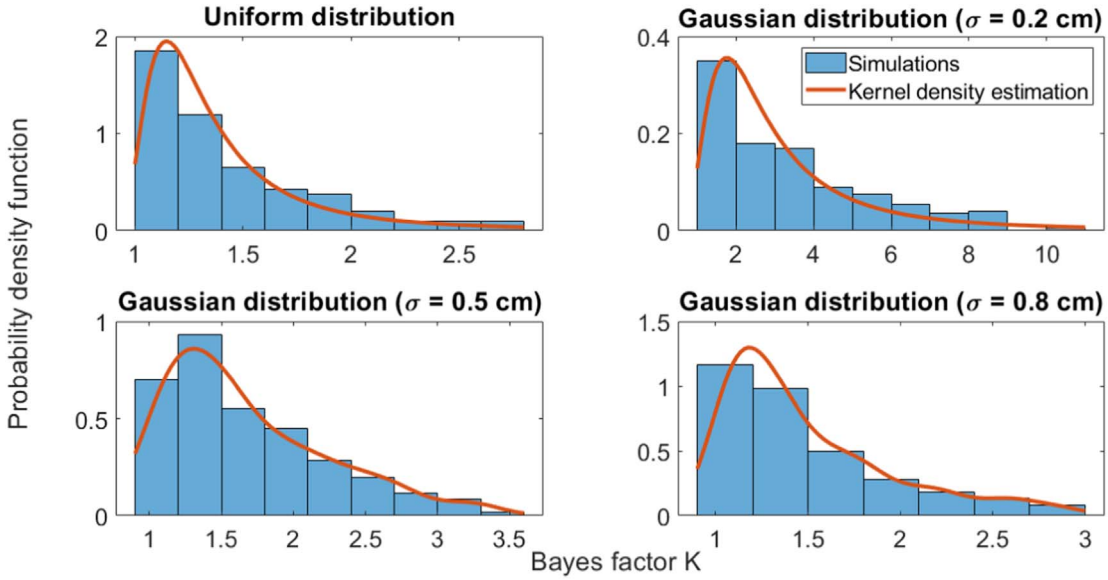


Figure 10. Histograms resulting from the 200 simulations for Uniform distribution and Gaussian distribution with three different variances.

and $AIC_2 = 7.84$ for model M2. The model with lowest AIC value is the chosen one according to this framework. Akaike's method penalizes more for the amount of parameters than our method and chooses directly Model 1 because of its simplicity. Note that in equation (14) the optimized likelihood and the amount of parameters are considered. The parameter k is fixed depending on the model, but the likelihood must be found through optimization methods or, as in this work, by exploring the parameter space. The reason behind that both criteria suggest that it is better to use M_1 instead of M_2 can be found by exploring the likelihood, given that both frameworks use it. The optimized log-likelihood for the experiment using M_1 is $\nu_{M_1} = -0.9286$ and using M_2 is $\nu_{M_2} = -0.9246$, suggesting that the optimized likelihoods corresponds to very similar DTOF according to the square norm criteria, figure 11 shows that this is the case, indeed.

This suggests that, in this particular case, it is possible to replace M_2 with geometrical parameters in table 6, which are the values predetermined in the experimental setup, with a specific planar setup. In this case this corresponds to $L = 0.7$ cm, which would be the planar equivalent configuration that replaces the sinusoidal configuration.

4.3.2. Study of changes in absorption properties and thickness of the first layer

As an additional illustration of how the BMS approach can be used, we present in this section the case of a three layers geometry, with planar boundaries, and with the deepest layer being semi-infinite (see figure 1). For this new situation, BMS is used to decide between the following two models:

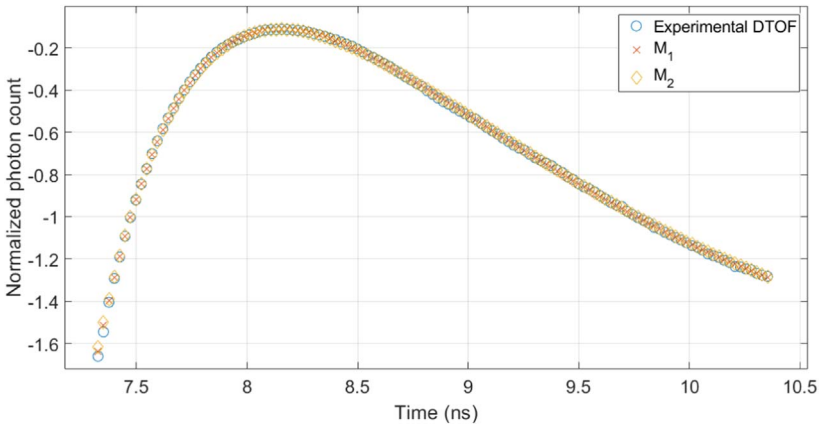


Figure 11. Comparison of the DTOF with optimized likelihood according to the square norm criteria.

Table 8. Set of parameters chosen for the 3D slab simulations considering only planar interfaces.

Planar layers			
Parameter	Values	Number of steps	Step size
$\mu_{a,1}$	$(0.02 \leqslant \mu_{a,1} \leqslant 0.26)$ cm^{-1}	20	0.013 cm^{-1}
$\mu_{a,2}$	$(0.02 \leqslant \mu_{a,2} \leqslant 0.108)$ cm^{-1}	20	0.004 cm^{-1}
L_1	$(0.15 \leqslant L_1 \leqslant 1.5)$ cm	20	0.0711cm

(i) Model 1, $M_1(\mu_{a,1})$, is a model for which only variations in the absorption coefficient of the most superficial layer can occur.

(ii) Model 2, $M_2(\mu_{a,1}, \mu_{a,2}, L_1)$, for which variations are allowed for the absorption coefficients of the first two layers, as well as as well as for the thickness of the first one.

Please note that, as in the analysis of the previous sections, we have chosen M_1 to be the simpler model and M_2 the one that depends on more parameters.

The data consisted of a DTOF, obtained from a simulated experiment using the optical and geometrical properties of table 1. It was compared to a set of 8000 simulations resulted of varying each of the three parameters $\mu_{a,1}$, $\mu_{a,2}$ and L_1, and for a fixed source-detector distance, $\rho = 2.5$ cm; details are shown in table 8. The goal is to determine if Model M_1 is both, descriptive and simple enough, to be able to replace Model M_2, which is more precise (by construction) but computationally more expensive.

Similarly as in the previous sections, we studied two different initial situations, namely: *(i)* using non informative prior information, i.e. a Uniform distribution for the parameters, representing complete ignorance of them and *(ii)* using a Gaussian initial distribution which is considered informative since it gives more weight to a given set of parameters, and its relevance can be adjusted by its deviation (uncertainty of the analyst). For the first, Uniform case (table 9) it is observed that the simpler model, M_1, dominates, what is indicated by a Bayes factor $K = 3.04$.

Table 9. Bayes Factor, K for the planar structure using a uniform prior.

Planar layers. **Bayes factor *K***. Uniform prior			
Prior pdf	K	Selected model	Conclusion
Uniform	3.04	Model 1	Barely worth mentioning

This means that the simpler model M_1 is precise enough to represent the data considering its low cost, but slightly underestimating the absorption of the first layer, as it can be seen from the maximum likelihood DTOF found with this model, shown in figure 12. Figure 13 shows the likelihood values obtained using model M_1 as function of $\mu_{a,1}$; as can be seen, the maximum likelihood occurs for $\mu_{a,1} = 0.12$ cm^{-1}.

For the case of the Gaussian prior, two situations were considered:

- when the prior information is reliable, meaning that the deviation of the distribution is small, even if its mean is not exactly known.
- when the deviation increases, meaning that there is greater uncertainty about the optical or geometrical properties.

Results are summarized in tables 10 and 11, for a correct initial guess and an incorrect one, respectively. The former case, correct guess, leads to the result that

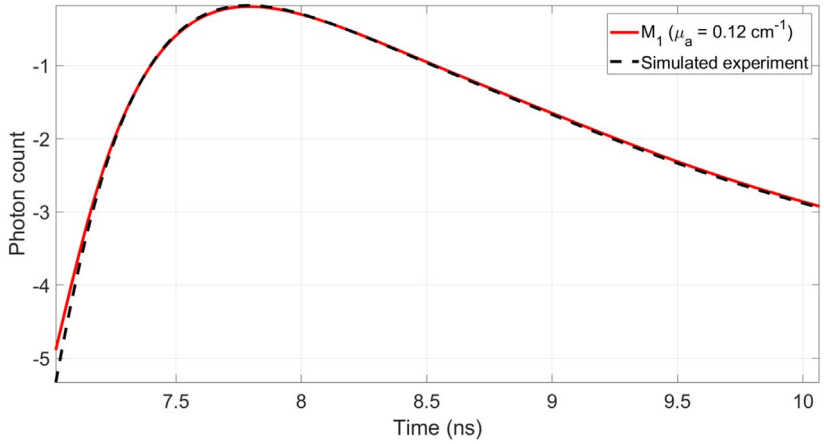


Figure 12. Maximum likelihood DTOF found by using a Uniform prior distribution. Model M_1 dominates, but the absorption coefficient of the superficial layer results underestimated.

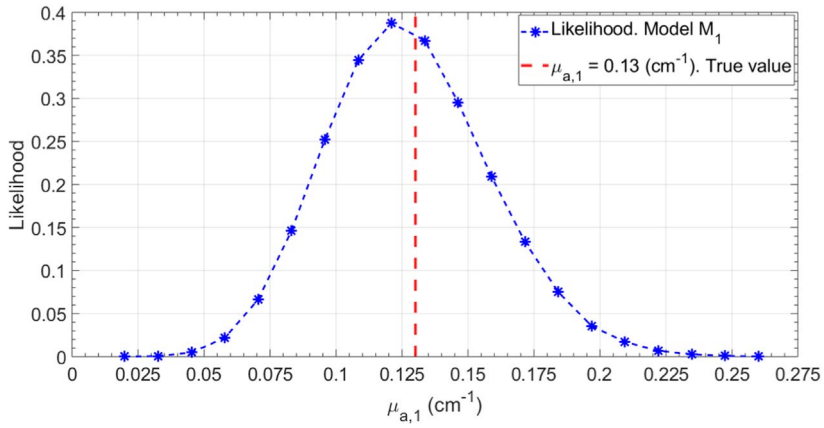


Figure 13. Likelihood values versus $\mu_{a,1}$ obtained using the simpler model, M_1. The maximum likelihood occurs for $\mu_{a,1} = 0.12$ cm^{-1}, but the true value in $\mu_{a,1} = 0.13$ cm^{-1} (vertical red line).

Table 10. K values taking a Gaussian prior for the parameters, with a correct guess and varying standard deviation. Results suggest that when the initial guess is near the true values of $\mu_{a,2}$ and L_1, M_1 dominates, specially for decreasing deviations.

Planar layers. **Bayes factor *K***. Gaussian prior, correct guess			
Correct guess. $\mu_{a,1} = 0.13$ cm^{-1}, $\mu_{a,2} = 0{,}054$ cm^{-1}, $L_1 = 0.9$ cm			
(σ_{L1}(cm), $\sigma_{\mu_{a,1}}$(cm^{-1}), $\sigma_{\mu_{a,2}}$(cm^{-1}))	K	Model	Conclusion
(0.8, 0.13, 0.054)	2.30	Model 1	Barely worth mentioning
(0.8, 0.065, 0.027)	4.06	Model 1	Substantial
(0.8, 0.043, 0.018)	5.99	Model 1	Substantial

model M_1 is substantially preferred when the deviation tends to decrease. This is reasonable because, if there is some knowledge about the correct and fixed parameters $\mu_{a,2}$ and L_1, the only property that can vary is $\mu_{a,1}$, and thus, the data can be represented by a DTOF obtained using the simpler model at a very low cost. Conversely, when the initial guess is incorrect, the simpler model, M_1, is not capable of representing the data with its only variable, $\mu_{a,1}$.

So it is necessary to combine the three parameters in order to find a DTOF that fits the data best. Note that this DTOF is not necessarily the one corresponding to the underlying optical properties because of the non-uniqueness of the light-propagation with limited perspectives. Moreover, this is a problematic situation, because it could be misinterpreted by the inexperienced operator who may think that those parameters are correct; this situation is illustrated in figure 14. It should be noted, again, that the BMS is able to show multiple local minima, i.e. several parameters can be used to reconstruct the simulated experiment in a quite good manner, but it is not a reconstruction method but an exploratory one.

5. Conclusions

In this work we implemented a Bayesian Model Selection methodology for selecting between

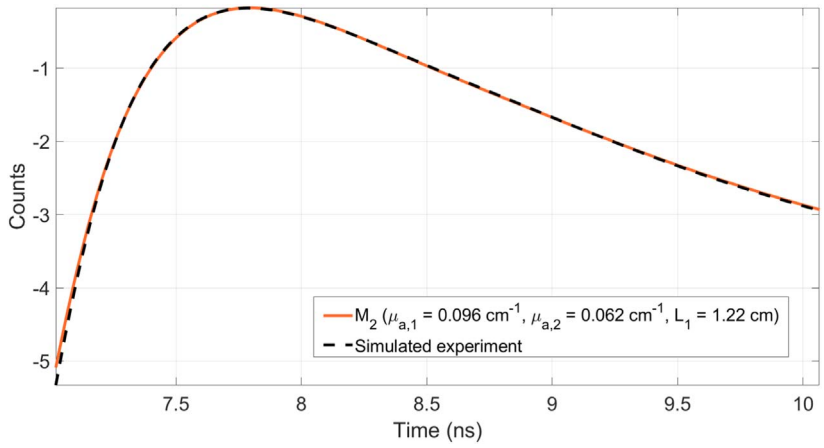


Figure 14. DTOF generated with Model M_2 compared with the simulated experiment. Although it is possible to find a new set of optical and geometrical properties that generate a DTOF that fits the simulated experiment, the parameters considered for model M_2 are not the same as the real ones.

Table 11. K values taking a Gaussian prior with an incorrect guess and varying standard deviation. Results suggest that when the initial guess is far away from the true values, M_2 dominates.

Planar layers. **Bayes factor *K***. Gaussian prior, incorrect guess.			
Incorrect guess. $\mu_{a,1} = 0.08\ \text{cm}^{-1}, \mu_{a,2} = 0{,}108\ \text{cm}^{-1}, L_1 = 0.12\ \text{cm}$			
$(\sigma_{L1}(\text{cm}), \sigma_{\mu_{a,1}}(\text{cm}^{-1}), \sigma_{\mu_{a,2}}(\text{cm}^{-1}))$	K	Model	Conclusion
(0.4, 0.04, 0.054)	1.03	Model 2	Barely worth mentioning
(0.4, 0.027, 0.036)	1.26	Model 2	Barely worth mentioning
(0.4, 0.02, 0.027)	1.41	Model 2	Barely worth mentioning

competitive models, based on a Bayesian scheme. Particularly, it was applied to a three-layered turbid medium with one non-planar interface which is intended to introduce, in a simplified and more accurate way, some geometrical aspects of the structure of a human head than the usual planar boundaries medium. Simulations were carried on 2D and 3D turbid media, and time-resolved reflectance measurements were performed on a three layered agarose phantom for the experimental validation. The Kullback-Leibler divergence was employed to quantify how reflectance changes when optodes orientation is changed. This was necessary to reduce complexity when analyzing the BMS methodology by reducing to a single orientation. The results obtained show that the information lost using different optode orientations, in this configuration, is negligible. This analysis was repeated through simulations, supporting these results.

In the simulations, the analysis of Bayes factor shows the capabilities of the method used, and suggests that it is not worth to consider a non-planar boundaries medium when using a Uniform prior distribution which can be seen as a representation of ignorance when the distribution is bounded and there is no need to perform transformations to the parameter space. When there is available information, a Gaussian prior distribution is used and, when the information regarding the geometrical parameters is more precise, BMS suggests to use the more complex model depending on the quality of our prior information.

In the experimental case, it was shown that the model M_2 is a good approach only if there is a good knowledge about the medium geometry, i.e. using a Gaussian prior distribution with mean near the true configuration and small variance. When a Uniform prior distribution is preferred, both models are equally likely to replace the given data. Even when the precise model is capable of reproducing the data correctly, the uncertainty increases with more number of parameters, resulting in preferring a simplified model or to consider it equally appropriate. Using a Gaussian prior distribution with a wrong knowledge leads to the preference of the simpler model M_1 as the appropriate one. On the other hand, if the knowledge is appropriate, i.e. the mean closer to the real model and decreasing variance, the value of K tends to increase towards M_2. The use of Bayesian model selection analysis supports that it is not worth to replace the more used planar boundary media model by a sophisticated sinusoidal approach unless the information is highly precise.

To compare our results with existing model selection methods the Akaike information criteria (AIC), which is based on a Kullback-Leibler criterion and uses the optimized likelihood and the number of parameters to make a recommendation, was used. It agrees with our results even after penalizing for the increased amount of parameters, though BMS is more flexible because it allows for different prior distributions for the parameters. This is particularly useful when the user has some knowledge about the underlying parameters of the models.

Given the results presented above for the geometrical aspects of the interfaces, we employed the same BMS formalism to decide between two three-layered models with planar boundaries, considering variations on the absorption coefficient of the first layer on the simpler case, and changes on layers one and two as well as in the thickness of the first layer of the complex one. It turns out that, when the fixed parameters are correctly guessed, the simpler model results representative enough of the data at a very low cost, and thus, it is worth to use it. Conversely, if the initial guess is wrong, the complex model results in a best representation of the data. Should the more complex model not be used, the simpler model could lead to a sub-estimation of the absorption in the superficial layer.

The methodology presented here has the potential of becoming a useful tool for choosing between different models to decide whether or not it is worth replacing complexity for precision. In this work, first results suggest that a planar model is a fairly good approximation of a brain structure, in a light propagation sense. As additional result, it is shown that a simple model with variable absorption of the first layer describes the data correctly and at a low cost, but only when the fixed initial parameters are correctly guessed. It is important to note that, although the BMS analysis was applied to two particular situations that relates only two pairs of models, it can be used in more general ways, including more sophisticated approaches. For example, cerebrospinal fluid is a critical component of the brain structure, and thus, using models that allows for void, or non-scattering regions, it could be considered using this same formalism.

Acknowledgments

Authors want to acknowledge financial support from Grants FCCIC 2016-2018 (CICPBA, Argentina) and ANPCyT PICT 2018 Nr. 1295. Support from NVIDIA® is also acknowledged for donation of the *GPU* card (*GeForce GTX Titan Xp*). Dr. Héctor O. Di Rocco is gratefully acknowledged for his critical reading of the manuscript.

ORCID iDs

Demián A Vera https://orcid.org/0000-0002-6788-888X
Guido R Baez https://orcid.org/0000-0002-3735-8589

References

[1] Jöbsis F F 1977 Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters *Science* **198** 1264–7
[2] Martelli F, del Bianco S, Ismaelli A and Zaccanti G 2010 *Light Propagation Through Biological Tissue and Other Diffusive Media* (Bellingham, Washington USA: SPIE)
[3] Ishimaru A 1989 Diffusion of ligth in turbid material *Appl. Opt.* **28** 2210–5
[4] Patterson M S, Chance B and Wilson B 1989 Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties *Appl. Opt.* **28** 2331–6
[5] Contini D, Martelli F and Zaccanti G 1997 Photon migration through a turbid slab described by a model based on diffusion approximation *Appl. Opt.* **36** 4587–99
[6] Strangman Z L G E and Zhang Q 2013 Depth sensitivity and source—detector separations for near infrared spectroscopy based on the colin27 brain template *PLoS One* **8** e66319
[7] Strangman G, Franceschini M A and Boas D A 2003 Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters *NeuroImage* **18** 865–79
[8] Essenpreis M, Elwell C E, Cope M, van der Zee P, Arridge S R and Delpy D T 1993 Spectral dependence of temporal point spread functions in human tissues *Applied Optics* **32** 418–25
[9] Kienle A, Patterson M S, Dögnitz N, Bays R, Wagnières G and van den Bergh H 1998 Investigation of two-layered turbid media with time-resolved reflectance *Appl. Opt.* **37** 6852–62
[10] Liemert A and Kienle A 2010 Light diffusion in a turbid cylinder. ii. layered case *Opt. Express* **18** 9266–79
[11] García H A, Iriarte D I, Pomarico J A, Grosenick D and Macdonald R 2017 Retrieval of the optical properties of a semi infinite compartment in a layered scattering medium by single-distance, time-resolved diffuse reflectance measurements *Journal of Quantitative Spectroscopy & Radiative Transfer* **189** 66–74
[12] García H, Baez G and Pomarico J 2018 Simultaneous retrieval of optical and geometrical parameters of multilayered turbid media via state-estimation algorithms *Biomedical Optics Express* **9** 3953–73
[13] Okada E and Delpy D T 2003 Near-infrared light propagation in an adult head model. i. modeling of low-level scattering in the cerebrospinal fluid layer *Appl. Opt.* **42** 2906–14
[14] Hallacoglu B, Sassaroli A and Fantini S 2013 Optical characterization of two-layered turbid media for non-invasive, absolute oximetry in cerebral and extracerebral tissue *PLoS One* **8** e64095
[15] Dehghani H and Delpy D T 2000 Near-infrared spectroscopy of the adult head: effect of scattering and absorbing obstructions in the cerebrospinal fluid layer on light distribution in the tissue *Appl. Opt.* **39** 4721–9
[16] Wang X 2017 Time domain transport equations of light in multilayered rectangular biological tissue with semi-infinite medium *Optik* **154** 67–73
[17] Liemert A and Kienle A 2013 Exact and efficient solution of the radiative transport equation for the semi-infinite medium *Sci. Rep.* **3** 2018
[18] Okada E, Firbank M, Schweiger M, Arridge S R, Cope M and Delpy D T 1997 Theoretical and experimental investigation of near-infrared light propagation in a model of the adult head *Appl. Opt.* **36** 21–31

[19] Wu X, Eggebrecht A T, Ferradal S L, Culver J P and Dehghani H 2015 Fast and efficient image reconstruction for high density diffuse optical imaging of the human brain *Biomedical Optics Express. OSA* **6** 4567–84
[20] Eggebrecht A T, White B R, Ferradal S L, Chen C, Zhan Y, Snyder A Z, Dehghani H and Culver J P 2012 A quantitative spatial comparison of high-density diffuse optical tomography and fmri cortical mapping *NeuroImage* **61** 1120–8
[21] Sikora J, Zacharopoulos A, Douiri A, Schweiger M, Horesh L, Arridge S R and Ripoll J 2006 Diffuse photon propagation in multilayered geometries *Phys. Med. Biol.* **51** 497–516 IOP Science (2006)
[22] Leamer E E 1978 *Specification Searches: Ad Hoc Inference with Nonexperimental Data* (New York City, United States: John Wiley and Sons)
[23] Ranyimbo A O and Held L 2006 Estimation of the false negative fraction of a diagnostic kit through bayesian regression model averaging *Statistics in Medicine* **25** 653–67
[24] Phelan S, Liu T, Gorin A, Lowe M, Hogan J, Fava J and Wing R R 2009 What distinguishes weight-loss maintainers from the treatment-seeking obese? analysis of environmental, behavioral, and psychosocial variables in diverse populations *Annals of Behavioral Medicine* **38** 94–104
[25] Visweswaran S, Angus D C, Hsieh M, Weissfeld L, Yealy D and Cooper G F 2010 Learning patient-specific predictive models from clinical data *Journal of Biomedical Informatics* **43** 669–85
[26] Conti D V, Cortessis V, Molitor J and Thomas D C 2003 Bayesian modeling of complex metabolic pathways *Hum Hered* **56** 83–93
[27] Geweke J 1999 Using simulation methods for bayesian econometric models: inference, development,and communication *Econometric Reviews* **18** 1–73
[28] Raftery A E, Gneiting T, Balabdaoui F and Polakowski M 2005 Using bayesian model averaging to calibrate forecast ensembles *Mon. Weather Rev.* **133** 1155–74
[29] Liddle A R, Mukherjee P, Parkinson D and Wang Y 2006 Present and future evidence for evolving dark energy *Phys. Rev. D* **74** 123506
[30] Parkinson D and Liddle A R 2013 Bayesian model averaging in astrophysics: a review *Statistical Analysis and Data Mining* **6** 3–14
[31] Fragoso T M, Bertoli W and Louzada F 2018 Bayesian model averaging: A systematic review and conceptual classification *International Statistical Review* **86** 1–28
[32] Royackkers N, Desvignes M, Fawal H and Revenu M 1999 Detection and statistical analysis of human cortical sulci *NeuroImage* **10** 625–41
[33] Arridge S R 1999 Optical tomography in medical imaging *Inverse Prob.* **15** R41–93 IOP Science (1999)
[34] Wang L, Jacques S L and Zheng L 1995 Monte Carlo modeling of light transport in multi-layered tissues in standard C *University of Texas M. D. Anderson Cancer Center* **47** 131–5
[35] Carbone N A, Iriarte D I and Pomarico J A 2017 Gpu accelerated monte carlo simulation of light propagation in inhomogeneous fluorescent turbid media: application to whole field cw imaging *Biomed. Phys. Eng. Express* **3** 4
[36] Arridge S R 1995 Photon-measurement density functions. part i: Analytical forms *Applied Optics.* **34** 7395–409
[37] Frieden R 2004 Science from Fihser information: . unification *Optical Sciences Center* (Tucson, AZ: The University of Arizona)
[38] Bishop C M 2006 *Pattern recognition a Machine Learning* (New York, NY 10013, USA: Springer Science + Business Media)
[39] Kullback S and Leibler R A 1951 On information and sufficiency *Ann. Math. Statist.* **22** 79–86
[40] Rudin W 1987 *Real and Complex Analysis* (New York: McGraw-Hill)
[41] Chipman H, George E I and McCulloch R E 2001 The practical implementation of bayesian model selection *IMS Lecture Notes —Monograph Series* **38** 65–116
[42] Burnham K P and Anderson D R 2004 Multimodel inference: understanding aic and bic in model selection *Sociological Methods &Research* **33** 261–304
[43] Akaike H 1974 A new look at the statistical model identification *IEEE Trans. Autom. Control* **19** 716–23
[44] Pardini P, Serra M W, Ranea-Sandoval H, Pomarico J and Iriarte D 2015 Study of inks used in biomedical optics phantoms: stability and ageing *J. Near Infrared Spectrosc.* (https://doi.org/10.1255/jnirs.1171)
[45] Bevilacqua F, Piguet D, Marquet P, Gross J D, Tromberg B J and Depeursinge C 1999 In vivo local determination of tissue optical properties: applications to human brain *Appl. Opt.* **38** 4939–50
[46] Vo-Dinh T (ed) 2003 *Biomedical photonics Handbook* (Boca Raton: CRC Press)
[47] Haeussinger F B, Heinzel S, Hahn T, Schecklmann M, Ehlis A-C and Fallgatter A J 2011 Simulation of near-infrared light absorption considering individual head and prefrontal cortex anatomy: implications for optical neuroimaging *PLoS One* **6** 10
[48] Di Rocco H O, Carbone N, Iriarte D, Pomarico J and Ranea Sandoval H 2013 Modeling transition diffusive-nondiffusive transport in a turbid media and application to time-resolved reflectance *J. Quant. Spectrosc. Radiat. Transfer* **120** 16–22
[49] Baez G R, García H, Grosenick D and Wabnitz H 2020 Implementation of the extended kalman filter for determining the optical and geometrical properties of turbid layered media by time-resolved single distance measurements *Biomed. Opt. Express* **11** 251–66
[50] Bowman A W and Azzalini A 1997 *Applied Smoothing Techniques for Data Analysis* (New York: Oxford University Press Inc.)
[51] Jaynes E T 1968 Prior probabilities *IEEE Trans. Syst. Sci. Cybern.* **4** 227–41