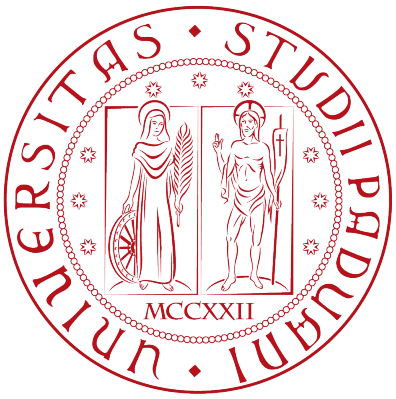


UNIVERSITÀ DEGLI STUDI DI PADOVA

Facoltà di Ingegneria
Corso di Laurea in Bioingegneria

Tesi Di Laurea

Multi-level data-driven approaches for voxel-wise quantification of dynamic PET images: application to $[^{11}C]$WAY100635 data

Laureanda
Veronica Ceccato
Matricola 620483

Relatore
Prof.ssa Alessandra Bertoldo

Correlatore
Ing. Gaia Rizzo

Correlatore
Ing. Mattia Veronese

Anno Accademico 2011/2012

I dedicate this thesis to those who have supported, encouraged and heartened me over the years, whenever things did not go as hoped. Thank you very much.

Dedico questa tesi alle persone che in questi anni mi hanno sostenuta, incoraggiata e rincuorata ogni qualvolta le cose non andavano nel modo sperato. Grazie di cuore.

Contents

List of Figures

List of Tables

Introduction

Modern medicine relies increasingly on massive technological solutions to meet the health needs expressed by the population. Drugs, devices, medical equipment along with diagnostic and therapeutic procedures identify an exceptionally broad and diversified concept of "health technology". The advancement of scientific knowledge, combined with the development of skills in the field of medicine, now offer the foundation for seemingly unstoppable technological advancement. With such innovation, we are witnessing increasing development of imaging techniques related to physiological activity. In the last fifty years, attention has focused on the analysis of the behavior of the human body, not at the level of movement or architecture, but under the point of view of its inner functionalities. This remarkable improvement is due to the birth of PET (Positron Emission Tomography) equipment that allows analysis of parameters of physiological interest by detecting the radiation emitted by a radiopharmaceutical injected into the patient's body, that would otherwise not be accessible and therefore not analyzed.

The data provided by this analysis are of fundamental importance but they have to be selected because only some of them are of interest (the signal acquired needs to be processed). Then it is necessary to correlate the data obtained with the actual, functional mechanism of the area of the body under examination. This is described by the branch of pharmacology called pharmacokinetics, which investigates the kinetics of the processes through which the action of the organism on drugs is carried out (i.e. absorption, distribution, metabolism and excretion). The final goal is to develop useful models to predict the evolution in time of concentrations in plasma and tissue, following the administration of a drug.

To achieve this goal one develops compartmental approaches, which are a type of modeling that combines the use of mathematical equations correlated to physical phenomena occurring in the body parts affected by the radiopharmaceutical. The limitation of this approach is due to the fact that it is necessary to define the type of model suitable to describe the kinetics of the drug in advance. Therefore a new approach was proposed (Cunningham and Jones, 1993; Turkheimer et al., 1994) which identifies the components of data on the radioactivity, without any *a priori* information on the kinetics of the model and on how the tracer reaches equilibrium or is dispersed. This technique

is known as Spectral Analysis. Spectral Analysis characterizes the systems impulse response function ($h(t)$) as a positive sum of exponentials; Schmidt showed that in many plasma input models the observation that all compartments lead to only positive coefficients is valid (Schmidt, 1999). However, it is straightforward to deduce that for reference tissue input models negative coefficients can be encountered and that the Spectral Analysis approach is not valid (Gunn et al., 2001). In this thesis we will refer to plasma input models, so it will be possible to work with the Spectral Analysis technique; this transparent modeling technique has application to a wide range of PET radiotracers but the emphasis in this thesis is for the tracer $\left[^{11}C\right]$WAY100635, which is a selective antagonist with affinity and selectivity for serotonin 5–HT_{1A} receptors. The analysis of the tracer was performed both at region of interest level and at voxel level plus, to extract the useful information from the data, different approaches were used.

Despite the great advantage of not requiring *a priori* information, there are some aspects of the Spectral Analysis technique which, if they are not studied with particular attention, may lead to inaccurate results. First of all it has been demonstrated that the choice of the grid, on which the linear Spectral Analysis relies its basis to be solved with a linear estimator, is critical for the goodness of the results obtained (Turkheimer et al., 1994). Secondly, this technique can be heavily affected by the presence of noise in the data. Observing that this technique permits obtaining results very close to those obtained with the non linear analysis, which when the analysis is performed at voxel level it becomes inadequate for its computational cost and for its high rate of failure, the goal of this thesis is the definition of a new linear method capable of improving the existent linear Spectral Analysis.

The methodologies studied are oriented specifically to the improvement of the grid that, as previously stated, make it possible to estimate the parameters of the model using a linear estimator. What is proposed is the automatic creation of grids that will be customized for the input data. This should led to a more robust method of analysis.

A Matlab®'s environment was utilized to implement these concepts. The following sections illustrate the theoretical basis necessary for conducting this study, the various steps executed for the implementation and the results obtained. Some comparisons between the new methods and the methods already existing (linear and non linear Spectral Analysis) are then proposed and proven. The goal is to verify the plausibility of new methods and their eventual ability to provide superior results from those found in literature.

Chapter 1

PET: Positron Emission Tomography

One of the most important imaging systems that has been created in the last fifty years for the use in biomedical researches and clinical applications is the PET scanner (part of those non optical systems that map object properties associated with invisible radiation into visible images). This particular system makes use of the electromagnetic spectrum capturing gamma rays and providing an extension to the range of vision into the realm of object properties that are totally inaccessible to natural, unaided vision.

Examples of these properties are the local value of blood volume, the tissue metabolism, the receptor binding, all of which can be obtained from images of the distribution of an administered radiotracer. In the case of PET analysis, it is the spatiotemporal distribution *in vivo* of the radioactive material that is imaged; it may be described by the local concentration of the material averaged over the observation period. This is possible today thanks to the advance in digital computer technology: now one can deal with any image as an array of numbers and make quantitative measurements of object properties from their images.

It has just been said that PET images can represent the spatial distribution of parameters such as glucose metabolism (Reivich et al., 1979), blood flow (Kety, 1951) and receptor concentration (Mintun et al., 1984); for this reason this imaging system can be used to detect tumors, locate areas affected by disease and identify regions influenced by drugs. This method is identified as a *functional imaging* approach to distinguish it from methods that study the structure of the body, intended as the architecture (some examples are conventional radiography, computed tomography (CT) and magnetic resonance (MR) imaging systems); this form of imaging, the functional one, is the result of the union of two basic principle: the *tracer principle* (Hevesy, 1962) explains how it uses the gamma-ray emission and the *tomography* that refers to the volumetric imaging of the body's interior. The tracer principle

explains how radioactive compounds behave exactly in the same way as non-radioactive materials when speaking of an organism's physiological process. Thus, radioactive materials can be detected as they emit gamma rays, they can be used to track the flow and distribution of different substances in the body. The agents, used to create the images, are called *radiopharmaceuticals* or *radiotracers* and they can be designed to function as markers for a large variety of substances that participate in the body's natural processes.

1.1 How it works

Figure 1.1 provides a schematic representation of the steps made to obtain a PET image. It shows that there is an annihilation process during which two photons are emitted in diametrically opposite directions. The PET registers these photons as soon as they arrive at the detector ring. After this step, the data is forwarded to a processing unit which decides if two registered events are, so-called, coincidence event. All coincidences are then sent to the image processing unit where the final image data is produced by mathematical reconstruction procedures.

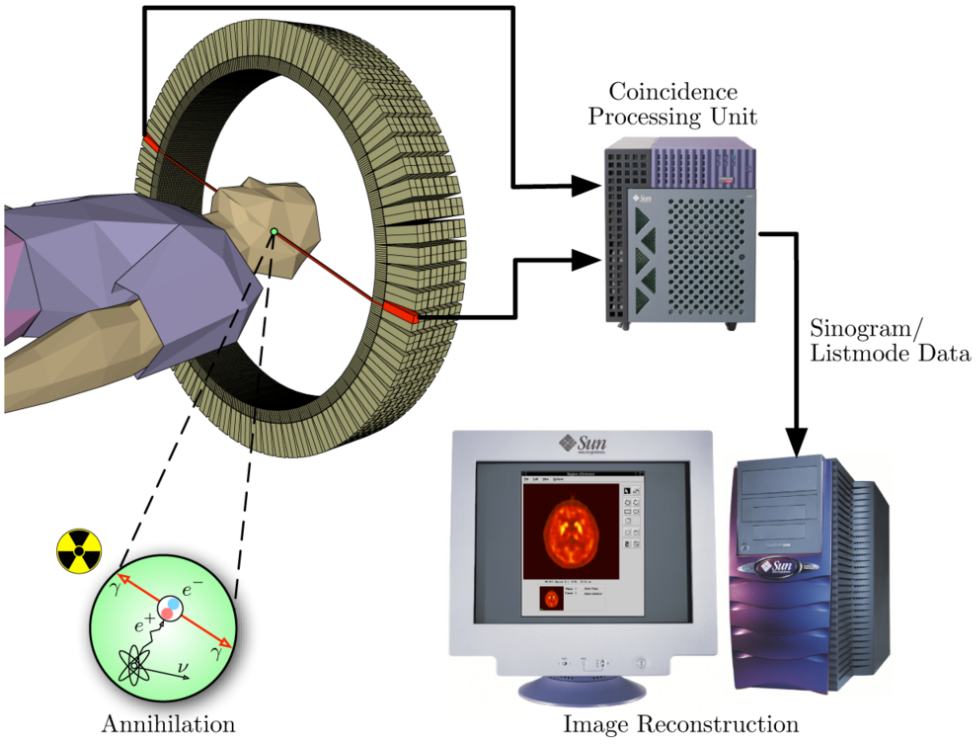


Figure 1.1: Diagram of a PET acquisition process

1.1.1 Radiopharmaceutical production

The first thing that is needed in a PET study is the radiotracer: it must be appropriate for imaging the disease (or the physiological process) in question and it needs to be produced. A limitation of this imaging system is due precisely to this last observation, in fact most of the positron-emitting radioisotopes have a short half-life and need to be produced using a cyclotron in close proximity to the PET imaging facility.

A radiotracer is a radioactive isotope that replace a stable chemical element in a compound or, more specifically, into a biologically active molecule, such as glucose (or its analogues), water, ammonia or into molecules that bind receptors or other sites of drug action, normally used by the body. The typical isotopes that are used in this kind of studies are those with short half-lives such as $[^{11}C]$ ($\sim$20 min), $[^{13}N]$ ($\sim$10 min), $[^{15}O]$ ($\sim$2 min), and $[^{18}F]$ ($\sim$110 min). It is important to notice that any kind of molecule can be radiolabeled with a PET isotope; this implies that we can trace the biological pathway of any compound in living humans (specific processes that can be tested with PET are then virtually limitless).

1.1.2 Administration of radiotracer and data acquisition

The next step is the introduction of the radiotracer into the body; usually this is done by injection but sometimes can be done by inhalation (it depends on the radiotracer and on the purpose of the exam). Speaking of the dose of radiotracer that can be used thee are two limitations to take on account: the first that impose a safety limit on the dose so the internal organs won't be exposed to a dangerously high radiation (Budinger, 1998) and the second that there is an inner limit of the scanner, it is not able to detect information over a certain count/rate. The time lapse between the administration of the radiotracer and the beginning of the acquisition has to be established accordingly to the purpose of the study and the type of tracer used.

The theory behind the acquisition of the data is that at a certain moment (that can be precisely known) the radioisotopes distributed inside the body start their decay process and emit positrons (positive beta decay). A positron is a form of antimatter; it has the same mass as an electron but an opposite charge. When a positron encounters an electron the two completely annihilate and release energy; in order for this to occur the positron must travel into the surrounding tissue, losing kinetic energy until it reaches a velocity that allows it to interact with the electron. This interaction produces a pair of gamma photons, that are anti-parallel (emitted at 180° relative to one another) and so move in opposite directions.

After this, those photons travel through the tissue since they reach the scintillator where the incident ones strike the surface called photocathode; here, due to the photoelectric effect, there is an emission of electrons. Following

their passage through the scintillator, they lose energy and excite other electrons in the process, that decay back to their ground state and give off light. The scintillator is coupled to a photomultiplier tube (PMT) which generates an electrical signal when light hits it. Then there is an electronic circuit that allows one to determine the position the crystal has undergone with scintillation. The last circuit permits recognition of the energy of the incident radiation, to select the signals that are useful to reconstruct the image and to localize the precise coordinates of the point where the event took place.

1.1.3 Image reconstruction

The next necessary step is to reconstruct the image starting from the projection data acquired; the PET machine provides a list of events that took place at almost the same time (the timing window allowing to affirm that two events are coincident is typically assumed to be of a few nanoseconds) and so have been registered by a pair of detectors in the same Line Of Response (LOR), a straight line of coincidence. With this data it is possible to obtain the *sinograms*, groups of projection images organized by the angle of each view and inclination (when required).

Algorithms of reconstruction can be principally of two types:

- *analytic strategies*: The approach used with higher frequency is based on the filtered backprojection (FBP) principles, or its modification (Ollinger and Fessler, 1997). It is based on the inversion of the Radon transform through the central slice theorem (Deans, 1983); it has the advantage of speed, since it operates linearly on the data but it has two main disadvantages. The first is that statistical noise in the data manifests as high-frequency components and, since it uses the ramp filter, it amplifies high-frequency components in the back-projection, so the process of FBP amplifies noise in the image. The second is that the images produced are of a poor quality. This poor quality is due to the fact that it doesn't use the information about the spatially-variant response and it treats the measurement noise in a post-hoc manner.

- *iterative statistical methods*: Two of the main examples of this approach are the maximum likelihood −expectation maximization (ML−EM) (Shepp and Vardi, 1982) and the maximum a posteriori (MAP) (Hebert and Leahy, 1989) algorithms. Those methods are based on models that describe the physical systems and the statistics of the measurements, giving the possibility to improve the performance of the bias-variance.

Once PET machines were designed with a single ring of detector and this was a limitation because it didn't allow acquisition of data from other planes. Today PET systems are composed of multiple rings: any ring create a tomographic image of the correspondent plane and assembling all the images

together it is possible to reconstruct a 3D distribution of the radiotracer in the organ under study. Data can be reconstructed utilizing two approaches:

- *2D*: There are partitions of lead or tungsten that are interposed between the different rings of crystal, that allow coincidences to be registered only between detectors situated on the same ring or on rings very near to the other one.

- *3D*: There is no requirement for any type of partitions and so coincidences can be detected from any combination of rings, this allows reconstruction of the entire volume.

One problem of the 2D approach is that the partitions create a shadow effect that implies a reduction of over 50% of the events that are truly utilized. The other approach, where the partitions are removed, allows the utilization of a higher number of LOR, increasing the sensitivity to true coincidences but making the system more sensitive to the effects of scatter and random coincidences (it detects scatter photons with an higher angle of deviation). True coincidences occur when both photons from an annihilation event are detected by detectors in coincidence, scattered coincidences are those in which at least one of the detected photons has undergone one Compton scattering event prior to detection and random coincidences occur when two photons, not arising from the same annihilation event, are incident on the detectors within the coincidence time window of the system (Figure 1.2 represents these three types of coincidences). Another limitation of this last method is that increasing the FOV (Field of View) for single events it also increase the sensitivity to random events.

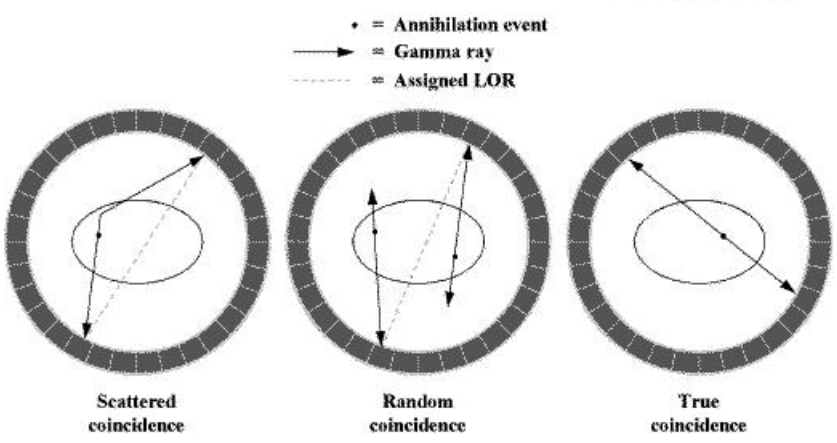


Figure 1.2: Types of coincidences in PET

1.1.4 Image analysis

The amount of data collected from thousands of coincidence events allows reconstruction of the total activity for each voxels (a volume element) and so obtain a map of radioactivity as a function of space distribution. A computer uses the data obtained to create multi-dimensional (normally 3-dimensional volumetric or 4-dimensional time-varying) images that show the distribution of the radiotracer in the part of the body under study. PET studies can be categorized by the manner in which time variations of the radiotracer distribution are treated in the data-acquisition process, and the type that we are going to analyze is a *dynamic* study, which aims is to capture the radiotracer distribution as a function of time.

1.2 Applications

PET imaging systems have assumed a fundamental role in many clinical applications in different fields of interest, like for example cardiology, oncology and almost every neurologic disease or disorder (Volkow et al., 1996). The study of physiological parameters and their pathological variation has been carried out also to investigate brain function in normal volunteer with the goal of developing experimental models for various human diseases, studying the *in vivo* pharmacokinetics of drugs and the efficacy of disease treatment.

Concerning the brain, which is the organ of interest for our work, we can see that the application of PET images is wide; this type of analysis is part of the broader discipline called neuroimaging (Nilsson and Markowitsch, 1999). The advent of sophisticated radiotracer, that can be specifically targeted, has enabled us to study the biochemistry of the brain in terms of blood flow, oxygen, glucose metabolism and many other parameters, and to also study tissues of the working brain. These measurements reflect the amount of brain activity in the various regions of the brain and, thanks to an improved resolution, it permits to study precisely which area of the brain is activated by a particular task. This is still limited from the fact that the radioactivity decays rapidly permitting one to monitor only short tasks (Nilsson and Markowitsch, 1999).

As previously stated, PET scanning is used for the diagnosis of brain disease and, in particular, it has notable use in brain tumors, strokes and neurodamaging diseases which cause dementia. All those diseases have a thing in common: they cause changes in brain metabolism which in turn causes detectable changes in PET scans. This is one of the reasons why in early cases of certain dementias PET is more useful than CT and MRI in which there are too few differences in brain volume and gross structure to detect any differences from the normal subject (Pietrzyk et al., 1996).

Chapter 2

I/O Quantitative methods

When speaking of tracer kinetic modeling there are two main approaches, which return biological based parameter estimates, that can be taken in consideration:

- *The model-driven approaches*:

  Different model-driven approaches have been proposed over the years (Kety, 1951; Sokoloff et al., 1977; Phelps et al., 1979; Gunn et al., 2001) and they use a particular compartmental structure to describe the behavior of the tracer and allow for an estimation of either micro or macro system parameters. In order to interpret the observed PET data over time, it has been assumed that there are physiologically separate pools of tracer substance known as *compartment*. A compartment is characterized by one variable, the concentration of the tracer in it, which is described as a function of time. These parameters are defined by ordinary differential equations which show the equilibrium between the input quantity and the output quantity for each compartment. Solving these equations it is possible to obtain the quantities of interest for our study. In the studies that have been done with PET systems, the models that described the activity of the tracers were in many cases with one, two or three compartment (Gunn et al., 2002). In many cases, the model's input function was assumed to be the measured blood curve or, with other words, the concentration of the radiotracer in the blood as a function of time. With the right assumptions it is possible to calculate the output of the model and make a comparison with the data obtained by the PET system. This comparison will result in an estimate of the kinetic parameters and, with it, it will be possible to extract information about delivery, binding or any process that is still unknown, or not completely known.

- *The data-driven approaches*:

  Data-driven analysis, on the other hand, require no *a priori* decision about the most adequate model structure (and the connections between the different compartments); instead, this information, is obtained directly from the data. This permits to avoid problems like for example if the data for all subjects should be forced into a single model, equal to all, or if every subject should be modeled upon the fit of their own plasma concentration data (Nilsson and Markowitsch, 1999). Some examples of these methods are graphical approaches such as the Patlak and Logan plots (Logan et al., 1990; Patlak et al., 1983), Spectral Analysis (Cunningham and Jones, 1993; Turkheimer et al., 1994) and Basis Pursuit method (Gunn et al., 2002)

Let us examine, with more detail, the second approach by referring specifically to the Spectral Analysis technique, that is a powerful tool for the detection of components in PET dynamic data. In the early '90s, Cunningham and Jones (Cunningham and Jones, 1993) developed a new technique able to identify the components in PET tissue radioactivity data without requiring any prior assumptions, like for example information about the kinetic model or the tissue equilibration or the product loss. Because the technique provides a spectrum of the kinetic components in the data, it has been called "Spectral Analysis" (Turkheimer et al., 1994). Spectral analysis is applicable to heterogeneous as well as homogeneous tissues and, despite the fact that it has been created to determinate local metabolic rate of glucose in the brain, it is now applicable to a wide range of tracer compounds and so to a large number of different physiological systems (Schmidt, 1999).

It is possible to observe that kinetic modeling techniques are used to estimate biological parameters by fitting a mathematical model to the time-activity curve (TAC) either for an anatomical region of interest (ROI) drawn on the image or otherwise for each image voxel. Calculations at the voxel level combined to the classical non linear least−squares (NLS) is impractical because the computational speed is not adequate (Feng et al., 1995) and convergence of the global optimum may be problematic at high noise levels typical of pixel time-course (Maltz, 2002). For these reasons are now presented both approaches, linear and non linear SA, in terms of model equations and parameters estimation.

2.1 Linear Spectral Analysis

The linear spectral analysis technique allows the description of the radioactivity curve in the tissue in terms of a subset of kinetic components. Instead of basing it on the *a priori* definition of a kinetic model and on the non linear least-squares algorithms to estimate the parameters, this new approach describes the measured data by choosing, from a large set, a linear combination

of components. The larger set is chosen in order to cover the range of possible components that may be detectable in the data. To relate the system under consideration to an appropriate compartmental model, the set is composed by convolution integrals of the input function with decaying exponentials (see Equation (2.1)). This allows use of the compartmental model for interpreting the kinetic data and the estimation of the parameters of interest (Turkheimer et al., 1998).

2.1.1 Equations and parameters

In most kinetic models used with PET tracer, the exchange of material among different compartments is assumed to follow a first-order kinetics. What happens is that the tracer delivered to the tissue, exchanges, by the arterial blood, with one or more components and there can be an exchange of material between different components as well. The model in the linear SA for the total concentration of the tracer in the tissue can be written as the convolution integral:

$$C_T(T) = \sum_{j=1}^{n} a_j \int_0^T C_p(t) e^{-b_j(T-t)}\, dt \tag{2.1}$$

where $C_T(T)$ is the tissue activity at time T, $C_p(t)$ denotes the plasma concentration (t minutes after injection of tracer), n is the number of compartments in the tissue, and a_j and b_j are the parameters describing exchange of tracer among the compartments, the values of which have to be estimated. a_j and b_j are assumed to be positive or zero; this constrain derives from the assumption that we are dealing with first-order tracer kinetics (Turkheimer et al., 1998).

The estimation problem is nonlinear in the parameters b_j but, fixing their number, i.e. $N = 100$, and their value, we can build the same number of basis functions

$$f_j(T) = \int_0^T C_p(t) e^{-b_j(T-t)}\, dt, \quad j = 1, 2, ..., N \tag{2.2}$$

and with them obtain a linear estimation problem that can be solved in the linear coefficients a_j. The model, therefore, becomes

$$C_T(T) = \sum_{j=1}^{N} a_j f_j(T) \tag{2.3}$$

where the mathematical process now consists of estimating the N-vector $\mathbf{a} = [a_1, a_2, ..., a_N]$. We can observe that if the basis function $f_j(T)$ is not included in the model the relative coefficient a_j is zero.

The selection of the best set of exponents (b_j) consists of choosing an upper and a lower bound for the values of b_j, as well as their distribution within the chosen interval. The distribution which is commonly used is the one suggested

by DiStefano (DiStefano, 1981)(used also in part of this work) and it is given by

$$b_j = \frac{1}{\tau_j}, \quad \tau_j = \tau_{j-1}\left[\frac{\tau_N}{\tau_1}\right]^{[1/(N-1)]} \tag{2.4}$$

where τ_1 is the fastest time constant of the distribution, τ_k is the slowest, and N is the number of the points in the distribution. Generally speaking we can affirm that the equilibrating component of tissues with very high rates of efflux and/or metabolism, i.e., $b_j \to \infty$, may become indistinguishable from the plasma compartment itself. That is,

$$\int_0^T C_p(t)e^{-b_j(T-t)}\,dt \to \frac{1}{b_j}C_p(T) \tag{2.5}$$

as b_j becomes large. On the other hand, compartments that are equilibrating very slowly may become indistinguishable from the integrated plasma activity, i.e.,

$$\int_0^T C_p(t)e^{-b_j(T-t)}\,dt \to \int_0^T C_p(t) \tag{2.6}$$

as $b_j \to 0$. The last component can be considered the "low-frequency" one because it accounts for the irreversible trapping of the tracer. There are components that assume intermediate values of b_j, "intermediate frequency", which reflect the extravascular activity of the tracer (Turkheimer et al., 1998).

Returning to the equation of the total concentration of the tracer (Eq. (2.3)), one notices that, due to the impossibility of continuous time measurements of the concentration, the equation for $C_T(T)$ is replaced by the discrete time counterpart. If the tissue data are sampled at times $t_1, t_2, ..., t_M$, the model equation becomes:

$$C_T(t_k) = \sum_{j=1}^{N} a_j f_j(t_k), \quad k = 1, 2, ..., M \tag{2.7}$$

where M is the number of data points (frames). It is now important to be aware of the fact that measured tissue concentrations always includes noise, or errors in the measurements; if $X(t_k)$ represents the measured tissue concentration at time t_k, and ϵ_k is the error in the measurement, we have

$$X(t_k) = C_T(t_k) + \epsilon_k = \sum_{j=1}^{N} a_j f_j(t_k) + \epsilon_k \tag{2.8}$$

(the errors are assumed to have zero mean). Estimation of the vector $\mathbf{a}$ is usually carried out with the non-negative least squares (NNLS) algorithm: for a given set of weights, w_k, the NNLS determines the linear coefficients that minimize

$$\|\mathbf{r}(a_1, a_2, ..., a_N)\|_2^2 = \sum_{k=1}^{M}\left[\frac{X(t_k) - C_T(t_k)}{w_k}\right]^2 \tag{2.9}$$

where r is the residual vector and the constrain $a_j \geq 0, j = 1, 2, ..., N$ still exists. The result of minimization of that is a N-vector of estimated coefficients $\hat{\mathbf{a}} = [\hat{a}_1, \hat{a}_2, ..., \hat{a}_N]$ which usually contains only a few values different from zero.

Components detected with spectral analysis can be combined to obtain parameters of physiologic interest. We will use, as proposed by Cunningham et al. (Cunningham et al., 1993), the detected components $(\hat{a}_j, b_j)$ to obtain an estimate of the impulse response function $(h(t))$ of the tissue:

$$\hat{h}(t) = \sum_{j=1}^{N} \hat{a}_j e^{-b_j t} \tag{2.10}$$

From this equation, the parameters K_1, the unidirectional clearance of tracer from blood to tissue, and V_d, the volume of distribution of the tracer in tissue relative to blood, were determined as:

$$K_1 = h(t = 0) = \sum_{j=1}^{N} \hat{a}_j \tag{2.11}$$

$$V_d = \int_0^{\infty} h(t)\, dt = \sum_{j=1}^{N} \hat{a}_j / b_j \tag{2.12}$$

(Turkheimer et al., 1998).

This method shows both advantages and disadvantages. It is capable of fitting well the data and, using a linear estimator, is really fast computationally speaking. But it has some problems too: the first limitation is the requirement of imposing non-negativity constraints on the parameters a_j. Secondly, to the a_j which show very poor precision, it rises their values and even if one can improve the estimation by using new grids of b_j to avoid line doubling, some bias may occur in any case. This because the "true" b_j is not necessarily the arithmetic mean of the doubled line eigenvalues (Bertoldo et al., 1998). So it can be heavily penalized by the choice of the grid and by the noise in the data.

2.2 Non Linear Spectral Analysis

In this work the nonlinear approach has been implemented both as a reference method and as the starting point for the implementation of a new method.

2.2.1 Equations and parameters

The model in the NLSA (Non Linear Spectral Analysis) for the total concentration of the tracer in the tissue can be rewritten as for the linear SA (see Eq. (2.1)) and then solved with a non linear estimator. This approach is proposed by revisiting the classical SA as an exponential impulse response

identification problem. From compartmental theory it is known that the coefficients of the impulse response do not have to be positive, so the restriction that a_j have to be positive is not necessary (Carson et al., 1983). Then since SA doesn't want to impose any structure but has the aim of finding the number of exponential components of the system impulse, also the definition of the grid is unnecessary. This new approach proposes to estimate the number of exponentials necessary for the impulse response

$$h(t) = \sum_{j=0}^{N} a_j e^{-b_j t} \tag{2.13}$$

to develop a good fit to the data; this can be obtained by using models of increasing order and, once that all the parameters have been estimated, one can choose the best one by using criteria of parsimony and other considerations.

In this work the models tested are the following, where the term V_b (unitless) accounts in the brain data for the vascular volume present in the tissue ROI and the term $C_b(t)$ for the whole-blood time activity (kBq/ml):

- model with only trapping component

$$C_i(t) = V_b C_b(t) + a_0 \int_0^t C_p(\tau)\, d\tau \tag{2.14}$$

- one-exponential model

$$C_i(t) = V_b C_b(t) + a_1 \int_0^t C_p(\tau) e^{-b_1(t-\tau)}\, d\tau \tag{2.15}$$

- one-exponential model plus trapping component

$$C_i(t) = V_b C_b(t) + a_0 \int_0^t C_p(\tau)\, d\tau + a_1 \int_0^t C_p(\tau) e^{-b_1(t-\tau)}\, d\tau \tag{2.16}$$

- two-exponential model

$$C_i(t) = V_b C_b(t) + a_1 \int_0^t C_p(\tau) e^{-b_1(t-\tau)}\, d\tau + a_2 \int_0^t C_p(\tau) e^{-b_2(t-\tau)}\, d\tau \tag{2.17}$$

- two-exponential model plus trapping component

$$C_i(t) = V_b C_b(t) + a_0 \int_0^t C_p(\tau)\, d\tau + a_1 \int_0^t C_p(\tau) e^{-b_1(t-\tau)}\, d\tau + a_2 \int_0^t C_p(\tau) e^{-b_2(t-\tau)}\, d\tau \tag{2.18}$$

- three-exponential model

$$C_i(t) = V_b C_b(t) + a_1 \int_0^t C_p(\tau) e^{-b_1(t-\tau)}\, d\tau + a_2 \int_0^t C_p(\tau) e^{-b_2(t-\tau)}\, d\tau + a_3 \int_0^t C_p(\tau) e^{-b_3(t-\tau)}\, d\tau \tag{2.19}$$

For each model it has been assumed the same input function and the parameters were estimated by weighted non linear least squares (WNLLS); weights were chosen optimally as

$$w_k = \frac{\Delta t_k}{C_i^{obs}(t_k)} \tag{2.20}$$

where Δt_k is the length of the scanning interval and $C_i^{obs}(t_k)$ is the description of each tissue activity curves, defined as follow:

$$C_i^{obs}(t_k) = C_i(t_k) + e(t_k), \quad k = 1, 2, ..., M \tag{2.21}$$

where t_k is the midscan time, $e(t_k)$ is the measurement error at time t_k, and M is the number of data points (frames). So, the cost function that has to be be minimized is

$$WRSS(p) = \sum_{k=1}^{M} w_k \left[C_i^{obs}(t_k) - C_i(\mathbf{p}, t_k) \right]^2 \tag{2.22}$$

with $WRSS$ that denotes the weighted residual sum of squares, w_k is the weight of the kth datum and $\mathbf{p}$ is the vector of unknown model of dimension P. The measurement error was assumed to be addictive, uncorrelated, Gaussian, zero mean and with a variance equal to

$$\sigma^2(t_k) = \gamma \frac{C_i^{obs}(t_k)}{\Delta t_k} \tag{2.23}$$

where Δt_k is the length of the scanning interval relative to $C_i^{obs}(t_k)$ and γ is an unknown proportionality constant. The scale factor γ was estimated *a posteriori* as

$$\gamma = \frac{WRSS(\hat{\mathbf{p}})}{M - P} \tag{2.24}$$

where $WRSS(\hat{\mathbf{p}})$ is the value of the cost function evaluated at the minimum, i.e., for $\mathbf{p}$ equal to the vector of estimated model parameters $\hat{\mathbf{p}}$. With this new SA model it was possible, in an easy way, to obtain also the precision (or uncertainty) of the parameter estimates calculated in the form of Coefficient of Variation (CV) as

$$CV = 100 * \frac{\sigma(\hat{\mathbf{p}})}{\hat{\mathbf{p}}} \tag{2.25}$$

where $\sigma(\hat{\mathbf{p}})$ is the standard deviation defined as the square root of the variance of the estimates of the parameters, calculated *a posteriori*. The residuals and the weighted residuals have also been calculated because they allowed us to check two things: that the model describes the data correctly and that the assumptions made on the measurement error (i.e. zero mean and uncorrelated samples) are verified; at time t_k these parameters are defined as:

$$res(t_k) = C^{obs}(t_k) - C(\hat{\mathbf{p}}, t_k), \quad k = 1, ..., M \tag{2.26}$$

$$wres(t_k) = \frac{C^{obs}(t_k) - C(\hat{\mathbf{p}}, t_k)}{\sigma(t_k)} \quad k = 1, ..., M \tag{2.27}$$

Then the Akaike Information Criterion (AIC) was evaluated, as a parsimony criterion useful to compare different models. Its definition is:

$$AIC = M \cdot ln(WRSS(\hat{\mathbf{p}})) + 2P \tag{2.28}$$

Also this method have advantages and disadvantages; it gives informations about the model that better describes the data but it requires a non linear estimator that makes the time of computation inadequate for the analysis at voxel level. It has also an high tax failure rate since convergence of the global optimum is problematic at high noise levels.

Chapter 3

Novel approaches to improve I/O model performances

The goal that this study wants to achieve is to propose new methods capable of overcome the limitations of the existent methods. Since is not possible to use non linear approaches when the analysis has to be done at voxel level, for the computational time necessary and for the high failure rate, the linear approaches are the only possible choice. The existent linear SA approach presents high sensitivity to the noise and its results are really dependent on the grid that has been defined. Are now proposed four alternative methods having the purpose of making the existent SA more robust, improving the results at voxel level.

3.1 SA with impulse response linearized

The first idea that we tried to develop was based on the linearization of the impulse response

$$h(t) = \sum_{j=0}^{N} a_j e^{-b_j t} \tag{3.1}$$

using the Series of Taylor. This mathematic tool is a representation of a function as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{3.2}$$

where $n!$ denotes the factorial of n and $f^{(n)}(a)$ denotes the nth derivative of f evaluated at the point a. We sought to obtain models (starting from the easiest model with one convolution term and arriving to the more complex model with

two or three convolution terms) solvable with linear least of squares. Once estimated the parameters a and b, we would take advantage of the information given by the b values by defining a new grid, that can be specifically constructed around the expected value for that voxel.

The reasons why we wanted to linearize the model in the first instance are that we couldn't afford the analysis at voxel level, computationally speaking, with the nonlinear approach and also because with this method the convergence of the global optimum may be problematic at high noise levels typical of pixel time-course. Then, the reason why once that we obtain the parameters with the linear estimation we wanted to modify the grid, instead of using them directly as parameters of the model, is that the linear method performance shows poor results at high noise levels (i.e. as is typical for the TACs derived for each single voxel). Knowing that the choice of the grid can heavily penalize the final results, the best choice was to find a way of defining the grid based on *a priori* information.

The calculations performed are the following:

- We started with the easiest model with just one convolution term: $h(t) = a_1 e^{-b_1 t}$, its development with the series of Taylor approximated at the second order around the point $t = 0$ is: $a_1 - a_1 b_1 t + a_1 b_1^2 t^2 = A - Bt + Ct^2$. The model that needs to be estimated with the Spectral Analysis method is: $C_T = h(t) * C_p(t)$ that becomes:

$$C_T = [A\ B\ C] \begin{bmatrix} \int C_p(t) \\ -C_p(t) * t \\ C_p(t) * t^2 \end{bmatrix} \tag{3.3}$$

  Using the liner least squares it is possible to estimate the observable parameters $A,\ B,\ C$ from which it is possible to derive the parameters of the model:

$$\begin{cases} A = a_1 \\ B = a_1 b_1 \\ C = a_1 b_1^2 \end{cases} \rightarrow \begin{cases} a_1 = A \\ b_1 = B/a_1 \end{cases} \tag{3.4}$$

  The model linearized in this way is *a priori* identifiable.

- The second model that we tried to linearize was one with one convolution term plus the component of trapping; the impulse response is: $h(t) = a_0 + a_1 e^{-b_1 t}$ which becomes $a_0 + a_1 - a_1 b_1 t + a_1 b_1^2 t^2 = A - Bt + Ct^2$ after the linearization. The model that requires estimation this time is:

$$C_T = [A\ B\ C] \begin{bmatrix} \int C_p(t) \\ -C_p(t) * t \\ C_p(t) * t^2 \end{bmatrix} \tag{3.5}$$

  Using the liner least squares it is possible to estimate the observable parameters $A,\ B,\ C$ from which it is possible to derive the parameters

of the model:

$$\begin{cases} A = a_0 + a_1 \\ B = a_1 b_1 \\ C = a_1 b_1^2 \end{cases} \rightarrow \begin{cases} b_1 = C/B \\ a_1 = B/b_1 \\ a_0 = A - a_1 \end{cases} \tag{3.6}$$

Also this model, linearized in this way, is *a priori* identifiable.

- When we linearized the model with higher complexity, that is the one with two convolution terms $h(t) = a_1 e^{-b_1 t} + a_2 e^{-b_2 t}$, we needed to use the series of Taylor approximated at the third order (always around the point $t = 0$); the result is $a_1 \left[1 - b_1 t + b_1^2 t^2 - b_1^3 t^3\right] + a_2 \left[1 - b_2 t + b_2^2 t^2 - b_2^3 t^3\right] = A - Bt + Ct^2 - Dt^3$. The model becomes:

$$C_T = [A\ B\ C\ D] \begin{bmatrix} \int C_p(t) \\ -C_p(t) * t \\ C_p(t) * t^2 \\ -C_p(t) * t^3 \end{bmatrix} \tag{3.7}$$

Also in this case it is possible to estimate the observable parameters $A,\ B,\ C,\ D$

$$\begin{cases} A = a_1 + a_2 \\ B = a_1 b_1 + a_2 b_2 \\ C = a_1 b_1^2 + a_2 b_2^2 \\ D = a_1 b_1^3 + a_2 b_2^3 \end{cases} \tag{3.8}$$

and from these estimated parameters it is possible to obtain the kinetic micro-parameters (equations not presented).

Despite that the linearized models are *a priori* identifiable, a short simulation (not presented) demonstrated how is not possible to properly describe a decaying exponential function with a Taylor series approximation. Thus, we decided not to proceed with this method and tried other approaches.

3.2 SA with prior from ROI

As a second alternative approach, we chose a hierarchical approach, in which the image is clustered into K clusters then, on the curves obtained, we applied the non linear Spectral Analysis thus obtaining the optimal model and estimates for the parameters a_j and b_j. At this point we decided to keep the dense grid of beta (number of $b = 100$), as in the original version, but imposing *a priori* information on the exponents, obtained with the nonlinear SA. This method relies then on the hypothesis that the parameter estimates derived at cluster level are representative of the expected model parameters values of the voxels composing the cluster itself.

The first step was the implementation of a clustering on the PET image; the meaning of which was to use functional information, instead of anatomical,

to segment the image, allowing us to obtain kinetically homogeneous regions. The algorithm chosen to define the clusters was the K-Means partitioning method (MacQueen, 1967), implemented using the Euclidean distance as a metric to separate the TACs (Time Activity Curve) into different clusters. The algorithm is based on the principle of maximizing the distance between the various clusters and to minimize the variance within each cluster. It consists of the following steps:

1. Place K points into the space represented by the objects that are being clustered; these points represent initial group centroids.

2. Assign each object to the group that has the closest centroid.

3. When all objects have been assigned, recalculate the positions of the K centroids.

4. Repeat Steps 2 and 3 until the centroids no longer move; this produces a separation of the objects into groups from which the metric to be minimized can be calculated.

Matlab provides a function that allows to implement this method (*kmeans*) and gives the opportunity of specifying many parameters, like:

- The data: we first tried to use the TACs but the result was not physiologically corrected, so instead of passing the TACs directly to the kmean, we passed to it some parameters that represent the TACs: we chose the area under the curve (AUC) and the slope from the peak to the midpoint (where the middle point is the point at the time in the middle between the time of the peak and the end time);

- The number K of clusters: we use $K = 8$ because we were able to describe all significantly different kinetics within the brain;

- The number n of replicates: this number, we chose $n = 5$, indicates that the algorithm has to repeat the clustering n times with different starting points and choose the replica where it benefits from the best cost function's minimization;

- The type of distance: in this work we chose the Euclidean distance to measure the distance between each element of the "population";

- The maximum number i of iterations: this number, fixed in our work at the value $i = 200$, indicates the maximum number of iterations that the algorithm can compute.

The function gave as results a vector that for each voxel associates a number that identifies the cluster to which it has been assigned and also a matrix

that contains the medium trend of each cluster. Figure 3.1 shows the kmeans clustering where each color is representative of one cluster.

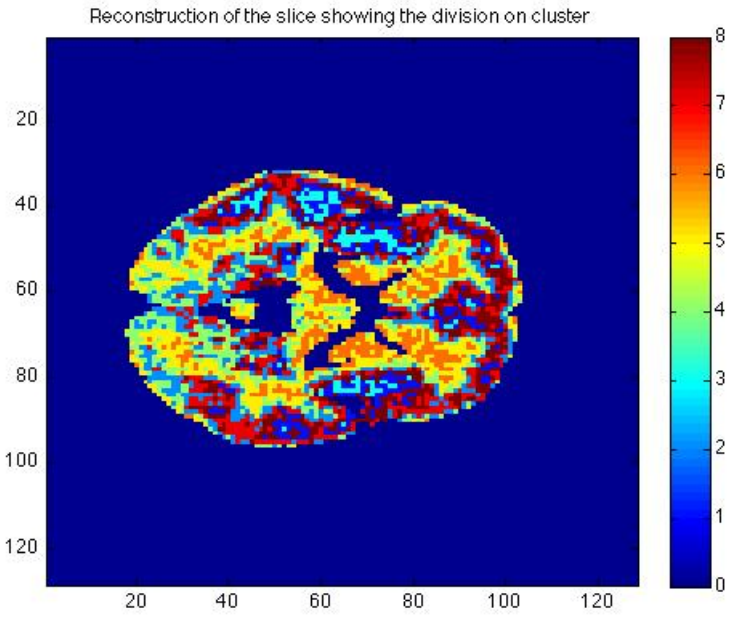


Figure 3.1: Slice 45: Image showing the 8 different clusters

Once we obtained the TACs that correspond to the centroid of the eight clusters, we implemented the nonlinear SA (explained in the previous chapter, section 2.2) in order to obtain parameters that better describe those curves. The method tested all the six models previously presented and then, based on parsimony criterion (AIC), best fit and best coefficient of the estimates, we decided which one gave the best result for each curve. With the b estimated and their standard deviations (SD) we built 8 new grids of 100 elements: the idea for constructing each grid was to generate two Gaussian distribution, one with mean equal to the value of b_1 specific of the cluster under study and standard deviation equal to its SD, composed by 50 elements. The second with mean equal to the value of b_2 and SD equal to that obtained for the parameter on that cluster, with the remaining 50 elements (the example for one cluster is shown in Figure 3.2).

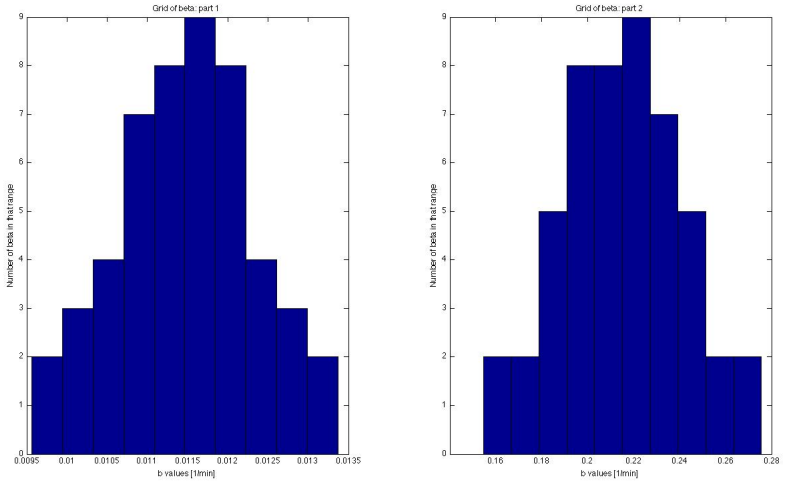


Figure 3.2: Grid of b obtained for one specific cluster

In this way grids were generated densely distributed around the values that should gave the best fit of the data. The procedure used after obtaining these grids was to evaluate each voxel within the mask with the linear SA (explained in the former chapter, section 2.1), using the grid correspondent to the cluster of belonging of the voxel and so obtaining the parameters V_b, a_j, b_j which best describe its curve of activity.

3.3 SA with prior from voxel

Still interested in creating a grid that is more specific for the problem under study, we examined the distribution in the whole brain of the chosen b, previously obtained with linear SA referring to a grid with logarithmic distribution (instead of the DiStefano distribution used in the previous method). So, a code was implemented that counted the number of times that a specific value of b was found to have a nonzero value of a, and that was developed for each voxel within the mask. We obtained the histogram in Figure 3.3 (removing the first and the last b value for reasons explained in chapter 2 where we addressed $b \to \infty$ and $b \to 0$ Eq. (2.5) and Eq. (2.6)), where it is easy to recognize two Gaussian distributions. The concept under review was to extrapolate the information about mean and standard deviation of the two Gaussians directly from the histogram obtained.

This goal was achieved thanks to the Matlab tool *cftool* (Figure 3.4). So we exported the data in this tool and fitted them with a Gaussian curve that

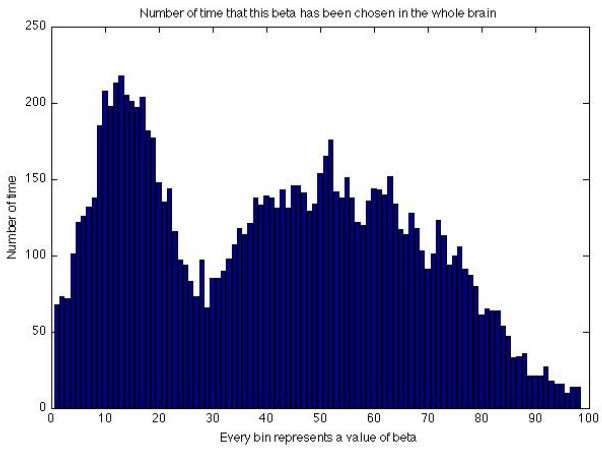


Figure 3.3: Distribution of b in the wholebrain

is best described as a sum of two terms:

$$a_1 * e^{\left(-\left(\frac{(x-b_1)}{c_1}\right)^2\right)} + a_2 * e^{\left(-\left(\frac{(x-b_2)}{c_2}\right)^2\right)} \tag{3.9}$$

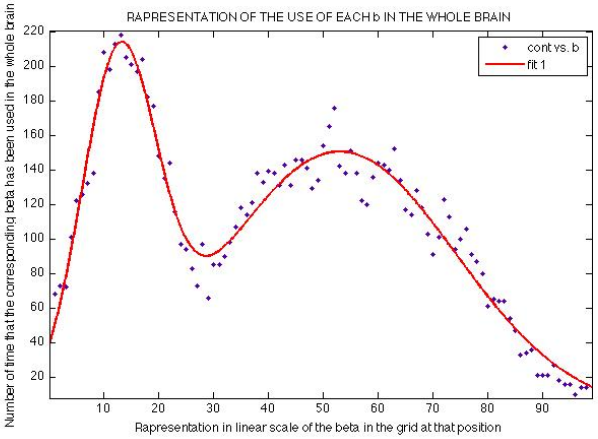


Figure 3.4: Fit of the Gaussian distributions

Knowing that a normal (or Gaussian) distribution is a continuous probability distribution, that has a bell-shaped probability density function, described as follow:

$$f(x : \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.10}$$

where the parameter μ is the mean and σ^2 is the variance, we derived the mean and the SD of our two curves with those equations:

$$\begin{cases} \mu_1 = b_1 \\ \sigma_1 = \frac{c_1}{\sqrt{2}} \end{cases} \qquad \begin{cases} \mu_2 = b_2 \\ \sigma_2 = \frac{c_2}{\sqrt{2}} \end{cases} \tag{3.11}$$

The observation required is that these histograms have on the x-axis a linear scale, where each number represents the position of the correspondent b in the grid. Due to this reason values of mean and SD of the two Gaussians were obtained in terms of positions in the grid; the next step was to correlate these values with the correspondent values of b, then obtaining the consequent lognormal distribution following these steps:

1. We refer to the aleatory variable as "position" so described: $pos \in N(\mu, \sigma)$ and we describe b as $b = i \cdot e^{j \cdot pos}$: we can obtain the system:

$$\begin{cases} b_{min} = i \cdot e^{j \cdot pos_{min}} = i \cdot e^{j \cdot 1} = i \cdot e^{j} = 0.005 \\ b_{max} = i \cdot e^{j \cdot pos_{max}} = i \cdot e^{j \cdot 100} = i \cdot e^{100j} = 10 \end{cases} \rightarrow \begin{cases} i = 0.0046 \\ j = 0.0768 \end{cases} \tag{3.12}$$

   where the values 0.005 and 10 are the lower and the upper bounds, respectively, of the logarithmic distribution.

2. We call $x = j \cdot pos$ and we define its expectation and its variance as: $E[j \cdot pos] = \mu_x$, $var[j \cdot pos] = \sigma_x^2$; if we elaborate this equations, remembering that pos is an aleatory variable, we obtain:

$$\mu_x = j \cdot E[pos] = j \cdot \mu_{pos} \tag{3.13}$$

$$\begin{aligned} \sigma_x &= \sqrt{E[(x - \mu_x)^2]} = \sqrt{E[x^2 - 2\mu_x x + \mu_x^2]} \\ &= \sqrt{E[x^2] - 2\mu_x E[x] + \mu_x^2} = j \cdot \sigma_{pos} \end{aligned} \tag{3.14}$$

3. Now that we know the mean and SD of x we can easily see that defining $b = i \cdot e^{j \cdot pos} = i \cdot e^{x}$ as the curve that represents the distribution of the b back to the logarithmic grid, it is possible to calculate these distributions just passing to the Matlab function *lognrnd* μ and σ, that are the mean and standard deviation, respectively, of the associated normal distributions. Thus, the function returns an array of random numbers (as many as desired; we sought 50 elements for each curve) generated from the lognormal distribution. We used this approach for both curves, obtaining a final grid of 100 elements composed of two lognormal distributions; the histogram of the final distribution is shown in Figure 3.5.

Once this new grid was generated, we implemented the linear SA, obtaining the parameters V_b, a_j, b_j and all the variables required to compare the different methods.

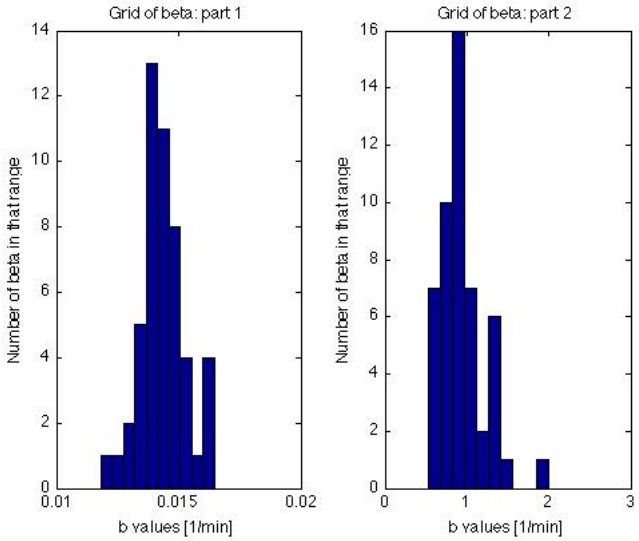


Figure 3.5: Grid of b composed by two lognormal distributions

3.4 SA using an empirical Bayesian estimation approach

In the fourth method, it was decided to change procedure, so we used compartmental analysis and identification of the model obtained using an empirical Bayesian estimation approach. A Bayesian estimator considers as known the *a priori* probability of the parameters and then, proceeding with the implementation, it updates the prior probability using the measured data. In this approach, the prior was extrapolated from the nonlinear SA executed on the eight different ROI; thus, using the parameters estimated for each cluster and their SD, we created the information *a priori* for the vectors of the parameters $p_1 \in N(\mu_{p_1}, \Sigma_{p_1})$, $p_2 \in N(\mu_{p_2}, \Sigma_{p_2})$, etc.. as:

$$\mu_{p_1} = \begin{bmatrix} \mu_{a_1} \\ \mu_{b_1} \end{bmatrix} \quad \Sigma_{p_1} = \begin{bmatrix} \sigma^2_{a_1} & 0 \\ 0 & \sigma^2_{b_1} \end{bmatrix} \tag{3.15}$$

$$\mu_{p_2} = \begin{bmatrix} \mu_{a_2} \\ \mu_{b_2} \end{bmatrix} \quad \Sigma_{p_2} = \begin{bmatrix} \sigma^2_{a_2} & 0 \\ 0 & \sigma^2_{b_2} \end{bmatrix} \tag{3.16}$$

The procedure follows these steps:

- The linear SA was implement obtaining the vector of parameters a, while the b grid was obtained with the DiStefano distribution.
- Observing that our data are describable with two components, was decided to identify the information about the cutoff b, i.e. the b value

that represents the point from which lower b values can be associated to the first component while higher b values can be associated to the second component. This was done by exploiting the curve of Figure 3.4 and identifying the value of b_{cutoff} as the point of minimum between the two Gaussians. So, once its value was obtained, it has been used to reconstruct the contribution of the two components as follows:

$$C_{T_1}(t) = \sum_{j=1}^{b_{cutoff}} a_j \cdot e^{-b_j t} \otimes C_p(t) \tag{3.17}$$

$$C_{T_2}(t) = \sum_{j=b_{cutoff}+1}^{N} a_j \cdot e^{-b_j t} \otimes C_p(t) \tag{3.18}$$

- Deciding to consider and estimate the two component (C_{T_1} and C_{T_2}) separately, one sees that to describe the kinetic behavior of each one, the model with one exponential of equation (2.15) can be used (a schematic representation is showed in Figure 3.6) where there are the arterial

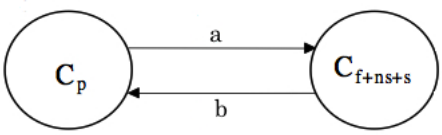


Figure 3.6: Model describable with one exponential

plasma compartment (C_p) and a tissue compartment (C_T) that incorporated the free fraction plus a non specifically bound fraction and a specific bound fraction. The equation that describes this kinetic behavior is:

$$\dot{C}_T(t) = aC_p(t) - bC_T(t) \quad C_T(t) = 0 \tag{3.19}$$

where $\dot{C}_T$ is the first time derivative of the tissue compartment; the measurement equation is given by:

$$C_i(t) = (1 - V_b)C_T(t) + V_b C_b(t) \tag{3.20}$$

where V_b is the vascular volume in the ROI, C_b is the concentration of the tracer in the whole blood, and a and b are the constant rates for transport of the tracer from plasma to tissue and back, respectively. Since the data were affected by a measurement error, also in this case it has been considered additive, uncorrelated, coming from a Gaussian distribution with zero mean and a matrix of covariance obtained as the inverse of the weights vector. It has been previously explained why, when

the analysis is conducted at voxel level, it is appropriate to use a linear approach, thus we linearize the model, obtaining the following equation:

$$C_T(t) = a \int_0^t C_p(\tau)\, d\tau - b \int_0^t C_T(\tau)\, d\tau \tag{3.21}$$

After these considerations, it was decided to use a MAP (Maximum A Posteriori) estimator, which was the preferred choice since it allowed us to take account of the bias introduced by the linearization, for obtaining the parameters (we implemented the MAP two times, once for each component). The cost function to be minimized was:

$$\mathbf{J}(\mathbf{p}) = [\mathbf{C}_T - \mathbf{G}(\mathbf{p})]^T \Sigma_v^{-1} [\mathbf{C}_T - \mathbf{G}(p)] + (\mathbf{p} - \mu_p)^T \Sigma_p^{-1} (\mathbf{p} - \mu_p) \tag{3.22}$$

where $\mathbf{C}_T$ is the vector obtained before for each component (respectively Eq. (3.17) and Eq. (3.18)), $\mathbf{G}(\mathbf{p})$ is the predicted model output, $\mathbf{\Sigma}_v$ is the covariance matrix of the measurement error and $\mathbf{p}$ is the unknown parameter vector that has a prior normally distributed with mean $\boldsymbol{\mu}_p$ and matrix of covariance $\mathbf{\Sigma}_p$. The information *a priori* is updated from the observed data, generating the *a posteriori* probability density function, that is the probability density function of the parameters given the data. The estimate has an expression that is analytically solvable as:

$$\hat{p}_{MAP} = \left(G^T \Sigma_v^{-1} G + \Sigma_p^{-1}\right)^{-1} \left(G^T \Sigma_v^{-1} C_T + \Sigma_p^{-1} \mu_p\right) \tag{3.23}$$

where G is defined as:

$$G = \begin{bmatrix} \int_0^{t_1} C_p(\tau)\, d\tau & -\int_0^{t_1} C_T(\tau)\, d\tau \\ \int_0^{t_2} C_p(\tau)\, d\tau & -\int_0^{t_2} C_T(\tau)\, d\tau \\ \vdots & \vdots \\ \int_0^{t_M} C_p(\tau)\, d\tau & -\int_0^{t_M} C_T(\tau)\, d\tau \end{bmatrix} \tag{3.24}$$

where M is the number of the tissue measures, (Rizzo et al., 2012).

- Once the MAP estimation for both components was executed , we derived the total concentration of the tracer, for each voxel, as sum of the contributions:

$$C_{T_{tot}} = G_1 \cdot \hat{p}_{MAP_1} + G_2 \cdot \hat{p}_{MAP_2} \tag{3.25}$$

remembering that the blood component was considered negligible at voxel level. We were also able to assess the CV of the estimates remembering that for linear models the covariance of the estimates has this expression:

$$cov(\hat{p}_{MAP}) = \left(G^T \Sigma_v^{-1} G + \Sigma_p^{-1}\right)^{-1} \tag{3.26}$$

with which we can obtain:

$$CV = 100 * \frac{\sqrt{cov(\hat{p}_{MAP})}}{\hat{p}_{MAP}} \tag{3.27}$$

Chapter 4

Data Set

In the first chapter of this work we described the general principles and functions of Positron Emission Tomography explaining how and why it is such an important imaging system; one of its main applications is that it allows the illustration of some particular brain functions. Depending on which functions the investigator is interested in, it is necessary to utilize different kinds of tracers: for example $\left[^{18}F\right]$FDG is commonly used to obtain the regional metabolic rate of glucose by estimating the tissue fractional uptake of $\left[^{18}F\right]$FDG and then by using a scale factor between it and glucose metabolism (Bertoldo et al., 1998). There is then the $[^{15}O]H_2O$ that has been used to estimate the cerebral blood flow, and so on, with many other tracers that are used to investigate different activities. One of these is the $\left[^{11}C\right]$WAY100635 that permits evaluation of specific receptors implicated in the pathophysiology of neuropsychiatric conditions and that we are going to describe in more detail since is the tracer used in this work.

4.1 Tracer

We start this dissertation by presenting the theory concerning the receptors to which the radiotracer is capable of binding; these receptors are specific for the serotonin ($5\text{–}HT$) considered a modulatory neurotransmitter with general inhibitory effects and with an important role in the regulation of psychobiological processes. Serotonin projections, arising from the dorsal and median raphe nuclei, innervate virtually all regions of the brain (Steinbusch, 1981). There are seven known receptors, one of which is the $5\text{–}HT_{1A}$ that is a somatodendritic autoreceptor in the raphe nuclei and a postsynaptic receptor in the neocortex, hippocampus and other limbie structures (Wright et al., 1995). This type of receptors is of central interest in research on the pathophysiology and treatment of psychiatric disorders, for example they are thought to be involved in phenomena as anxiety, depression, dementia and schizophrenia (Gurevich and Joyce, 1997).

Initially, WAY100635 (a selective antagonist with affinity and selectivity for serotonin 5–HT_{1A} receptors) was labeled in the methoxy position with carbon−11 obtaining $\left[O-methyl-^{11}C\right]$WAY100635 that successfully delineated 5–HT_{1A} receptors. This compound, however, rapidly forms labeled metabolites which cross the BBB (Blood Brain Barrier) (Osman et al., 1996); to solve this problem, WAY100635 was labeled in the carbonyl position obtaining $\left[carbonyl-^{11}C\right]$ WAY100635 (from now on we will refer to it just with $\left[^{11}C\right]$ WAY100635) which showed a strong signal in humans, consistent with reduced uptake of labelled metabolites in the brain (Lammertsma et al., 1996).

4.1.1 PET studies

A dataset of four healthy male subjects was made available by the Imperial College. Each subject underwent a 90-min dynamic PET study in a CTI ECAT EXACT3D tomography after a bolus injection $\left[^{11}C\right]$WAY100635. All PET data were acquired in 3-dimensional mode, corrected for attenuation, detector efficiency, random events, and scatter, and reconstructed into tomographic images using filtered back-projection. Acquisition was performed in list-mode (event by event) and scans were rebinned into 23 time frames of increasing duration (two variable length background frame, 3x5s, 2x15s, 4x60s, 7x300s, 5x600s). The reconstructed voxel sizes were 2.096 x 2.096 x 2.43 mm^3. After injection of $\left[^{11}C\right]$WAY100635, the radioactivity concentration in blood was measured continuously and, in addition, serial discrete blood samples were taken at increasing time intervals throughout the study for the measurement of the radioactivity in blood and plasma. Nine of these samples were also used for quantification of the fraction of radioactivity attributable to unmetabolized parent radiotracer, generating the metabolite-corrected arterial plasma input function.

4.2 Image

For the subject under examination, PET image was imported in Matlab®'s environment, obtaining a 4D matrix, 128x128x95x23 where 95 is the number of slices obtained by the PET scan and 23 are the time frames of increasing duration. Since the image was already corrected for movement we applied just a decay correction in each region of interest, generating the tissue-time activity curves (TACs); this was done using the formula:

$$A_0 = A_t \times e^{\lambda t} \tag{4.1}$$

where A_t and A_0 are respectively the uncorrected and corrected value of concentration at time t, which is expressed in minute, $\lambda = \frac{ln2}{T_{1/2}}$ with $T_{1/2}$ half-life of the radioactive isotope (in our case $20.4min$). All blood signals were corrected for the delay, which is caused by the difference in time between the

tracer arrival in the brain and the arterial sampling site. The input function, moreover, was fitted using a sum of three exponentials from the time of the peak to the last data point, whereas the rising part of the curve was fitted as a straight line between the first point and the peak. After that we selected the voxel of interest by applying a mask (see the section below for details) (Parsey et al., 2000).

In Figure 4.1 we show in the left part the total blood curve ($C_b(t)$) and the parent plasma curve ($C_p(t)$), which are corrected for delay and decay as said before, while in the right part there are the decay corrected ROI TACs obtained by cluster analysis for a representative subject (subject 1).

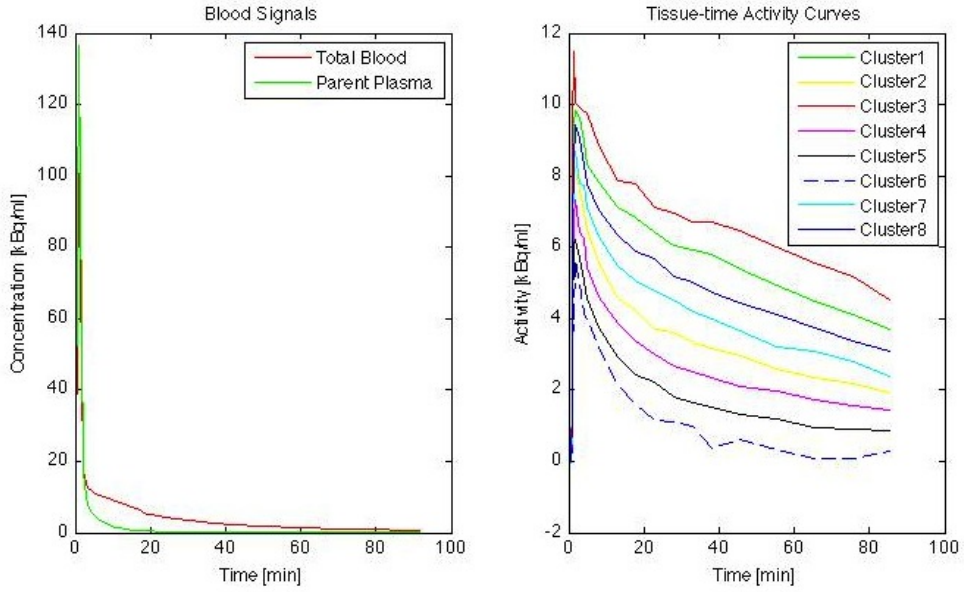


Figure 4.1: Left column: blood and unmetabolized plasma activity curves. Right column: tissue-time activity curves for 8 ROIs

4.2.1 Mask

In order to eliminate the noise present in the image, we applied a binary mask (i.e. the mask assumes the value 0 on the pixels that need to be eliminated because represent noise, and assumes value 1 on all the other pixel); the mask used in this work has been taken from the anatomical atlas made available by the PET centre, Division of Experimental Medicine, of the Imperial College of London (Hammers et al., 2003)

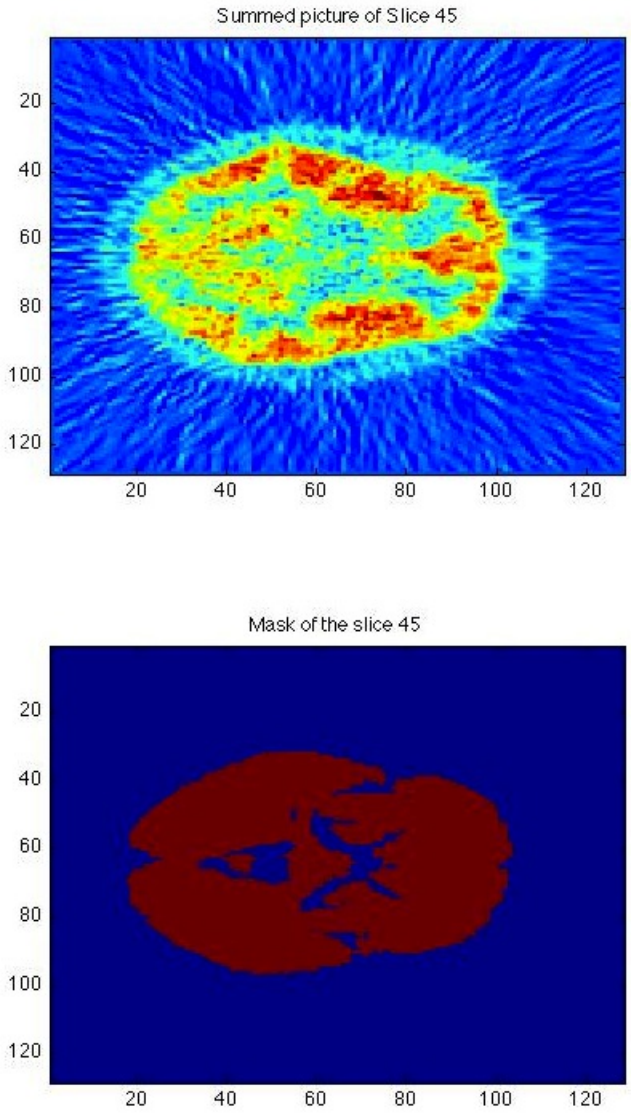


Figure 4.2: Summed image of slice 45 and the mask that has been applied to it, for a representative subject

Chapter 5

Results

This chapter presents the results of all the new proposed methods; it is then necessary to compare these results with those obtained by the existent methods (presented in Chapter 2) in order to verify if they are physiologically corrected and if they reached the established goal of improving the quantification at voxel level. As terms of comparisons are used the non linear SA in terms of the macro-parameter of interest, the distribution volume, and the results obtained by the linear SA in terms of fit, residuals, WRSS and parameters estimated. The results obtained for these methods will be firstly described separately using a representative voxel for each cluster, and then comparing them all in terms of medium values within the clusters.

All the methods will be analyzed separately but, since the same vector of weights has been used in all methods, its trend is now introduced. It has been calculated following the equation defined in Chapter 2 Eq. (2.20) but using as $C_i^{obs}(t_j)$ the description of the tissue activity curve obtained as mean of all those within the mask. Moreover its first four elements have been forced at the zero value. In Figure 5.1 the trend of the weights is shown, before and after modification.

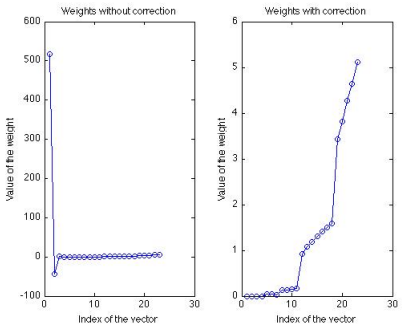
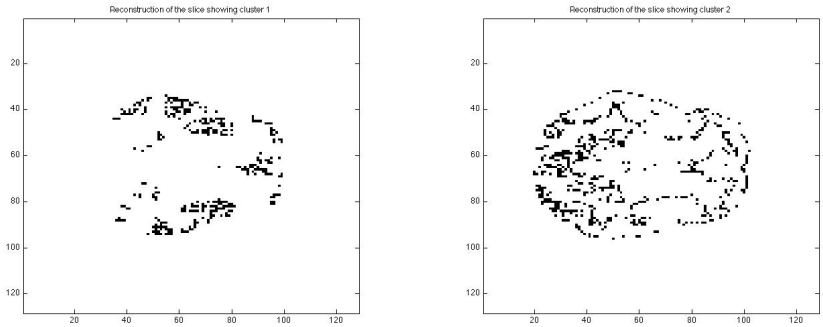
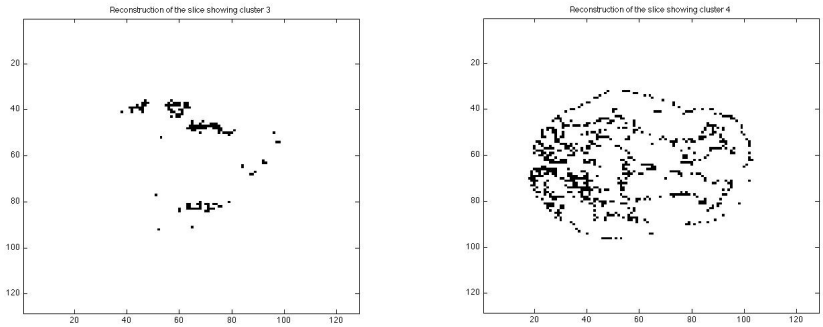


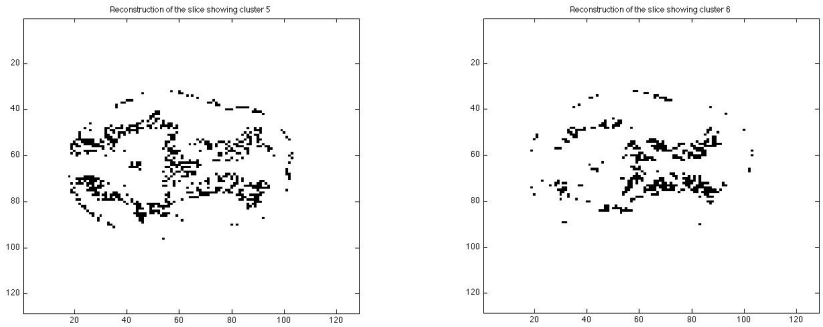
Figure 5.1: Vector of weights: left picture, weights without correction and right picture, weights with correction



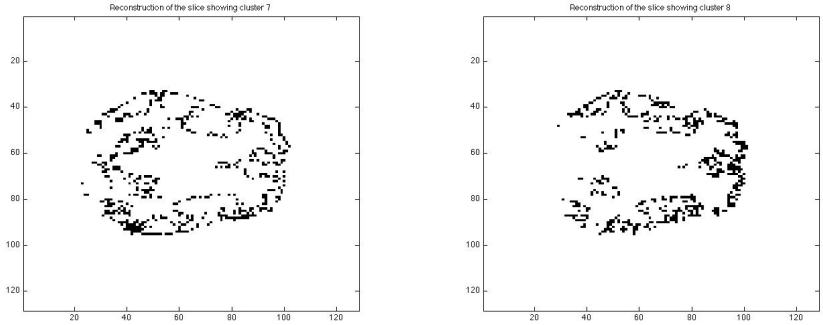
(a) *Cluster 1.* (b) *Cluster 2.*



(c) *Cluster 3.* (d) *Cluster 4.*



(e) *Cluster 5.* (f) *Cluster 6.*



(g) *Cluster 7.* (h) *Cluster 8.*

Figure 5.2: Results of clustering, for a representative subject

Figure 5.2 shows the eight clusters obtained with the *kmeans* tool; this allows association with the voxel, used as a reference for showing the results obtained for each method, with the anatomical region to which it belongs to.

5.1 I/O model results

5.1.1 Non linear SA

During this study the non linear SA, both at voxel level and at ROI level, was implemented, but for different reasons. The first one has been used as a reference to understand whether the new methods implemented were correct, in terms of signal amplitudes, referring specifically to the volume of distribution. The second one has been used for creating a new b grid developed starting from the values of its estimates, as already discussed in Chapter 3 (see Section 3.2. The implementation of the non linear SA has been executed for models of increasing order in both cases; then, using criteria of parsimony (Akaike Criterion), looking to the fits and to the precision of the estimates, a comparison has been made in order to understand which was the best model to describe our data. Both at ROI level and at voxel level, the NLSA identify model of Equation (2.2.1) as the best model. The following is a presentation of the results achieved following the steps presented in the previous chapters.

1. *Voxel analysis*

   The map of the volume of distribution obtained with the non linear analysis at voxel level is shown in Figure 5.3

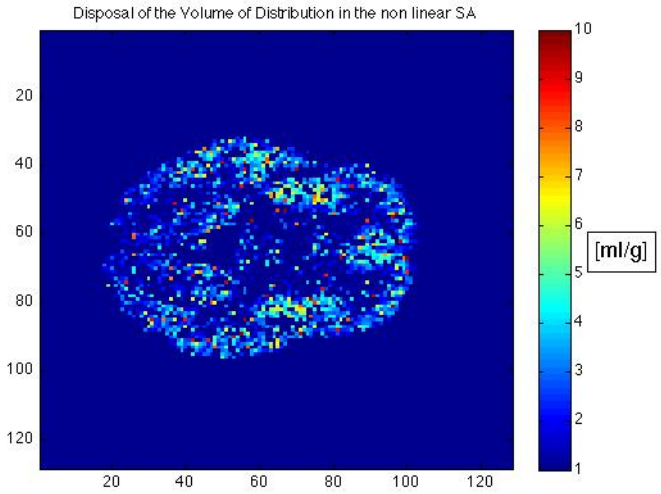
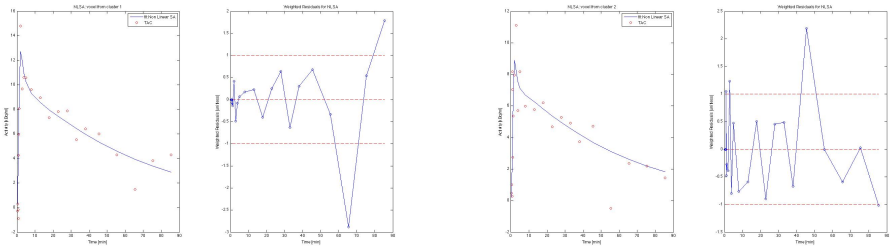


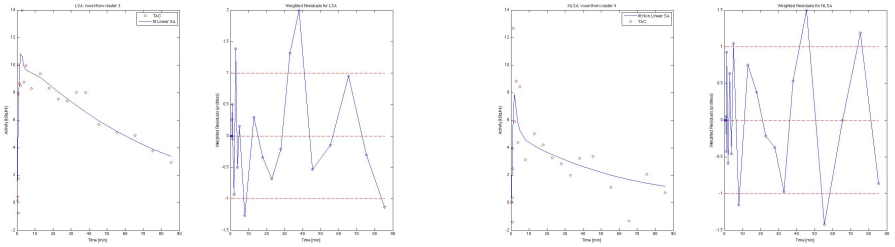
Figure 5.3: Slice 45: Image showing the volume of distribution, obtained with NLSA at voxel level

Figure 5.4 presents some fits and their relative weighted residuals obtained with non linear SA applied at voxel region. It was decided to represent one voxel belonging to each clusters. This leads to a better understanding of the results obtained.



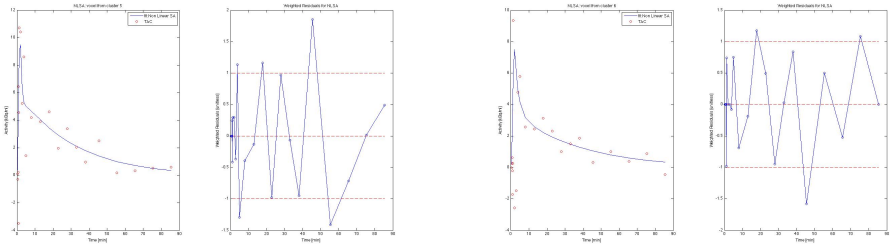
(a) *Fit vs Data for a voxel of Cluster 1.*

(b) *Fit vs Data for a voxel of Cluster 2.*



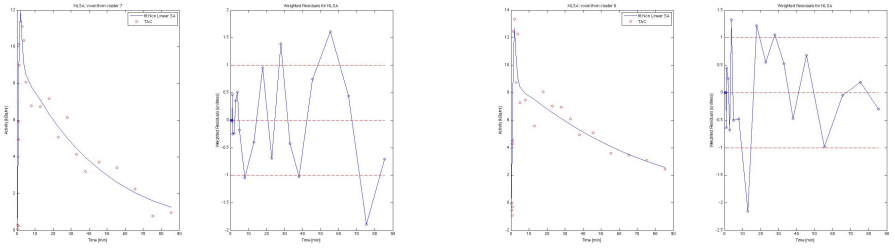
(c) *Fit vs Data for a voxel of Cluster 3.*

(d) *Fit vs Data for a voxel of Cluster 4.*



(e) *Fit vs Data for a voxel of Cluster 5.*

(f) *Fit vs Data for a voxel of Cluster 6.*



(g) *Fit vs Data for a voxel of Cluster 7.*

(h) *Fit vs Data for a voxel of Cluster 8.*

Figure 5.4: Model estimated curves with NLSA applied at voxel level and relative weighted residuals

In order to compare the results obtained with the different methods, it is useful to present other variables of interest. Those chosen for this study were formally presented in previous chapters, where the calculations made to determine their value are also presented. They are: V_d Eq. (2.12), K_1 Eq. (2.11), AIC Eq. (2.28) and WRSS Eq. (2.22). In this chapter they will be presented (Table 5.1) referring to each voxel chosen as representative of its cluster of belonging, while in the next chapter they will be presented as mean value of all voxel of each cluster. The units of the variables presented are $V_d[ml\ g^{-1}], K_1[ml\ g^{-1}\ min^{-1}], AIC[unitless]$ and $WRSS[unitless]$.

Belonging cluster	V_d $[ml\ g^{-1}]$	K_1 $[ml\ g^{-1}\ min^{-1}]$	AIC $[unitless]$	$WRSS$ $[unitless]$
1	3.227	0.104	97	44
2	2.127	0.077	63	10
3	3.694	0.075	67	12
4	1.414	0.069	78	19
5	1.279	0.043	97	45
6	0.665	0.071	56	7
7	1.888	0.107	69	13
8	2.855	0.128	66	11

Table 5.1: Table of variables calculated for one voxel representative of each cluster, obtained with NLSA at voxel level

For the same voxels for which have been presented fits and weighted residuals (Figure 5.4), is shown the value of the parameters estimated, presented with the following units $V_b[unitless], a_1\ and\ a_2[ml\ g^{-1}\ min^{-1}]$, $b_1\ and\ b_2[min^{-1}]$ in Table 5.2.

Belonging cluster	V_b $[unitless]$	a_1 $[ml\ g^{-1}\ min^{-1}]$	b_1 $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	b_2 $[min^{-1}]$
1	0.002	0.050	0.016	0.054	0.509
2	0	0.038	0.018	0.039	0.694
3	0.035	0.054	0.015	0.022	0.696
4	0.008	0.024	0.018	0.045	0.619
5	0.022	0.028	0.023	0.015	0.260
6	0	0.017	0.029	0.054	0.583
7	0.019	0.046	0.025	0.061	0.809
8	0	0.044	0.016	0.084	0.925

Table 5.2: Table of estimated parameters for one voxel representative of each cluster, obtained with NLSA at voxel level

2. *ROI analysis*

   The value of the parameters estimated for each cluster at ROI level, that generate the best fit for the data, are presented with the following units $V_b[unitless], a_1$ *and* $a_2[ml\ g^{-1}\ min^{-1}], b_1$ *and* $b_2[min^{-1}]$ in Table 5.3. For the parameters b_1 and b_2 are shown also their precision of estimates, since these are the values used in one of the new approaches proposed in this study (see Section 3.2).

Cluster	V_b $[unitless]$	a_1 $[ml\ g^{-1}\ min^{-1}]$	$b_1 \pm \sigma$ $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	$b_2 \pm \sigma$ $[min^{-1}]$
1	0.017	0.038	0.0092 ± 0.0003	0.036	0.358 ± 0.051
2	0.020	0.022	0.0112 ± 0.0006	0.036	0.264 ± 0.031
3	0.036	0.042	0.0074 ± 0.0005	0.029	0.284 ± 0.086
4	0.019	0.017	0.0115 ± 0.0009	0.031	0.214 ± 0.027
5	0.015	0.012	0.0153 ± 0.0014	0.029	0.209 ± 0.028
6	0.003	0.015	0.0466 ± 0.0072	0.027	0.284 ± 0.097
7	0.018	0.028	0.0109 ± 0.0005	0.033	0.283 ± 0.043
8	0.017	0.033	0.0098 ± 0.0003	0.035	0.301 ± 0.033

Table 5.3: Table of estimated parameters for each cluster, obtained with NLSA at ROI level

5.1.2 Linear SA

The map of the volume of distribution obtained with the linear analysis, the standard one presented on Chapter 2, is shown in Figure 5.5.

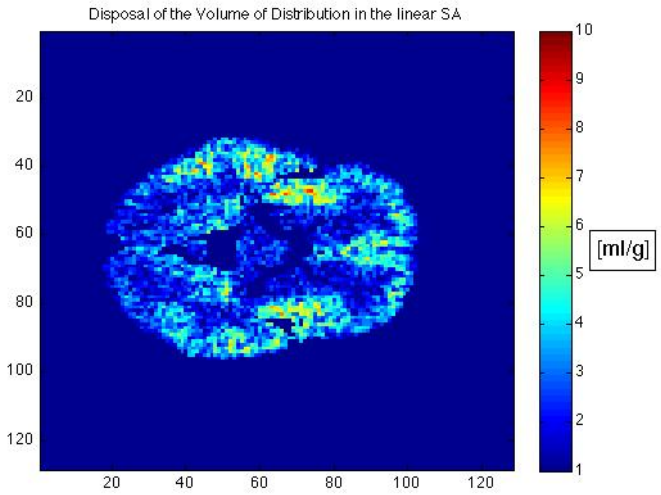


Figure 5.5: Slice 45: Image showing the volume of distribution, obtained with LSA

The value of the parameters estimated for each representative voxel is presented with the following units $V_b[unitless], a_1, a_2, a_3, a_4$ *and* $a_5[ml\ g^{-1}\ min^{-1}]$, b_1, b_2, b_3, b_4 *and* $b_5[min^{-1}]$ in Table 5.4 on the next page. The reason why there are up to five components is that this method presents a problem called "line doubling": some couples of b_j are actually next to each other on the grid. The poor precision of the a_j stems from the fact that the approach cannot attribute to a certain b_j the right a_j, and so it divides the a_j between two b_j next to each other.

Belonging cluster	V_b $[unitless]$	a_1 $[ml\ g^{-1}\ min^{-1}]$	b_1 $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	b_2 $[min^{-1}]$	a_3 $[ml\ g^{-1}\ min^{-1}]$	b_3 $[min^{-1}]$
1	0	0.021	0.006	0.029	0.038	0.008	0.041
2	0	0.014	0.017	0.024	0.018	0.051	0.804
3	0.016	0.042	0.015	0.011	0.016	0.001	0.964
4	0.001	0.008	0.017	0.017	0.019	0.058	0.712
5	0.027	0.021	0.022	0.007	0.024	0.018	0.269
6	0	0.003	0.006	0.016	0.049	0.020	0.630
7	0	0.041	0.025	0.005	0.027	0.094	1.025
8	0	0.028	0.015	0.016	0.016	0.092	0.964

Belonging cluster	a_4 $[ml\ g^{-1}\ min^{-1}]$	b_4 $[min^{-1}]$	a_5 $[ml\ g^{-1}\ min^{-1}]$	b_5 $[min^{-1}]$
1	0.076	1.157	0.002	1.230
2	–	–	–	–
3	0.054	1.025	–	–
4	–	–	–	–
5	–	–	–	–
6	0.029	0.670	–	–
7	0.007	2.399	–	–
8	0.006	1.025	–	–

Table 5.4: Table of estimated parameters for each representative voxel, obtained with LSA

The variables of interest are presented in Table 5.5 and their units are $V_d[ml\ g^{-1}], K_1[ml\ g^{-1}\ min^{-1}], AIC[unitless]\ and\ WRSS[unitless]$.

Belonging cluster	V_d $[ml\ g^{-1}]$	K_1 $[ml\ g^{-1}\ min^{-1}]$	AIC $[unitless]$	$WRSS$ $[unitless]$
1	4.699	0.136	95	41
2	2.125	0.089	57	9
3	3.665	0.108	65	11
4	1.422	0.082	76	19
5	1.288	0.045	95	45
6	0.961	0.068	54	7
7	1.897	0.147	66	12
8	2.868	0.142	65	12

Table 5.5: Table of variables calculated for one voxel representative of each cluster, obtained with LSA

The linear Spectral Analysis implemented with the standard method make use of a grid based on the DiStefano distribution (2.4) and gives this histogram showed in Figure 5.6:

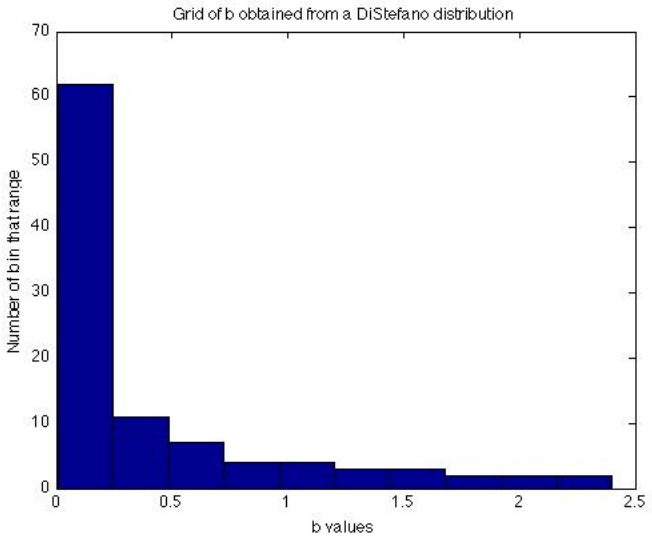
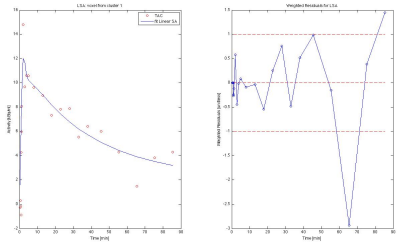
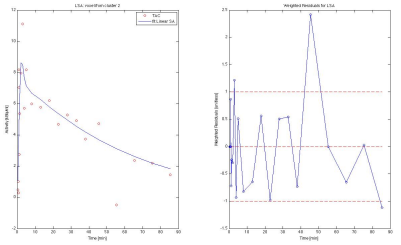


Figure 5.6: Image showing the DiStefano distribution of the b in the grid

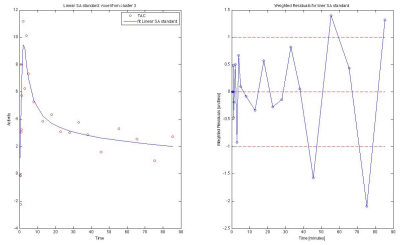
In Figure 5.7 the fits of the same voxel selected previously, as representative of each cluster, and their relative weighted residual are presented.



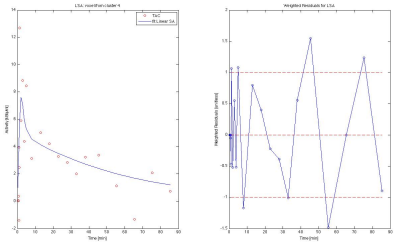
(a) *Fit vs Data for a voxel of Cluster 1.*



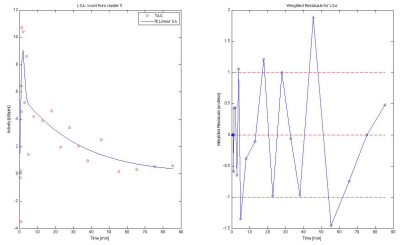
(b) *Fit vs Data for a voxel of Cluster 2.*



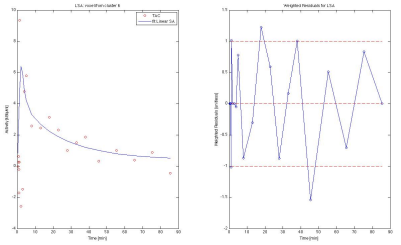
(c) *Fit vs Data for a voxel of Cluster 3.*



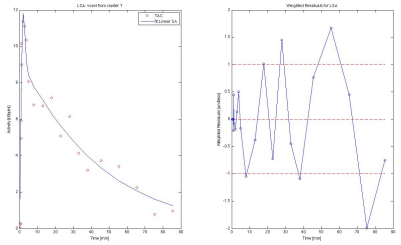
(d) *Fit vs Data for a voxel of Cluster 4.*



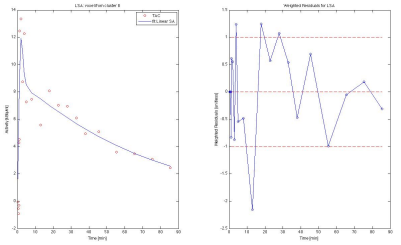
(e) *Fit vs Data for a voxel of Cluster 5.*



(f) *Fit vs Data for a voxel of Cluster 6.*



(g) *Fit vs Data for a voxel of Cluster 7.*



(h) *Fit vs Data for a voxel of Cluster 8.*

Figure 5.7: Model estimated curves with linear SA and relative weighted residuals

5.1.3 Linear SA with prior from ROI

The map of the volume of distribution, obtained with the linear analysis that uses the prior from ROI, is shown in Figure 5.8.

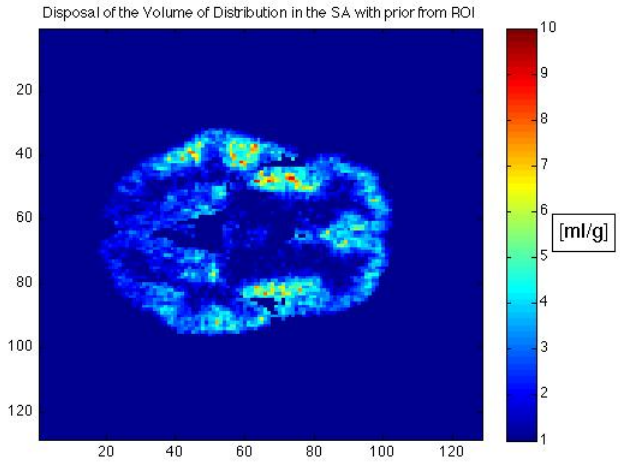


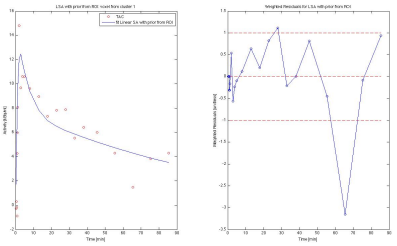
Figure 5.8: Slice 45: Image showing the volume of distribution, obtained with LSA with prior from ROI

The variables of interest are presented in Table 5.6 and their units are $V_d[ml\ g^{-1}], K_1[ml\ g^{-1}\ min^{-1}], AIC[unitless]\ and\ WRSS[unitless]$

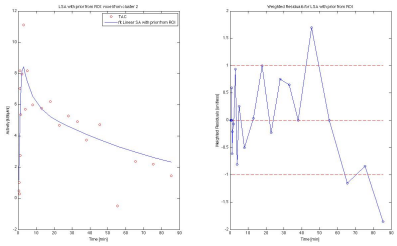
Belonging cluster	V_d $[ml\ g^{-1}]$	K_1 $[ml\ g^{-1}\ min^{-1}]$	AIC $[unitless]$	$WRSS$ $[unitless]$
1	4.116	0.094	97	52
2	2.523	0.065	70	15
3	4.872	0.064	77	22
4	1.596	0.049	78	21
5	1.387	0.047	95	45
6	0.673	0.061	54	8
7	2.454	0.075	85	32
8	3.531	0.089	75	20

Table 5.6: Table of variables calculated for one voxel representative of each cluster, obtained with LSA with prior from ROI

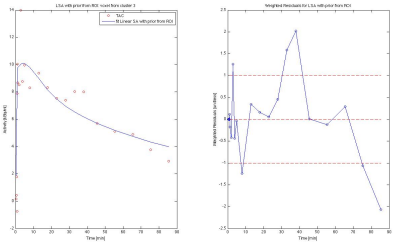
In Figure 5.9 we present the fits of the same voxel selected previously as representative of each cluster and their relative weighted residual.



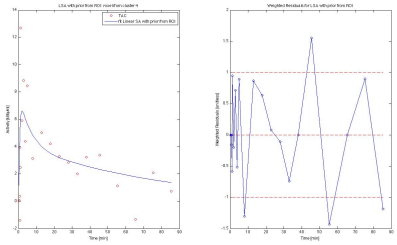
(a) *Fit vs Data for a voxel of Cluster 1.*



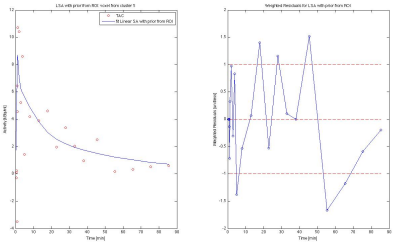
(b) *Fit vs Data for a voxel of Cluster 2.*



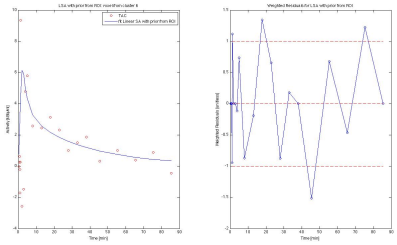
(c) *Fit vs Data for a voxel of Cluster 3.*



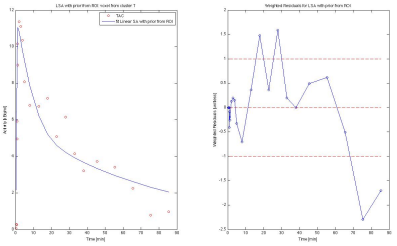
(d) *Fit vs Data for a voxel of Cluster 4.*



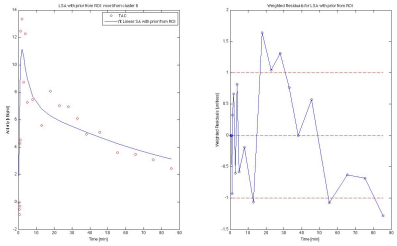
(e) *Fit vs Data for a voxel of Cluster 5.*



(f) *Fit vs Data for a voxel of Cluster 6.*



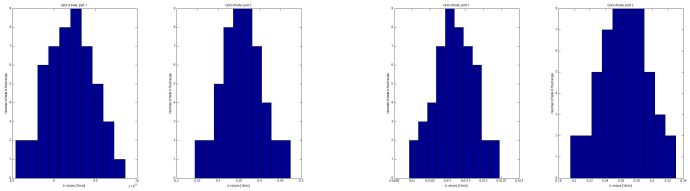
(g) *Fit vs Data for a voxel of Cluster 7.*



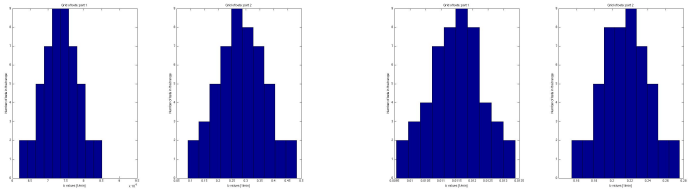
(h) *Fit vs Data for a voxel of Cluster 8.*

Figure 5.9: Model estimated curves with LSA obtained with prior from ROI and relative weighted residuals

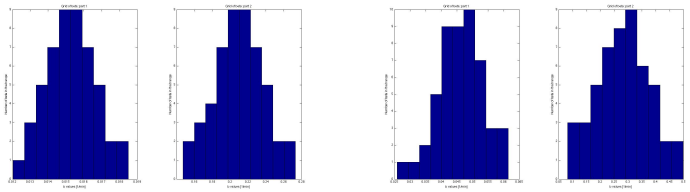
The linear Spectral Analysis implemented with the prior obtained by the information extracted from the non linear analysis made at ROI level (as described in Section 3.2), makes use of a grid specific for each cluster where the histograms are those showed in Figure 5.10:



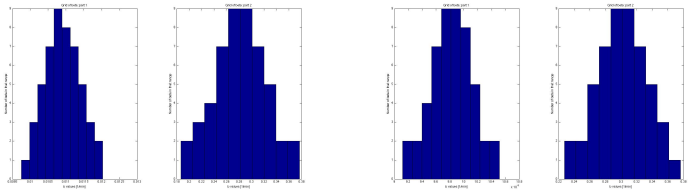
(a) *Distribution of b for cluster 1.* (b) *Distribution of b for cluster 2.*



(c) *Distribution of b for cluster 3.* (d) *Distribution of b for cluster 4.*



(e) *Distribution of b for cluster 5.* (f) *Distribution of b for cluster 6.*



(g) *Distribution of b for cluster 7.* (h) *Distribution of b for cluster 8.*

Figure 5.10: Histograms of the 8 different grids obtained one for each cluster in the LSA method with prior from ROI

The value of the parameters estimated for each representative voxel is presented with the following units $V_b[unitless], a_1, a_2, a_3[ml\ g^{-1}\ min^{-1}], b_1, b_2, b_3[min^{-1}]$ in Table 5.7 presented on the next page. In this case the number of exponential components are less then those found with the linear SA and, looking at the values obtained, it seems like if the problem is in the first component. This problem can be related with the standard deviation used to construct the grid that is very low (see Table 5.3 and Figure 5.10).

Belonging cluster	V_b $[unitless]$	a_1 $[ml\ g^{-1}\ min^{-1}]$	b_1 $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	b_2 $[min^{-1}]$	a_3 $[ml\ g^{-1}\ min^{-1}]$	b_3 $[min^{-1}]$
1	0.025	0.038	0.010	0.057	0.244	–	–
2	0.016	0.030	0.012	0.005	0.195	0.030	0.330
3	0.053	0.038	0.008	0.026	0.091	–	–
4	0.028	0.019	0.013	0.005	0.155	0.024	0.276
5	0.029	0.026	0.018	0.020	0.262	–	–
6	0	0.012	0.026	0.006	0.061	0.043	0.498
7	0.062	0.027	0.012	0.048	0.187	–	–
8	0.043	0.036	0.010	0.053	0.376	–	–

Table 5.7: Table of estimated parameters for each representative voxel, obtained with LSA with prior from ROI

5.1.4 Linear SA with prior from voxel

The map of the volume of distribution obtained with the linear analysis that uses the prior from voxel is shown in Figure 5.11

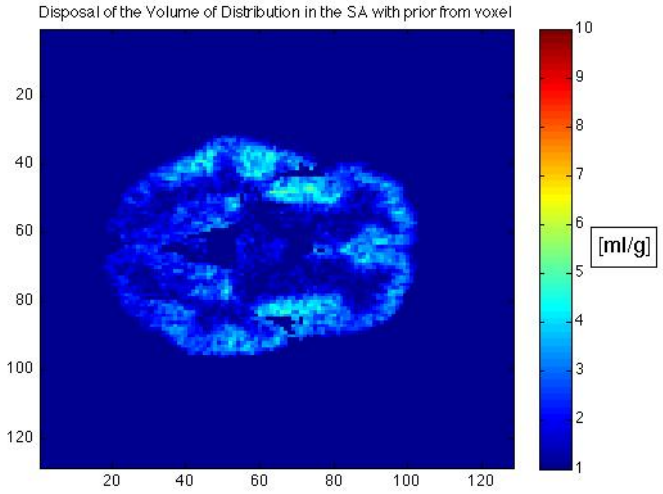


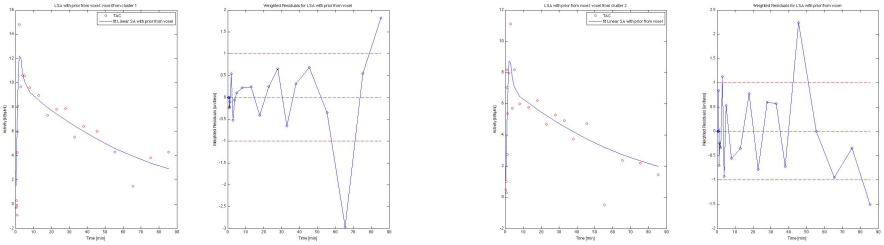
Figure 5.11: Slice 45: Image showing the volume of distribution, obtained with LSA with prior from voxel

The variables of interest are presented in Table 5.8 and their units are $V_d[ml\ g^{-1}], K_1[ml\ g^{-1}\ min^{-1}], AIC[unitless]\ and\ WRSS[unitless]$

Belonging cluster	V_d $[ml\ g^{-1}]$	K_1 $[ml\ g^{-1}\ min^{-1}]$	AIC $[unitless]$	$WRSS$ $[unitless]$
1	0.809	0.127	95	44
2	3.251	0.088	59	10
3	1.998	0.107	63	11
4	1.483	0.078	76	19
5	1.463	0.051	95	47
6	2.230	0.084	55	9
7	3.663	0.107	78	23
8	2.866	0.142	63	12

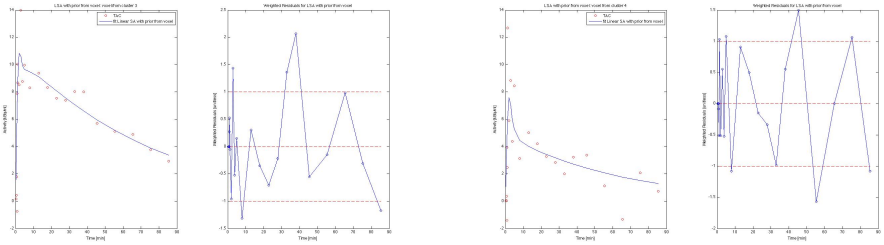
Table 5.8: Table of variables calculated for one voxel representative of each cluster, obtained with LSA with prior from voxel

In Figure 5.12 the fits of the same voxel previously selected as representative of each cluster and their relative weighted residual are presented.



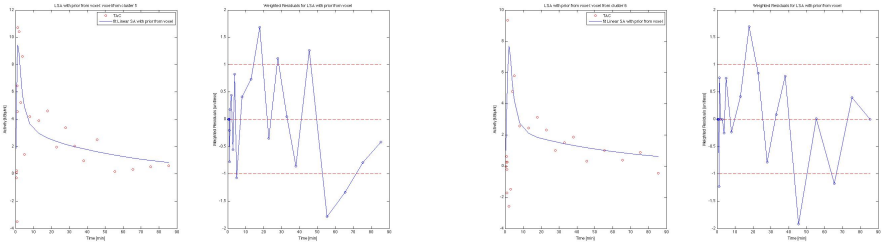
(a) *Fit vs Data for a voxel of Cluster 1.*

(b) *Fit vs Data for a voxel of Cluster 2.*



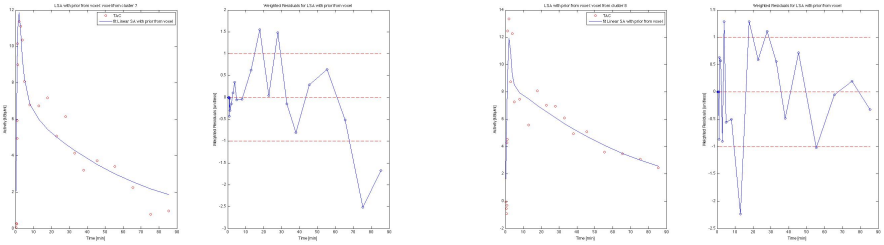
(c) *Fit vs Data for a voxel of Cluster 3.*

(d) *Fit vs Data for a voxel of Cluster 4.*



(e) *Fit vs Data for a voxel of Cluster 5.*

(f) *Fit vs Data for a voxel of Cluster 6.*



(g) *Fit vs Data for a voxel of Cluster 7.*

(h) *Fit vs Data for a voxel of Cluster 8.*

Figure 5.12: Model estimated curves with LSA with prior from voxel and relative weighted residuals

The value of the parameters estimated for each representative voxel is presented with the following units $V_b[unitless], a_1, a_2, a_3, a_4[ml\ g^{-1}\ min^{-1}], b_1, b_2, b_3, b_4[min^{-1}]$ in Table 5.9.

Cl	V_b $[unitless]$	a_1 $[ml\ g^{-1}\ min^{-1}]$	b_1 $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	b_2 $[min^{-1}]$	a_3 $[ml\ g^{-1}\ min^{-1}]$	b_3 $[min^{-1}]$	a_4 $[ml\ g^{-1}\ min^{-1}]$	b_4 $[min^{-1}]$
1	0	0.005	0.012	0.045	0.016	0.031	1.344	0.046	0.534
2	0	0.035	0.016	0.016	0.679	0.036	0.710	–	–
3	0.016	0.048	0.015	0.005	0.015	0.054	1.010	–	–
4	0.006	0.023	0.016	0.045	0.645	0.010	0.580	–	–
5	0.057	0.023	0.016	0.028	0.534	–	–	–	–
6	0	0.011	0.016	0.073	0.534	–	–	–	–
7	0.040	0.034	0.016	0.073	0.534	–	–	–	–
8	0	0.033	0.016	0.011	0.015	0.099	0.975	–	–

Table 5.9: Table of estimated parameters for each representative voxel, obtained with LSA with prior from voxel

The linear SA implemented with the prior obtained by the information extracted from the distribution of b in the whole brain (as explained in Section 3.3), makes use of the grid showed in Figure 5.13:

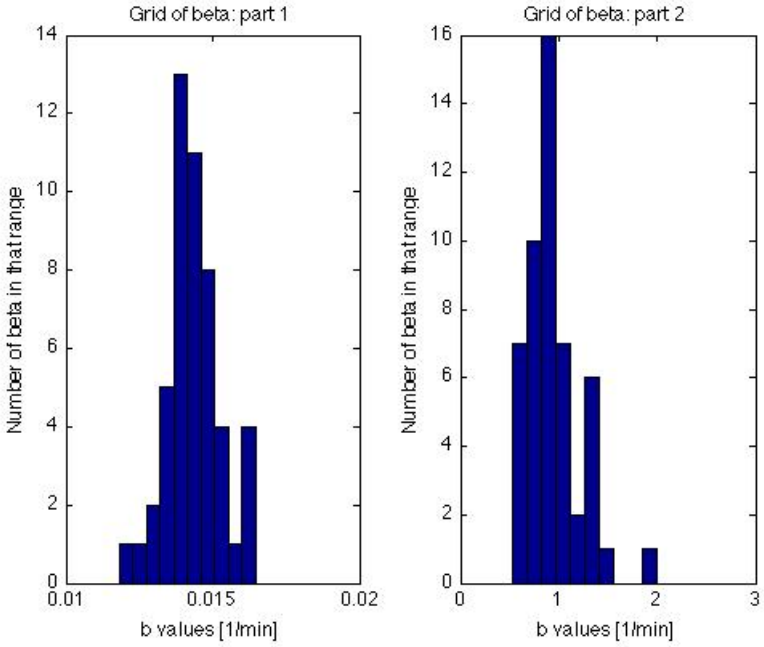


Figure 5.13: Image showing the distribution of the b in the grid for the method LSA with prior from voxel

5.1.5 Linear SA implemented with MAP estimator

After trying this method, an unusual behavior in the disposal of the b values and their relative a was noticed (shown in Figure 5.14), and, hypothesizing that it was due to an high correlation between the a and b estimated, it was decided to repeat the method again, fixing the prior just on the b value. The different disposal obtained is shown in Figure 5.15.

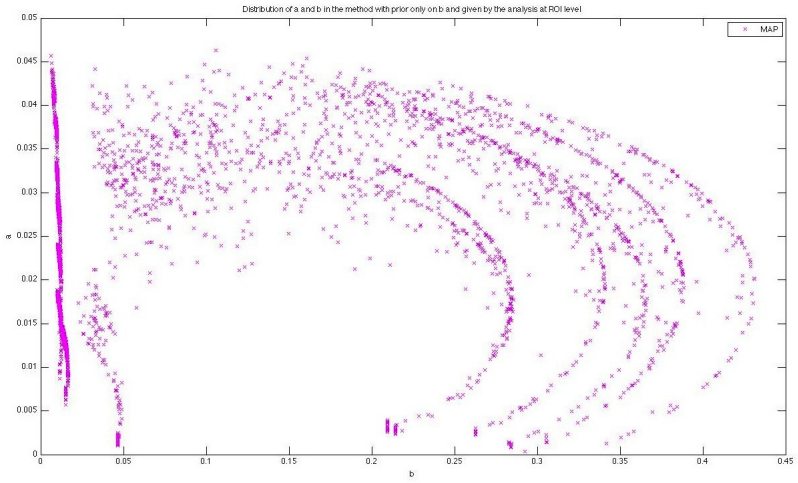


Figure 5.14: Distribution of the parameters a and b for the method with prior on both

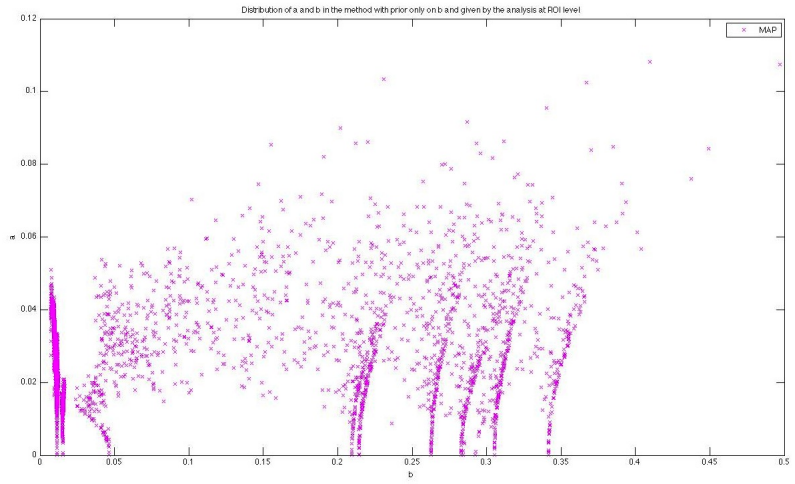


Figure 5.15: Distribution of the parameters a and b for the method with prior just on b

The modification made in order to achieve this goal was deleting the contribution of the prior on the a, raising the value of its covariance. In this way, when the term Σ_p of the equation was inverted, the weight associated at the prior of a became nearly zero. The results have improved significantly and the unusual behavior was no longer present.

The map of the volume of distribution obtained with the linear analysis is shown in Figure 5.16

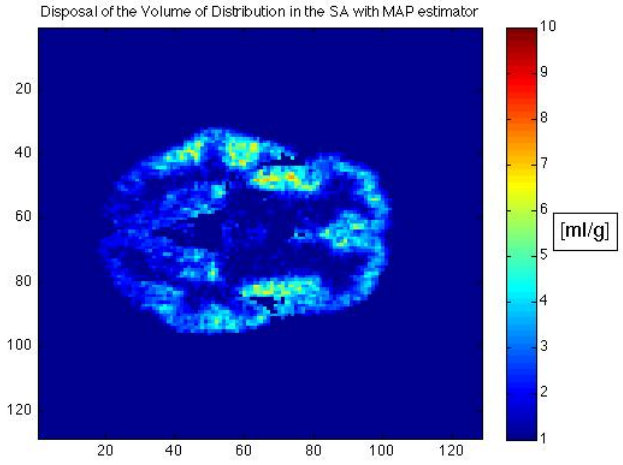


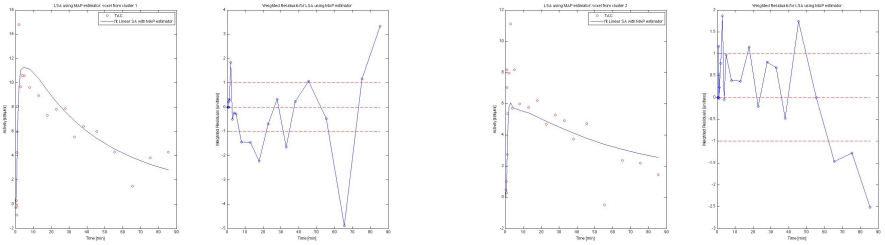
Figure 5.16: Slice 45: Image showing the volume of distribution

The value of the parameters estimated for each representative voxel is presented with the following units $a_1, a_2[ml\ g^{-1}\ min^{-1}], b_1, b_2[min^{-1}]$ in Table 5.10; parameter V_b is not present because it has been considered negligible.

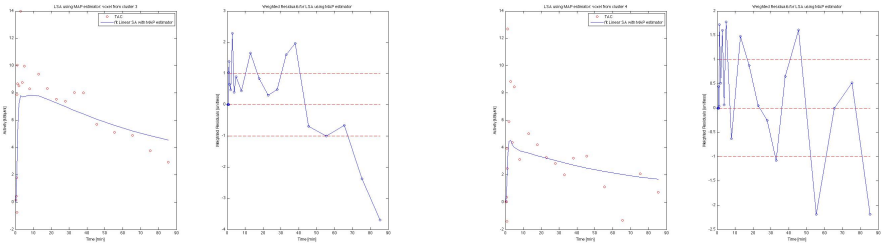
Cluster	a_1 $[ml\ g^{-1}\ min^{-1}]$	b_1 $[min^{-1}]$	a_2 $[ml\ g^{-1}\ min^{-1}]$	b_2 $[min^{-1}]$
1	0.025	0.009	0.045	0.045
2	0.030	0.012	0.018	0.267
3	0.042	0.008	0.017	0.303
4	0.020	0.012	0.018	0.219
5	0.023	0.016	0.014	0.211
6	0.008	0.041	0.025	0.062
7	0.029	0.012	0.028	0.297
8	0.035	0.010	0.032	0.311

Table 5.10: Table of estimated parameters for each representative voxel, obtained with LSA with MAP estimator

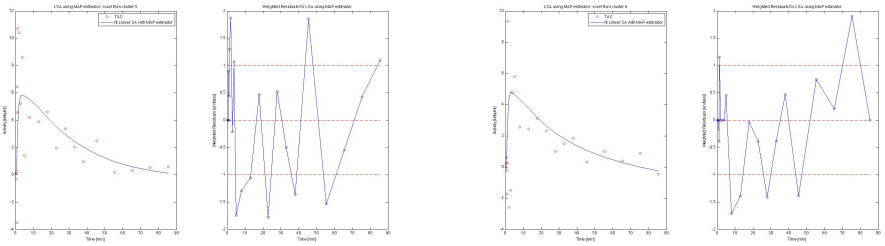
Figure 5.17 presents the fits of the same voxel chosen previously as representative of each cluster and their relative weighted residual.



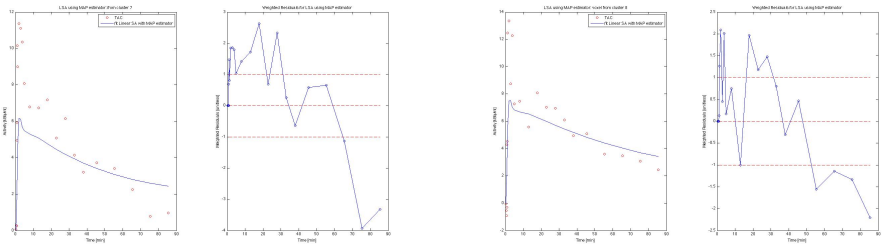
(a) *Fit vs Data for a voxel of Cluster 1.* (b) *Fit vs Data for a voxel of Cluster 2.*



(c) *Fit vs Data for a voxel of Cluster 3.* (d) *Fit vs Data for a voxel of Cluster 4.*



(e) *Fit vs Data for a voxel of Cluster 5.* (f) *Fit vs Data for a voxel of Cluster 6.*



(g) *Fit vs Data for a voxel of Cluster 7.* (h) *Fit vs Data for a voxel of Cluster 8.*

Figure 5.17: Model estimated curves with linear SA using a MAP estimator with prior on b and relative weighted residuals

The variables of interest are presented in Table 5.11 and their units are $V_d[ml\ g^{-1}]$, $K_1[ml\ g^{-1}\ min^{-1}], AIC[unitless]$ *and* $WRSS[unitless]$

Belonging cluster	V_d $[ml\ g^{-1}]$	K_1 $[ml\ g^{-1}\ min^{-1}]$	AIC $[unitless]$	$WRSS$ $[unitless]$
1	3.694	0.069	100	54
2	2.638	0.048	80	23
3	5.027	0.058	94	41
4	1.741	0.038	83	27
5	1.512	0.037	97	49
6	0.590	0.033	71	15
7	2.595	0.057	103	61
8	3.627	0.068	87	31

Table 5.11: Table of variables calculated for one voxel representative of each cluster

5.2 Results comparisons

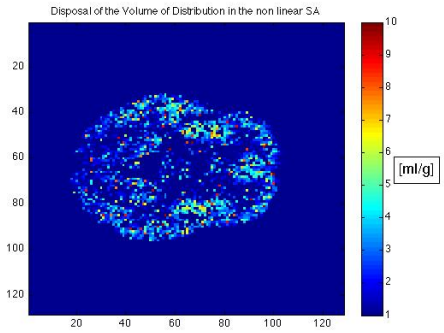
The previous section introduced all the results obtained for each method taken into consideration during this study both just as a reference and as new methods proposed. It is necessary to compare these results to understand if the methods proposed give results physiologically corrected and if they introduce improvements to those existent.

Due to high levels of noise presented by the tissue-time activity curves at voxel level, before making any comparisons, it was necessary to extrapolate only the voxel for which the estimated parameters have physiological meaning. This means that we did not consider voxel for which $V_d < 0$, $V_d > 10[ml\ g^{-1}]$ and the estimates for this parameter, in terms of Coefficients of Variation (introduced previously in chapter 2 by Eq. (2.25)), were higher than 50%. These values were decided considering that a negative V_d value has no biological explanation and V_d over 10 are, for this tracer, non-physiological (Gunn et al., 1998). This operation induced a reduction in the number of voxel, which can be used to compute the comparisons, of more than 50%.

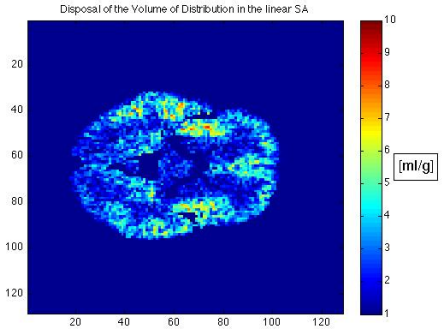
Distribution Volume V_d

As previously stated, the method selected as *gold standard*, in order to compare the volume of distribution, is the non linear analysis (see section 2.2) made at voxel level. Is in its results that the voxel having parameters with

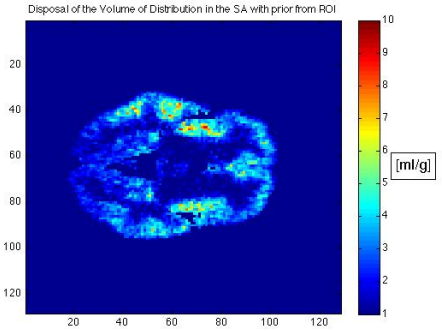
physiological plausibility have been selected and it is for these voxel that comparisons have been made between different methods. Once the voxel have been extrapolated the first parallel has been made in terms of volume of distribution, both using parametric maps (Figure 5.18) and using mean values for each cluster and for the whole brain (Table 5.12).



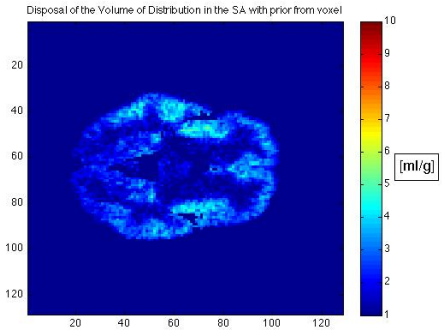
(a) *Map of the NLSA.*



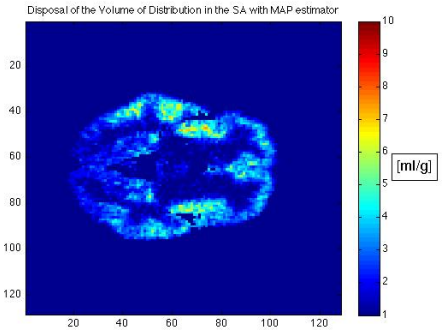
(b) *Map of the LSA.*



(c) *Map of the LSA with prior from ROI.*



(d) *Map of the LSA with prior from voxel.*



(e) *Map of the LSA implemented with MAP estimator.*

Figure 5.18: Maps showing the Volume of Distribution $[ml\ g^{-1}]$ in all methods tested, for a representative subject.

From the comparison of these maps is easy to observe that both the LSA methods with prior from ROI and with MAP estimator show good resemblance with the gold standard method. The LSA implemented in the classical way gives good results but shows an higher noise level. Instead the LSA method with prior from voxel shows the lowest values in terms of volume of distribution. In Table 5.12, for each cluster, the mean values and the standard deviations representative of the variability of the parameter V_d within the cluster, are shown.

Belonging cluster	NLSA $V_d \pm \sigma$ $[ml\ g^{-1}]$	LSA $V_d \pm \sigma$ $[ml\ g^{-1}]$	LSA with prior from ROI $V_d \pm \sigma$ $[ml\ g^{-1}]$
1	5.79 ± 1.21	4.15 ± 0.88	4.19 ± 0.31
2	2.95 ± 0.63	2.20 ± 0.54	2.17 ± 0.25
3	6.99 ± 1.29	5.02 ± 0.92	5.40 ± 0.56
4	2.33 ± 0.55	1.80 ± 0.47	1.72 ± 0.29
5	1.72 ± 0.47	1.35 ± 0.41	1.18 ± 0.24
6	1.37 ± 0.48	1.07 ± 0.40	0.77 ± 0.16
7	3.84 ± 0.85	2.90 ± 0.65	2.73 ± 0.25
8	4.71 ± 1.01	3.45 ± 0.75	3.41 ± 0.25
Whole Brain	3.57 ± 1.74	2.65 ± 1.24	2.58 ± 1.20

Belonging cluster	LSA with prior from voxel $V_d \pm \sigma$ $[ml\ g^{-1}]$	LSA with MAP estimator $V_d \pm \sigma$ $[ml\ g^{-1}]$
1	3.50 ± 0.27	4.14 ± 0.24
2	1.99 ± 0.23	2.17 ± 0.22
3	4.14 ± 0.33	5.47 ± 0.43
4	1.60 ± 0.24	1.72 ± 0.24
5	1.22 ± 0.23	1.16 ± 0.21
6	1.00 ± 0.24	0.77 ± 0.20
7	2.47 ± 0.24	2.69 ± 0.21
8	2.96 ± 0.26	3.39 ± 0.21
Whole Brain	2.29 ± 0.89	2.56 ± 1.19

Table 5.12: Table representing the mean values of the Volume of Distribution for each cluster and in the whole brain

For better understanding the differences in the macro-parameter V_d between the methods proposed, is shown now an explicative chart (Figure 5.19).

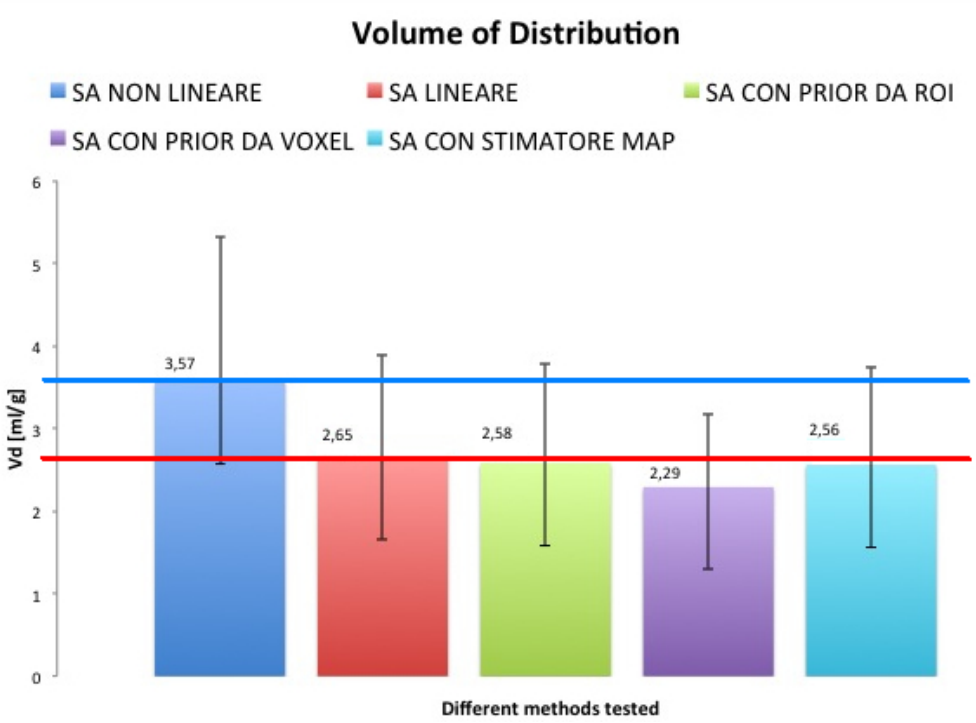


Figure 5.19: Chart of V_d values estimated using all the methods proposed.

The goal of this thesis was to define a new method that would give results more similar to those given by the NLSA, with respect to those given by the LSA already existent. From this chart is easy to see that (in terms of mean and variability) the new methods proposed don't introduce any improvements. This chart shows clearly that the linear methods underestimate the volume of distribution values when referring to the gold standard method (NLSA).

Another important way to evaluate the goodness of the parameter of interest V_d, with respect to the reference method, is obtaining the scatterplot. It aids the interpretation of the correlation coefficient or regression model, which are useful for us to numerically interpret the differences between methods. To show these results are utilized V_d calculated as mean of the V_d of all voxel belonging to the specific cluster. In Figure 5.20 there are four scatterplots of V_d values; each one was obtained with the values of the non linear SA analysis in the x axis and those obtained with the other methods in the y axis. There was good agreement and strong correlation among the NLSA and other methods estimates ($R^2 > 0.99$ for all models tested) but the regression line was quite far from the bisector almost in all cases (this result confirm the underestimation

given by the linear methods). These values are summarized in Table 5.13 as well.

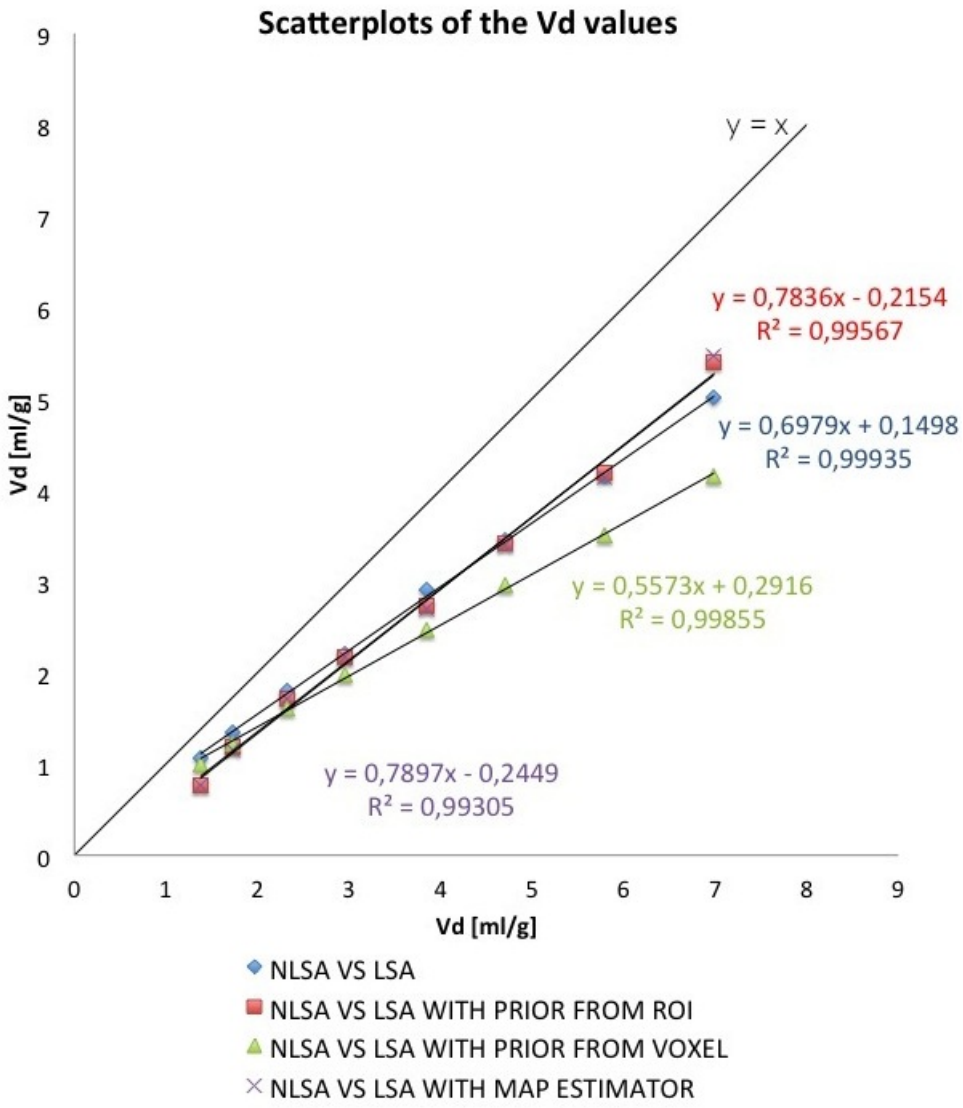


Figure 5.20: Scatterplots of V_d values estimated using NLSA (x axis) as gold standard method and all four other methods tested (y axis). In each scatterplot the value of slope and intercept of the fitted regression line and Pearson's R2 value are reported.

	Correlation coefficient (R^2)	*Regression line* (*slope*)
NLSA vs LSA	0.999	0.698
NLSA vs LSA with prior from ROI	0.996	0.784
NLSA vs LSA with prior from voxel	0.999	0.557
NLSA vs LSA with MAP estimator	0.993	0.790

Table 5.13: Table of statistics (Pearson's value R^2 and slope of the fitted regression line) for the NLSA versus the four methods tested

It is now possible to make some observations based on the results just presented. All the methods proposed underestimate the value of this parameter of interest but, even so, it is important to notice that they show results that are comparable and in agreement with the expectation (the regional distribution of V_d was consistent with known 5–HT_{1A} receptor distributions (Gunn et al., 1998)). Looking at the coefficient of correlation, the best method is the one with prior from voxel, but at the same time it is the one with slope furthest from unity. The other two methods show similar results, even if the parametric map of the method with prior from ROI seems the more consistent.

Fits and weighted residuals

Our goal is to evaluate the goodness of the methods proposed and this can be achieved also analyzing the fits that each method present for the data. To compare the fits, the best choice is to plot all the curves, referring to the same voxel, in one figure (as seen in Figures 5.21) and showing the correspondent weighted residuals together in another figure (see Figures 5.22). The figures presented here are representative for a voxel representative of cluster 1, the results for the other clusters are presented in Appendix A.

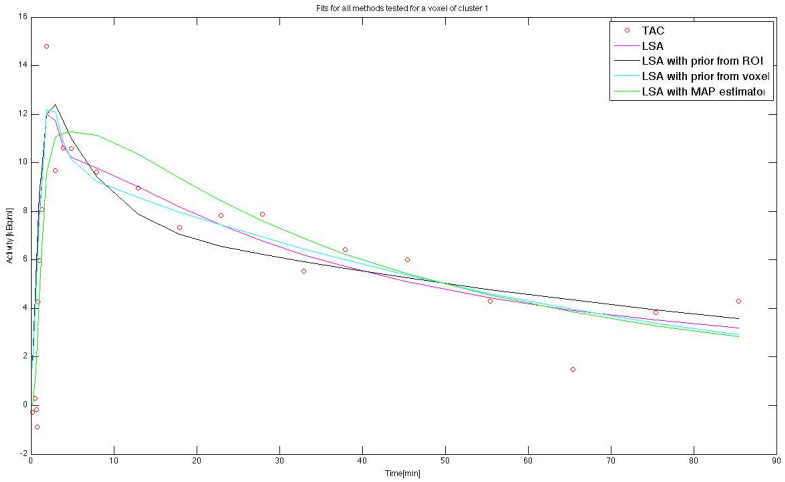
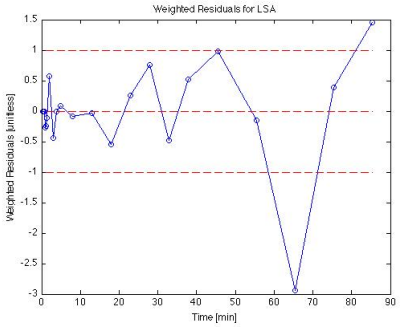
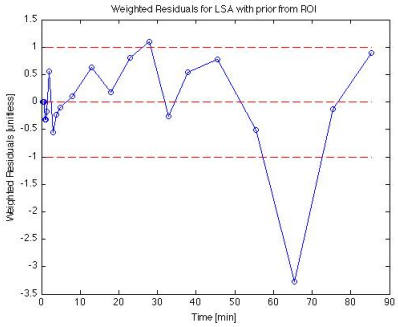


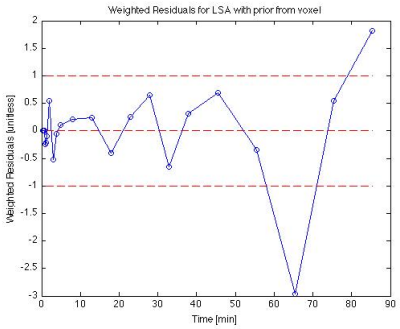
Figure 5.21: Fits for all the methods tested for the voxel selected as representative of the cluster 1



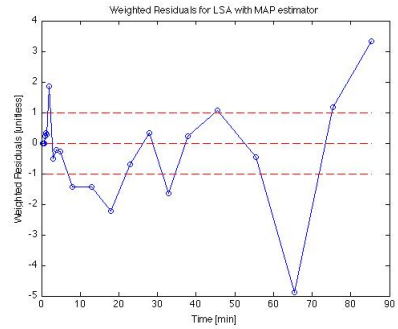
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure 5.22: Weighted residuals for all methods tested on this voxel representative of cluster 1.

Evaluating the comparisons of these results indicates that the method that makes use of the MAP estimator is the one that presents the worst fits. This behavior is confirmed reviewing the weighted residuals. In fact, thinking that the residuals are an estimate of measurement error, if the model (structure plus numerical value of the parameters) is "good" it is logical to expect that the residuals are consistent with the statistical properties of the measurement error. This means they should have zero mean, they should be uncorrelated (alternatively positive and negative) and they should have amplitude between -1 and 1. This is primarily true for all the methods, except for the method LSA with MAP estimator where, for some clusters, it presents residuals that are correlated.

WRSS and AIC

The final parameters used in this study to compare the proposed methods are WRSS (see Eq. (2.22)) and AIC (see Eq. (2.28)). The first can be considered as the weighted error between the observed and predicted value for each sample time. Therefore it is a number that represent the information given visually by the weighted residuals. The second is a parsimony criterion that can be useful to understand which method gives the best balance between accuracy and complexity of the model. The smaller its value, the better results in terms of no information loss when it is used. These parameters, shown in Table 5.14 and in Table 5.15, are calculated as medium values of all those belonging to each cluster and then as mean value of all voxel within the mask.

The standard deviations calculated for each value refer to the variability within the cluster and within the whole brain. The results confirm what was already shown with the plots of the fits and the weighted residuals, i.e. the method that uses the MAP estimator gives the worst description of the data (its WRSS are higher than the others).

Belonging cluster	LSA $\pm\,\sigma$	LSA with prior from ROI $\pm\,\sigma$	LSA with prior from voxel $\pm\,\sigma$	LSA with MAP estimator $\pm\,\sigma$
1	61 ± 30	66 ± 30	64 ± 30	75 ± 31
2	43 ± 21	46 ± 21	45 ± 21	56 ± 24
3	72 ± 32	75 ± 32	77 ± 33	87 ± 36
4	37 ± 18	40 ± 18	39 ± 18	49 ± 20
5	26 ± 12	28 ± 12	28 ± 12	36 ± 15
6	19 ± 10	21 ± 10	21 ± 10	26 ± 11
7	53 ± 26	56 ± 26	55 ± 26	65 ± 28
8	54 ± 24	59 ± 25	57 ± 24	69 ± 28
Whole Brain	45 ± 26	49 ± 27	48 ± 26	58 ± 30

Table 5.14: Table of WRSS for the method used as a reference and the new methods proposed

Belonging cluster	LSA $\pm \sigma$	LSA with prior from ROI $\pm \sigma$	LSA with prior from voxel $\pm \sigma$	LSA with MAP estimator $\pm \sigma$
1	100 ± 11	100 ± 11	99 ± 11	105 ± 10
2	92 ± 11	92 ± 11	91 ± 11	99 ± 10
3	103 ± 11	104 ± 11	103 ± 11	109 ± 10
4	88 ± 11	89 ± 10	88 ± 10	96 ± 9
5	81 ± 11	81 ± 10	81 ± 10	88 ± 9
6	72 ± 12	73 ± 12	73 ± 11	81 ± 10
7	97 ± 11	96 ± 11	96 ± 11	102 ± 10
8	97 ± 11	97 ± 11	97 ± 11	103 ± 10
Whole Brain	92 ± 13	92 ± 13	92 ± 13	98 ± 12

Table 5.15: Table of AIC for the method used as a reference and the new methods proposed

Chapter 6

Discussion

In the previous chapter (Chapter 5) we presented the results found with the I/O models, both the existent ones (linear and non linear Spectral Analysis) and the new methods proposed (LSA with prior from ROI, LSA with prior from voxel, LSA with MAP estimator). In this chapter the goal is to explain the strengths and weaknesses of all these methods.

- *NLSA*:

  The non-linear approach used in the analysis at voxel level is not appropriate since it requires computational time that is too high and there are problems of failure rate caused by the high noise that characterizes the curves of activities specific for each voxel. In fact, due to noise, the estimator is not able to converge and generates values of the parameters inappropriately.

- *LSA*:

  The linear analysis rectifies to the limitations just listed, not presenting any outliers and having a limited computing time. In turn, it introduces other problems. It remains in fact very sensitive to noise and provides results heavily dependent on the starting grid provided.

The goal of this thesis was to propose methods that remedy the limitation presented above of these existent methods. The proposed solutions are all linear and designed in such a way to make more robust the solution provided by this type of methods and as close as possible to the solution provided by the non-linear ones.

- *LSA with prior from ROI*:

  The results obtained with this method demonstrate a problem: using as inputs the results obtained by the NLSA method implemented on the ROI TACs, which mean to curves of activity characterized by a good

signal to noise ratio, the precision of estimates that are obtained have values very low. This is generally a very positive aspect, but has a negative impact on our results. This happened because we built the Gaussian with medium equal to the values of model parameters and standard deviation equal to the parameter estimates themselves. Having very good estimates, the Gaussians are generated with values very concentrated around the mean. This is found in the results where the parameters b tend to take the limit values of the Gaussian, as if trying to have values lower or higher respectively. This effect is very evident in the first of the two Gaussians, since, working on small values (around 0.01), it is more affected by the low standard deviation. The results provided are still very good, both in terms of fit and then weighted residuals, and in terms of macro-parameter of interest (the volume of distribution). This method appears more robust in terms of number of given components and therefore less sensitive to noise than the LSA.

- *LSA with prior from voxel*:

  This method does not present the same problem as the previous one. In fact the grid constructed according to the idea of analyzing the distribution of values on the grid of b used by the LSA, in order to create a specific grid which respects the typical aspects of it, covers a good range of values. In this case, however, there is a slight underestimate of the value of the parameter of interest (the volume of distribution). One limitation of this technique may be the use of the same grid for all voxels inside the mask, without distinction of their belonging to different ROIs. Since the TACs identified for the different regions of interest have values of the parameters that are also quite different, it might be wise to evaluate a solution taking account of these differences, thus creating different grids for different ROIs.

- *LSA with MAP estimator*:

  The last method provided the worst results in terms of adherence of the model to the data in question, even if it has supplied values of volume of distribution that were very good. A possible explanation for these results is due to the choice of the critical parameter, the cutoff b, which is used as the value for identifying the two components in the kinetics of the tracer and then to rebuild them as sums of convolution integrals. In the method investigated, this value is identified by analyzing the distribution of b in the whole brain in which we clearly distinguish two Gaussian distributions. Once extrapolated, the value of the minimum between the two distributions is used to reconstruct the two components. A different choice of this parameter, considering such distributions in terms of ROI instead of the whole brain, may provide different values of the cutoff b

and consequently different results in estimates of the parameters that best describe the TACs at voxel level.

All proposed methods require running the data analysis twice. Then, in spite of good results, this is something that must not be underestimated. It must be said that they are all linear methods and thus do not require high computational time, but this is still an evaluation that has to be done case by case. After analyzing the results and making all these observations it can be said that the proposed methods do not introduce substantial improvements to the linear SA currently in use.

Chapter 7

Conclusions and future developments

In the analysis of models in which one considers the plasma function operating as input, one can refer to a data-driven approach that takes the name of Spectral Analysis. It was showed that this method has two major problems: noise sensitivity and the high dependence of the goodness of the results from the grid. For this reason, and starting from the purpose of wanting to make more robust this method, four alternative methods have been proposed, only three of which were actually implemented since the first proved to be inadequate. Aware that the choice of the grid can be fundamental in these types of analysis, our work has focused on developing a new way to define it. Once these grids were obtained our goal became the study of the results provided in order to understand if they were physiologically correct and if they obtained better results compared to existent methods.

The tracer used for this work is $\left[^{11}C\right]$WAY100635, which is a selective antagonist with affinity and selectivity for serotonin 5–HT_{1A} receptors. The analysis of the tracer was performed both at region of interest level and at voxel level and, to extract the useful information from the data, different approaches were used (linear and non linear Spectral Analysis). These input-output techniques were used both to quantify some variables of interest and, especially, to have models on which basing our comparison to understand whether the methods proposed were correct. The non linear approach gave good results for the voxel where the convergence of the global optimum were not problematic, but the computational cost was really high and the failure rate was more than 50%. The linear SA approach instead, gave good results on all data, giving a good description of the data and values of the variables of interest with physiological meaning. But, as previously stated, we observed that the use of a fixed b grid and the presence of noise created some problems to this approach too, i.e. double lines close to the real value.

The first method evaluated (LSA with prior from ROI) extrapolates in-

formation from a preliminary nonlinear analysis performed at ROI level to generate a grid that has distribution densely distributed around the values that we expect as the parameter of the model. This is based on the concept that the parameter estimates derived at ROI level are representative of the expected model parameters' values of the voxels composing the ROI itself. The second method instead (LSA with prior from voxel) performs a first analysis based on the LSA and with the results obtained in terms of distribution of b values on the grid, generates a new grid that reflects the truly distribution assumed by b. The third proposed method (LSA with MAP estimator) follows a slightly different path: run first an LSA analysis and analyze the point that separates the two Gaussian distributions of b. Then obtains these two component separately as sums of the integrals of convolution from zero to the cutoff value and from the cutoff value onwards. Finally, it implements an MAP estimator for each of the two curves obtained, using as prior the information obtained from the non linear analysis made at ROI level.

The results of these three methods, physiologically speaking, are correct but they don't improve the results of the existent linear SA method. The third method, even if it provided values of distribution volume macro-parameter correlated with the physiological information about the selective distribution of 5–HT_{1A} receptors within the brain, was the method that provided the worst fits of the data. Instead the first and the second methods proposed, presented precise concordance with the results of LSA, both in terms of distribution volume and of fits of the data. One problem observed in the first method was due to the low value of the standard deviation obtained from the non linear analysis at ROI level. This is not a problem generally speaking since a low SD indicates a good estimate of the parameters, but in our study it resulted in a grid too concentrated around the two mean values. This implied values too near one to the other, showing that the parameters estimated tended to assume the two values of the bounds, as if they were trying to find a lower and a higher value respectively. Despite this, methods one and two are those that demonstrated behavior, generally speaking, closer to the one presented by the linear SA. Evaluating all the results and the performance of the various methods proposed, the preferred method to obtain a grid that is specific for the case under examination is the linear method that constructs a prior starting from the non linear analysis executed at ROI level.

Further studies, involving a larger number of subjects and different types of tracers, are required to confirm the reliability of these methods. Moreover, noting the good performance obtained it would be of interest to test these methods modifying the critical parameter chosen for each method. For example in the first method should be interesting to maintain the same idea of constructing the grid with Gaussian distributions, but find a way to enlarge their standard deviation. For the second method, it would be interesting to evaluate the distribution of b not at level of the whole brain but at ROI level, in order to obtain a more specific grid. For the last method, the critical value

is the b of cutoff, so it would be interesting to evaluate changes between the different regions of interest and determine if a different choice of it for each ROI can affect the results.

Appendix A

In this appendix are shown the results, in terms of fits (see Figures A.1, A.3, A.5, A.7, A.9, A.11, A.13) and weighted residuals (see Figures A.2, A.4, A.6, A.8, A.10, A.12, A.14), for the same voxel chosen as representative of each cluster in Chapter 5.

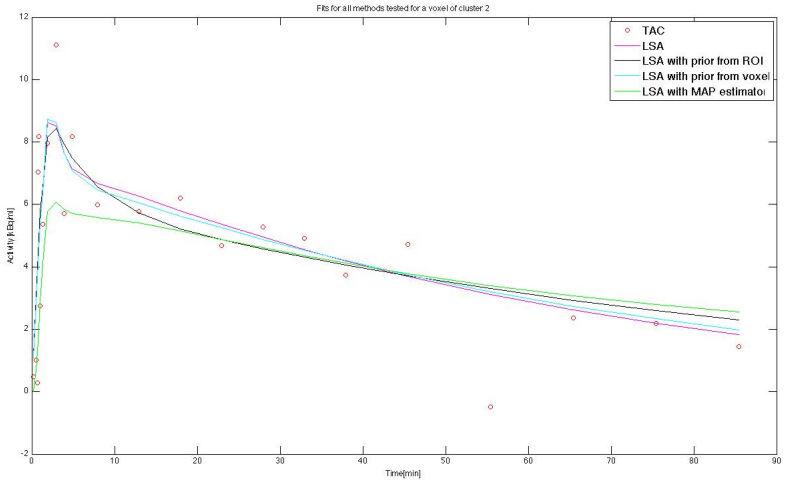
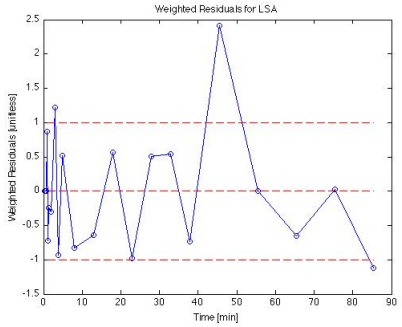
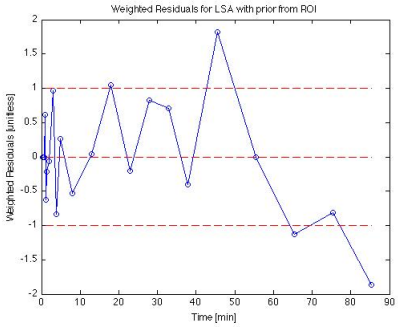


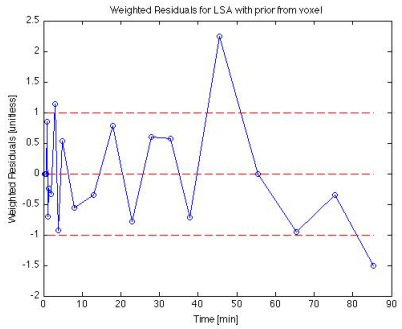
Figure A.1: Fits for all the methods tested for the voxel selected as representative of the cluster 2



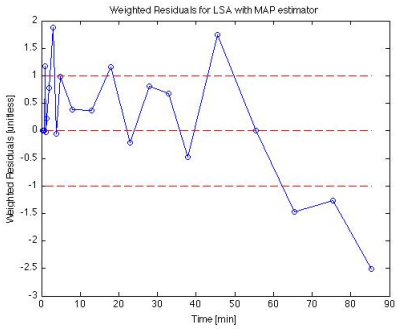
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.2: Weighted residuals for all methods tested on this voxel representative of cluster 2.

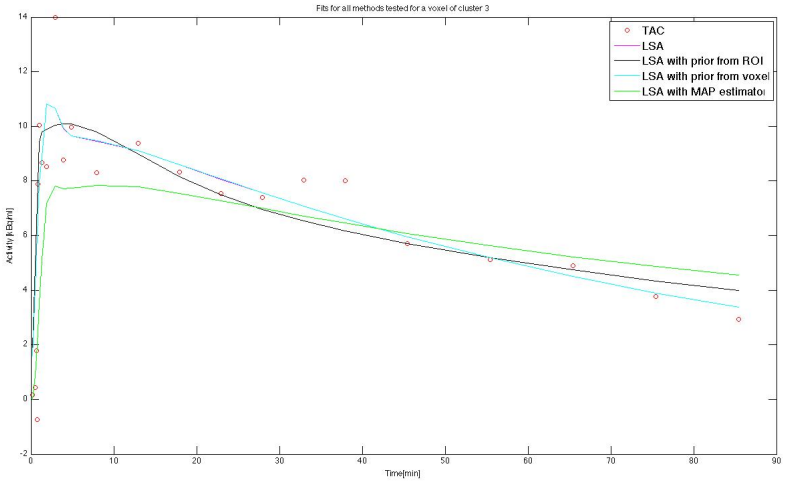
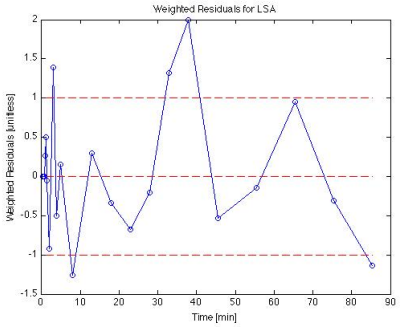
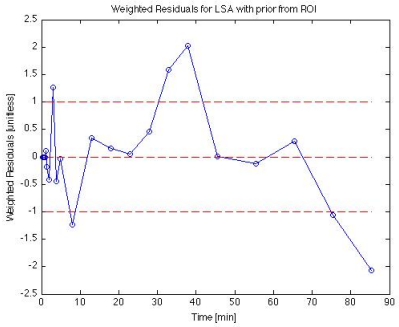


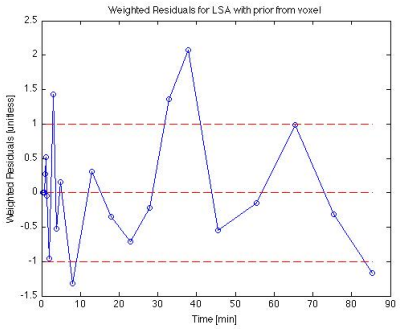
Figure A.3: Fits for all the methods tested for the voxel selected as representative of the cluster 3



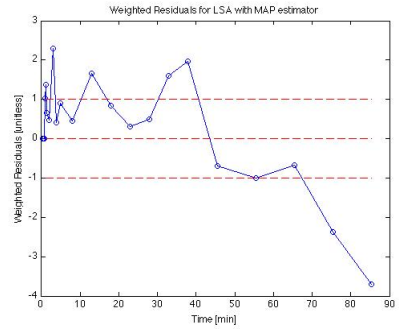
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.4: Weighted residuals for all methods tested on this voxel representative of cluster 3.

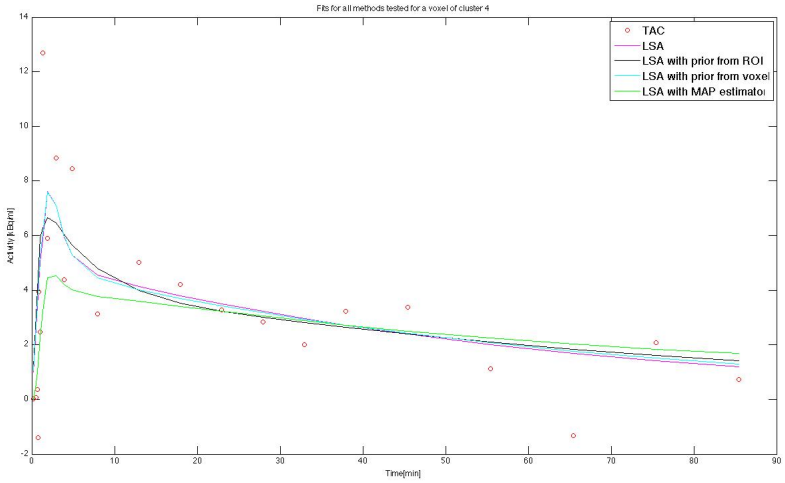
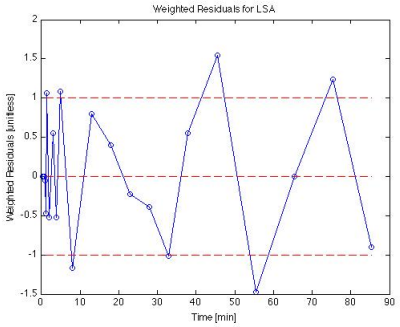
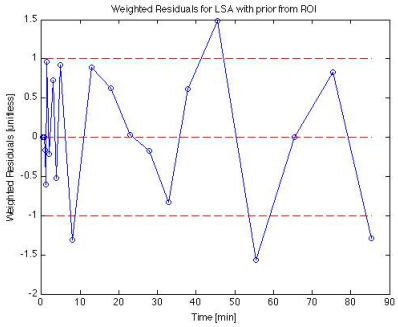


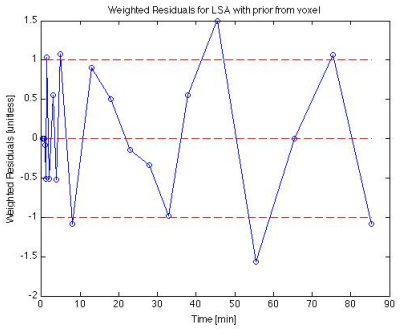
Figure A.5: Fits for all the methods tested for the voxel selected as representative of the cluster 4



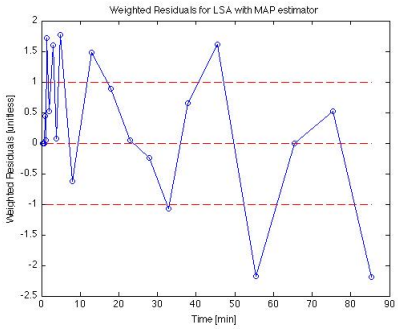
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.6: Weighted residuals for all methods tested on this voxel representative of cluster 4.

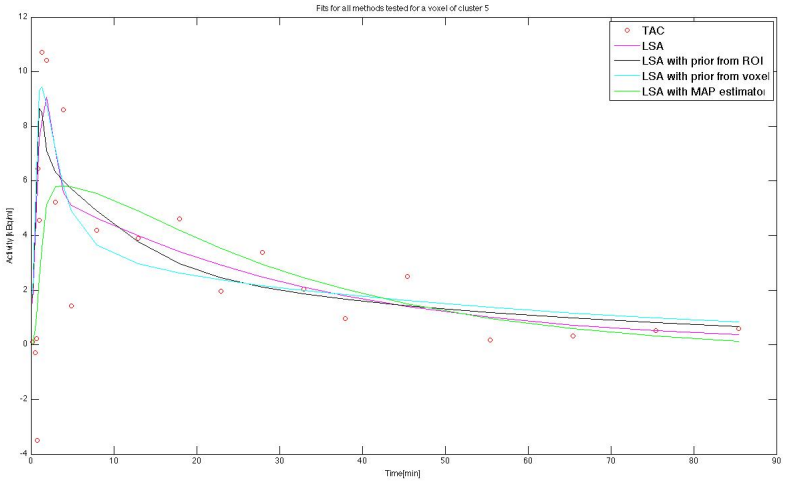
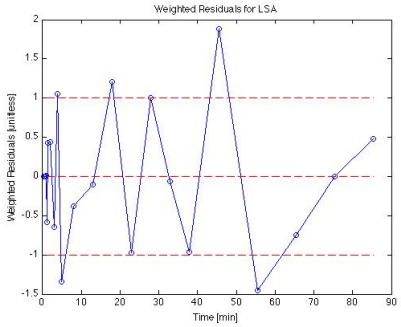
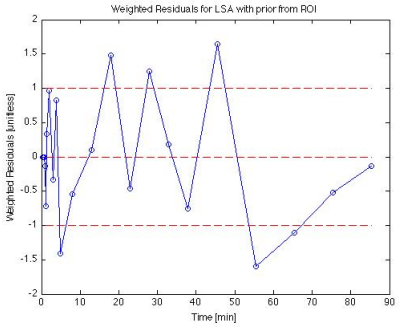


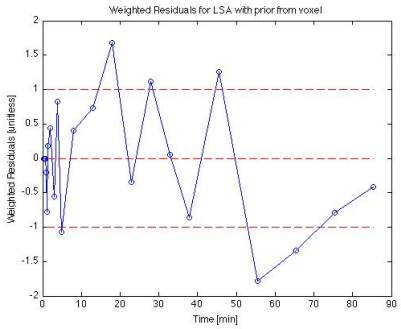
Figure A.7: Fits for all the methods tested for the voxel selected as representative of the cluster 5



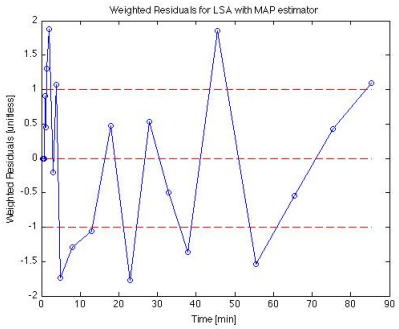
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.8: Weighted residuals for all methods tested on this voxel representative of cluster 5.

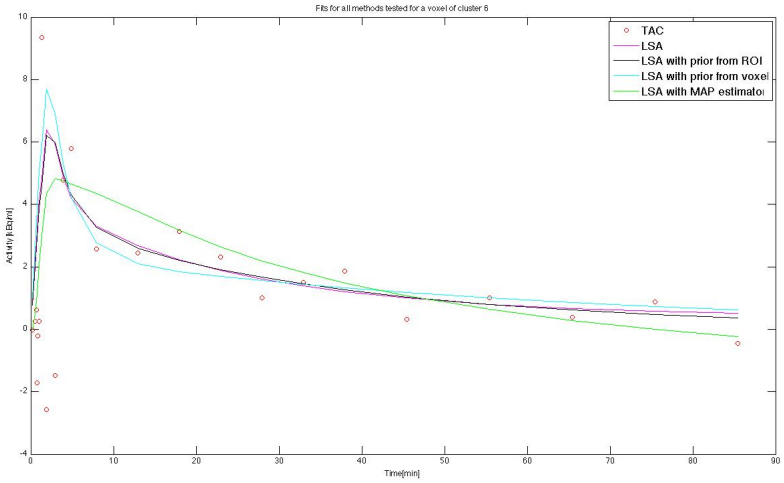
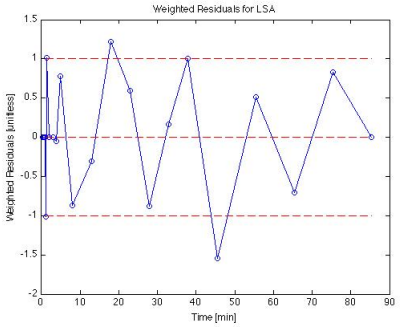
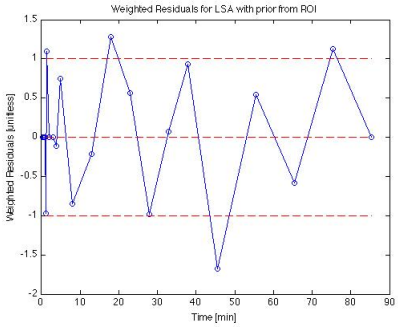


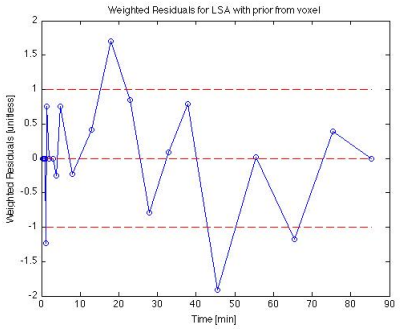
Figure A.9: Fits for all the methods tested for the voxel selected as representative of the cluster 6



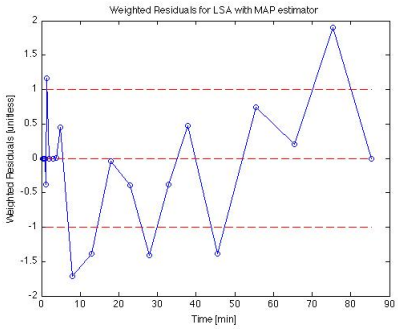
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.10: Weighted residuals for all methods tested on this voxel representative of cluster 6.

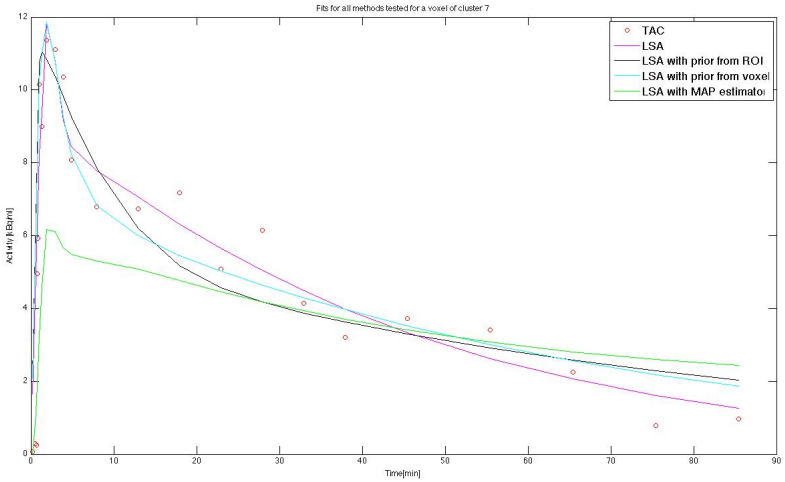
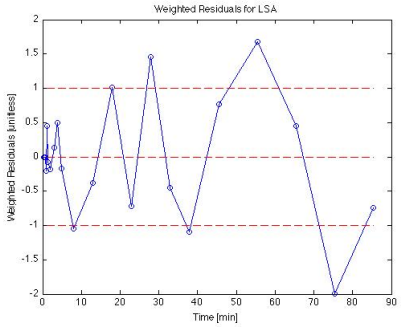
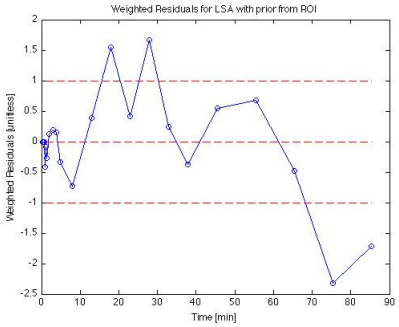


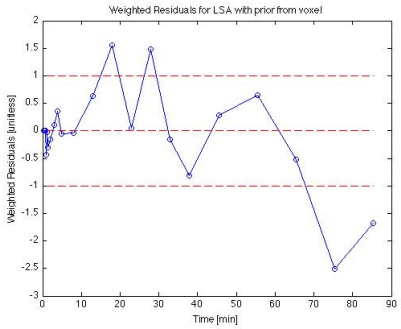
Figure A.11: Fits for all the methods tested for the voxel selected as representative of the cluster 7



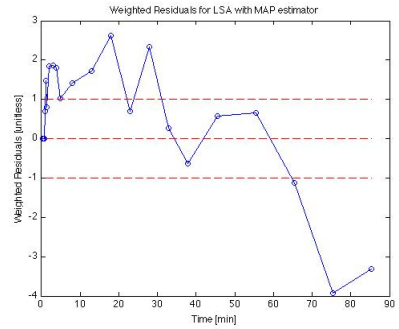
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.12: Weighted residuals for all methods tested on this voxel representative of cluster 7.

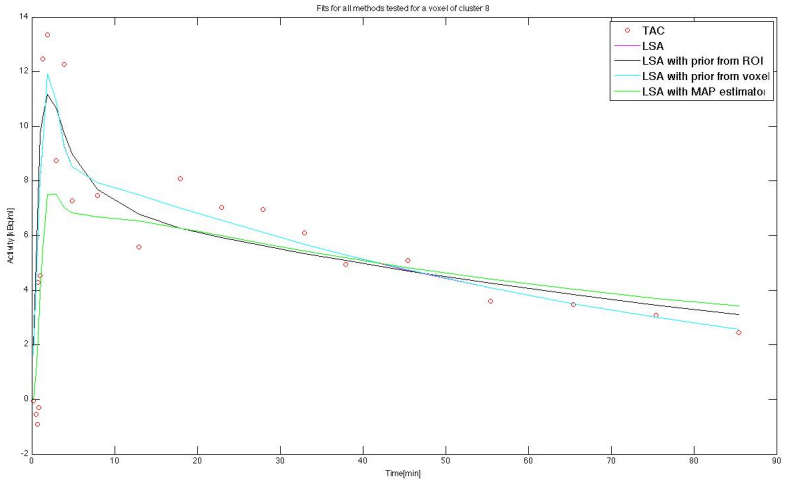
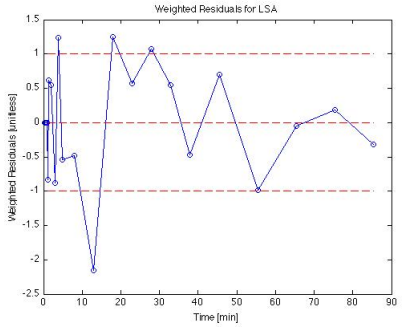
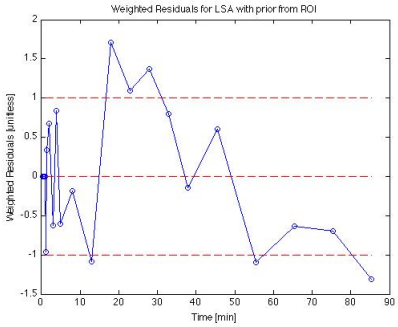


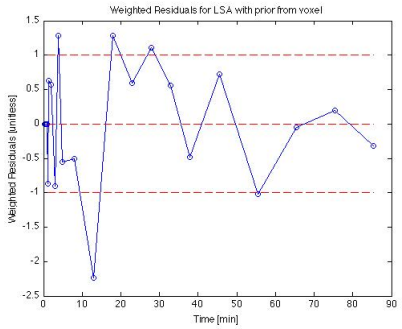
Figure A.13: Fits for all the methods tested for the voxel selected as representative of the cluster 8



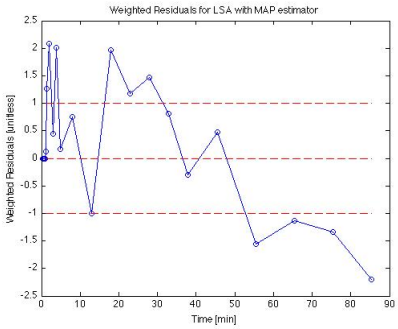
(a) *Weighted residuals for LSA standard.*



(b) *Weighted residuals for LSA with prior from ROI.*



(c) *Weighted residuals for LSA with prior from voxel.*



(d) *Weighted residuals for LSA implemented with MAP estimator.*

Figure A.14: Weighted residuals for all methods tested on this voxel representative of cluster 8.

Bibliography

Bertoldo A., Vicini P., Sambuceti G., Lammertsma A.A., Parodi O. and Cobelli C. Evaluation of compartmental and spectral analysis models of [18F]FDG kinetics for heart and brain studies with PET. IEEE transactions on bio-medical engineering, 1998.

Budinger T.F. PET instrumentation: What are the limits? Seminars in Nuclear Medicine, 1998.

Carson E.R., Cobelli C. and Finkelstein L. The mathematical modeling of metabolic and endocrine systems : model formulation, identification, and validation. J. Wiley, 1983.

Cunningham V., Ashburner J., Byrne H. and Jones T. Use of spectral analysis to obtain parametric images from dynamic PET studies. (In) Quantification of Brain Function. Tracer Kinetics and Image Analysis in Brain PET. Elsevier Science Publishers, 1993.

Cunningham V. and Jones T. Spectral analysis of dynamic PET studies. Journal of Cerebral Blood Flow & Metabolism, 1993.

Deans S.R. The Radon Transform and Some of Its Applications. John Wiley & Sons, 1983.

DiStefano J.J. Optimized blood sampling protocols and se quential design of kinetic experiments. American Journal of Physiology - Regulatory, Integrative and Comparative Physiology, 1981.

Feng D., Hu D., Chen K., Wu L.C., Wang J.K., Liu R.S. and Yeh S.H. An evaluation of the algorithms for determining local cerebral metabolic rates of glucose using positron emission tomography dynamic data. IEEE Transactions on Medical Imaging, 1995.

Gunn R.N., Gunn S.R. and Cunningham V.J. Positron emission tomography compartmental models. Journal of Cerebral Blood Flow & Metabolism, 2001.

Gunn R.N., Gunn S.R., Turkheimer F.E., Aston J.A.D. and Cunningham V.J. Positron emission tomography compartmental models: A basis pursuit strategy for kinetic modelling. Journal of Cerebral Blood Flow & Metabolism, 2002.

Gunn R.N., Sargent P.A., Bench C.J., Rabiner E.A., Osman S., Pike V.W., Hume S.P., Grasby P.M. and Lammertsma A.A. Tracer kinetic modeling of the 5-HT_{1A}receptor ligand [carbonyl-11C]WAY-100635 for PET. NeuroImage, 1998.

Gurevich E.V. and Joyce J.N. Alterations in the cortical serotonergic system in schizophrenia: A postmortem study. Biological psychiatry, 1997.

Hammers A., Allom R., Koepp M.J., Free S.L., Myers R., Lemieux L., Mitchell T.N., Brooks D.J. and Duncan J.S. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Human Brain Mapping, 2003.

Hebert T. and Leahy R. A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. IEEE Transactions on Medical Imaging, 1989.

Hevesy G.v. Adventures in Radioisotope Research. New York, Pergamon Press, 1962.

Kety S.S. The theory and applications of the exchange of inert gas at the lungs and tissues. Pharmacological Reviews, 1951.

Lammertsma A.A., Bench C.J., Hume S.P., Osman S., Sargent P., McCarron J.A., Pike V.W. and Grasby P.M. Kinetic analysis of [C-11]WAY-100635 studies. European Journal of Nuclear Medicine, 1996.

Logan J., Fowler J.S., Volkow N.D., Wolf A.P., Dewey S.L., Schlyer D.J., MacGregor R.R., Hitzemann R., Bendriem B., Gatley S.J. and Christman D.R. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [n- 11c-methyl]-(âĹŠ)-cocaine pet studies in human subjects. Journal of Cerebral Blood Flow & Metabolism, 1990.

MacQueen J.B. Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 1967.

Maltz J.S. Parsimonious basis selection in exponential spectral analysis. Physics in Medicine & Biology, 2002.

Mintun M.A., Raichle M.E., Kilbourn M.R., Wooten G.F. and Welch M.J. A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. Annals of Neurology, 1984.

Nilsson L.G. and Markowitsch H.J. Cognitive Neuroscience of Memory. Hogrefe & Huber Publishers, 1999.

Ollinger J.M. and Fessler J.A. Positron emission tomography. IEEE Signal Processing Magazine, 1997.

Osman S., Lundkvist C., Pike V.W., Halldin C., McCarron J.A., Swahn C., Ginovart N., Luthra S.K., Bench C.J., Grasby P.M., Wikström H., Barf T., Cliffe I.A., Fletcher A. and Farde L. Characterization of the radioactive metabolites of the 5-HT1A receptor radioligand, [o-methl-11c]WAY-100635, in monkey and human plasma by HPLC: Comparison of the behaviour of an identified radioactive metabolite with parent radioligand in monkey using PET. Nuclear Medicine and Biology, 1996.

Parsey R.V., Slifstein M., Hwana D.R., Abi-Dargham A., Simpson N., Mawlawi O., Guo N.N., Heertum R., Mann J.J. and Laruelle M. Validation and reproducibility of measurement of 5-HT1A receptor parameters with [carbonyl-11C]WAY-100635 in humans: comparison of arterial and reference tisssue input functions. Journal of Cerebral Blood Flow & Metabolism, 2000.

Patlak C.S., Blasberg R.G. and Fenstermacher J.D. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Journal of Cerebral Blood Flow & Metabolism, 1983.

Phelps M.E., Huang S.C., Hoffman E.J., Selin C., Sokoloff L. and Kuhl D.E. Tomographic measurement of local cerebral glucose metabolic rate in humans with (f-18)2-fluoro-2-deoxy-d- glucose: validation of method. Annals of Neurology, 1979.

Pietrzyk U., Herholz K., Schuster A., Stockhausen H.M.v., Lucht H. and Heiss W.D. Clinical applications of registration and fusion of multimodality brain images from PET, SPECT, CT, and MRI. European Journal of Radiology, 1996.

Reivich M., Kuhl D., Wolf A., Greenberg J., Phelps M., Ido T., Casella V., Fowler J., Hoffman E., Alavi A., Som P. and Sokoloff L. The [18F]fluorodeoxyglucose method for the measurement of local cerebral glucose utilization in man. Circulation Research, 1979.

Rizzo G., Turkheimer F.E., Keihaninejad S., Bose S.K., Hammers A. and Bertoldo A. Multi-scale hierarchical generation of PET parametric maps: Application and testing on a [11C]DPN study. NeuroImage, 2012.

Schmidt K. Which linear compartmental systems can be analyzed by spectral analysis of pet output data summed over all compartments? Journal of Cerebral Blood Flow & Metabolism, 1999.

Shepp L. and Vardi Y. Maximum likelihood reconstruction for emission tomography. IEEE Transactions on Medical Imaging, 1982.

Sokoloff L., Reivich M., Kennedy C., DesRosiers M.H., Patlak C.S., Pettigrew K.D., Sakurada O. and Shinohara M. The 14C-deoxyglucose method for the measurement of local cerebral glucose utilisation: theory, procedure and normal values in the conscious and anaesthetized albino rat. Journal of Neurochemistry, 1977.

Steinbusch H.W.M. Distribution of serotonin-immunoreactivity in the central nervous system of the rat—Cell bodies and terminals. Neuroscience, 1981.

Turkheimer F.E., Moresco R.M., Lucignani G., Sokoloff L., Fazio F. and Schmidt K. The use of spectral analysis to determine regional cerebral glucose utilization with positron emission tomography and [18F]Fluorodeoxyglucose: theory, implementation and optimization procedures. Journal of Cerebral Blood Flow & Metabolism, 1994.

Turkheimer F.E., Sokoloff L., Bertoldo A., Lucignani G., Reivich M., Jaggi J.L. and Schmidt K. Estimation of component and parameter distributions in spectral analysis. Journal of Cerebral Blood Flow & Metabolism, 1998.

Volkow N.D., Fowler J.S., Gatley S.J., Logan J., Wang G.J., Ding Y.S. and Dewey S. PET evaluation of the dopamine system of the human brain. Journal of nuclear medicine, 1996.

Wright D.E., Seroogy K.B., Lundgren K.H., Davis B.M. and Jennes L. Comparative localization of serotonin 1A, 1C, and 2 receptor subtype mRNAs in rat brain. The Journal of Comparative Neurology, 1995.

Acknowledgments

The first thanks, coming from the heart, goes to Professor Alessandra Bertoldo. Who beside the unquestionable preparation, has always shown me patience, attention and gave me the opportunity to develop part of my thesis at the Imperial College of London. This wonderful experience enabled me to enjoy a different lifestyle, expanding my professional knowledge and improving my English, while surrounded by great people with a huge passion for their work combined with an unquestioned ability to make you feel at " home". Therefore, special thanks to Federico Turkheimer, Gianpaolo Tomasi, Elisabetta Grecchi and Shazi Singh for their hospitality, for their help and friendship.

Gaia and Mattia, THANKS! Thanks for your patience, thanks for the help and thanks for the understanding you have shown me.

To my parents, Nori and Nino. You are the parents who all children should have! Your only concern is the welfare, health and success of your daughters. I think you can be proud of you, you succeeded. I sincerely hope to always be able to make you proud of me because, at any place in the world where I will be, your teachings will accompany me and help me build my future on healthy principles and important values. You have my gratitude forever.

To my sister, who over the years has put up with my changes, my dark moments and my wickedness and yet she has always been close to me! You're the best sister one could wish to have. I hope that despite the miles that may separate us in life, we still maintain a relationship of love, esteem and mutual support!

To Simone, who despite being entered into my life recently has already won my heart completely. I think the way we love each other is special, giving oneself totally to the other to receive everything in return. I hope this will be the theme of the future that awaits us. Thanks for your patience and understanding, you have been a key support in recent months that have been so difficult.

To Andrea, that in the college years spent together helped me to move on, to hold on and to make me understand many of my faults. If I became what I am is also because of you.

Jeannine! Dear Jeannine! She came into my life for some coincidences, but it was a gift to know her. She has supported, promoted and encouraged me, and now she is here to fix my thesis! Thanks will never be enough!

To all my relatives and my friends, people who believed in me and took me to this path, THANKS!

Il primo grazie che mi nasce dal cuore va alla Professoressa Alessandra Bertoldo, che oltre all'indiscutibile preparazione, mi ha sempre dimostrato pazienza ed attenzione e mi ha dato la possibilità di svolgere parte della mia tesi presso l'Imperial College di Londra. Quest'esperienza meravigliosa mi ha consentito di vivere una realtà diversa dalla nostra, di allargare il mio bagaglio culturale e di migliorare il mio inglese, il tutto circondata da persone fantastiche con una passione enorme per il loro lavoro ed un'indiscussa capacità di farti sentire a "casa". Un grazie speciale quindi anche a Federico Turkheimer, Gianpaolo Tomasi, Elisabetta Grecchi e Shazi Singh per la loro accoglienza, per il loro aiuto e per la loro amicizia.

Gaia e Mattia, GRAZIE! Grazie per la pazienza, grazie per l'aiuto e grazie per la comprensione che mi avete dimostrato.

Ai miei genitori, Nori e Nino. Siete i genitori che tutti i figli dovrebbero avere; l'unica vostra preoccupazione è il benessere, la salute ed il successo delle vostre figlie. Credo che ci siate riusciti alla grande. Spero di cuore di essere sempre capace di rendervi orgogliosi di me perché, in qualsiasi posto al mondo sarò, i vostri insegnamenti mi accompagneranno e mi aiuteranno a costruire il mio futuro su principi sani e valori importanti. Avrete per sempre la mia gratitudine.

A mia sorella, che negli anni ha sopportato i miei cambiamenti, i miei momenti bui e le mie cattiverie e nonostante questo mi è sempre stata vicina! Sei la sorella migliore che si possa desiderare e spero che nonostante i chilometri che probabilmente ci separeranno nella vita, riusciremmo sempre a mantenere un rapporto di amore, di stima e di aiuto reciproco!

A Simone, che nonostante sia entrato nella mia vita da poco tempo ha già conquistato interamente il mio cuore. Credo che il nostro modo di amare sia speciale, donarsi completamente all'altro per ricevere tutto in cambio. Spero sarà questo il filo conduttore del futuro che ci aspetta. Grazie della pazienza e della comprensione, sei stato un sostegno fondamentale in questi mesi cosí difficili.

Ad Andrea, che negli anni di università trascorsi assieme mi ha aiutata ad andare avanti, a tener duro e a capire tanti miei difetti. Se sono diventata ciò che sono è anche per merito tuo.

Jeannine! Cara Jeannine! Sei entrata nella mia vita per alcune coincidenze, ma è stato un dono conoscerti. Mi hai sostenuta, incoraggiata e spronata ed ora sei qui che mi correggi la tesi! Un grazie non sarà mai abbastanza!

A tutti i miei parenti ed i miei amici, alle persone che hanno creduto in me e mi hanno accompagnata in questo percorso, GRAZIE!