

# CHALMERS



## Accurate segmentation of brain MR images

*Master of Science Thesis in Biomedical Engineering*

ANTONIO REYES PORRAS PÉREZ

Department of Signals and Systems  
*Division of Biomedical Engineering*  
CHALMERS UNIVERSITY OF TECHNOLOGY  
Göteborg, Sweden, 2010  
Report No. EX028/2010



## **Abstract**

Full brain segmentation has been of significant interest throughout the years. Recently, many research groups worldwide have been looking into development of patient-specific electromagnetic models for dipole source location in EEG. To obtain this model, accurate segmentation of various tissues and sub-cortical structures is thus required.

In this project, the performance of three of the most widely used software packages for brain segmentation has been analyzed: FSL, SPM and FreeSurfer. For the analysis, real images from a patient and a set of phantom images have been used in order to evaluate the performance of each one of these tools.

**Keywords:** dipole source location, brain, patient-specific model, image segmentation, FSL, SPM, FreeSurfer.



## Acknowledgements

To my advisor, Antony, for his guidance through the project.

To my partner, Koushyar, for all the days we have spent in the hospital helping each other.

To the staff in Sahlgrenska hospital for their collaboration.

To MedTech West for this opportunity to learn.



# Table of contents

---

1.	Introduction .....	1
2.	Magnetic resonance imaging.....	3
3.	Brain images segmentation .....	8
3.1.	FSL.....	9
3.1.1.	Brain Extraction .....	9
3.1.2.	Image registration .....	10
3.1.3.	Brain tissue segmentation.....	11
3.1.4.	Sub-cortical segmentation .....	12
3.2.	SPM.....	12
3.3.	FreeSurfer .....	13
4.	Methodology .....	15
4.1.	FSL.....	16
4.1.1.	Brain extraction.....	16
4.1.2.	Image registration .....	17
4.1.3.	Brain tissue segmentation.....	18
4.1.4.	Sub-cortical segmentation .....	18
4.2.	SPM.....	19
4.3.	FreeSurfer .....	20
5.	Experiments with real data.....	22
5.1.	FSL.....	22
5.2.	SPM.....	26
5.3.	FreeSurfer .....	28
6.	Software comparison.....	30
6.1.	Comparison results .....	30
6.1.1.	FSL .....	30
6.1.2.	SPM .....	37
6.1.3.	FreeSurfer.....	43
6.2.	Images.....	50
6.2.1.	FSL .....	50
6.2.2.	SPM .....	52
6.2.3.	FreeSurfer.....	53
7.	Conclusions .....	55
8.	Bibliography.....	57



# Index of figures

---

Figure 1. Spherical model. ....	1
Figure 2. Patient-specific model. ....	1
Figure 3. Nuclei spinning around it axis. ....	3
Figure 5. Low energy state. ....	3
Figure 4. High energy state. ....	3
Figure 6. Precession. ....	4
Figure 7. Motion of Magnetization vector. ....	5
Figure 8. Generating spin echoes. ....	6
Figure 9. T1-weighted image: axial(top-left),sagittal(top-right) and coronal(bottom). ....	9
Figure 10. T2-weighted image: axial(top-left),sagittal(top-right) and coronal(bottom). ....	9
Figure 11. Brain mesh. ....	10
Figure 12. FAST's Gaussians model. ....	10
Figure 13. Bias field effect. ....	11
Figure 14. BET: -B working mode: axial(top-left),sagittal(top-right) and coronal(bottom). ....	22
Figure 15. BET: -R working mode: axial(top-left),sagittal(top-right) and coronal(bottom). ....	23
Figure 16. BET: -A2 working mode: axial(top-left),sagittal(top-right) and coronal(bottom). ....	23
Figure 17. BET: -A2 working mode with -c option: axial(top-left),sagittal(top-right) and coronal(bottom). ....	24
Figure 18. BET: -A2 working mode with -c and -f=0.6 options: axial(top-left),sagittal(top-right) and coronal(bottom). ....	24
Figure 19. FAST: segmented brain: axial(top-left),sagittal(top-right) and coronal(bottom). ....	25
Figure 20. FIRST: results (b) ....	25
Figure 21. FIRST: results (a) ....	25
Figure 22. SPM: CSF using T1: axial(top-left),sagittal(top-right) and coronal(bottom). ....	26
Figure 23. SPM: CSF using T2: axial(top-left),sagittal(top-right) and coronal(bottom). ....	27
Figure 24. SPM: white matter using T1: axial(top-left),sagittal(top-right) and coronal(bottom). ....	27
Figure 25. SPM: white matter using T2: axial(top-left),sagittal(top-right) and coronal(bottom). ....	27
Figure 26. SPM: grey matter using T1: axial(top-left),sagittal(top-right) and coronal(bottom). ....	28
Figure 27. grey matter using T2: axial(top-left),sagittal(top-right) and coronal(bottom). ....	28
Figure 28. FreeSurfer segmentation: axial(bottom-right),sagittal(top) and coronal(bottom-left). ....	29
Figure 29. FSL: White matter specificity. ....	31
Figure 30. FSL: White matter sensitivity. ....	31
Figure 31. FSL: White matter f-factor. ....	32
Figure 32. FSL: Grey matter specificity. ....	33
Figure 33. FSL: Grey matter sensitivity. ....	33
Figure 34. FSL: Grey matter f-factor. ....	34
Figure 35. FSL: CSF specificity. ....	34
Figure 36. FSL: CSF sensitivity. ....	35
Figure 37. FSL: CSF f-factor. ....	35
Figure 38. FSL: Error ratio (all voxels). ....	36
Figure 39. FSL: Error ratio (non-empty voxels). ....	37
Figure 40. SPM: white matter specificity. ....	38
Figure 41. SPM: white matter sensitivity. ....	38
Figure 42. SPM: white matter f-factor. ....	39
Figure 43. SPM: grey matter specificity. ....	39
Figure 44. SPM: grey matter sensitivity. ....	40

Figure 45. SPM: grey matter f-factor.....	40
Figure 46. SPM: CSF specificity.....	41
Figure 47. SPM: CSF sensitivity.....	42
Figure 48. SPM: CSF f-factor.....	42
Figure 49. SPM: Error ratio (all voxels).....	43
Figure 50. SPM: Error ratio (non-empty voxels).....	43
Figure 51. FreeSurfer: White matter specificity.....	44
Figure 52. FreeSurfer: White matter sensitivity.....	44
Figure 53. FreeSurfer: White matter f-factor.....	45
Figure 54. FreeSurfer: Grey matter specificity.....	46
Figure 55. FreeSurfer: Grey matter sensitivity.....	46
Figure 56. FreeSurfer: Grey matter f-factor.....	47
Figure 57. FreeSurfer: CSF specificity.....	47
Figure 58. FreeSurfer: CSF sensitivity.....	48
Figure 59. FreeSurfer: CSF f-factor.....	48
Figure 60. FreeSurfer: Error ratio (all voxels).....	49
Figure 61. FreeSurfer: Error ratio (non-empty voxels).....	49
Figure 62. FSL: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	50
Figure 63. FSL: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	51
Figure 64. FSL: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	51
Figure 65. SPM: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	52
Figure 66. SPM: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	52
Figure 67. SPM: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	53
Figure 68. FreeSurfer: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	53
Figure 69. FreeSurfer: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	54
Figure 70. FreeSurfer: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).....	54

# Index of tables

---

Table 1. BET working modes. ....	17
Table 2. BET options. ....	17
Table 3. FAST options. ....	18
Table 4. FIRST options. ....	18
Table 5. SPM segmentation parameters. ....	19
Table 6. FreeSurfer processing stages. ....	20
Table 7. FreeSurfer stage groups. ....	20
Table 8. Phantom parameters. ....	30
Table 9. FSL: White matter specificity. ....	31
Table 10. FSL: White matter sensitivity. ....	31
Table 11. FSL: White matter f-factor. ....	32
Table 12. FSL: Grey matter specificity. ....	32
Table 13. FSL: Grey matter sensitivity. ....	33
Table 14. FSL: Grey matter f-factor. ....	33
Table 15. FSL: CSF specificity. ....	34
Table 16. FSL: CSF sensitivity. ....	35
Table 17. FSL: CSF f-factor. ....	35
Table 18. FSL: Error ratio (all voxels). ....	36
Table 19. FSL: Error ratio (non-empty voxels). ....	36
Table 20. SPM: white matter specificity. ....	37
Table 21. SPM: white matter sensitivity. ....	38
Table 22. SPM: white matter f-factor. ....	38
Table 23. SPM: grey matter specificity. ....	39
Table 24. SPM: grey matter sensitivity. ....	40
Table 25. SPM: grey matter f-factor. ....	40
Table 26. SPM: CSF specificity. ....	41
Table 27. SPM: CSF sensitivity. ....	41
Table 28. SPM: CSF f-factor. ....	42
Table 29. SPM: Error ratio (all voxels). ....	43
Table 30. SPM: Error ratio (non-empty voxels). ....	43
Table 31. FreeSurfer: White matter specificity. ....	44
Table 32. FreeSurfer: White matter sensitivity. ....	44
Table 33. FreeSurfer: White matter f-factor. ....	45
Table 34. FreeSurfer: Grey matter specificity. ....	45
Table 35. FreeSurfer: Grey matter sensitivity. ....	46
Table 36. FreeSurfer: Grey matter f-factor. ....	46
Table 37. FreeSurfer: CSF specificity. ....	47
Table 38. FreeSurfer: CSF sensitivity. ....	48
Table 39. FreeSurfer: CSF f-factor. ....	48
Table 40. FreeSurfer: Error ratio (all voxels). ....	49
Table 41. FreeSurfer: Error ratio (non-empty voxels). ....	49



## 1. Introduction

In recent years, the study of the brain behavior has produced an increasing interest in the location of the source of the electrical signals generated. It is important in the research of some brain diseases (such as epilepsy) to know with more accuracy about the nature of the disease and the exact location of where the problem is in the brain.

Some of these diseases are studied using tools such as electro-encephalography (EEG), which is able to measure the signals generated in the brain by putting some electrodes in the patient's head. In this sense, EEG technique can be invasive or non-invasive. The invasive method needs surgery and consists in locating the electrodes directly in the patient brain. The non-invasive method, on the other hand, aims to locate the electrodes over the patient head. It is straightforward to conclude that the non-invasive method is more sensitive to the external noise and that the invasive method could be dangerous for the patient. In addition, the location of the signal is much more accurate when using the invasive method.

However, the invasive method should only be used in that cases when there are no other options. For that reason, many researchers in the world are working in order to locate accurately the exact location of the signals received by using non-invasive EEG.

In the context of dipole source localization, nowadays, a spherical model for the human head and brain is being used in most studies (see Figure 1). In practice, however, people do not have spherical heads and the head size is different for everyone, there are many in-accuracies when trying to find out the exact point in the brain where one signal has been generated. For that reason, it is necessary to have a patient-specific model which let locate exactly the point where signals are generated, by using the information acquired with the electrodes located over the patient head (inverse problem).

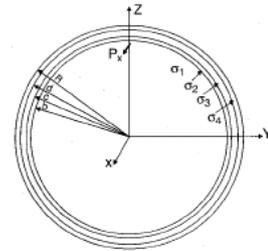


Figure 1. Spherical model.

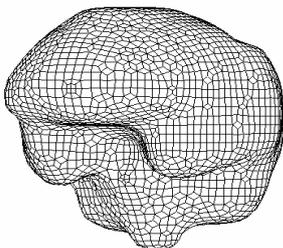


Figure 2. Patient-specific model.

Now the need of having a model for the patient head and brain has been set, it is necessary to build this model. For that purpose, we have to obtain the information from inside the patient head with a non-invasive technique. Magnetic Resonance Imaging (MRI) is a good alternative, because it lets obtain images of the patient head and brain, and it doesn't produce secondary effects. Thus, we need to apply some methods which allow obtaining as much information as possible from inside the patient head in order to perform its reconstruction. In [1] we can see that the most important parts which should be extracted from the brain are the scalp, bone, cerebrospinal fluid, grey matter and white matter.

In the last few years, many tools for reconstructing the human brain from magnetic resonance images have been developed. These tools have a very high complexity and needed many years for their development. In addition, they can obtain good results for our objective, so it would not be very useful to start developing new software and it would be better to use

one of them. The goal of this project is, therefore, to analyze the performance of these tools to obtain the best possible reconstruction of the patient head.

For this purpose, we will first make an introduction about the basics of MRI of how to obtain images. This could help us to better understand the nature of the images we have as input. After that, a brief explanation of what to segment an image is will be shown and the most popular brain segmentation tools will be analyzed in detail.

## 2. Magnetic resonance imaging

Magnetic Resonance Imaging (MRI) is a technique used to obtain high quality images of the inside of the human body. It is based on the radio-frequency waves that the protons in the examined tissues emit when exposed to an external magnetic field. Each signal is then processed by advanced computer programs, which transform it into high quality images. Unlike conventional x-ray systems and procedures of nuclear medicine, this kind of technique does not emit ionizing radiations.

In order to produce images without using ionizing radiations, the patient is undergone to a magnetic field thousands times bigger than the one produced by our planet. Thus, a powerful magnet attracts the protons inside the (usually) hydrogen atoms in the human tissues. These protons, when stimulated with radio-frequency waves, get out from their normal alignment. However, when this stimulus is stopped, protons come back to their original position, releasing energy that we can transform in radio signals in order to detect them with a computer. Then, we can transform these signals in images which describe the shape of organs.

Magnetic resonance imaging is based on a physical phenomenon called *nuclear magnetic resonance*, which is related to the nuclei of atoms. Since it is composed by neutrons and protons, a nucleus has a positive charge. Moreover, all nuclei with an odd atomic number or mass number have an angular momentum, so they are said to have spin. These nuclei which have spin are the ones we consider in this technique. The most used nuclei for magnetic resonance imaging are <sup>1</sup>H, <sup>13</sup>C and <sup>19</sup>F, because they are the most relevant in biological systems.

We can think of a nucleus with spin as a magnetic moment vector that causes the nucleus to behave just like a magnet, with one north pole and one south pole. We can explain this fact because a nucleus has a positive charge and, if it is spinning around an axis, then we have circulating charges which produce a magnetic field. We can calculate the magnetic moment vector with the following expression:

$$\mu = \gamma\Phi$$

where  $\gamma$  is the gyromagnetic ratio in radians per second per Tesla.

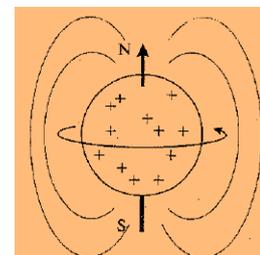


Figure 3. Nuclei spinning around its axis.

In normal conditions, all nuclei are randomly oriented in a given sample. Even though each nucleus has a microscopic spin, the total macroscopic spin has no magnetic field, since each microscopic spin cancel each other.

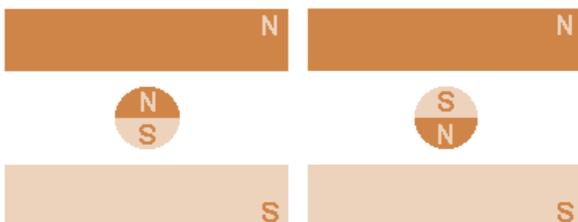


Figure 5. High energy state.

Figure 4. Low energy state.

all the microscopic spins align with this, causing the spin system to become magnetized. For each one of the nuclei, there are a high energy configuration and a low energy configuration, depending on the poles alignment, as seen in Figure 4 and Figure 5 respectively.

The applied external field can be represented as a  $B_0$  vector, which can be calculated with the product of the magnitude of the magnetic field and a vector pointing in the positive Z direction within a fixed frame.

$$B_0 = B_0 \cdot \hat{z}$$

Each one of the nuclei we consider in MRI has a spin quantum number, which takes a positive value multiples of  $\frac{1}{2}$ . For this systems, also called spin  $\frac{1}{2}$  systems, there are two possible orientations for the microscopic magnetizations: 54 degrees off  $\hat{z}$  (low energy state) or 126 degrees off  $\hat{z}$ , with a preference for the low energy state. In addition, the fact that the orientation of the magnetic moment vector  $\mu$  around the Z axis is random together with everything exposed above make the system to become magnetized in the  $\hat{z}$  direction. Thus, this system can be modeled with a magnetization vector M, which is the sum of all the individual nuclear magnetic moments.

$$M = \sum_{n=1}^N \mu_n$$

If the applied external magnetic field is constant for a period of time, the magnetization vector will reach an equilibrium value  $M_0$  parallel to  $B_0$ :

$$M_0 = \frac{B_0 \gamma^2 \hbar^2}{4\kappa T} P_D$$

where  $\kappa$  is the Boltzmann's constant, T is the temperature in Kelvin and  $P_D$  is the proton density. Here we can point that the larger is  $M_0$ , the larger is the received NMR signal.

As we have seen, the magnetization vector M is a function of time and can be defined in a three-dimensional coordinate system. So, as occurs with the microscopic angular momentum, we can define an angular momentum J associated to the system, which can be related to the magnetization as follows:

$$M = \gamma J$$

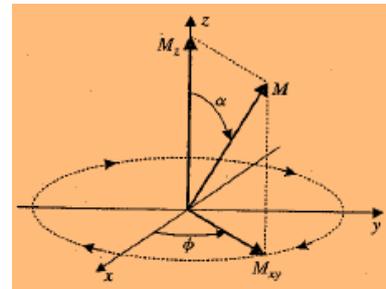


Figure 6. Precession.

If we apply a time-varying external magnetic field, we can set a relation between the external field, the magnetization vector and the angular momentum associated with the magnetization vector:

$$\frac{dJ(t)}{dt} = M(t) \times B(t)$$

From this expression, we can derive the equations of precession, which can be written as:

$$M_x(t) = M_0 \sin(\alpha) \cos(-2\pi\nu_0 t + \phi)$$

$$M_y(t) = M_0 \sin(\alpha) \sin(-2\pi\nu_0 t + \phi)$$

$$M_z(t) = M_0 \cos(\alpha)$$

where  $\nu_0$  is the Larmor frequency in hertz and can be calculated as  $\nu_0 = \gamma B_0$ . The precession of the magnetization vector around the external magnetic field  $\hat{z}$  axis can be shown in the Figure 6.

We could think that the Larmor frequency should be constant for a spin system. However, since the magnetic field is not constant due to magnetic field inhomogeneities, magnetic susceptibility and chemical shift, it's not really constant.

In Figure 6 we can define two components for the magnetization vector. We can call longitudinal component to the one that is oriented in the axis defined by the external magnetic field ( $M_z$ ), and transverse component to the other component, which is oriented in the XY plane ( $M_{XY}$ ). We can see the transverse magnetization as a complex number defined by  $M_{XY}(t) = M_x(t) + jM_y(t)$ , where the phase is given by  $\phi(t) = \tan^{-1} \frac{M_x(t)}{M_y(t)}$ .

If we place a coil of wire outside the sample, the rotating transverse magnetization will induce a voltage (Faraday's law), which can be measured and recorded for use in MRI. If we assume that the object is homogeneous, the field produced by the coil when excited is uniform and the time derivative of the z component of magnetization can be ignored, we can reach the following expression for the voltage induced in the coil:

$$V(t) = -2\pi\nu_0 V_s M_0 \sin(\alpha B^r) \sin(\omega t - 2\pi\nu_0 t + \phi + \theta_r)$$

Sometimes, it's better to analyze the evolution of the magnetization vector in a frame of reference which rotates at the Larmor frequency. So, the new rotating vector would be stationary in this new frame, called rotating frame. Its coordinates could be defined as follows:

$$\begin{aligned} x' &= x \cos(2\pi\nu_0 t) - y \sin(2\pi\nu_0 t) \\ y' &= x \sin(2\pi\nu_0 t) + y \cos(2\pi\nu_0 t) \\ z' &= z \end{aligned}$$

Consequently, the magnetization vector can be expressed as  $M_{x'y'}(t) = M_0 \sin\alpha e^{j\phi}$ .

As explained before, for M to precess it has to be oriented away from  $B_0$ . Thus, we can use a RF current through an antenna placed surrounding the sample, which causes the systems to be excited. In this way, we can generate circularly polarized RF excitations using quadrature RF coils. The field generated can be modeled as follows:

$$B_1(t) = B_1^e(t) e^{-j(2\pi\nu_0 t - \varphi)} \quad (\text{Fixed frame})$$

$$B_1(t) = B_1^e(t) e^{j\varphi} \quad (\text{Rotating frame})$$

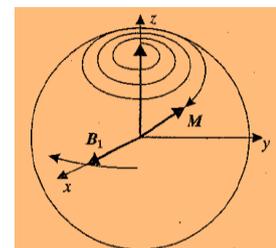


Figure 7. Motion of Magnetization vector.

With this approach, the magnetization vector will precess in the  $y'-z'$  plane (forced precession). Adding this precession to the one around the Z axis, we have that the evolution of  $M(t)$  is a spiral as shown in Figure 7. There, the tip angle and phase depend on the amplitude and duration of  $B_1(t)$ . So, if we turn off the pulse when M has precessed into the transverse plane, the pulse will be called a  $\frac{\pi}{2}$  pulse. Otherwise, if we turn off the pulse when M has precessed along the  $-z$  axis, we call it a  $\pi$  pulse. Anyway, the final tip angle can be calculated as  $\alpha = \gamma \int_0^{\tau_p} B_1^e(t) dt$ .

After a  $\alpha$ -pulse, M will precess because of the presence of  $B_0$ . But this precession is not infinite; there are two relaxation processes that cause the signal to vanish:

- Transverse relaxation: is caused by the perturbations in the magnetic field due to the closest spins. Because of these perturbations, a dephasing appears so that there is a loss in the coherence of the wave produced by the spin system and the signal is lost. The received signal is called Free Induction Decay (FID) and decreases exponentially with a time constant  $T_2$ , which is different for each tissue.
- Longitudinal relaxation: it's related to the recovery of  $M_z(t)$  to its equilibrium value  $M_0$ . It can be modeled as an increasing exponential with a time constant  $T_1$ .

Because of the dephasing between spins due to transverse relaxation, another kind of signals called spin echoes appear. These are signal generated by the transverse spin while recovering their coherence, followed by a deliberate  $\pi$ -pulse. The time between the  $\frac{\pi}{2}$ -pulse and the  $\pi$ -pulse is what we call echo time. We can control it because the difference in time between the application of both pulses is exactly the half of the echo time.

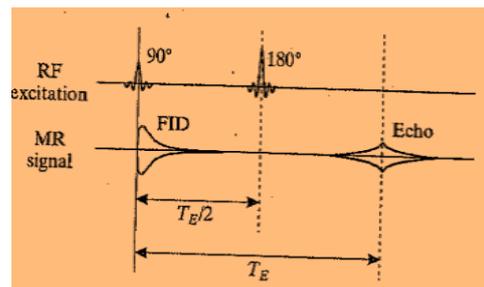


Figure 8. Generating spin echoes.

It is possible to use an equation, called Bloch equation, which let us predict the behavior of the magnetic spin system:

$$\frac{dM(t)}{dt} = \gamma M(t) \times B(t) - R(M(t) - M_0)$$

where R is called relaxation matrix and can be defined as  $R = \begin{pmatrix} 1/T_2 & 0 & 0 \\ 0 & 1/T_2 & 0 \\ 0 & 0 & 1/T_1 \end{pmatrix}$ .

We can use three parameters in order to distinguish different tissues when we want to image a sample with more than one tissue:

- Proton density: the image intensity is proportional to the number of hydrogen nuclei in the sample. We start with the sample in equilibrium by applying a RF pulse and image before transverse relaxation.  $T_R$  can be defined as the time between excitations and must be long enough to make sure the system to be in equilibrium before the pulse. It's better to use  $\frac{\pi}{2}$ -pulses to have a larger signal.

- T<sub>2</sub>-weighted contrast: echoes can also be used to obtain images, taking profit of the different T<sub>2</sub> for each tissue. We have to set a T<sub>E</sub> almost equal to T<sub>2</sub> for the tissues imaged.
- T<sub>1</sub>-weighted contrast: images are obtained by exciting the tissues repeatedly before they can recover their longitudinal magnetizations. If we set T<sub>R</sub> = T<sub>1</sub>, the transverse component vanishes but the longitudinal component doesn't. By applying  $\frac{\pi}{2}$ -pulses, we can detect a larger signal for those tissues with a shorter T<sub>1</sub>.

In this project, we have used T1 and T2-weighted images to evaluate the accuracy of the software packages we have studied. Therefore, our starting point is a set of magnetic resonance images which have been used to perform brain segmentation.

### 3. Brain images segmentation

Once we have obtained all data we need from the patient brain, the next step is about using these data to extract the desired information. At this point, it is very important to have a clear idea about what outcome we really want to obtain. As we will see later, the selection of the segmentation method (and the software to use) will depend on the results we need.

Since all we have as input is a series of images of the brain, we should define accurately what an image is. An image can be defined as a collection of values in a 2-dimensional or 3-dimensional space. Thus, when talking about magnetic resonance images, these values represent radio-frequency signal intensities. In addition, due to all images are obtained in a discrete domain, we can call the position of every one of these values as pixel (in 2D imaging) or voxel (in 3D imaging).

Image segmentation is about partitioning an image in non-overlapped regions, which are homogeneous regarding to a specific characteristic such as an intensity level or a texture. Therefore, if the image domain is considered to be  $I$ , then the segmentation problem falls on finding out a set  $S_k \subset I$  which can be used to reconstruct the  $I$  image so that:

$$I = \bigcup_{k=1}^K S_k$$

where  $S_k \cap S_j = \phi$ ,  $k \neq j$  and all points in  $S_k$  are connected. Ideally, a good method for brain segmentation finds those sets which correspond to different regions or structures that form the brain.

When removing the restriction for all points of the same set to be connected we talk about classification instead of segmentation, and sets are also called classes. To determine the number can be a very complex task and sometimes it is set basing on previous knowledge about the images. Furthermore, the process of assigning one label to a segment can be carried out separately from the segmentation process and it's called labeling.

At this point, we should ask ourselves about which sets we could want to identify in brain images. As we concluded from the previous chapters, it is very important for us to distinguish between brain white matter, grey matter and cerebrospinal fluid. It would also be desirable to extract the bone and skin/scalp from the patient brain images so that the patient specific model can be built.

Since it can be very difficult to obtain accurate brain tissue segmentation, there has been an intense research in this area in the last two decades. As result, many universities and research groups worldwide have developed different software packages which can perform brain segmentation and so extract tissues and structures from patient images. In the next sections we will briefly analyze how the three most popular software packages work: FSL [2], SPM [3] and FreeSurfer [4].

### 3.1. FSL

FSL is a software package developed by members of the Oxford Centre for Functional MRI of the Brain (Oxford University) which is composed by a series of independent tools that can be used separately or together. Although FSL has many different libraries and tools, in this section we are only going to focus in these ones which are useful for our goal of segmenting between the different tissues that can be found in a human brain.

In order to differentiate several tissues in magnetic resonance brain images, we have to perform some steps which are described below.

#### 3.1.1. Brain Extraction

When acquiring a brain MR image, we have a result similar to the one shown in Figure 9 and Figure 10.

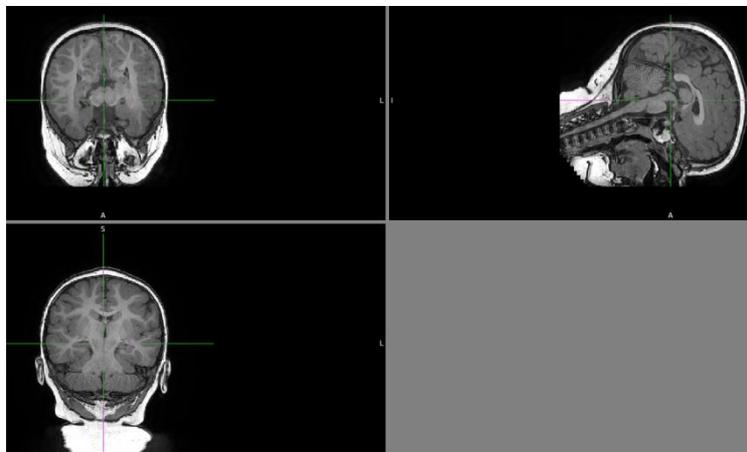


Figure 9. T1-weighted image: axial(top-left),sagittal(top-right) and coronal(bottom).

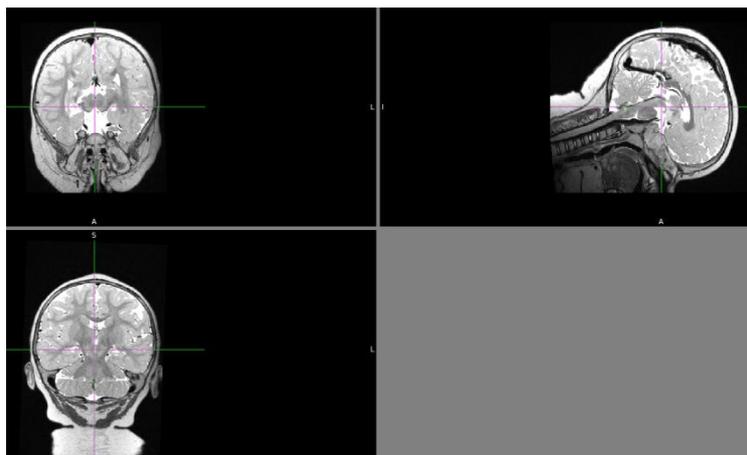


Figure 10. T2-weighted image: axial(top-left),sagittal(top-right) and coronal(bottom).

In both T1-weighted and T2-weighted images we can see all structures and elements in the head of the patient, not only the brain. Therefore, it is necessary to obtain only the region of interest from these images (the brain in this case).

In order to separate the brain from the other structures, FSL supplies a brain extraction tool (also called BET). The main advantage of using this kind of tool is that everything can be automated, so that we only have to wait until the process finishes. Furthermore, since every brain image can be different, there are some parameters which can be tuned to have a better result.

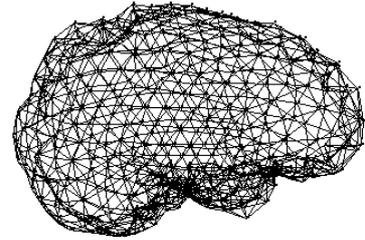


Figure 11. Brain mesh.

The main steps performed by the brain extracted tool in FSL are the following:

1. Triangular tessellation of icosahedrons.
2. Split each triangle until achieve required complexity.
3. Surface deformation to fit the brain surface.

This algorithm takes as input the T1 brain image and uses a brain model to try to fit the mesh for extracting the brain to the image. Although the results of the brain extractions use to be good, sometimes parts of the non-brain structures remain. However, it's not very important for our purpose that is to segment different tissues in the brain.

In addition, the brain extraction tool is also able to generate a model of the skin and skull from the patient images. This point is very important because, as we explained in the chapters before, these are very important parts that we need to know in order to build the patient specific model.

### 3.1.2. Image registration

Sometimes, when we are going to work with different images from the same patient, we need to align them so that we can work with them altogether. Furthermore, we can have misaligned slices in the same T1 or T2 image, so we should also align them in that case. To achieve that, FSL has two tools called FLIRT (FMRIB's Linear Image Registration Tool) and FNIRT (FMRIB's Non-linear Image Registration Tool).

The first one, FLIRT, is used to align images or slices only by using linear transformations in 2D and 3D (with several degrees of freedom) basing only in pixel/voxel intensities. Here, the registration process is finished when finding a global maximum of a similarity function. However, it can take too much time to find a global maximum, so methods to find a local maximum are normally used. In the case of FLIRT, instead of looking for a maximum for the similarity function, it looks for the minimum of the cost function.

One important characteristic of FLIRT is that we can set the method we want to use for finding this minimum, i.e. least squares, correlation ration, normalized correlation... So, once we chose this method, the tool will try to find this cost

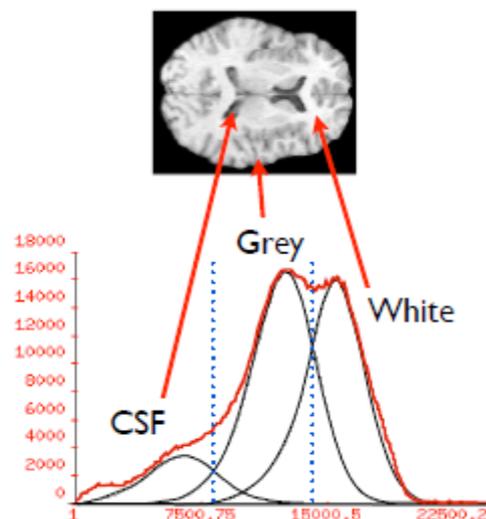


Figure 12. FAST's Gaussians model.

function minimum by performing repeated trials and searching for the best solution. It is important to have a starting point close to the global minimum, so that the algorithm can converge faster. Moreover, this tool uses low-resolution versions of the initial images to increase the speed. Thus, FLIRT provides a trade-off between speed and robustness. It is important to highlight that the input of FLIRT should be a brain extracted image instead of a normal T1 or T2 image.

On the other hand, FNIRT let us align images by using not only linear transformations. The way it works is similar to FLIRT, but it also tries to perform some non-linear transformations to reach the minimum of the cost function.

### 3.1.3. Brain tissue segmentation

FSL also has a tool called FAST (FMRIB's Automated Segmentation Tool), which let us perform tissue segmentation over the brain images. The input of this tool consists on brain-extracted images, which can be obtained as the output of the FSL's BET.

FAST is able to perform segmentation basing on an intensity model (histogram based), viewed as a mix of Gaussians. As they can be overlapped, FAST uses a probability model. In addition, this overlapping is even worsened because of the bias field (RF inhomogeneities),

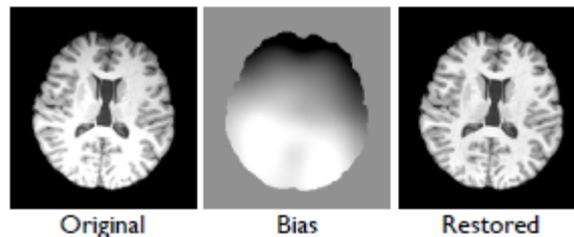


Figure 13. Bias field effect.

blurring, low resolution, head motion and noise. Specially, it is important to remove the bias field, since it causes intensity variations across space. In the Figure 13 we can see the effect of the bias field.

FAST is very robust to noise, because it does not only use the information of one voxel to classify it, but also uses its neighbors' information. Because of this, there must be a balance between believing the information from the pixel itself, or from its neighbors. Therefore, the final probability of a voxel to be part of a tissue can be computed as:

$$\text{Final probability} = \log p(\text{intensity}) + \beta \log p(\text{neighbors})$$

As output, FAST can obtain one probability volume model for every tissue, and also a binary segmentation instead of a probability volume model.

As a summary, we can see the behavior of the FAST algorithm as a series of steps as shown below:

- Initial segmentation (using tree-K-means)
- Iterate:
  - Estimate bias field.
  - Estimation segmentation. Iterate:
    - Update segmentation (intensity + neighbors).
    - Update tissue class parameters.
  - Apply partial volume model:

- MRF on mixed type (how many tissues)
- PV estimation.

### 3.1.4. Sub-cortical segmentation

Apart from obtaining segmentation between white, grey matter and CSF, it is possible to extract some of the sub-cortical structures which are present in the human brain, as hippocampus or brain stem. This can be performed with FSL using FIRST (FMRIB's Integrated Registration and Segmentation Tool). FIRST is able to extract 17 different brain sub-cortical structures by using a training model of 336 complete data set, which have been segmented manually by experts. So, the input of the algorithm must be a T1-weighted image, which is then compared with the data models.

Since all training data have been registered to a standard coordinates space (MNI-152), the first step of the algorithm is to align the input image to this space. Once it's done, it tries to fit every structure model (meshes) to the image and a boundary correction is applied finally. Since this project it is not very focus on sub-cortical structures segmentation, we will not explain too much about this.

## 3.2. SPM

SPM (Statistical Parametric Mapping) is a software package that works under MatLab and it's designed to perform segmentation between grey matter, white matter and CSF in a human brain. It's based on a mixture model clustering algorithm which has been extended to include probability maps and corrections for image problems common in MRI.

This tool is working under the following assumptions about the brain composition and distribution:

- Brain images are composed by tissue types (clusters) from which every voxel can be drawn.
- Voxel intensities can be modeled as a normal distribution.
- We know the number of classes in which voxels can fit.
- Some knowledge about special distribution in form of probability images obtained from the Montreal Neurological Institute.
- The intensity and noise associated to each voxel has been multiplied with a smooth scalar field

Since there are many unknown and related parameters, we need to know some information about some of them to find the value of the others. That's why SPM uses an iterative algorithm. It starts with a value of 1 for the modulation field and it also gives starting probabilities for every voxel to belong to grey matter, white matter or CSF, basing on prior probability maps. Thus, the algorithm tries to converge in each iteration by assigning new values to the cluster parameters by means of the values computed in the iterations before.

In order to make the segmentation between different tissues, SPM performs the following stages iteratively:

1. Estimating the cluster parameters:
  - a. Compute the number of voxels belonging to each class from the probability values of every voxel and the multiplicative correction estimation (basing on probability maps in the first step).
  - b. Later, mean and variance of voxel intensities for each cluster are computed.
2. Recalculating the belonging probabilities using the parameter values obtained before and the prior probability images. The fact of having this prior probabilities information will make the algorithm to converge faster.
3. Estimating the modulation function: to reduce the number of parameters describing an intensity modulation field, it is modeled by a linear combination of low frequency discrete cosine transform basic function.

### 3.3. FreeSurfer

FreeSurfer is a set of tools developed to study the brain cortical and sub-cortical anatomy. When focusing on cortical anatomy, the software can build a model of the boundary between white matter, grey matter and pial surface. As long as we know these surfaces, the software is also able to obtain information about surfaces, volumes, curvature, thickness, etc.

The software let us study the brain anatomy by means of two different streams:

- Surface-based stream.
- Volume-based stream.

The surface-based stream pipeline is performed by several stages, which are listed below:

1. An affine registration with a Talairach atlas [5] [6] is performed: this allows FreeSurfer set some seed point which are used in later stages.
2. B1 bias estimation by measuring variations in white matter intensity. This estimation can be done by using the changes in the intensity of the white matter across the volume. The white matter points used to measure these variations are chosen basing on their location on Talairach space as well as their intensity and its neighbors intensity.
3. Each voxel intensity is divided by the estimated bias field, as a attempt to remove its effects.
4. The skull is stripped by using a deformable template model.
5. All voxels are classified in white matter and other tissues basing on their intensity and their neighbors intensity.
6. Basing on the Talairach space, both brain hemispheres are separated with cutting planes.
7. Finally, the surface between grey and white matter, and the pial surface are generated and refined, as well as the brain surface.

Once the surface-based stream has finished, FreeSurfer has a volume-based stream designed to preprocess MRI volumes and label sub-cortical tissue classes. The stream is composed by the following stages:

1. Registration a Talairach space in order to perform a more accurate segmentation. Both cortical and sub-cortical labeling use the same algorithm so that the final result is computed by using the two outcomes. The atlas is built from a training set of subjects' brains which have been labeled by experts.
2. Initial volumetric labeling: labeling is performed by using a probabilistic model, setting value for each point basing of the training data set.
3. B1 bias correction: the algorithm used here is different to the one used in the surface-based stream.
4. High-dimensional non-linear volumetric alignment to Talairach atlas, to improve previous results.
5. An iterative algorithm is performed to compute final probabilities for every voxel until the probabilities do not change between two consecutive iterations.

As we can deduce from all the steps described above and the outcome we have as output, FreeSurfer is more oriented to study the anatomy of the brain than to distinguish between grey matter, white matter and CSF.

## 4. Methodology

Since the goal of this thesis, as explained before, is to evaluate different software packages in order to have an accurate segmentation, we need to have a ground truth which can be used to assess the quality of a performed segmentation. For that reason, we have 18 phantoms obtained from Brainweb [7] with a known distribution of white matter, grey matter and CSF. Together with the outcome from a segmentation performed by a software package, we can know accurately how good it is.

In order to compare the segmentation results, we have used MatLab together with a library created by Jimmy Shen [8] that let us load Nifti format images. As long as we can load the image data, we just have to compare the values of 3-D matrices to measure the difference between two images. In that way, we are able to count the total number of voxels in the image and, for every tissue, the true-positives, false-positives, true-negatives and false-negatives. The concepts can be explained as follows:

- True-positive: voxel of a tissue classified as part of this tissue.
- False-positive: voxel which is not part of a tissue but is classified as if it was.
- True-negative: voxel which is not part of a tissue and is right classified.
- False-negative: voxel which is part of a tissue and is classified as if it was not.

However, only to count the number of voxels well or bad-classified is not enough. Thus, we need to use some parameters which can give a more accurate idea of how good a segmentation process can be. For this reason, we have used 3 different parameters: specificity, sensitivity and f-factor [9] [10].

Specificity can be defined as the probability of classifying properly one voxel which doesn't belong to a tissue. It can be calculated as:

$$Specificity = \frac{true\ negatives}{true\ negatives + false\ positives}$$

On the other hand, sensitivity can be defined as the probability of classifying properly one voxel as part of a tissue. It can be computed as:

$$Sensitivity = \frac{true\ positives}{true\ positive + false\ negatives}$$

However, sometimes it is difficult to compare the classification made by a software package if we have more than one parameter. To solve that, we only define the f-factor as the harmonic mean of specificity and sensitivity:

$$F\ factor = 2 \cdot \frac{specificity \cdot sensitivity}{specificity + sensitivity}$$

If the f-factor is close to one, then the classification will be better than if it is close to zero.

Finally, in addition to these three measures, we have also used an error ratio. This has been used because we have three different classes and it would be desirable to have one measure which lets have a general idea for the result. In this sense, the error ratio has been computed as:

$$Error = 1 - \frac{\text{well classified voxels}}{\text{number of voxels}}$$

where “well classified voxels” are the voxels which have been classified right when comparing with the ground truth.

But this error ratio can be used in two different ways:

- We could consider all the voxels in the image so that “number of voxels” is the total number of voxels in the images (the number of voxels in the ground truth is the same than in the segmented images, because they have the same size and coordinate space).
- We could use only those voxels which are not empty (which are not the background in either the ground truth or the segmented images). This last approach will probably give us a better idea of the real performance for a segmentation method.

However, before comparing all segmentation results from different softwares, we first need to understand how they work and how we can obtain the best possible results with each one of them. Thus, the first part of the experiments we have carried out is focused on trying to set a standard procedure to perform segmentation with each set of tools, as can be seen below.

## 4.1. FSL

FSL has different tools in order to perform the steps which are necessary in the segmentation process. Thereby, we are going to explain the way we need to use each one of them. We have only used all these tools from a command-line console because, even though some of them have graphical user interfaces (GUI), we cannot use all available options from these interfaces.

### 4.1.1. Brain extraction

The first step in the process is to extract the brain from the patient MR images we have obtained. Thus, the “BET” tool allows doing this and supplies a list of options and configurable parameters that can be used to improve the results. All this parameters and their possible values can be found in the FSL’s official documentation. Here we are not going to focus in all these parameters, but in those ones we have considered are important for our goal.

In addition to the configurable parameters, “BET” also let us change its functionality with a list of mutual exclusive options. Therefore, if we want to specify a working mode and some configurable parameters values, then we have to write in a command-line console the following:

```
bet input output <mode> <parameters>
```

In Table 1 and Table 2, we can see a list of working modes and parameters we have changed in order to obtain a more accurate result for the brain extraction.

Mode	Description
<b>-R</b>	This option runs more "robust" brain centre estimation; it repeatedly calls bet2, each time using the same input image and the same main options, except that the -c option (which sets the starting centre of the brain estimation) is set each time to the centre-of-gravity of the previously estimated brain extraction. The primary purpose is to improve the brain extraction when the input data contains a lot of non-brain matter - most likely when there is a lot of neck included in the input image. By iterating in this way the centre-of-gravity should move up each time towards the true centre, resulting in a better final estimate. The iterations stop when the centre-of-gravity stops moving, up to a maximum of 10 iterations.
<b>-B</b>	This attempts to reduce image bias, and residual neck voxels. This can be useful when running SIENA or SIENAX, for example. Various stages involving FAST segmentation-based bias field removal and standard-space masking are combined to produce a result which can often give better results than just running bet2.
<b>-A</b>	This runs both bet2 and betsurf programs in order to get the additional skull and scalp surfaces created by betsurf. This involves registering to standard space in order to allow betsurf to find the standard space masks it needs.
<b>-A2 &lt;T2&gt;</b>	This is the same as -A except that a T2 image is also input, to further improve the estimated skull and scalp surfaces. As well as carrying out the standard space registration this also registers the T2 to the T1 input image.

Table 1. BET working modes.

Option	Description
<b>-f &lt;f&gt;</b>	Fractional intensity threshold (0->1); default=0.5; smaller values give larger brain outline estimates
<b>-g &lt;g&gt;</b>	Vertical gradient in fractional intensity threshold (-1->1); default=0; positive values give larger brain outline at bottom, smaller at top
<b>-c &lt;x y z&gt;</b>	Centre-of-gravity (voxels not mm) of initial mesh surface.

Table 2. BET options.

As long as we have chosen the working mode and the values for those parameters which are configurable, we can run the brain extraction program and wait for the results.

#### 4.1.2. Image registration

Once we have the brain-extracted images, we should be able to start the segmentation process. However, if we want to use more than one input image (as T1 and T2 for example), we will need to align them to the same coordinate space.

FSL let us align two different images from the same subject by using the "FLIRT" tool. We have used this tool in order to align T1 and T2 images from the same subject. It can be used from a command-line console by typing the following:

```
flirt -in input -ref reference -out output -omat transformation.mat <-dof X>
```

where “input” should be the image we want to align, “reference” is the reference image we want to use and “transformation.mat” is the filename for the saved ASCII transformation matrix.

FSL gives some low-resolution images that let us align images to standard coordinate spaces, as MNI152 for example. Therefore, we should align both T1 and T2 images to the same space in order to use them in the segmentation process.

#### 4.1.3. Brain tissue segmentation

Once we have extracted the brain from the images we want to use for segmentation and we have aligned them (if we have more than one), we can start the segmentation process. For that purpose we will use the “FAST” tool supplied by FSL. This tool can be used just by writing the following in a command-line console:

*fast [options] file(s)*

where “files(s)” is the list of input images we are going to use to perform segmentation. The options we have used in this project are listed below:

Option	Description
<b>-t &lt;t&gt;</b>	Indicates if the input is a set of T1-weighted images (1), T2-weighted (2) or proton density (3).
<b>-o &lt;o&gt;</b>	Sets the base name for the output.
<b>-n &lt;n&gt;</b>	Sets the number of classes for segmentation.
<b>-B</b>	Estimated bias field as output.
<b>-b</b>	Bias corrected image as output
<b>-g</b>	Gets a separate binary file for each tissue class as output.
<b>-H &lt;h&gt;</b>	MRF beta value for main segmentation phase (increasing this gives spatially smoother segmentations)

Table 3. FAST options.

In the results chapter we will analyze how these parameters can affect to the segmentation process outcome. In this case, since we want to distinguish between grey matter, white matter and CSF, the number of classes we are going to work with is three (one for each tissue).

#### 4.1.4. Sub-cortical segmentation

In the case we also want to obtain a sub-cortical segmentation from the patient’s brain, we can use the “first” tool. For that purpose, we have to write in a command-line console the following:

*run\_first\_all [options] -i input -o output*

where we have chosen the following option:

Option	Description
<b>-b</b>	Input is already brain-extracted. If we don’t use this option, the “bet” tool is called so that the brain is extracted.

Table 4. FIRST options.

For a complete list of available options, they can be all consulted in the official FSL's documentation. As output, we will have one image per sub-cortical structure segmented.

## 4.2. SPM

As we highlighted in chapters before, SPM is able to perform segmentation between brain tissues (i.e. grey matter, white matter and CSF). Since this package is written and designed to run in MatLab environment, we have to run the "spm.m" file to open the program. Unlike FSL or FreeSurfer, SPM only has a graphical user interface, which has to be used to perform all steps in the segmentation process.

Same as FSL, SPM also has some options that let us re-align those images we want to use for segmentation. However, the segmentation process in SPM does not allow performing segmentation over more than one image yet. Because of that reason, the only goal of image re-alignment is to correct movements in time-series images (4D images). In our case, since we don't have this kind of image, we will skip this step.

Furthermore, although FSL has different tools for extracting the brain and for performing segmentation, SPM integrates everything in the same process. Therefore, we cannot see and correct the brain-extracted image, but we have to perform all segmentation and then check the results.

Thus, once we have started SPM in "fmri" mode, we need to press the "segment" button inside the "spatial pre-processing" box. After that, one configuration window will pop up, where we will be able to choose the values of the configurable parameters for the whole process. Between these options we can find the following:

Option	Description
<b>Data</b>	In this option we have to choose the input file.
<b>Output files</b>	Here we can choose what kind of output we want for grey matter, white matter and CSF. The possible options are "native", "modulated" and/or "unmodulated". The native output gives the segmentation results in the same space as the input image. The modulation option is an attempt to compensate for the effect of spatial normalization (volumetric differences as result of warping an image to match a template). Finally, the unmodulated option uses a kind of smoothing procedure to try to correct the projected data, with worse results than modulated.
<b>Bias corrected</b>	This option can be used to also save the bias corrected image.
<b>Tissue probability maps</b>	We can change here the images we want to use as prior probability maps for the segmentation algorithm.
<b>Gaussians per class</b>	Number of Gaussians to be used for every class. The typical values are 2 for grey matter, 2 for white matter, 2 for CSF and 4 for everything else. These values could work properly and we shouldn't have to change them.
<b>Affine regularization</b>	We can change the standard space which the input is registered to.

Table 5. SPM segmentation parameters.

Once we have chosen the desired values for all the parameters, we just have to press then “run batch” button at the top of the window and SPM will generate all results in the MatLab working directory.

### 4.3. FreeSurfer

FreeSurfer supplies a tool called “recon-all” that can be used to perform automatically all the segmentation process. Thus the process can be divided in a list of stages shown below:

Number	Task
1	Motion Correction and Conform
2	NU (Non-Uniform intensity normalization)
3	Talairach transform computation
4	Intensity Normalization 1
5	Skull Strip
6	EM Register (linear volumetric registration)
7	CA Intensity Normalization
8	CA Non-linear Volumetric Registration
9	Remove Neck
10	LTA with Skull
11	CA Label (Volumetric Labeling, ie Aseg) and Statistics
12	Intensity Normalization 2 (start here for control points)
13	White matter segmentation
14	Edit WM With ASeg
15	Fill (start here for wm edits)
16	Tessellation (begins per-hemisphere operations)
17	Smooth1
18	Inflate1
19	QSphere
20	Automatic Topology Fixer
21	Final Surfs (start here for brain edits for pial surf)
22	Smooth2
23	Inflate2
24	Spherical Mapping
25	Spherical Registration
26	Spherical Registration, Contralateral hemisphere
27	Map average curvature to subject
28	Cortical Parcellation - Desikan_Killiany and Christophe (Labeling)
29	Cortical Parcellation Statistics
30	Cortical Ribbon Mask
31	Cortical Parcellation mapping to Aseg

Table 6. FreeSurfer processing stages.

All these stages are divided in three big groups:

Group	Stages
1	1-5
2	6-23
3	24-31

Table 7. FreeSurfer stage groups.

In this way, FreeSurfer let us perform only the stages in one of the previous groups, so that we can divide the process and look at the outcome of every stage group in order to fix possible errors. Moreover, there are many parameters and options that can be used together with the “recon-all” command. If we want to change the way of how segmentation is performed, there are some parameters we could use and which can be found in the FreeSurfer documentation.

In order to solve some problems that can appear when using “recon-all”, we also have some tools supplied with FreeSurfer. A complete list of possible failure modes and how we can try to solve them can also be found in the official documentation. In this project, having the goal of comparing the performance of automatic segmentation with every software package, we have only used automatic options to obtain different structures from the human brain.

Finally, it is important to highlight that FreeSurfer works with “mgz” format images. Most of the magnetic resonance machines use to work with DICOM or Nifti files so, in order to solve the format conversion problem, FreeSurfer includes a tool called “mri\_convert”, which let us convert between different image formats. This tool can also be run automatically when calling “recon\_all” if necessary.

In our experience, the easiest way to run FreeSurfer is writing the following in a command-line console:

```
recon-all -subject subjectName -i input1 <-i input2> -all
```

This command will create a new folder in the subjects directory called “subjectName”, will convert all input files to “mgz” format and will run all processing steps. Therefore, when using this command everything will be done automatically, from the format conversion to the segmentation process.

## 5. Experiments with real data

As explained before, the first step before doing any comparison between the software packages is to try to set the best values for those parameters which can be configured. To achieve that, we have used both T1-weighted and T2-weighted images obtained from a 3 years-old boy. In the following sub-sections, the possible values for these configurable parameters in each software are explained. After this, the results obtained for the images from this 3-years old patient are shown.

### 5.1. FSL

As we could see in previous chapters, FSL has different tools to perform different stages in the segmentation process. In addition, each one of these tools is configurable by means of some parameters than can be changed. In this section, we will expose which values we have chosen for these parameters and we will try to explain why.

As explained before, the first step is to extract the brain from the patient head. For this extraction we have to choose between different working modes for the “BET” tool. Throughout the studies in this thesis, we have chosen only the T1-weighted image as input. The reason for that is that in T2-weighted images the contrast between grey matter and white matter is really poor; thereby we will use the T1-weighted image to have better segmentation results. Regarding to “BET”, we tried some of the working modes with the following results.

#### *-B working mode:*

This working mode attempts to reduce image bias and residual neck voxels. The result is not very good, because the contrast in the image is much lower than in the original image, as we can see in Figure 14.

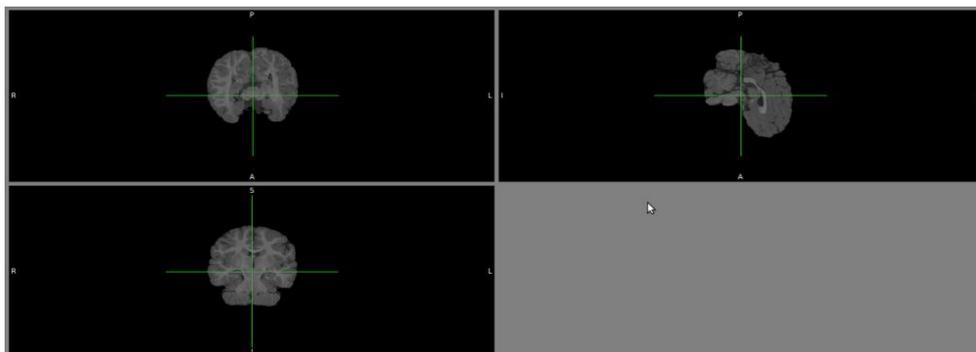


Figure 14. BET: -B working mode: axial(top-left),sagittal(top-right) and coronal(bottom).

#### *-R working mode*

As we can see in the Figure 15, the result is very good, since the whole brain has been extracted. In addition, the contrast between tissues remains almost the same as in the original image.

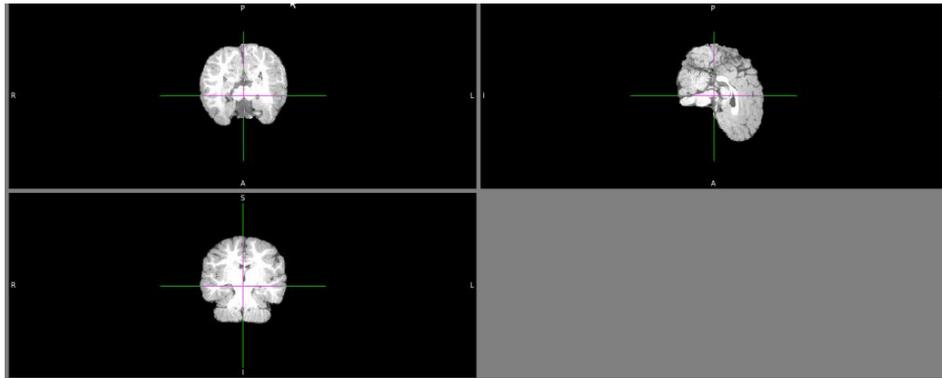


Figure 15. BET: -R working mode: axial(top-left),sagittal(top-right) and coronal(bottom).

### *-A2 working mode*

In this working mode, we can use the T2-weighted image together with the T1-weighted image to extract the brain from the last one. The brain extraction using this working mode should be better because we have more information. For that reason, “bet” is also able to extract the skull and skin, which are very useful for the goal of construction a patient

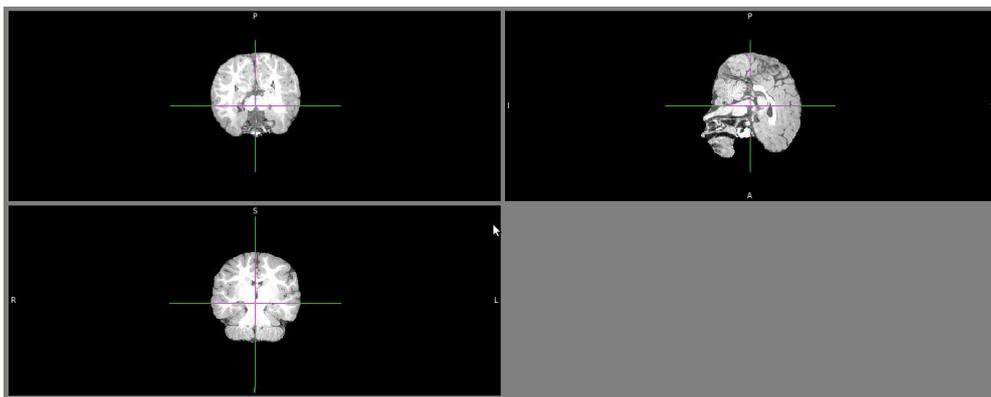


Figure 16. BET: -A2 working mode: axial(top-left),sagittal(top-right) and coronal(bottom).  
specific model. In the Figure 16 we can see the results.

The result is worse than it was with the “-R” working mode. The reason is that the “-R” working mode performs some more iterations in the algorithm in order to find the centre of gravity for the brain, so that the extraction can be better. To solve this we could include some information about the centre of gravity, so that we can obtain the same good results than before, also extracting the skull and scalp.

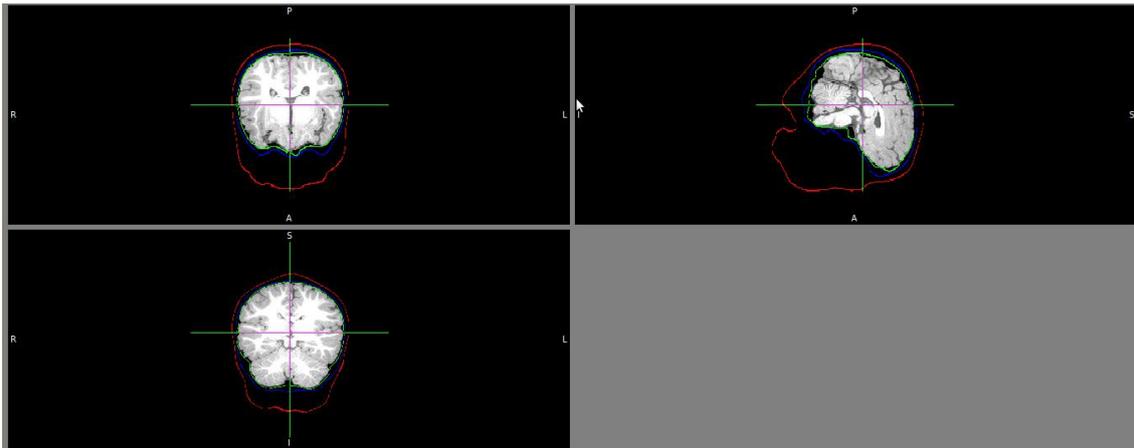


Figure 17. BET: -A2 working mode with -c option: axial(top-left),sagittal(top-right) and coronal(bottom).

In order to add this information, we can use the “-c” option in order to add a centre of gravity (in voxels) for the image. This centre doesn’t have to be exact, but it will improve the results much. The result of including this information is shown in Figure 17.

As we can see, the result is much better now and similar to the one obtained with the “-R” working mode. Now we could ask, can this result be improved even more? Although this result is good to perform segmentation, we can obtain a better result by changing the value of another parameter.

We say in previous chapter that “BET” tool in FSL let us change the value for the intensity threshold, so that smaller values give larger brain outlines estimates. This value has to be between 0 and 1 and has a default value of 0.5. Thus, if we increase a little this value (to 0.6) we can even obtain a smoother result. Therefore, the result we obtained is the one shown in Figure 18.

Well, now we have a good result for the brain-extracted image, we can start the segmentation process using the “fast” tool in FSL.

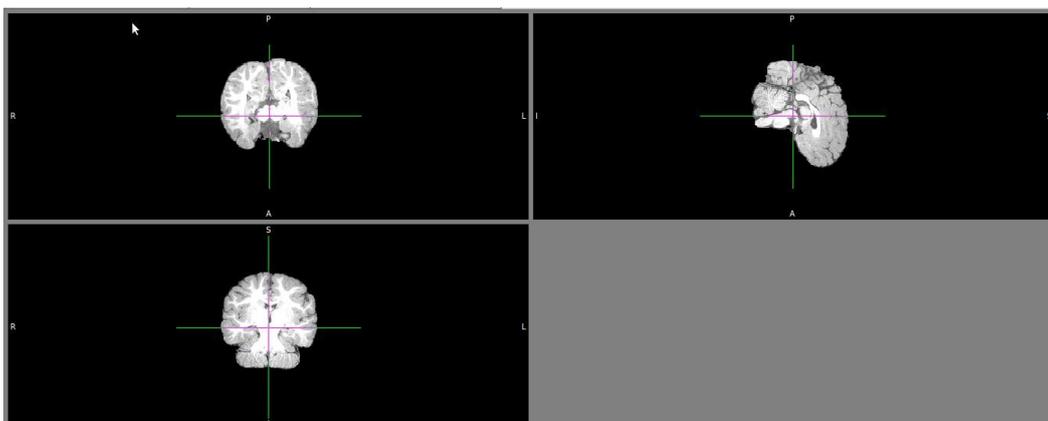


Figure 18. BET: -A2 working mode with -c and -f=0.6 options: axial(top-left),sagittal(top-right) and coronal(bottom).

Between the available options for “fast” we are going to choose the following:

- -b: we want to obtain a bias-corrected image.
- -B: we want to see the estimated bias field to evaluate possible errors.

- -g: we also want to obtain one binary image per classified tissue, in order to make easier the comparison later.
- -t 1: we tell “FAST” we are going to use the T1 image as input. We only use the T1 image because in our experiments we have seen that the results when using T1 and T2 together are very bad. This happens because the T2 image uses to decrease the contrast between grey and white matter, so that the segmentation is much worse.
- H 0.5: this is the importance of the neighbors when segmenting one voxels (values from 0 to 1). We put an intermediate value because we want smooth results but also giving importance to the intensity of the voxel to be segmented.

After running “FAST” tool, we have obtained the following results. We show only the segmented image, but we also have one binary image per class.

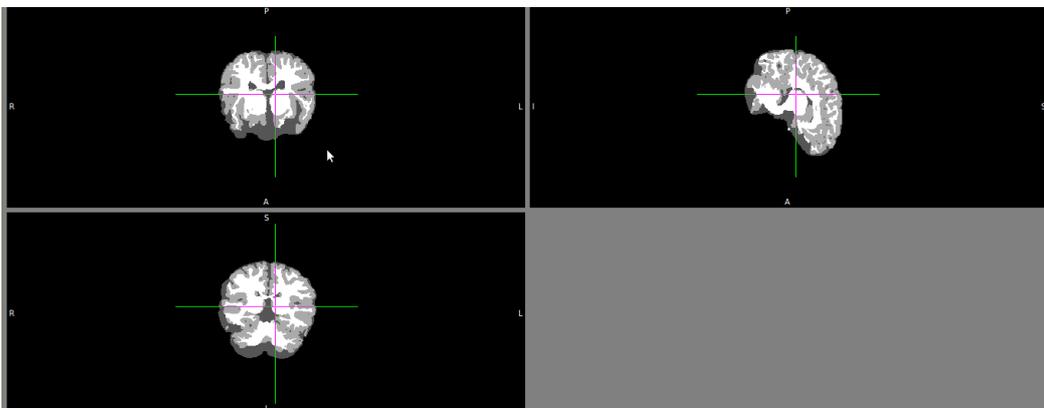


Figure 19. FAST: segmented brain: axial(top-left),sagittal(top-right) and coronal(bottom).

In this image, we can see white matter in white color, grey matter in grey color and CSF in a darker color (black voxels belong to empty space).

Finally, if we want to perform sub-cortical segmentation, we have to use the “first” tool. With it, we have used the “-b” option, which let us use the brain-extracted image instead of the original image. By doing this, the process will be faster. In this tool, almost everything is automatic, so we just have to run it and wait for the results, which are shown in the Figure 20 and the Figure 21.

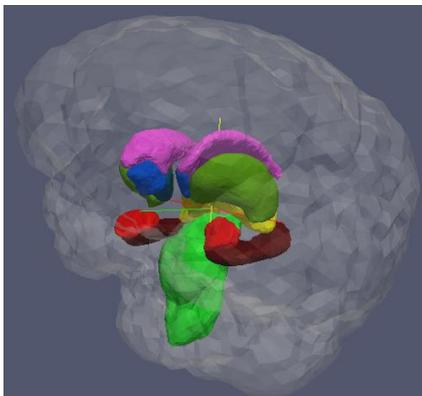


Figure 21. FIRST: results (a)

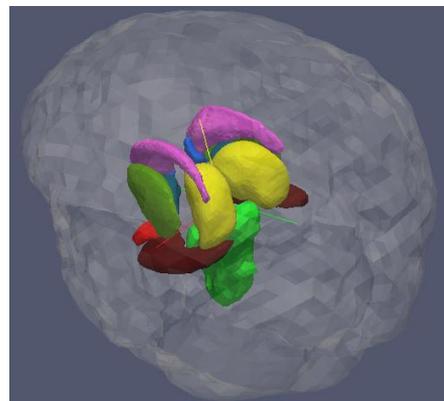


Figure 20. FIRST: results (b)

## 5.2. SPM

SPM, as it is concerned, is not as flexible as FSL can be. This tool only separates between image registration (when talking about 4D images) and image segmentation.

In this case, as explained previously, we are not going to use the image re-alignment option included in SPM because we don't need it (as long as we don't have time series images).

Regarding to image segmentation, there are no many parameters we could change apart from the number of Gaussians per class (which shouldn't be changed because the results are good with its default values) or the space we used for affine transformation (which only should change the orientation of the image or the coordinate system). For that reason, since SPM does not allow performing a segmentation basing on more than two images, we tried to segment tissues using only the T1-weighted image, and alter using only the T2-weighted image.

Moreover, we chose to obtain native, modulated and unmodulated images. However, since we want to compare these results with the ground truth, we should only use the native images, because they are in the same coordinate system than the original ones. Therefore, the result obtained for T1 and T2 images separately are the following.

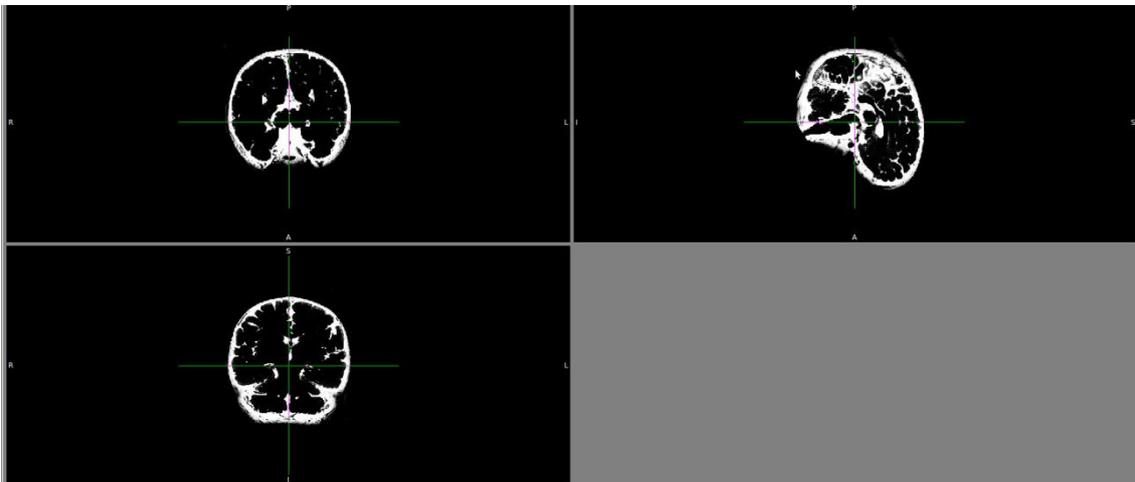


Figure 22. SPM: CSF using T1: axial(top-left),sagittal(top-right) and coronal(bottom).

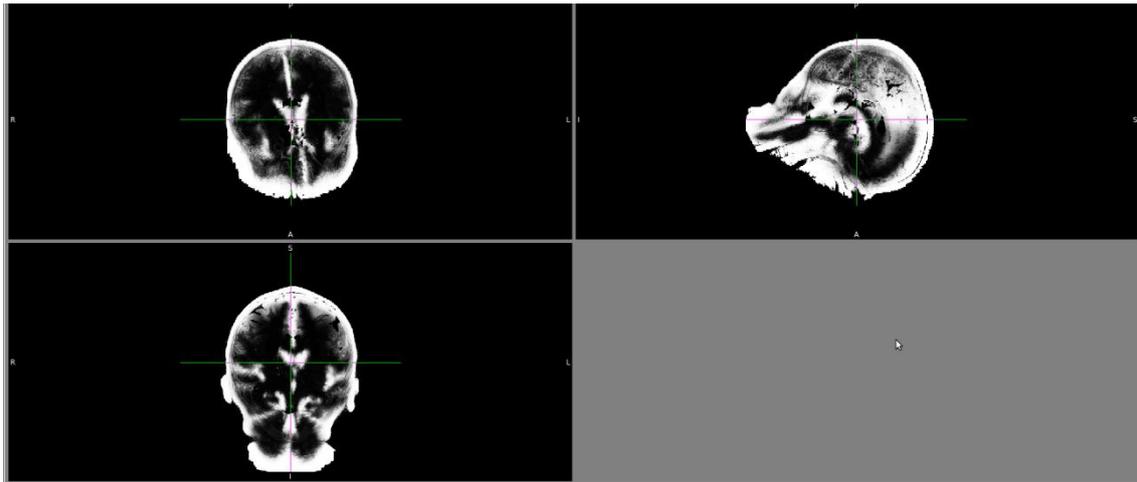


Figure 23. SPM: CSF using T2: axial(top-left),sagittal(top-right) and coronal(bottom).

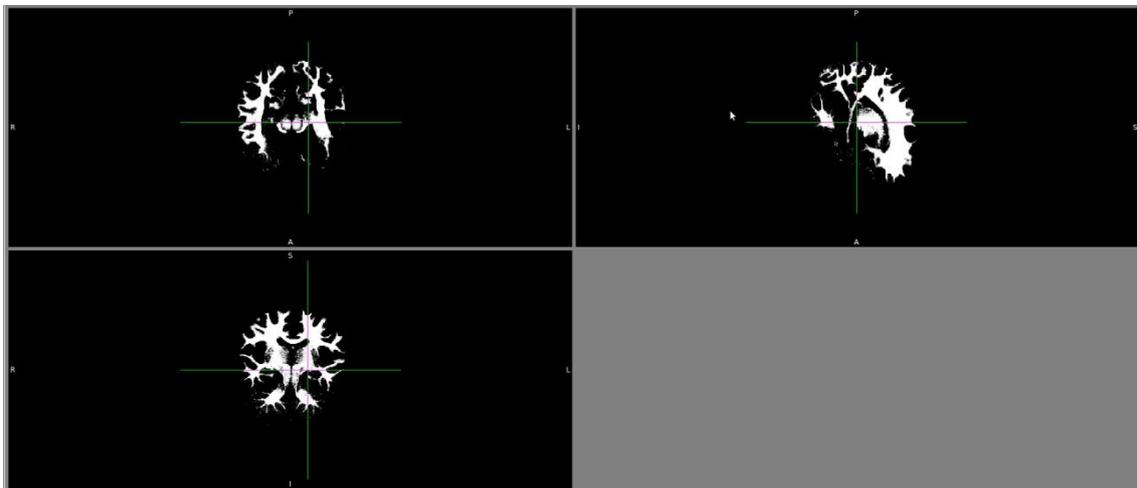


Figure 24. SPM: white matter using T1: axial(top-left),sagittal(top-right) and coronal(bottom).

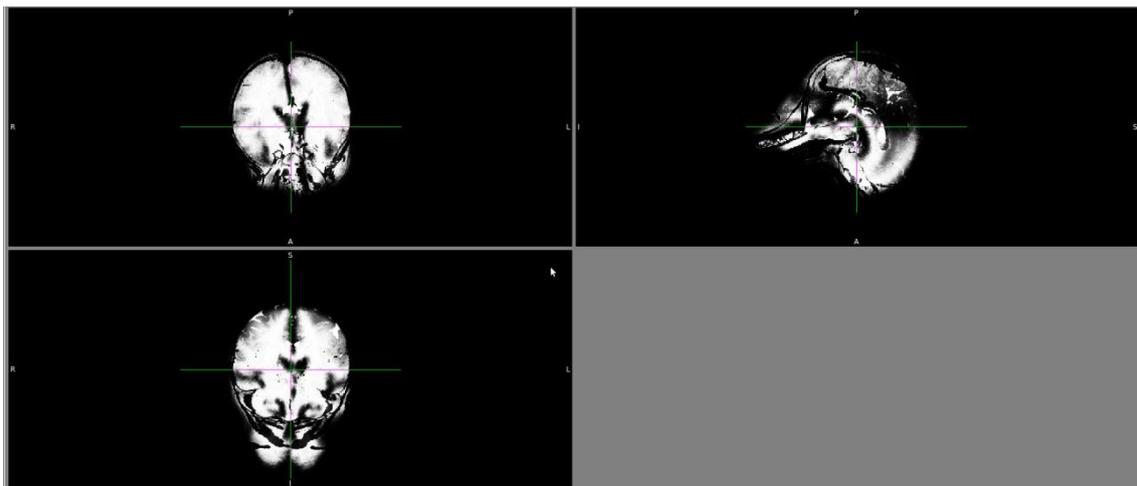


Figure 25. SPM: white matter using T2: axial(top-left),sagittal(top-right) and coronal(bottom).

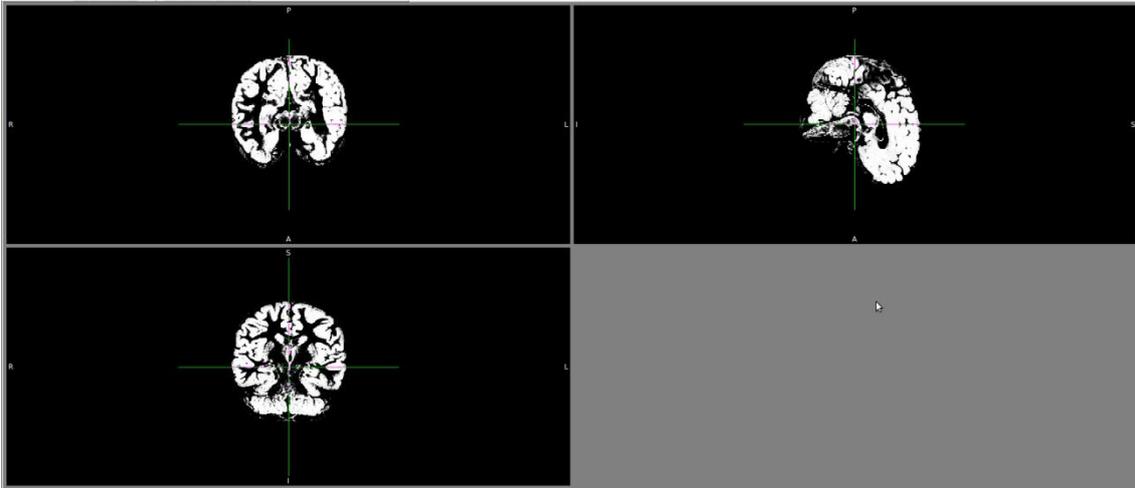


Figure 26. SPM: grey matter using T1: axial(top-left),sagittal(top-right) and coronal(bottom).

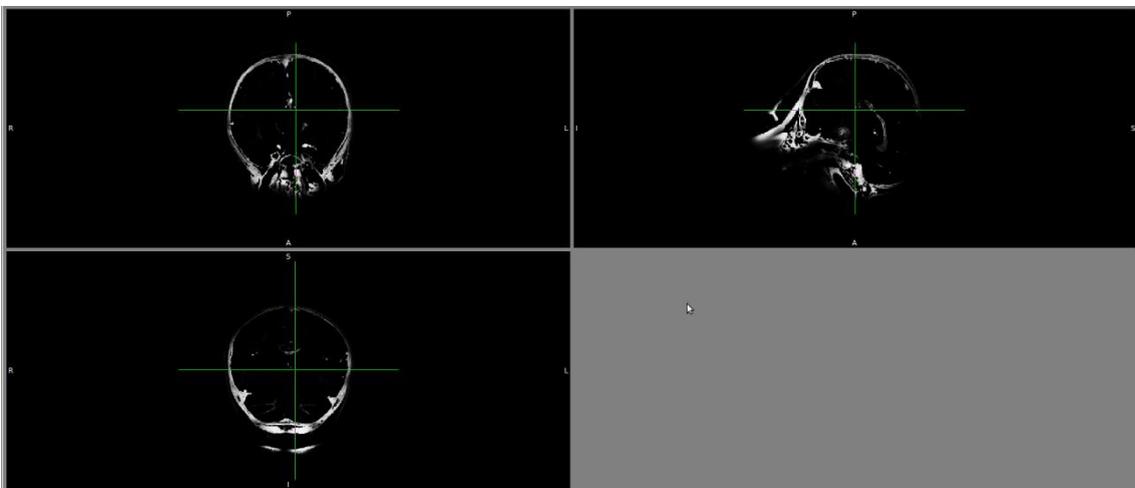


Figure 27. grey matter using T2: axial(top-left),sagittal(top-right) and coronal(bottom).

When comparing the results using the T1-weighted and the T2-weighted images, we can see that the second one will give completely wrong results. For that reason, we should only use T1-weighted images when working with SPM.

In addition, we can observe in Figure 22 that the results for CSF are also bad, because it takes part of the skull when it tries to extract the cerebro-spinal fluid. However, the results for grey and white matter seem to be good, so we will see how good they are when comparing to the ground truth later.

### 5.3. FreeSurfer

FreeSurfer, as also explained before, is able to segment between different structures we can find in the human brain. If compared to FSL and SPM, it is the most automatic tool because everything can be performed just with one command (“recon\_all”). In addition, stages can also be performed individually so that we can fix manually the error which can appear in the process.

Since in this project we are looking for a half-automatic method to perform brain segmentation, we have only used the automatic procedure for FreeSurfer. In addition, we are not experts in medicine so we aren't able to perform a manual segmentation or we can't correct all errors which can be present in the results. For that reason, when comparing with results from FSL and SPM, we will only wait for the automatic results from FreeSurfer.

Once we have launched the automatic tool "recon\_all", we obtain as result many files for volumes and surfaces segmented. However, we can use one of these files, which contains all the segmented structures. The name of this file is "aseg.mgz" and it's shown below in Figure 28.

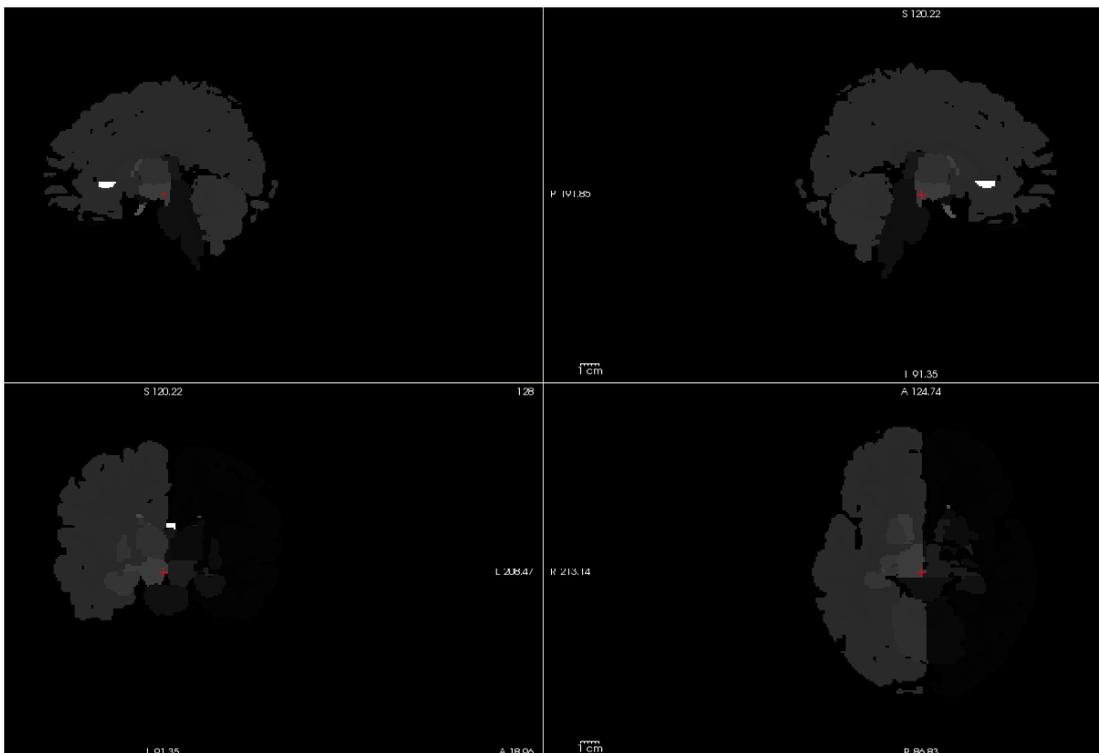


Figure 28. FreeSurfer segmentation: axial(bottom-right),sagittal(top) and coronal(bottom-left).

As we can see in the results, unlike FSL and SPM, FreeSurfer does not segment between grey matter, white matter and CSF. However, it segments between the structures in the brain, as brain step, hippocampus, thalamus, etc. In that sense, the output from FreeSurfer is different than the one from SPM, and could be considered as a mixture between "fast" a "first" tools in FSL.

Basing on our experience, FreeSurfer should be used mostly when we are interested in brain structures instead of segmenting between brain tissues.

## 6. Software comparison

At the moment we have set a value for all configurable options using one software package, we can start performing segmentation over other images. In this chapter, the results obtained when comparing the performance of each software package we have studied are presented. In order to know the accuracy of segmentation results we must have a ground truth, which can be used to make the comparison. For this reason, we have obtained a set of images with a known ground truth from Brainweb.

The set of images we are going to use is composed by 18 phantoms with different conditions regarding to the noise and radio-frequency in-homogeneities which are present in the images. In this sense, this data-set can be considered as a good one to compare the behavior of each package, because some of them can be more sensitive to the noise of RF in-homogeneities.

The noise level present in the images and calculated relatively to the brightest tissue can vary between 0% (no noise) and 9% (high-intensity noise), with values of 1%, 3%, 5% and 7%. Moreover, the RF in-homogeneities can have an intensity of 0%, 20% or 40%. Therefore, we have 18 different T1-weighted and T2-weighted images to make the comparison.

The following table shows the parameters used for the simulation which has generated the phantom images.

Phantoms simulation data	
<b>Slice thickness</b>	1mm
<b>Scan technique</b>	Spoiled FLASH
<b>Flig angle</b>	30°
<b>Echo time</b>	10ms
<b>Image type</b>	Magnitude

Table 8. Phantom parameters.

### 6.1. Comparison results

When the voxel-by-voxel comparison is finished, we can represent graphically the obtained results using the measures explained in Chapter 4 in order to evaluate how accurate a segmentation is, and how the noise and radio-frequency in-homogeneities can affect the result. In the graphics shown below each represented line corresponds to a percentage of RF in-homogeneities, so that different percentages are shown in each graphic's legend.

#### 6.1.1. FSL

In this section, the result for specificity, sensitivity, f-factor and error ratio when using FSL are shown .

#### *White matter results*

In FSL, as can be seen in the following tables and graphics, specificity for white matter is really good and close to 1, but its value is worse when adding noise to the image.

On the other hand, sensitivity is lower than specificity and has the opposite behavior; when adding noise to the image, the result is better.

The f-factor, as it is concerned, has its higher value when adding noise to the image. In the results, we can see that the results are not very influenced by RF in-homogeneities.

		Specificity					
		Noise					
RF		0	1	3	5	7	9
0		1,00000	1,00000	1,00000	0,99993	0,99958	0,99885
20		1,00000	1,00000	1,00000	0,99995	0,99967	0,99914
40		1,00000	1,00000	1,00000	0,99996	0,99972	0,99924

Table 9. FSL: White matter specificity.

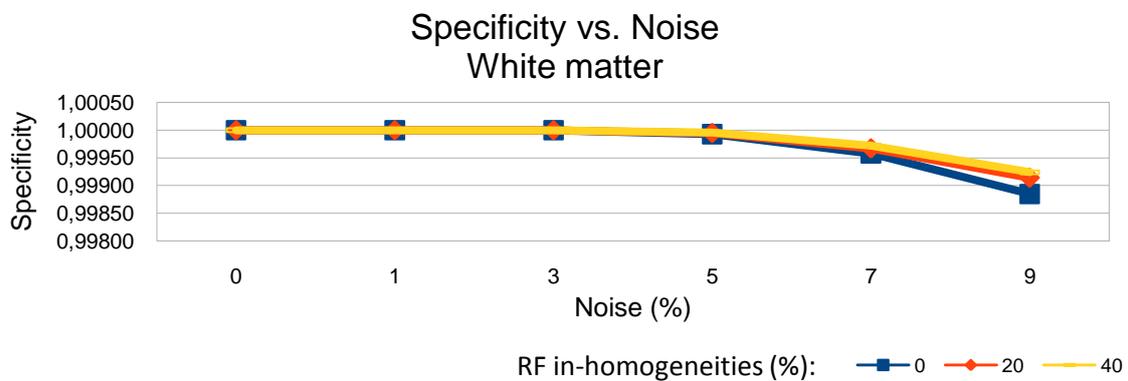


Figure 29. FSL: White matter specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,47736	0,62513	0,73657	0,76508	0,77815	0,78137
20		0,62901	0,65702	0,73080	0,76104	0,77437	0,77532
40		0,63143	0,65243	0,71857	0,75030	0,76752	0,77205

Table 10. FSL: White matter sensitivity.

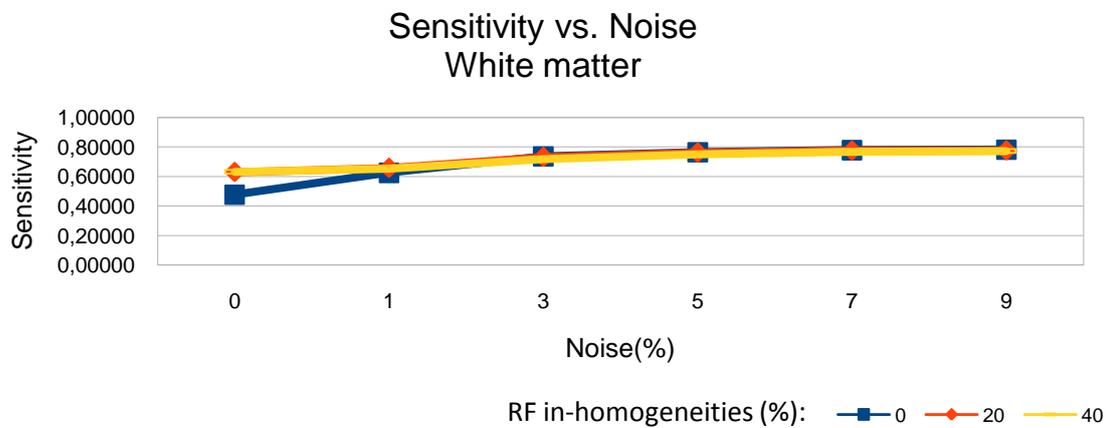


Figure 30. FSL: White matter sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,64623	0,76933	0,84830	0,86688	0,87507	0,87683
20		0,77226	0,79302	0,84446	0,86429	0,87272	0,87311
40		0,77408	0,78966	0,83624	0,85732	0,86837	0,87107

Table 11. FSL: White matter f-factor.

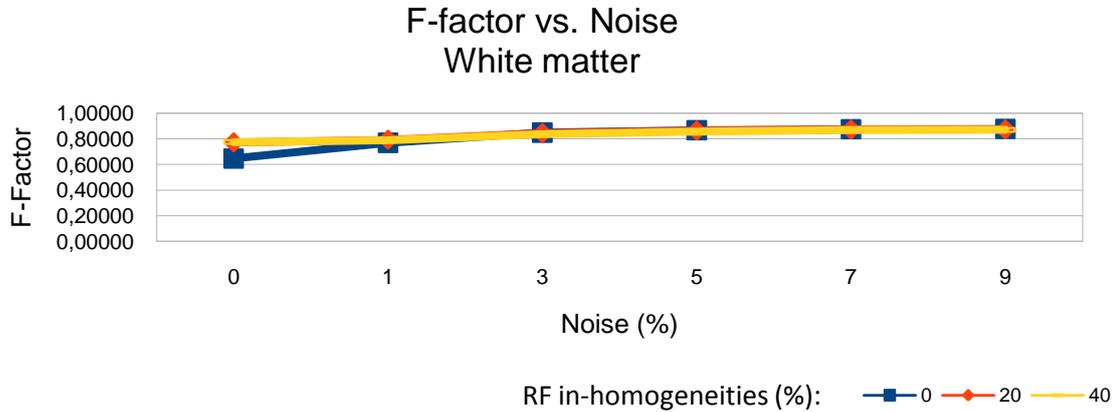


Figure 31. FSL: White matter f-factor.

### Grey matter results

Regarding to grey matter results, we can see below the tables and graphics representing the values obtained for the parameters we have used. We can see that specificity is very close to 1 and its value doesn't change very much when adding noise to the image.

Sensitivity, on the other hand, has different values between 83% and 67%. As can be seen below, the best result is obtained when there is no noise in the image.

Since the f-factor is the harmonic mean between specificity and sensitivity, it has its highest values when the noise in the image is low.

		Specificity					
		Noise					
RF		0	1	3	5	7	9
0		0,99029	0,99813	0,99876	0,99880	0,99874	0,99853
20		0,99831	0,99847	0,99877	0,99887	0,99878	0,99862
40		0,99839	0,99852	0,99883	0,99895	0,99896	0,99874

Table 12. FSL: Grey matter specificity.

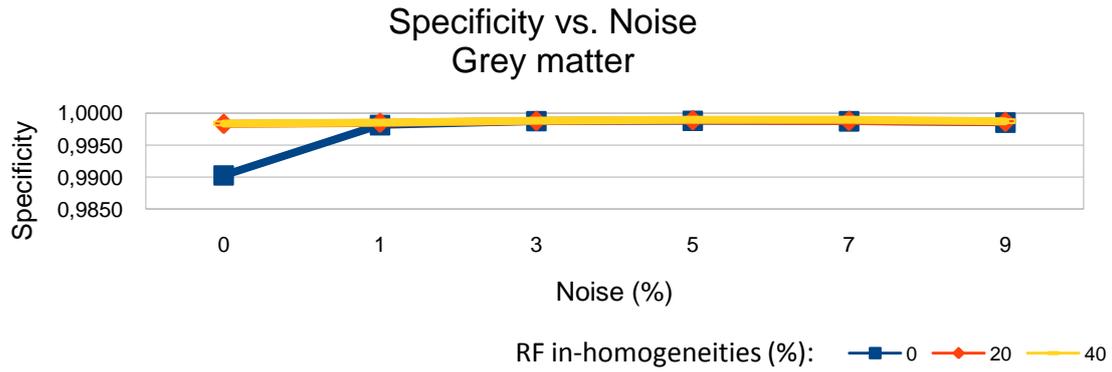


Figure 32. FSL: Grey matter specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,83261	0,76266	0,68421	0,67745	0,68375	0,69301
20		0,75447	0,73547	0,68632	0,67839	0,68546	0,69652
40		0,75089	0,73619	0,69193	0,67311	0,68099	0,68931

Table 13. FSL: Grey matter sensitivity.

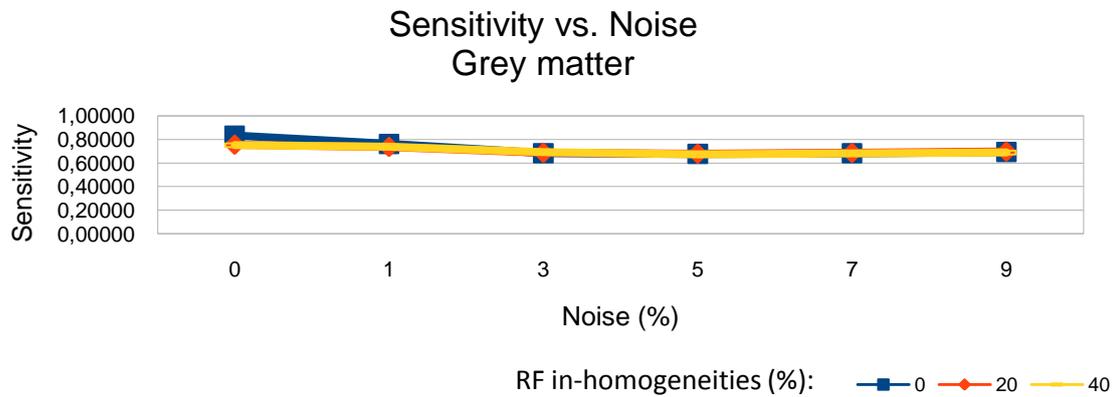


Figure 33. FSL: Grey matter sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,90463	0,86465	0,81209	0,80732	0,81176	0,81818
20		0,85943	0,84702	0,81358	0,80802	0,81298	0,82065
40		0,85713	0,84752	0,81753	0,80428	0,80988	0,81566

Table 14. FSL: Grey matter f-factor.

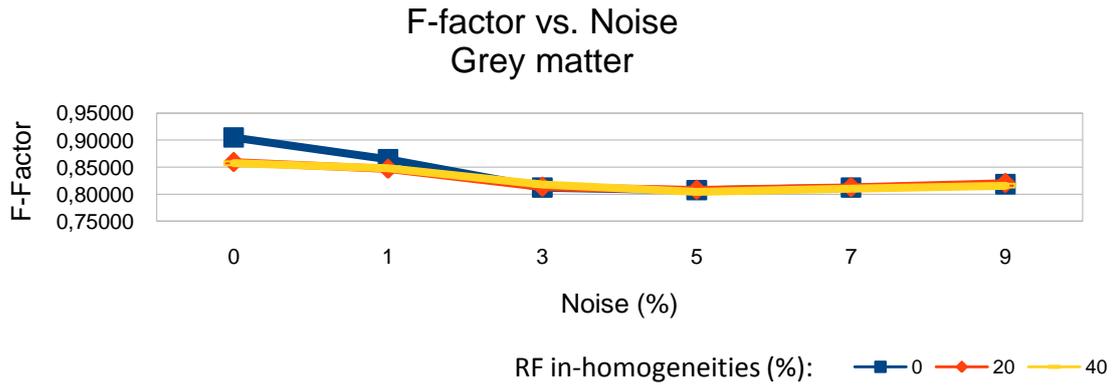


Figure 34. FSL: Grey matter f-factor.

### Cerebro-spinal fluid

For CSF segmentation, FSL obtains good values regarding to specificity. In addition, these values are not very influenced by the noise and RF in-homogeneities in the image.

If we focus on sensitivity, we can see that the results are not that good, and are even worse when increasing the noise level in the image.

Regarding to the f-factor, we can see that the best result obtained in FSL for CSF is achieved when the noise in the image is low (1%).

RF	Specificity						
	Noise						
	0	1	3	5	7	9	
0	0,99707	0,99693	0,99711	0,99705	0,99620	0,99583	
20	0,99727	0,99711	0,99729	0,99719	0,99645	0,99579	
40	0,99707	0,99708	0,99720	0,99717	0,99670	0,99601	

Table 15. FSL: CSF specificity.

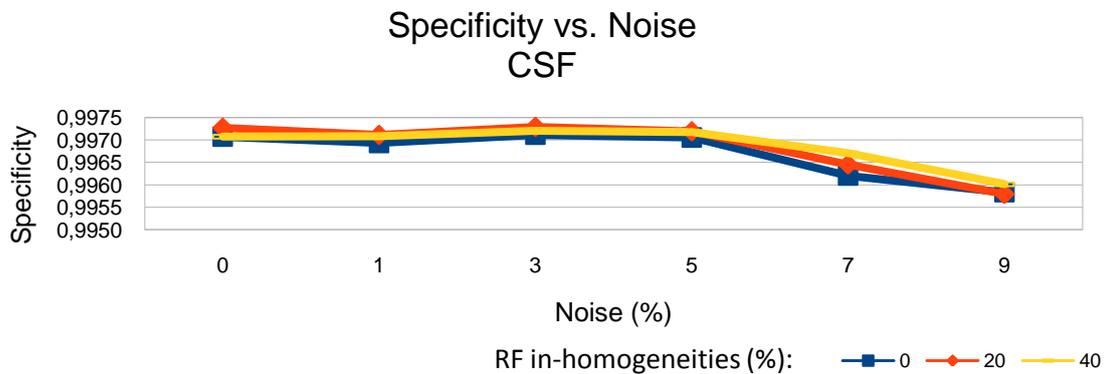


Figure 35. FSL: CSF specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,53593	0,55470	0,53489	0,50119	0,47922	0,46354
20		0,55528	0,55862	0,53260	0,49707	0,48366	0,46761
40		0,55502	0,55593	0,52753	0,48678	0,47066	0,45418

Table 16. FSL: CSF sensitivity.

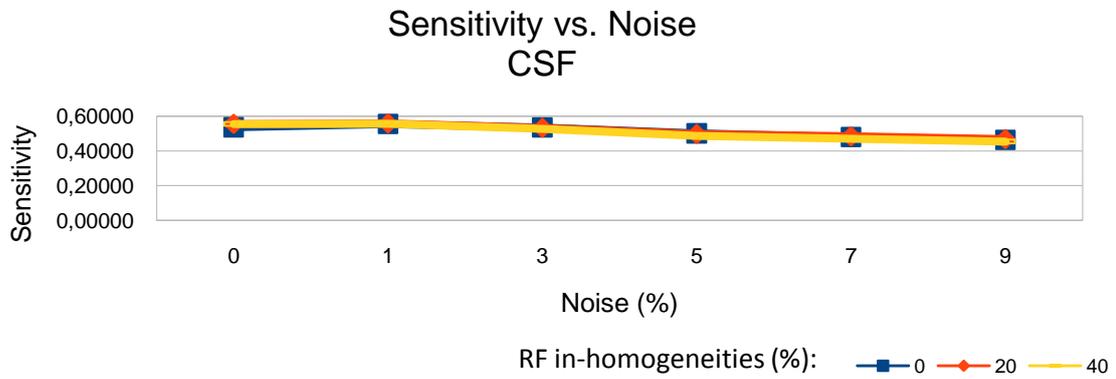


Figure 36. FSL: CSF sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,69714	0,71280	0,69628	0,66706	0,64714	0,63261
20		0,71336	0,71607	0,69437	0,66344	0,65123	0,63638
40		0,71310	0,71385	0,69003	0,65421	0,63939	0,62387

Table 17. FSL: CSF f-factor.

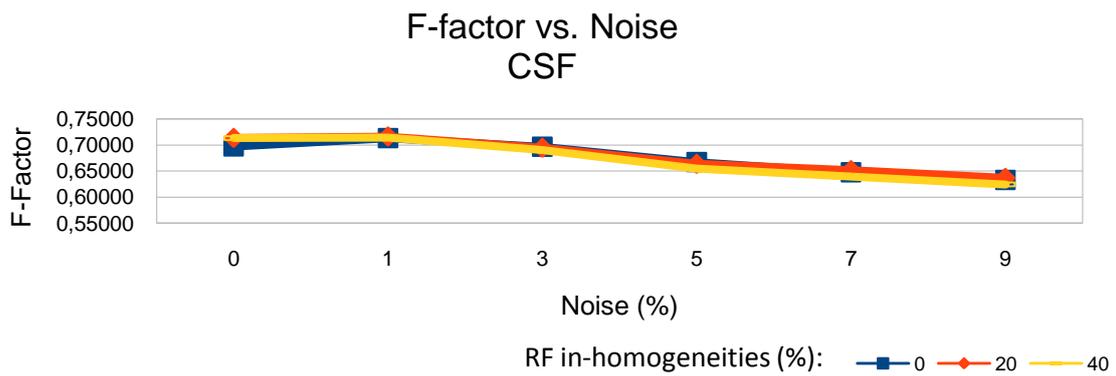


Figure 37. FSL: CSF f-factor.

**Overall result**

If we focus on the general result for white matter, grey matter and CSF, we can compute the error ratio. As explained in chapters before, it can be obtained counting all voxels in the image, or only those voxels which are not empty.

The results obtained are shown below. We can see in the tables and graphics that the lowest error ratio (and the best result) is obtained when the noise in the image is 1% and RF in-homogeneities are low. In this sense, the result is going worse when increasing one of these two parameters.

Error ratio (all voxels)		Noise					
RF	0	1	3	5	7	9	
0	0,0288	0,0222	0,0242	0,0248	0,0248	0,0244	
20	0,0228	0,0227	0,0245	0,0253	0,0243	0,0242	
40	0,0233	0,0233	0,0254	0,0283	0,0267	0,0267	

Table 18. FSL: Error ratio (all voxels).

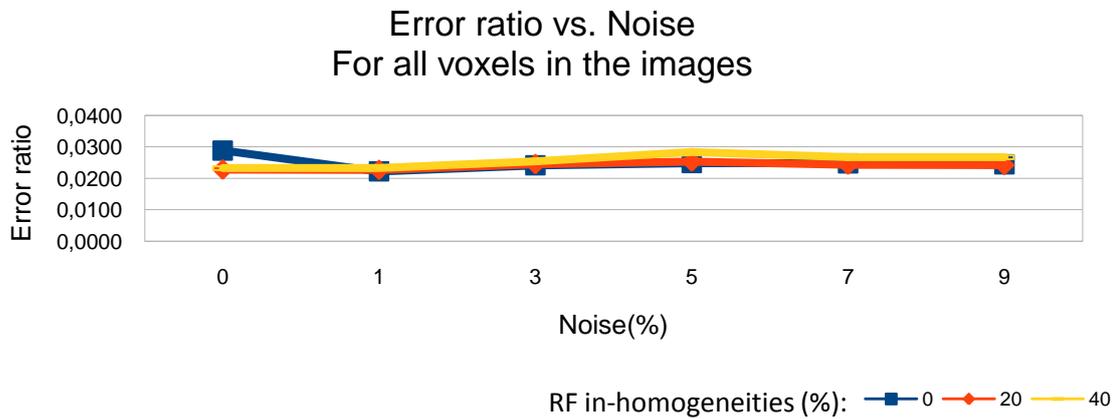


Figure 38. FSL: Error ratio (all voxels).

Error ratio (non-empty voxels)		Noise					
RF	0	1	3	5	7	9	
0	0,1033	0,0795	0,0866	0,0890	0,0886	0,0872	
20	0,0817	0,0812	0,0879	0,0908	0,0870	0,0864	
40	0,0834	0,0834	0,0911	0,1017	0,0955	0,0954	

Table 19. FSL: Error ratio (non-empty voxels).

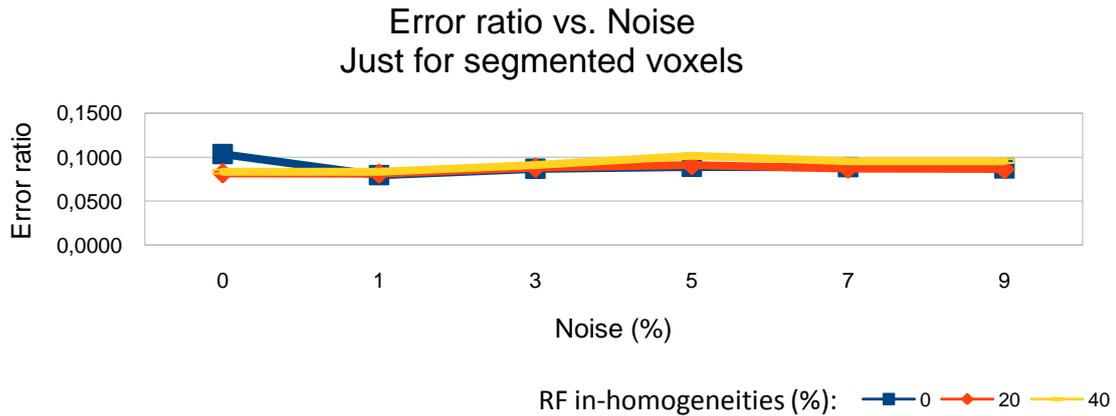


Figure 39. FSL: Error ratio (non-empty voxels).

### 6.1.2. SPM

The results for the parameters we used to evaluate the performance of SPM are shown in this section.

#### White matter

When classifying white matter voxels, we can see in the following tables and graphics that specificity is very good (close to 1) and it gets worse when increasing the noise level. In addition, the RF in-homogeneities don't affect very much to the results.

Regarding to sensitivity, the results are not that good when there is no noise in the image, but they are getting much better when increasing the noise level.

Finally, f-factor values show that SPM classifies better white matter when the noise level is intermediate (around 3%). RF in-homogeneities don't influence the results in this case.

	Specificity					
	Noise					
RF	0	1	3	5	7	9
0	0,99981	0,99958	0,98581	0,97307	0,93947	0,91752
20	0,99982	0,99959	0,98995	0,97701	0,94220	0,92270
40	0,99981	0,99963	0,98844	0,97890	0,95019	0,92797

Table 20. SPM: white matter specificity.

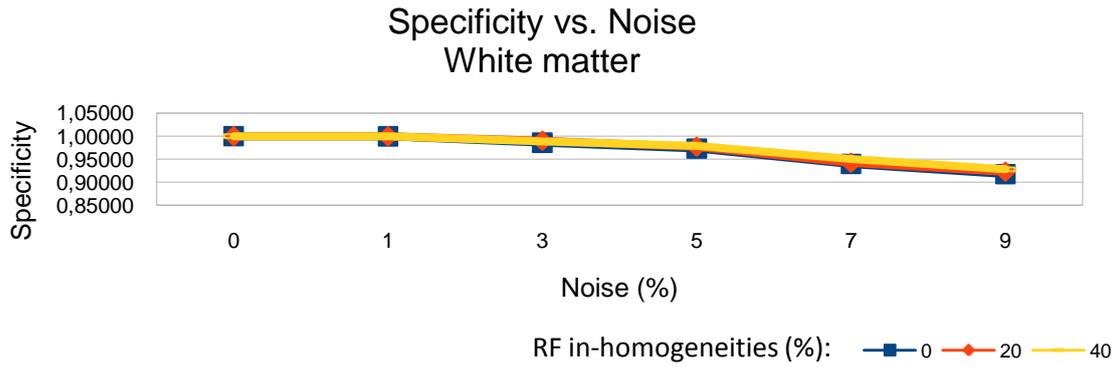


Figure 40. SPM: white matter specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,71197	0,89956	0,96608	0,95963	0,98360	0,99099
20		0,71046	0,89922	0,95635	0,95371	0,98199	0,98852
40		0,71332	0,88739	0,96064	0,94964	0,97453	0,98512

Table 21. SPM: white matter sensitivity.

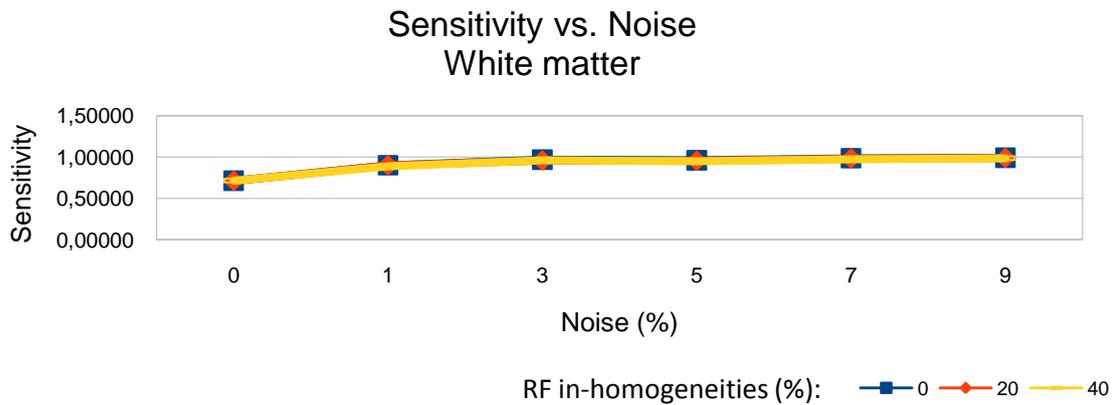


Figure 41. SPM: white matter sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,83169	0,94694	0,97585	0,96630	0,96103	0,95284
20		0,83066	0,94675	0,97286	0,96522	0,96168	0,95448
40		0,83261	0,94017	0,97434	0,96405	0,96221	0,95569

Table 22. SPM: white matter f-factor.

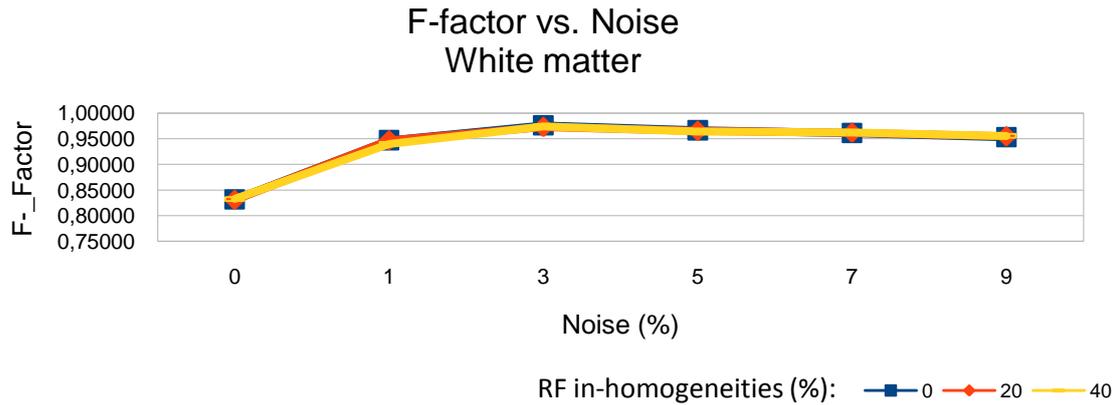


Figure 42. SPM: white matter f-factor.

### Grey matter

When looking at specificity for grey matter, we can see that its values are lower than they were for white matter. In addition, they get their best values for 1% and 3% noise level.

Sensitivity, as it is concerned, has its highest value with 3% noise and its worst value with 5% noise.

Finally, f-factor shows that SPM has a better result when the noise level is between 0% and 3%, and it is not influenced by RF in-homogeneities.

		Specificity					
		Noise					
RF		0	1	3	5	7	9
0		0,86241	0,86606	0,86328	0,85852	0,85367	0,84859
20		0,86301	0,86647	0,86644	0,86112	0,85395	0,84925
40		0,86346	0,86639	0,86580	0,86370	0,85811	0,85027

Table 23. SPM: grey matter specificity

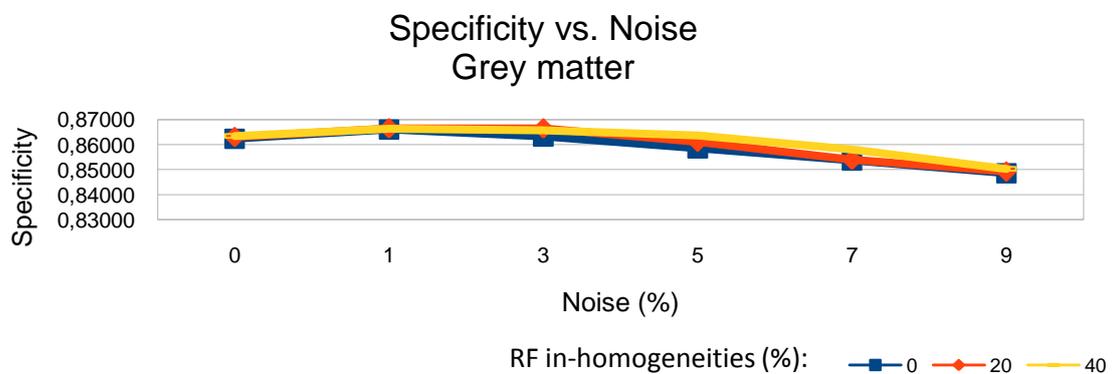


Figure 43. SPM: grey matter specificity

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,69957	0,68754	0,67795	0,57504	0,62729	0,68351
20		0,69974	0,67995	0,70179	0,56928	0,62567	0,68624
40		0,69825	0,68965	0,70131	0,57860	0,61876	0,66213

Table 24. SPM: grey matter sensitivity.

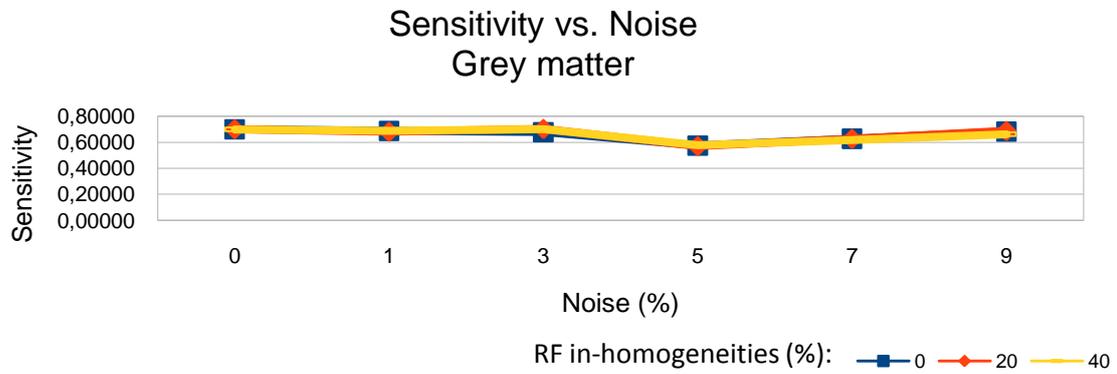


Figure 44. SPM: grey matter sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,77250	0,76654	0,75947	0,68875	0,72318	0,75716
20		0,77284	0,76196	0,77547	0,68543	0,72220	0,75909
40		0,77212	0,76798	0,77492	0,69298	0,71904	0,74450

Table 25. SPM: grey matter f-factor.

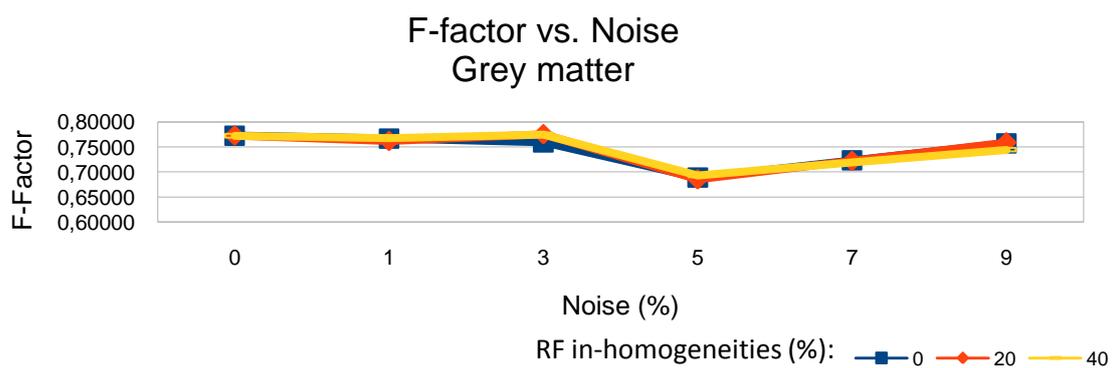


Figure 45. SPM: grey matter f-factor.

*Cerebro-spinal fluid*

When focusing on CSF, we can see that the specificity has its best value when there is no noise in the image.

Regarding to sensitivity, the values are really bad, and we can conclude with then that the classification in SPM for CSF is wrong. In addition, we can prove it by looking at the f-factor because its best value is 3% (when the noise intensity is 1%).

		Specificity					
		Noise					
RF		0	1	3	5	7	9
0		0,84368	0,82411	0,82509	0,83008	0,81984	0,81464
20		0,84012	0,82192	0,82110	0,82797	0,81762	0,81506
40		0,83827	0,82486	0,82221	0,82527	0,82919	0,81474

Table 26. SPM: CSF specificity.

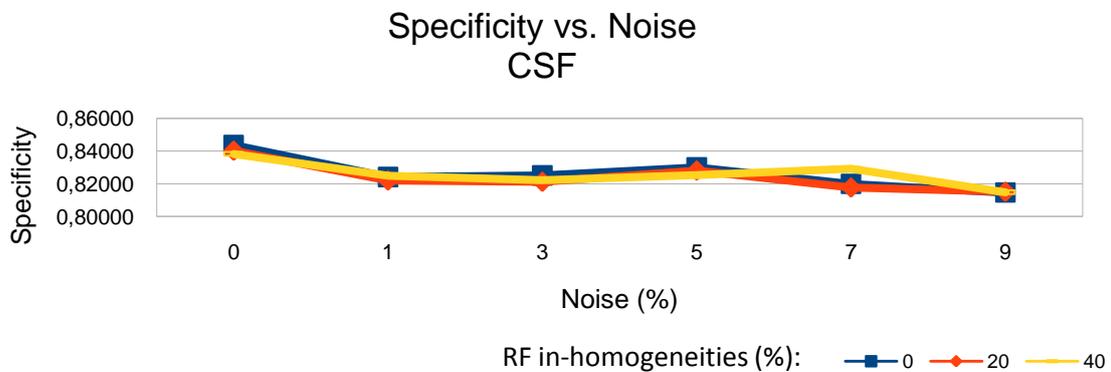


Figure 46. SPM: CSF specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,00083	0,01629	0,01053	0,00723	0,01008	0,01444
20		0,00182	0,01954	0,01465	0,01044	0,01115	0,01334
40		0,00426	0,01865	0,01297	0,01350	0,00877	0,01412

Table 27. SPM: CSF sensitivity.

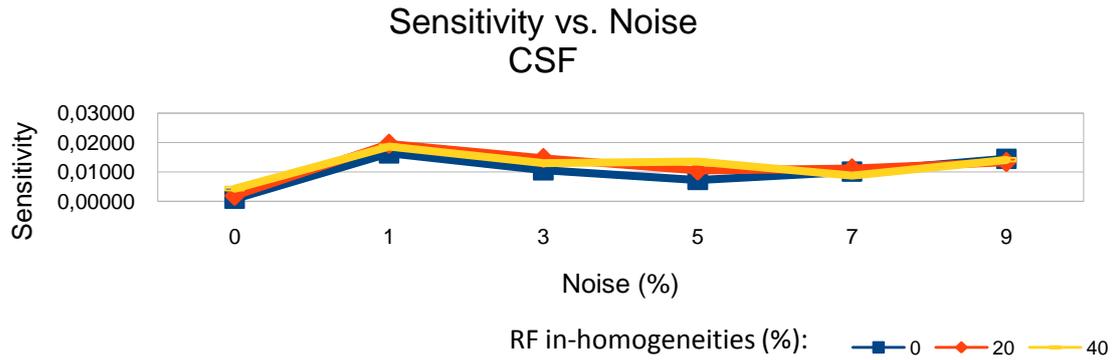


Figure 47. SPM: CSF sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,00166	0,03195	0,02080	0,01434	0,01991	0,02838
20		0,00364	0,03817	0,02879	0,02062	0,02200	0,02625
40		0,00847	0,03648	0,02553	0,02657	0,01736	0,02776

Table 28. SPM: CSF f-factor.

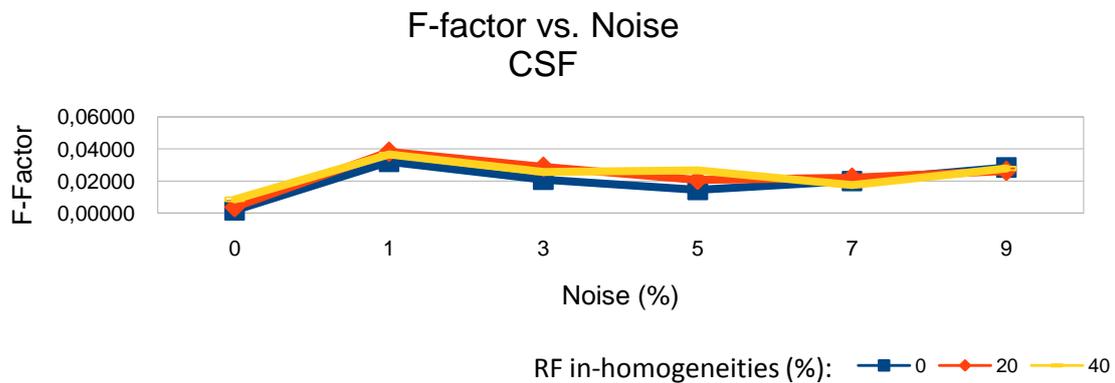


Figure 48. SPM: CSF f-factor.

### Overall result

If we look at the error ratio in SPM, we can see it obtains a better classification (lower values) when the noise in the image is intermediate (3%). RF in-homogeneities are not important for SPM performance regarding to these results.

		Error ratio (all voxels)					
		Noise					
RF		0	1	3	5	7	9
0		0,05007	0,05094	0,04725	0,05512	0,06855	0,07639
20		0,05002	0,05078	0,04703	0,05129	0,06804	0,07616
40		0,04850	0,04877	0,04731	0,04945	0,06128	0,07477

Table 29. SPM: Error ratio (all voxels).

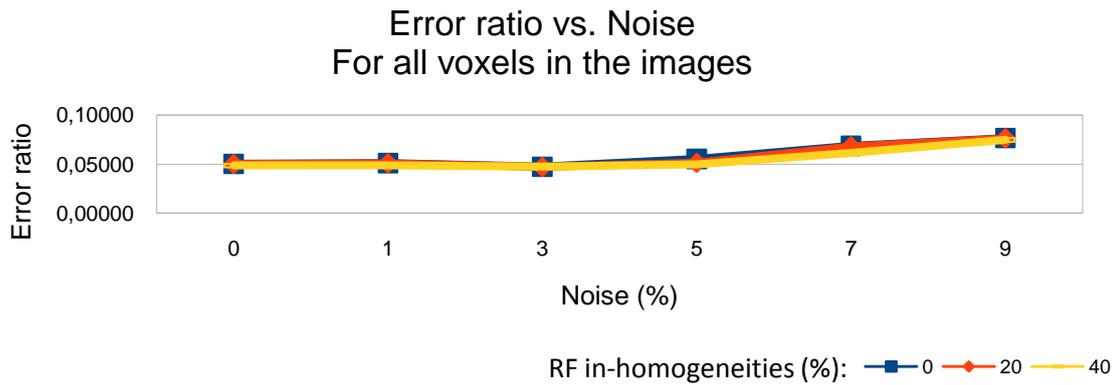


Figure 49. SPM: Error ratio (all voxels).

		Error ratio (non-empty voxels)					
		Noise					
RF		0	1	3	5	7	9
0		0,15387	0,15596	0,14636	0,16662	0,19913	0,21710
20		0,15375	0,15557	0,14578	0,15695	0,19796	0,21660
40		0,14980	0,15038	0,14652	0,15222	0,18196	0,21356

Table 30. SPM: Error ratio (non-empty voxels).

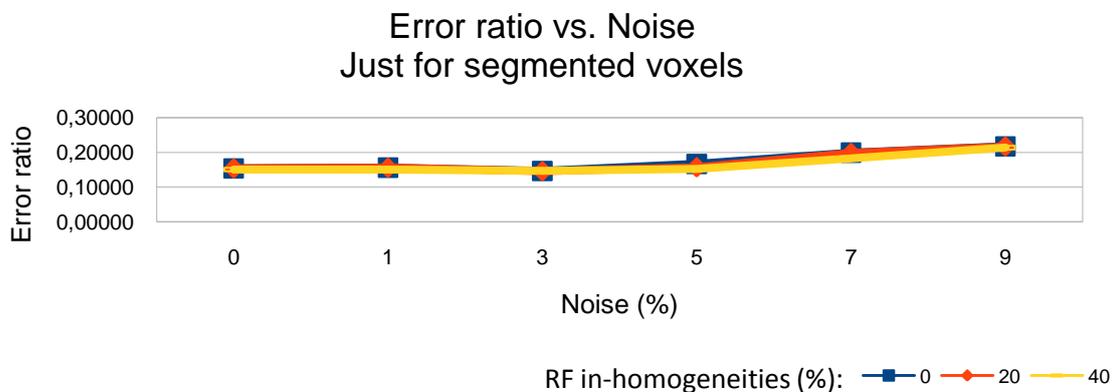


Figure 50. SPM: Error ratio (non-empty voxels).

### 6.1.3. FreeSurfer

In this section, the results obtained when using FreeSurfer to classify the different brain tissues are shown.

#### White matter

Specificity for white matter is very good, with values close to 1, having always a value of 98% regardless of the noise and RF in-homogeneities.

Sensitivity, as it is concerned, has values close to 70% and the best result is obtained when the noise level is intermediate (3% and 5%).

Finally, f-factor shows that the best classification is performed when the noise level is between 3% and 5%.

		Specificity					
		Noise					
RF		0	1	3	5	7	9
0		0,98619	0,98584	0,98481	0,98549	0,98568	0,98578
20		0,98553	0,98585	0,98543	0,98545	0,98555	0,98536
40		0,98556	0,98590	0,98499	0,98463	0,98609	

Table 31. FreeSurfer: White matter specificity.

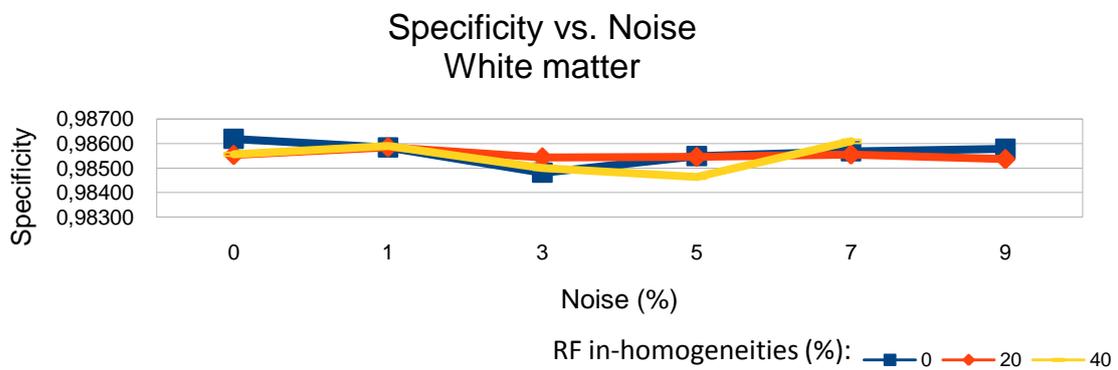


Figure 51. FreeSurfer: White matter specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,70517	0,71013	0,72381	0,71180	0,70973	0,70793
20		0,71450	0,71067	0,71792	0,71235	0,70938	0,71388
40		0,71215	0,71174	0,72003	0,72329	0,70440	

Table 32. FreeSurfer: White matter sensitivity.

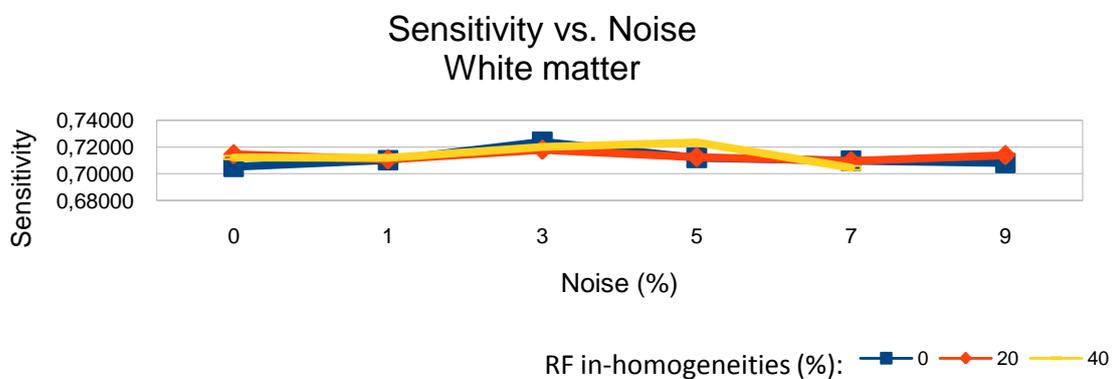


Figure 52. FreeSurfer: White matter sensitivity.

RF	F-factor					
	Noise					
	0	1	3	5	7	9
0	0,82233	0,82557	0,83437	0,82658	0,82525	0,82406
20	0,82841	0,82594	0,83067	0,82694	0,82496	0,82793
40	0,82684	0,82668	0,83193	0,83396	0,82177	

Table 33. FreeSurfer: White matter f-factor.

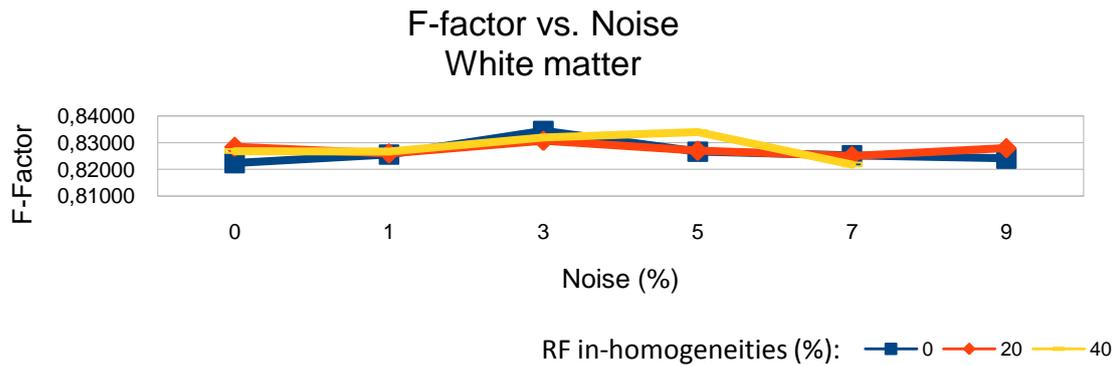


Figure 53. FreeSurfer: White matter f-factor.

### Grey matter

If we look at grey matter, the specificity in SPM is high and close to 1. Its value does not vary much with changes in the noise level or RF in-homogeneities, and it is 98%.

Regarding to sensitivity, the result obtained is not very good and it is worse then the noise level is intermediate (5%).

F-factor, as it is concerned, shows the same evolution that sensitivity, but with values close to 64%. These results can't be considered as very good.

RF	Specificity					
	Noise					
	0	1	3	5	7	9
0	0,98797	0,98795	0,98861	0,98802	0,98799	0,98743
20	0,98824	0,98823	0,98842	0,98825	0,98788	0,98799
40	0,98828	0,98837	0,98850	0,98867	0,98771	

Table 34. FreeSurfer: Grey matter specificity.

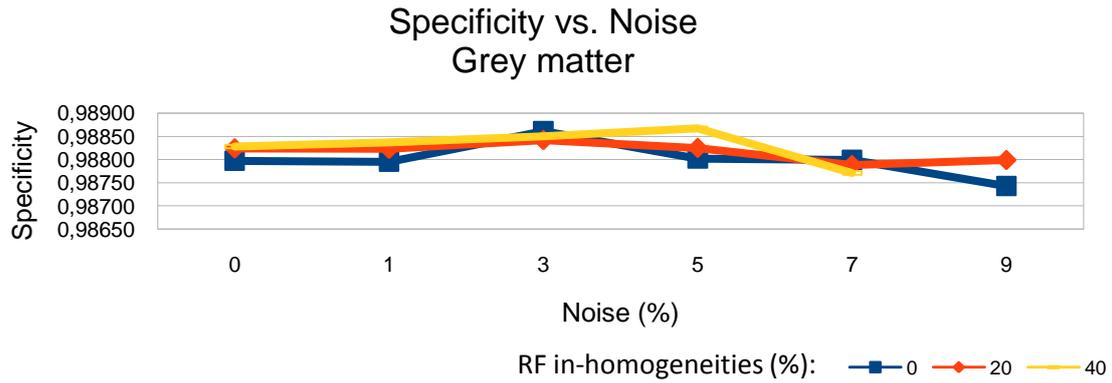


Figure 54. FreeSurfer: Grey matter specificity.

RF	Sensitivity					
	Noise					
	0	1	3	5	7	9
0	0,48156	0,48040	0,47432	0,47829	0,47842	0,48352
20	0,47849	0,47912	0,47565	0,47596	0,48177	0,48056
40	0,47910	0,47989	0,47488	0,47164	0,48479	

Table 35. FreeSurfer: Grey matter sensitivity.

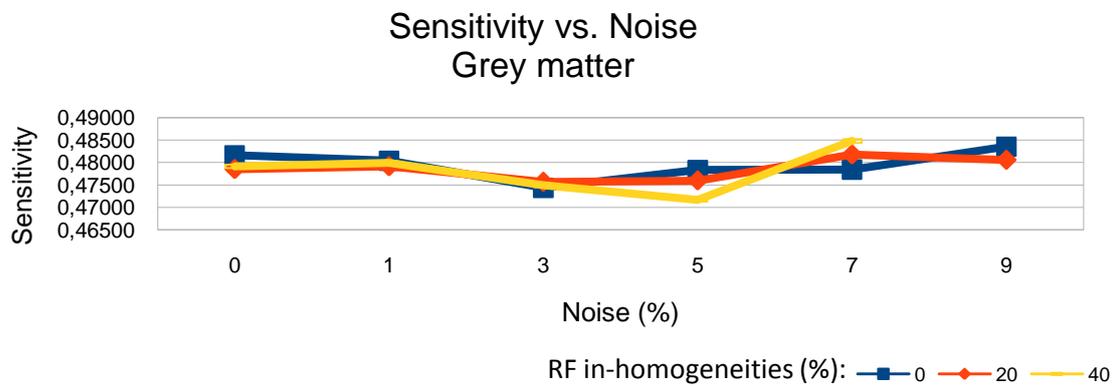


Figure 55. FreeSurfer: Grey matter sensitivity.

RF	F-factor					
	Noise					
	0	1	3	5	7	9
0	0,64750	0,64645	0,64107	0,64456	0,64467	0,64916
20	0,64479	0,64536	0,64224	0,64248	0,64768	0,64661
40	0,64534	0,64608	0,64155	0,63863	0,65037	

Table 36. FreeSurfer: Grey matter f-factor.

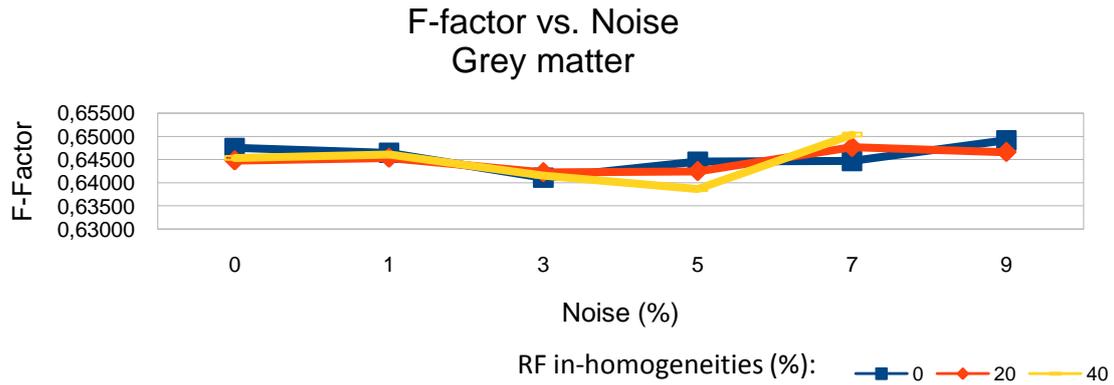


Figure 56. FreeSurfer: Grey matter f-factor.

### Cerebro-spinal fluid

If we focus on specificity for CSF, we can see that its value is very close to 1 and is very good, regardless of different noise or RF in-homogeneity levels.

On the other hand, sensitivity is almost zero, which means that FreeSurfer is not able to find CSF in the image. It is only able to find a little part of CSF.

Finally, the f-factor shows that the performance of FreeSurfer when classifying CSF is very bad and it can't be used to extract CSF.

RF	Specificity					
	Noise					
	0	1	3	5	7	9
0	0,99993	0,99994	0,99990	0,99990	0,99990	0,99988
20	0,99993	0,99994	0,99994	0,99989	0,99987	0,99989
40	0,99990	0,99991	0,99990	0,99989	0,99989	

Table 37. FreeSurfer: CSF specificity.

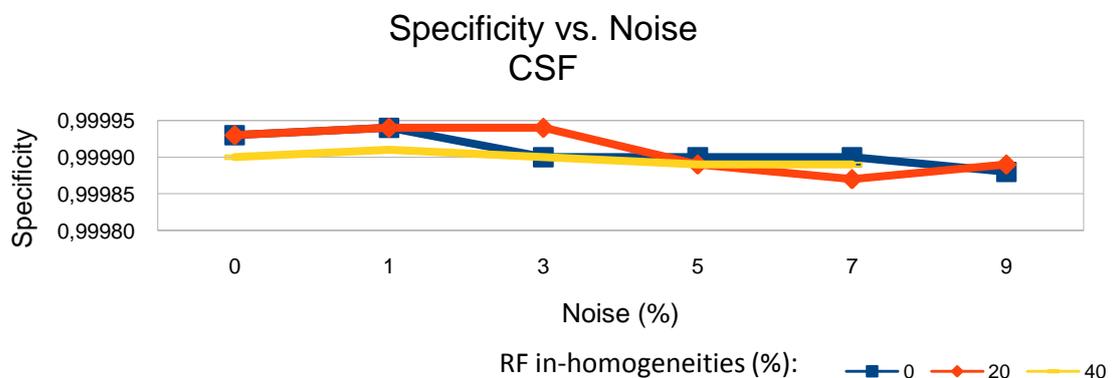


Figure 57. FreeSurfer: CSF specificity.

		Sensitivity					
		Noise					
RF		0	1	3	5	7	9
0		0,00184	0,00193	0,00194	0,00199	0,00195	0,00200
20		0,00190	0,00183	0,00193	0,00191	0,00200	0,00194
40		0,00193	0,00189	0,00196	0,00189	0,00194	

Table 38. FreeSurfer: CSF sensitivity.

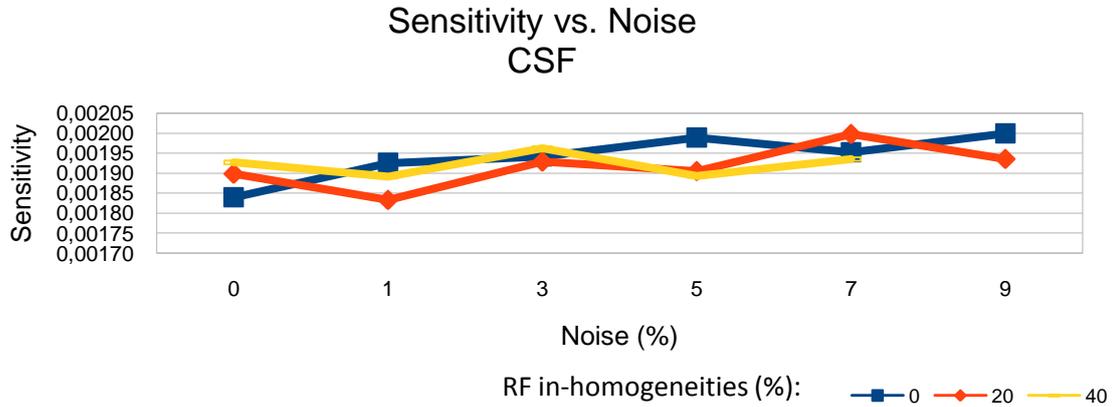


Figure 58. FreeSurfer: CSF sensitivity.

		F-factor					
		Noise					
RF		0	1	3	5	7	9
0		0,00367	0,00384	0,00388	0,00397	0,00390	0,00399
20		0,00379	0,00366	0,00385	0,00380	0,00399	0,00386
40		0,00385	0,00378	0,00392	0,00378	0,00386	

Table 39. FreeSurfer: CSF f-factor.

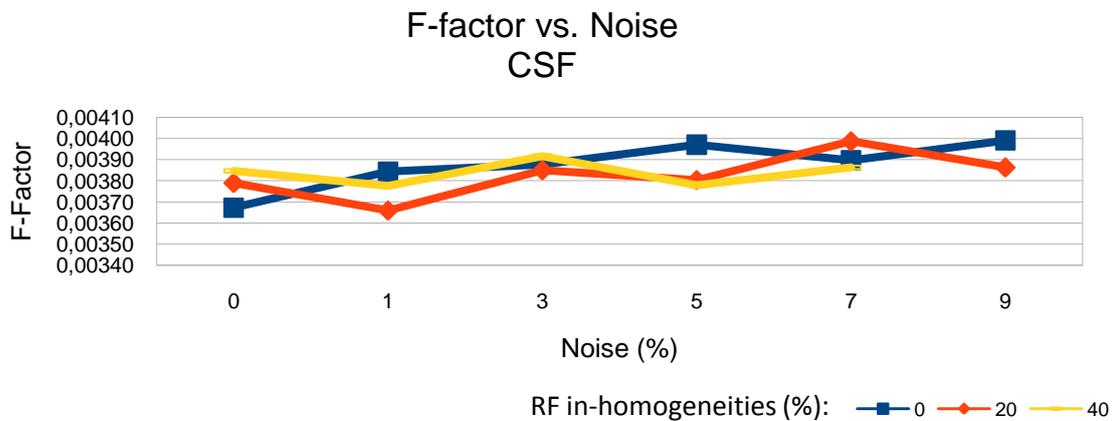


Figure 59. FreeSurfer: CSF f-factor.

**Overall result**

If we compute the error ratio for FreeSurfer, we can see that the results are very bad, with values higher than 50% error when considering only non-empty voxels. However, it does not mean that the result is wrong because of the way FreeSurfer works. It extracts all structures in the brain, which are composed by grey and white matter, but it does not classify them as grey matter or white matter. For that reason, the results could be good but it is not what we are interested in, as explained when defining the problem.

Error ratio (all voxels)						
	Noise					
RF	0	1	3	5	7	9
0	0,15672	0,15672	0,15674	0,15618	0,15737	0,15604
20	0,15661	0,15688	0,15701	0,15719	0,15655	0,15633
40	0,15658	0,15618	0,15697	0,15704	0,15601	

Table 40. FreeSurfer: Error ratio (all voxels).

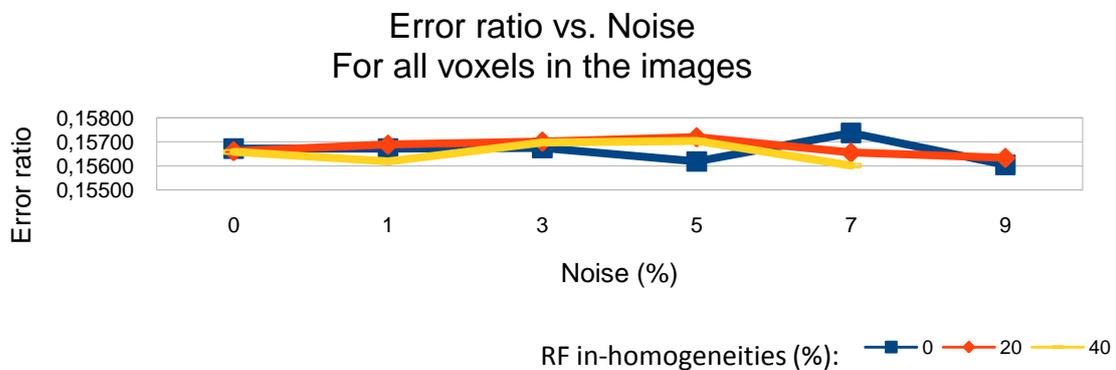


Figure 60. FreeSurfer: Error ratio (all voxels).

Error ratio (non-empty voxels)						
	Noise					
RF	0	1	3	5	7	9
0	0,56324	0,56319	0,56315	0,56112	0,56538	0,56029
20	0,56270	0,56376	0,56413	0,56466	0,56230	0,56156
40	0,56272	0,56157	0,56415	0,56433	0,56063	

Table 41. FreeSurfer: Error ratio (non-empty voxels).

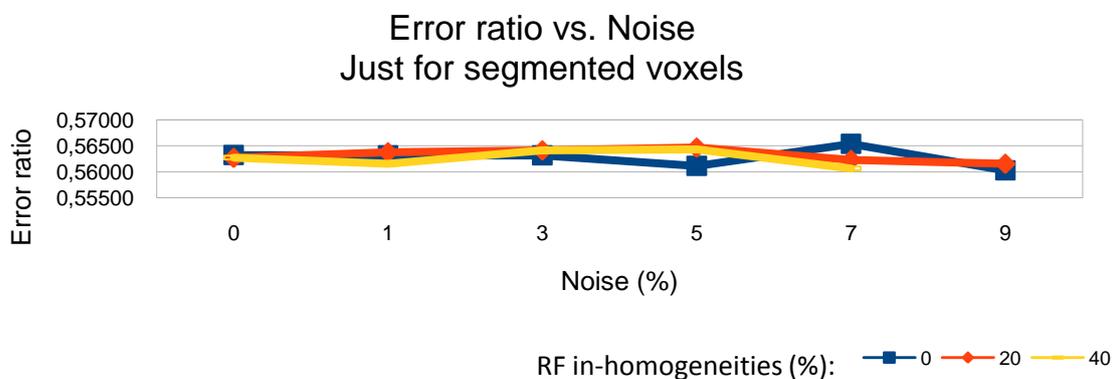


Figure 61. FreeSurfer: Error ratio (non-empty voxels).

## 6.2. Images

Once segmentation is performed and we have the comparison results, we can see the output obtained from each software. Since there are a lot of images, we are going to show the best result obtained for FSL, SPM and FreeSurfer.

### 6.2.1. FSL

We can see in the following figure the results obtained in FSL for white matter, grey matter and CSF segmentation respectively. The images were obtained for 1% noise and 20% RF in-homogeneities. In these pictures, the ground truth is shown in green and the segmented result in red.

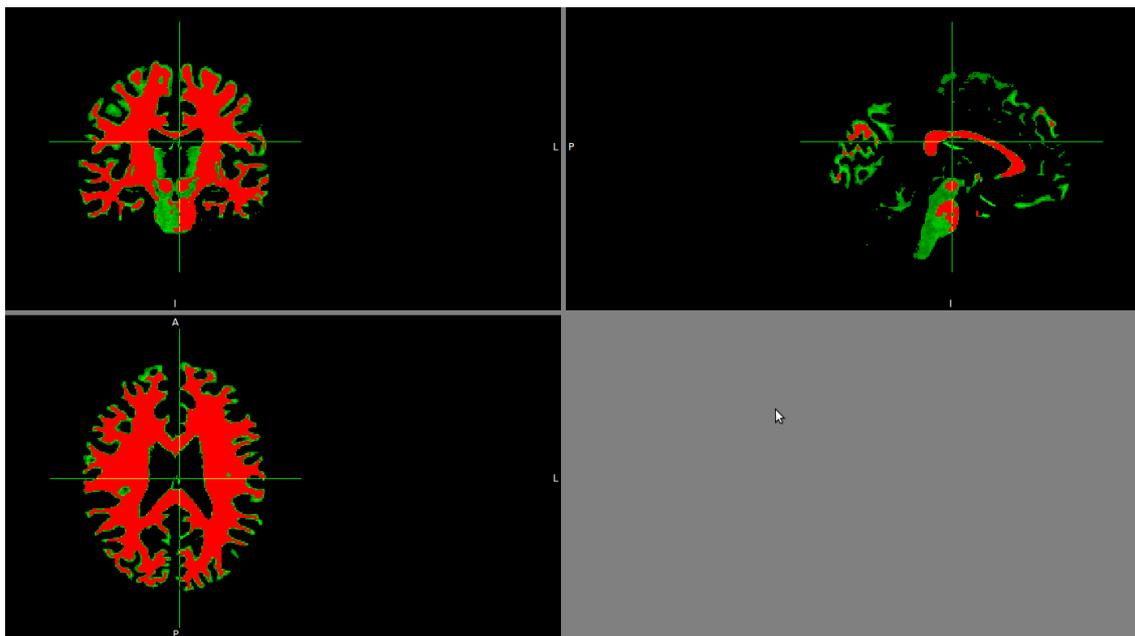


Figure 62. FSL: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

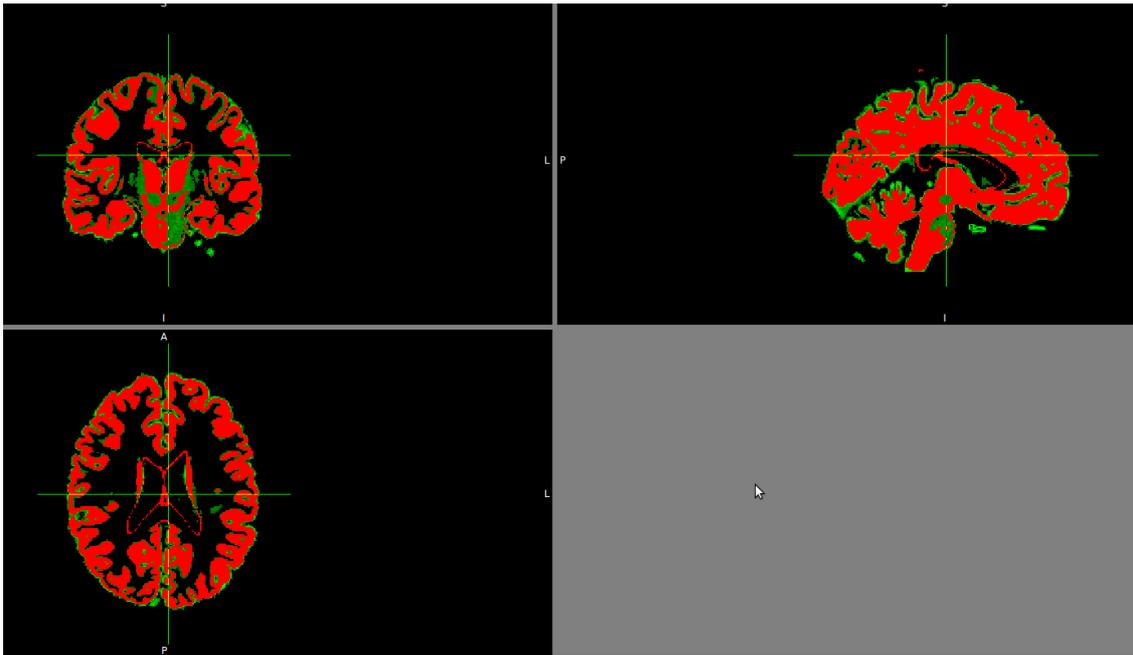


Figure 63. FSL: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

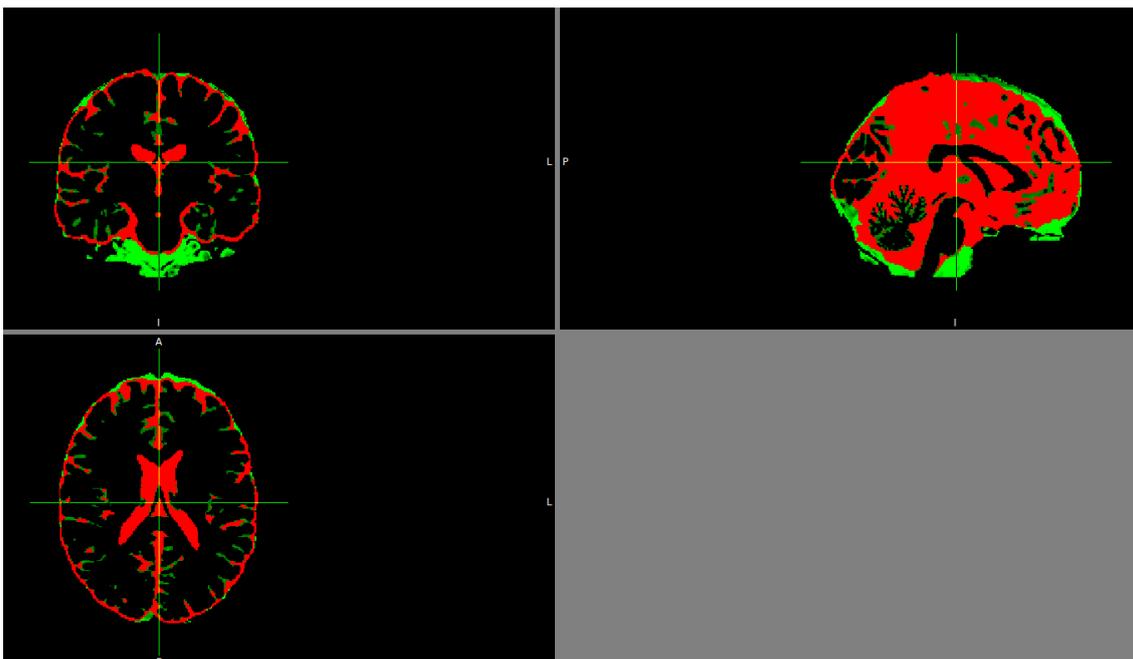


Figure 64. FSL: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

As can be observed in images, most of the problems found in FSL to classify all tissues are found in the borders between them. Anyway, we can see that the results we have obtained for white and grey matter are good. In the case of CSF, the result is not that good because part of the CSF was removed in the brain extraction process.

### 6.2.2. SPM

Regarding to SPM the results for the images with 3% noise and 20% in-homogeneities are shown below:

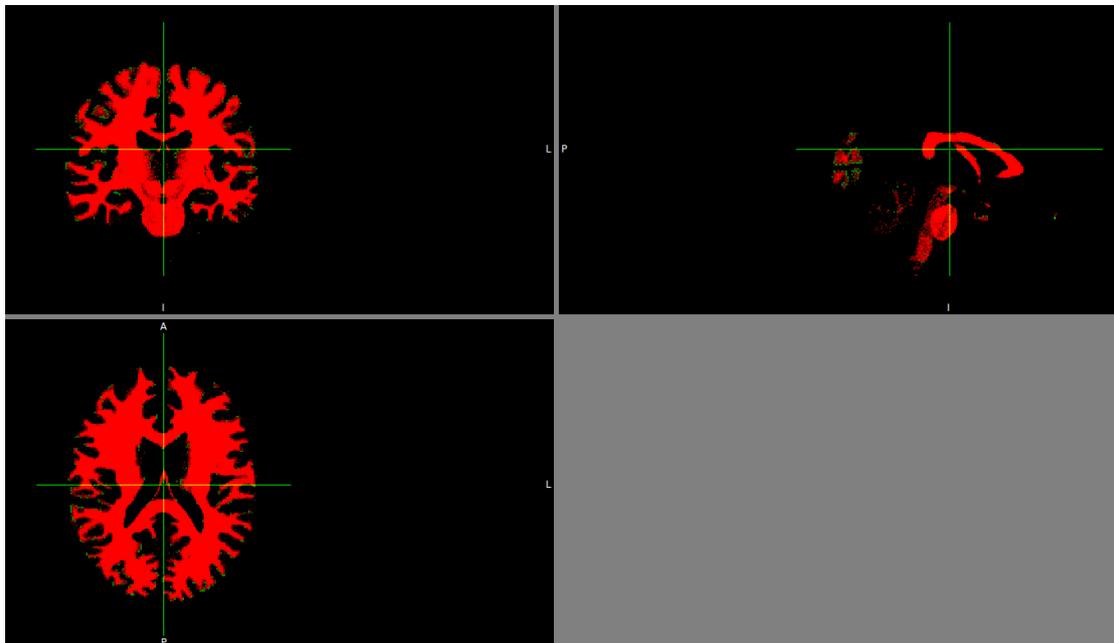


Figure 65. SPM: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

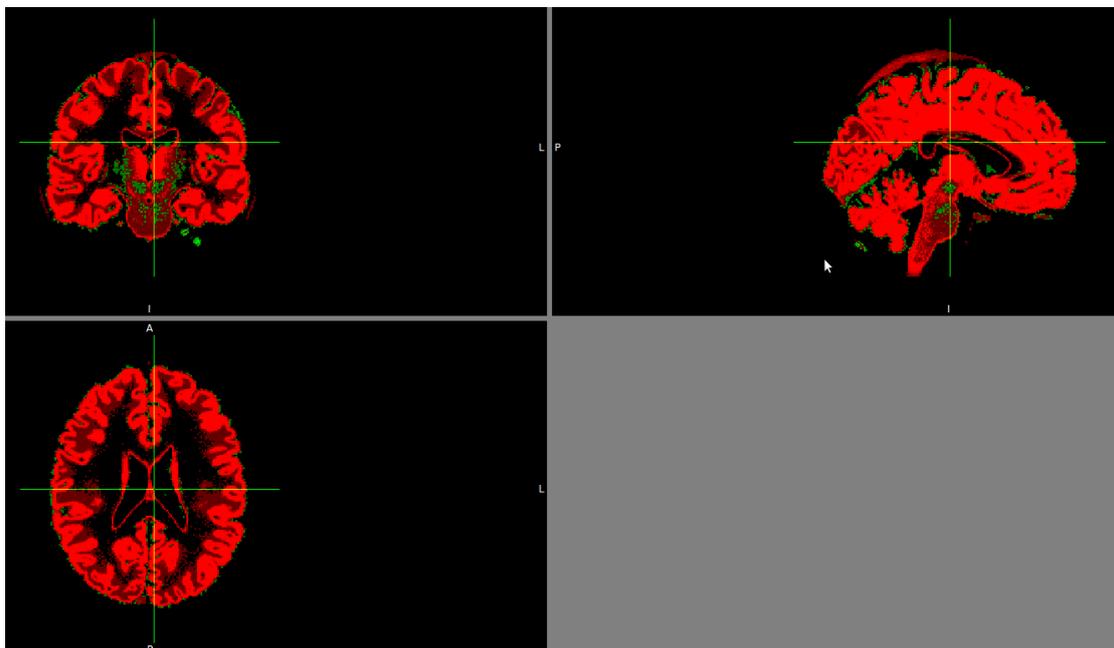


Figure 66. SPM: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

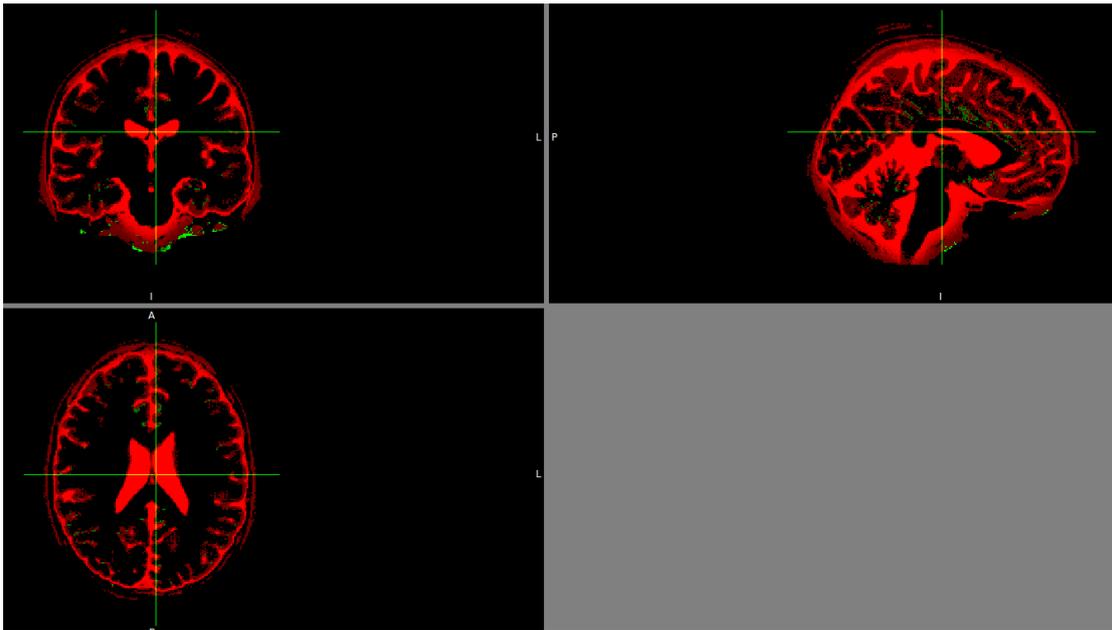


Figure 67. SPM: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

The results obtained for SPM are also good for white and grey matter, even though some errors can be found. Moreover, the result for CSF is wrong, so that part of the skull and scalp are classified as CSF.

### 6.2.3. FreeSurfer

Finally, for FreeSurfer, the results for 9% noise and 20% RF in-homogeneities are shown below:

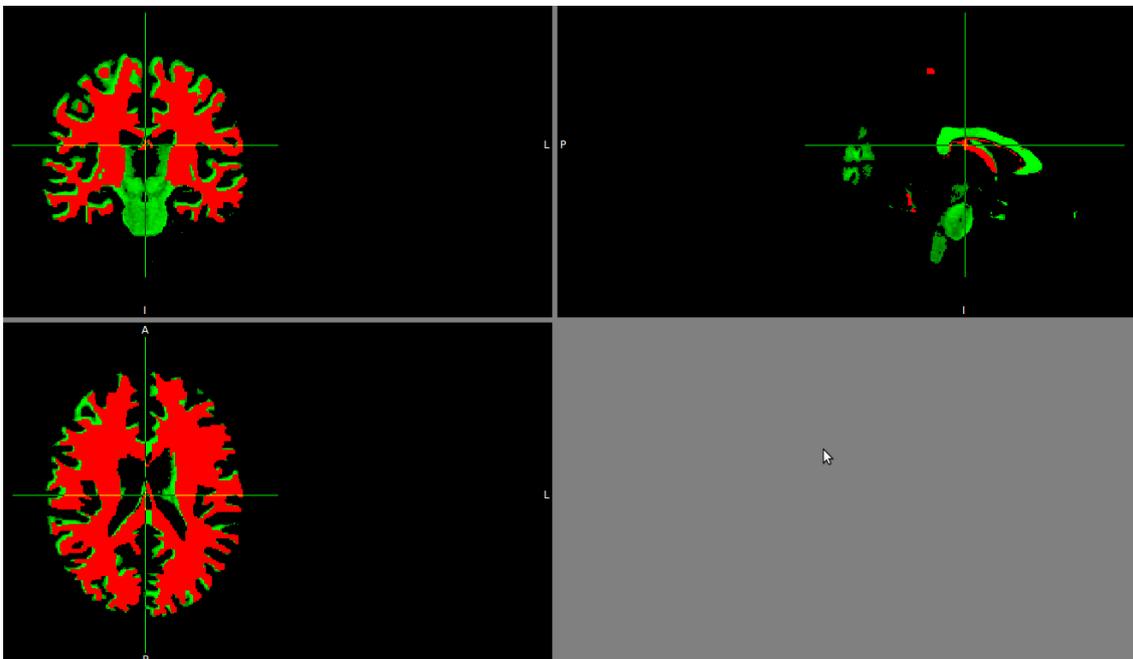


Figure 68. FreeSurfer: white matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

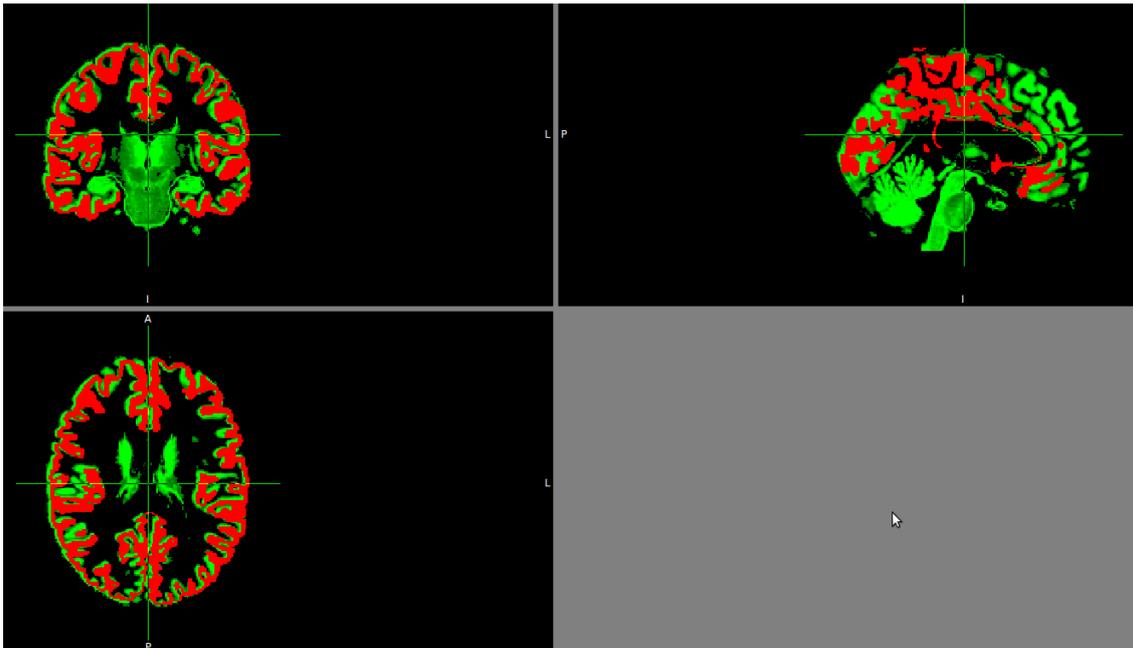


Figure 69. FreeSurfer: grey matter vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

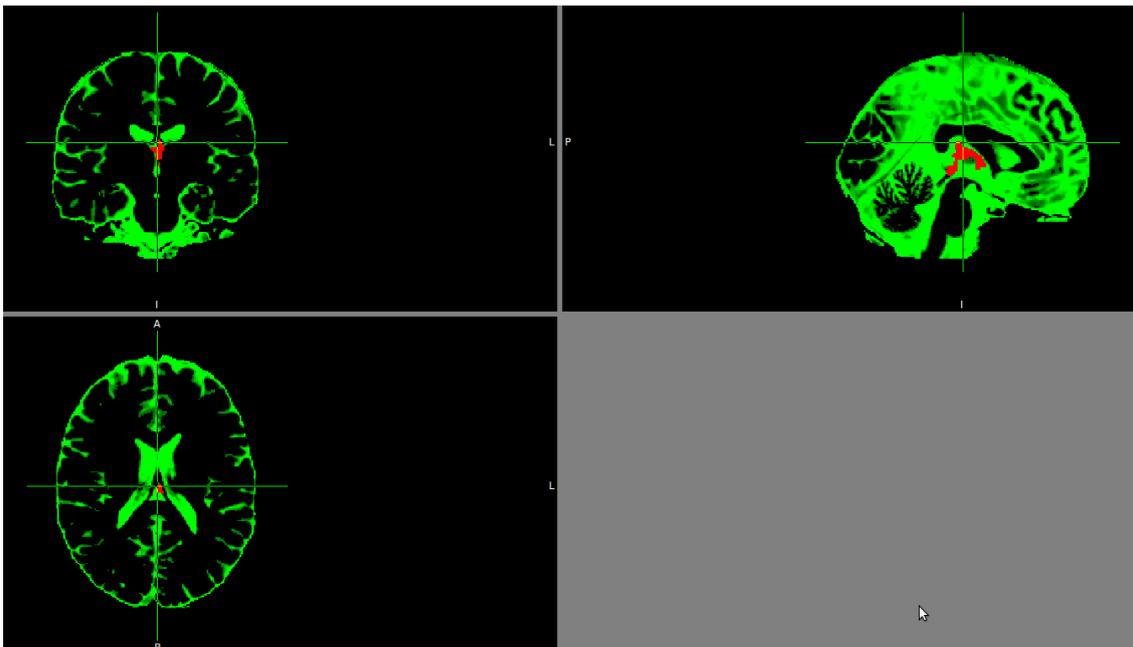


Figure 70. FreeSurfer: CSF vs. ground truth: coronal(top-left),sagittal(top-right) and axial(bottom).

Looking at these results, we can see that FreeSurfer is giving incomplete results as output. The reason is that it is classifying structures instead of types of tissue. In that way, we can't consider these results as wrong,.

## 7. Conclusions

The goal of this project at the beginning was to evaluate the performance of FSL and FreeSurfer in order to segment between different brain tissues and sub-cortical structures. This was necessary because our research-group was trying to develop patient-specific electromagnetic models for both dipole source location in electroencephalography and stroke detection (via microwave helmet, with Midfield Diagnostic).

While thinking on the problem, we considered it was also necessary to include SPM as another software package we had to analyze, because it's one of the most used software tools worldwide and we read much information about its good results regarding to cortical segmentation.

Once we could evaluate the segmentation performed by each one of the tools we have used (having a ground truth from Brainweb phantoms), the results we have found have been very different depending on what we wanted to obtain. The reason is that these three software packages used have been designed for a slightly different purpose, even though they all perform brain tissue segmentation. The differences we have found, both in their goal and their performance, are exposed below.

FSL is a software package that has been designed for a general purpose if we talk about brain segmentation. The meaning of general purpose here is that it is able to perform both brain tissue segmentation (white matter, grey matter and cerebro-spinal fluid) and sub-cortical segmentation (brain stem, amygdale, hippocampus...) in separated processes. Between other reasons, it's this modularity what gives FSL the capability to perform so many different tasks, since it's composed by different tools with different goals.

Regarding to FSL performance for brain tissue segmentation, it is able to classify grey matter with very good results (f-factor over 80%) and has also good results for white matter (f-factor over 75%, excepting one case). The result for CSF is not very good (f-factor over 62%) and is very influenced by the noise in the image. In general, we can see that segmentation performed by FSL is very influenced by the noise, and just a little by RF in-homogeneities, having better results when both have an intermediate value (the best result is obtained with 1% noise, with an error ratio close to 8% for non-empty voxels).

SPM, as it is concerned, is only designed to perform segmentation between white matter, grey matter and CSF. Unlike FSL, it is not able to perform sub-cortical segmentation. One advantage of SPM could be that, since it's designed to run in MatLab, it's not platform-dependent and it can be run in both Linux and Windows operating systems.

The performance of SPM, if we look at the graphics in the results section, is not influenced (almost) by RF in-homogeneities but, as it happened with FSL, it's very influenced by the noise. It can give a better result for white matter segmentation when comparing to FSL (f-factor over 83%) but its result for grey matter segmentation is slightly worse (f-factor over 68%). Regarding to CSF, the result is very bad (the highest f-factor is 3.8%) and can't be considered if we want to segment CSF from a MR image.

Finally, FreeSurfer was designed to perform segmentation between brain structures, both cortical and sub-cortical. However, the comparison with FSL or SPM when segmenting between white matter, grey matter and CSF is not fair. If we look at the results, the error ratio is always over 50% so we could think that this result is really bad. The problem is that where FSL or SPM see white or grey matter (or a combination of both), FreeSurfer can see a brain structure, like brain stem. The reason is that these structures are composed by both grey and white matter, so that the result from FreeSurfer can't be considered as wrong. What we can say is that FreeSurfer, at least, is affected by the noise and RF in-homogeneities because it couldn't perform segmentation when they both had their maximum value in the experiments.

Regarding to all the results we have obtained, we can conclude that FSL is the most complete package to perform brain segmentation, since it has better results for tissue segmentation than SPM and it can also perform sub-cortical segmentation. However, if we are really interested in sub-cortical structures, we should consider FreeSurfer as an alternative, because it's able to extract more sub-cortical structures than FSL.

## 8. Bibliography

1. Program for Imaging and Cognitive Sciences. Columbia University Medical Center. [Online] [Cited: March 20, 2010.] <http://www.fmri.org/index.html>.
2. Hornak, Joseph P. The basics of MRI. [Online] [Cited: March 10, 2010.] <http://www.cis.rit.edu/htbooks/mri/>.
3. e-MRI: MRI physics interactive tutorial (Online remote education for health professionals). [Online] [Cited: March 20, 2010.] <http://www.imaio.com/en/e-Courses/e-MRI/>.
4. Sociedad Española de Neuroimagen. [Online] [Cited: March 20, 2010.] <http://www.neuroimagen.es>.
5. Jerry L. Prince, Jonathan M. Links. *Medical Imaging, Signals and Systems*. 0-13-065353-5.
6. BrainWeb: Simulated Brain Database. [Online] [Cited: April 1, 2010.] <http://mouldy.bic.mni.mcgill.ca/brainweb/>.
7. Shen, Jimmy. Tools for NIFTI and ANALYZE image. [Online] April 13, 2010. [Cited: MAY 3, 2010.] <http://www.mathworks.com/matlabcentral/fileexchange/8797-tools-for-nifti-and-analyze-image>.
8. Baeza-Yates, Ricardo and Ribeiro-Neto, Berthier. *Modern information retrieval*. s.l. : ACM Press, 1999. 020139829X.
9. Olson, David L. and Delen, Dursun. *Advanced Data Mining Techniques*. s.l. : Springer, 2008. 3540769161.
10. FMRIB Software Library. [Online] [Cited: March 1, 2010.] <http://www.fmrib.ox.ac.uk/fsl/>.
11. FreeSurfer. [Online] [Cited: March 1, 2010.] <http://surfer.nmr.mgh.harvard.edu/>.
12. SPM. [Online] [Cited: March 1, 2010.] <http://www.fil.ion.ucl.ac.uk/spm/>.
13. Ramon, Ceon, Schimpf, Paul H. and Haueisen, Jens. Influence of head models on EEG simulations and inverse source. *Biomedical Engineering Online*. [Online] February 8, 2006. <http://www.biomedical-engineering-online.com/content/5/1/10>.
14. J. Talairach, P. Tournoux. Co-planar Stereotaxic Atlas of the Human Brain: 3-Dimensional Proportional System - an Approach to Cerebral Imaging. New York : Thieme Medical Publishers, 1988.
15. J. Talairach, P. Tournoux. Referentially Oriented Cerebral MRI Anatomy: An Atlas of Stereotaxic Anatomical Correlations for Gray and White Matter. New York : Thieme Medical Publishers, 1993.
16. On Tsang, Ali Gholipour, Nasser Kehtarnavaz, et al. Comparison of tissue segmentation algorithms in neuroimage analysis software tools. Vancouver. 30th Annual International IEEE EMBS Conference, 2008.