MASTER THESIS

Surface-based mapping of the serotonin transporter binding in cerebral cortex

- filtering and modeling of PET/MRI data

Mikael Agn

INTERNAL SUPERVISOR, DTU
Koen Van Leemput

EXTERNAL SUPERVISOR, NRU
Claus Svarer

DTU COMPUTE
TECHNICAL UNIVERSITY OF DENMARK

NEUROBIOLOGY RESEARCH UNIT
COPENHAGEN UNIVERSITY HOSPITAL

Kongens Lyngby 2013
M.Sc-2013-79

DTU Compute
Technical University of Denmark
Matematiktorvet, building 303B
DK-2800 Kgs. Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@compute.dtu.dk
www.compute.dtu.dk M.Sc-2013-79

Abstract

The aim of this project was to improve the filtering and modeling of PET data, to better handle the high noise level present in such data. It should lessen the artifacts of conventional volume-based spatial filtering with a Gaussian kernel. The focus was on the cerebral cortex due to its highly-folded and thin structure, which makes it particularly unsuited for the conventional approach to filtering.

A surface-based approach was developed, which took the highly folded intrinsic geometry of cerebral cortex into account in the filtering and modeling by the multilinear reference tissue method MRTM2. A surface representation of the cerebral cortex was obtained by the software package FreeSurfer. By smoothing across the surface of the cortical layer the data was less affected by edge artifacts, as the smoothing is done between more functionally connected regions with similar neuronal density. The approach was contrasted with the conventional volume-based approach with good results. In addition, the surface-based statistical tools of FreeSurfer for group analysis between brains was evaluated with the use of this model.

Furthermore, a Bayesian framework was used to directly incorporate the data filtering into the mathematical model. This model was based on MRTM2 and assumed that close regions have similar parameters and regularized the data on this assumption. This model was shown to handle high levels of noise better than the ordinary surface-based approach, while at the same time retaining a higher resolution and detail. In addition, it resulted in a higher repeatability between scans on a vertex level in a test-retest setting. Furthermore, attempts were made to treat the data in a fully Bayesian approach by including optimization of the hyperparameters of the model.

keywords: positron emission tomography, magnetic resonance imaging, multimodality, FreeSurfer, Bayesian modeling, regularization, cerebral cortex, MRTM2, serotonin transporter, reference tissue.

Preface

This thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark in fulfillment of the requirements for acquiring a M.Sc. in Mathematical Modeling and Computation. It was written in collaboration with the Neurobiology Research Unit (NRU) at the Copenhagen University Hospital.

The work was carried out from February to August 2013. The project was supervised by Ph.D. Koen Van Leemput, Associate Professor at DTU Compute, and Ph.D. Claus Svarer, Chief Engineer and Research Associate at NRU.

The thesis deals with the filtering and modeling of functional positron emission tomography data. The aim is to improve the handling of the high noise level present in such data. It focuses on the special challenges of the highly-folded cerebral cortex and to further develop the well-known and often used multilinear reference tissue model by the use of a Bayesian framework and structural magnetic resonance imaging.

Lyngby, August 19, 2013

Mikael Agn

Acknowledgments

I would like to thank my internal supervisor Koen Van Leemput for his competent supervision and the very giving theoretical discussions about mathematical modeling approaches, especially within the field of Bayesian modeling. Futhermore, I would like to thank my external supervisor Claus Svarer for his always engaging daily supervision, his guidance into the field of neuroimaging and for allowing me to carry out this project at NRU.

I want to direct special thanks to Vibe G. Frøkjær for being able to use her study in my thesis and for all the help and encouragement. I also want to thank Gitte Moos Knudsen for valuable comments on the progress of the project and all other colleagues at NRU for their friendly and helpful attitude.

I want to thank Douglas N Greve for the discussions and inspiration about the implementation in FreeSurfer and his always fast and relevant answers to any question I had about FreeSurfer. In addition, I want to thank the professors and fellow students at the image analysis section at DTU Compute for the constructive weekly meetings.

Lastly, I would like to thank my family and friends for their love, patience and support throughout the project.

Contents

1. Introduction

To enable communication between nerve cells or neurons in the human body, a chemical substance called neurotransmitter is necessary. One of the most important neurotransmitters is serotonin. Serotonin is involved in a wide range of processes in the brain such as behavior, cognition, mood, memory processing and response to stress. One important component of the serotonergic transmitter system is the serotonin transporter, which transports released serotonin back to the releasing cell. It is the most important element in controlling the strength and duration of serotonergic neurotransmission. The serotonin transporter is used as target in treating major depression and anxiety disorder and has been linked to many other disorders.

A widely used technique to measure functional processes in the brain is positron emission tomography (PET). The image acquisition process of an individual starts by injecting a radioactive substance called a radiotracer into the blood stream. [^{11}C]DASB is a radiotracer which binds selectively to the serotonin transporter. At each radioactive decay event, two rays are released in opposite directions to each other. The rays are detected by sensors in the PET scanner and the origin of the signal will thus be somewhere on the line between the two detecting sensors. With enough lines collected, a reconstruction process can reconstruct a 3-dimensional image of the signals in the brain. In a typical PET scan several such frames are reconstructed in a time interval. PET includes many sources of noise, e.g. rays can be scattered or even lost before reaching the sensor and multiple rays can reach the sensors at the same time. Also, the injected dose is usually small. This means that the signal-to-noise ratio is low.

The frames show counts of radioactive events, but a desirable measurement should be more directly linked to concentration of serotonin transporters. Therefore, the time frames are modeled to obtain a measurement called binding potential. This is called parametric imaging, as the binding potential will be obtained through the estimated model parameters. One such modeling method is the multilinear reference tissue method MRTM2 which can be solved by linear least squares, on a voxel level or a larger regional level. Most of these models are sensitive to high noise levels because they involve noisy predictors, which means that they will be unstable when using the highly noisy data from PET at voxel level.

The traditional way of improving the signal-to-noise ratio in PET data and making the noise more normally distributed is to spatially smooth the data with an Gaussian filter. The motivation for smoothing in space is that close regions in the brain often exhibit similar characteristics and brain functions are assumed to be clustered in the brain. There are however some significant drawbacks, with the most important one being edge artifacts. Several functional regions of the brain have abrupt borders. Especially at the outer border of the brain, the change is rapid between the cerebral cortex with a high concentration of neurons and cerebrospinal fluid with

no neurons. Cerebral cortex is also connected to white matter that mainly contains axions connecting neurons from the cortex to other neuron rich areas of the brain. Also in the folds of the highly-folded cortex, close regions will lack physiological connections. If the filter smooth across these borders, the resulting signal will be inaccurate. The Gaussian filter also reduces the spatial resolution. Furthermore, there is a risk at high noise levels that some noise will not be effectively reduced but instead will spread to surrounding areas.

The aim of this project is to improve the filtering and modeling of PET data. It should lessen the mentioned artifacts and improve the statistical power of the brain mapping of the serotonin transporter. The focus will be on the outer layer of the brain – the cerebral cortex. A reconstruction of the surface of cerebral cortex is obtained by the software package FreeSurfer which has been developed with this task in mind. FreeSurfer uses structural magnetic resonance scans to segment the brain and reconstruct the cortical surface. A surface-based approach will be developed, which will take the highly folded intrinsic geometry of cerebral cortex into account in the filtering and MRTM2 modeling process. By smoothing across the surface of the cortical layer the data should be less affected by edge artifacts, as the smoothing is done between more functionally connected regions with similar neuronal density. The approach will be contrasted with the conventional voxel-based approach. Furthermore, a Bayesian framework will be used to address the other mentioned issues of spatial smoothing and to directly incorporate the data filtering into the mathematical model. This model will be based on MRTM2 and will assume that close regions have similar parameters and regularize the data on this assumption. The approach can potentially filter out the noise more effectively than Gaussian pre-filtering, because it assumes that the physiologically interpretable parameters themselves should be similar instead of assuming that the event count rate at each time frame should be similar. In addition, the surface-based statistical tools of FreeSurfer for group analysis between brains will be evaluated.

The project is a collaboration between the image analysis section at DTU Compute at the Technical University of Denmark and the Neurobiology Research Unit (NRU) at Copenhagen University Hospital. A human in-vivo brain study of 61 subjects has been provided from NRU. It is a woman sex hormone fluctuation study, with 31 subjects given sex hormone treatment and 30 subjects given placebo. The study contains brain PET scans and external mood measurements before and after treatment. The time between the baseline and followup scans are for most subjects around one month. The PET scans are performed with the [^{11}C]DASB radiotracer and a structural high resolution T1-weighted MR scan is also provided for each subject. This study will be statistically analyzed to evaluate the tools of FreeSurfer and the placebo group will be used as a test-retest data set when comparing different filtering and modeling approaches.

2. Theoretical background

In section 2.1, the serotonergic transmitter system will be explained in more details and in section 2.2. the functionality and structure of cerebral cortex will be described. In section 2.3, theory about the positron emission tomography will be presented including the conventional ways of handling the high noise level present in such data. Lastly, in section 2.4 some theory on magnetic resonance imaging will be explained.

2.1. Serotonergic transmitter system

Neurons are the core cells of the nervous system in the human body. Their main function is to send fast electrical signals as communication throughout the body and especially in the brain. A simplified illustration of a neuron cell is shown in figure 2.1.1. It involves a cell body from which dendrites and an axon extend. Other neurons connect to the neuron at the dendrites or less commonly directly to the cell body or the axon. The axon can extend to up to a meter and the primary direction of the electrical signal is from the cell across the axon to another cell's dendrites. The electrical signal will however not go directly from one neuron to the other. The transmission of the signal between two neurons happens in a synapse, by a chemical compound called a neurotransmitter. When an electrical signal reaches the presynaptic part, the neurotransmitter will be released into the synaptic cleft and travel to the postsynaptic part which often is a dendrite. There, receptors will receive the neurotransmitter and thus be activated and trigger an electrical signal. Unused neurotransmitters can be removed from the cleft back to the presynaptic part by transporters. The neurotransmitter which is studied in this report is serotonin.

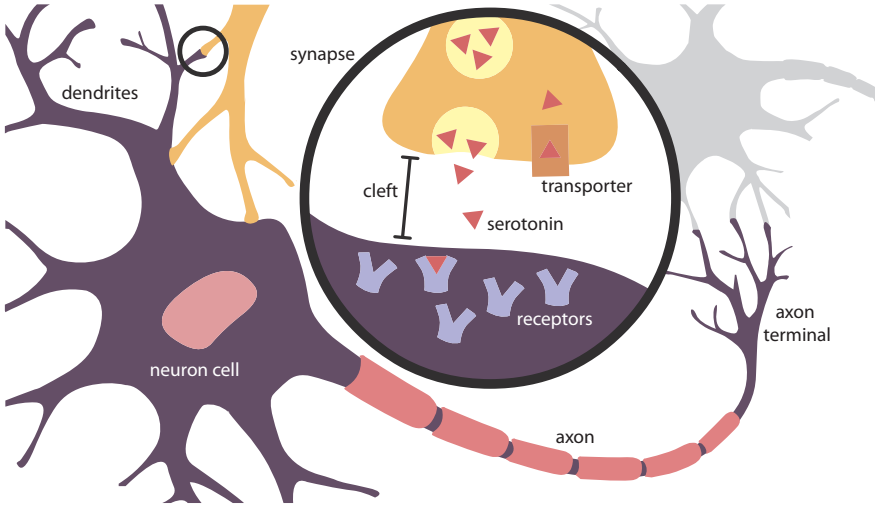

Figure 2.1.1. Neuron cell and synapse.

Serotonin or 5-hydroxytryptamine (5-HT) is mainly located in the gut, where it regulates intestinal movements, but there are also large concentrations in the brain. It regulates functions related to mood, appetite, memory processing, learning, behaviour, cognition and sleep. As seen in figure 2.1.2, it is produced in the raphe nuclei in the brain stem and from there transported within axons to other parts of the body. The upper raphe nuclei transports serotonin to most parts of the brain (Bue Klein, 2010).

Figure 2.1.2. Pathways of serotonin transmission.

There are one serotonin transporter (SERT) and a number of different serotonin receptors with similar but somewhat differing functions. This thesis concentrates on the serotonin transporter. As its function is to transport serotonin back to the presynaptic part, it has a key role in regulating the duration and strength of transmission between serotonergic neurons. Therefore, it is commonly used as the target of antidepressant drugs. Is has been linked to many neurological disorders such as major

depression, bipolar disorder, obesity, substance abuse, eating disorders, obsessive compulsive disorder, cognitive impairment, primary insomnia, Parkinson's disease and autism (Charnay and Léger, 2010; Daws and Gould, 2011).

The concentration of serotonin transporters is highest in subcortical regions such as the raphe nuclei, thalamus, hypothalamus, amygdala, putamen and caudate. Cerebral cortex is also a significant region with the highest concentrations in the cingulate cortex and insular cortex and moderate to low concentrations in other areas. The lowest concentrations are found in cerebral white matter and cerebellum, both the cerebellar cortex and cerebellar white matter (Saulin et al., 2012; Kish et al., 2005).

2.2. Cerebral cortex

The cerebral cortex is the outermost layer of each hemisphere of the human brain. It contains gray matter with a high density of neural cell bodies. The cerebral cortex is connected to white matter, which mainly contains axions connecting cortical neurons to neurons deeper within the brain and to neurons in other parts of cerebral cortex. Cerebrospinal fluid outside of the brain acts as a buffer and protection for the cortex, and contains no neural activity. The cortical layer is important for many functional processes of the brain. It plays a key role in most higher functions such as how we think, speak, associate, move, process memory and perceive the world around us. It is divided into six major lobes: the frontal lobe, the temporal lobe, the parietal lobe, the occipital lobe, the insular lobe and the limbic lobe. The five first lobes are together called the neocortex, the newest developed part of the brain.

The functions of some areas of the cortical layer are well-known, such as the primary visual cortex in the occipital lobe, the primary auditory cortex in the temporal lobe, primary motor cortex in the posterior part of the frontal lobe and many more. However, much knowledge is yet to be discovered about many cortical functions and the complex interconnections between areas and this is intensely researched upon.

The cortical layer is continuous and highly folded with a thickness of 1 to 5 mm. The ridge of a fold is called a gyrus and the valley between two gyri is called a sulcus. The many folds and complicated geometry of the cortical layer makes it hard to visualize and compute statistics on this layer. Regions physiologically far apart with few connections lies spatially close together and the extent of the layer can not simply be measured in one direction in voxel space. The thin structure of the layer in combination with the rather low resolution of most imaging techniques makes partial volume effects an important issue. Furthermore, the variability of the geometry between brains is rather large. Some larger gyri are present in every brain, but many smaller structures varies considerably. This makes the cerebral cortex a particular challenge in a statistical analysis between brains.

2.3. Positron emission tomography

Positron emission tomography (PET) is a technique used to acquire three dimensional images of functional processes in the body. The procedure starts by injecting a radioactive substance into the blood stream of the subject. The substance is called a radioligand or a radiotracer and contains a type of molecules which are used in the functional process of interest. To measure the serotonin transmission in the brain, it could be a compound which binds to a serotonin receptor or transporter. A positron-emitting isotope is attached to these molecules. One such isotope is carbon-11 (^{11}C). It is unstable and decays to stable boron-11 with a half-life of around 20 minutes. In the decay process a proton in the nucleus converts into a neutron. In this process a positron is released with a kinetic energy. It loses the energy when interacting with atoms in the surroundings and eventually, when it is essentially at rest, it combines with an electron also at rest. In this combination, called annihilation, they cease to exist and two gamma ray photons are emitted in opposite direction to each other, at 180 degrees. This is the basis of PET, where several rings of gamma sensors measures the impact of both photons and the line of response (LOR) can be computed. In some cases, the kinetic energy of the positron and electron will not be exactly zero and the photon pair will in that case not be emitted exactly at 180 degrees. This fact and the short distance traveled by the positron before annihilation (under 1 mm) contributes to an uncertainty in PET.

After the annihilation event, the photons might in some cases interact with the tissue, which can result in a change of direction or that the photon do not reach the sensors. The change of direction causes a scatter event. Photons from different events can also reach the sensors at the same time, this is then called a random event. The event detection in PET relies on electronic collimation and collects events with (1) two photons detected within a predefined time window, (2) the line of response between the photons are within a valid angle for the tomograph and (3) the energy of the photons are within a selected energy window. All events are collected in an ordered list over the whole time of the PET scan, which is in the case of the study used in this report 90 minutes. The duration of a scan depends on the characteristics of the kinetic process of the radiotracer and the functional system that is being scanned. It also depends on the decay rate of the radioactive isotope. At some point the detected events will be too low for a reliable measurement.

The events are then reconstructed to 3-dimensional frames. For the 90 minutes scan time, 38 dynamic frames will be reconstructed with varying time intervals between the frames. In the beginning of the scan the event detection rate will be high and a short interval between the reconstructed frames is possible, but as the radioactive isotope decays fewer and fewer events will be collected. There are several methods for reconstruction, but in all methods the data has to be corrected for geometrical distortion because of the circular shape of the detector rings, attenuation of the photons and dead time of the detectors. The approach can either be analytical or iterative. The iterative approach usually employs a type of Expectation Maximization algorithm (Bailey et al., 2005). Each reconstructed frame will be a 3-dimensional volume with voxels as discrete units. The voxel size equals the full width half maximum of the estimated spatial resolution, which depends on the properties of the scanner, tracer and the reconstruction algorithm in use. One

of the more advanced PET scanners on the market, the high resolution research tomograph (HRRT), can have a spatial resolution of full width half maximum 1.2 mm in all three directions. The 38 values of one voxel is called a time activity curve (TAC). An example of a TAC can be seen in figure 2.3.1. It should be noted though that this is an averaged TAC on a larger region, with significantly less noise than can be observed at voxel level. The noise level of a PET scan at voxel level is high due to the many sources of noise, the low dose of radioactive tracer and the complicated reconstruction procedure. Furthermore, the reconstruction procedure often results in an unpredictable noise distribution.

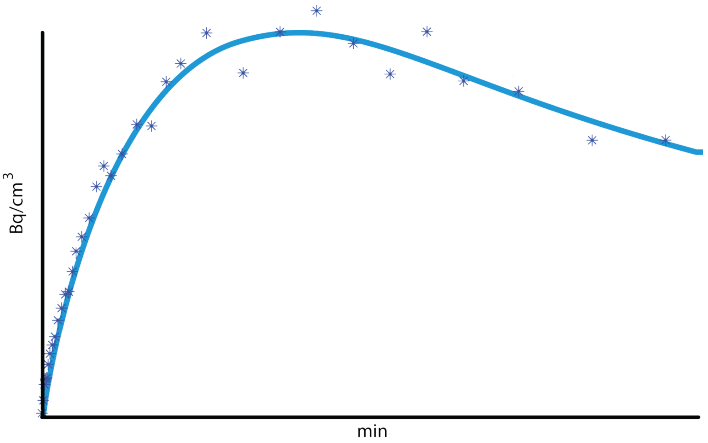

Figure 2.3.1. Example of a time activity curve (TAC).

2.3.1. Regional averaging to improve signal-to-noise ratio

Voxels in a region of interest is for many analyses averaged to a regional TAC, which will improve the noise properties of the data. Averaged TACs of pre-defined regions can give a good overview of the data and is quite often used as the main input in a statistical analyses between brains. Cortical regions are usually defined anatomically by main gyral structures, such as the middle temporal gyrus or superior frontal gyrus. An analysis can also be performed on the cortical lobes. The selected regions differ between studies, but are usually quite consistent. The advantages of such an analysis are that it is easy to compare it between studies and only a few tests needs to be performed. With just a few test, it is easier to get an overview and the question of correction for multiple comparisons is easier to handle. Instead of correcting for individual tests on a large set of voxels, one can correct only for a few regions of interest. It is also easier to account for the variability between individual brains. The main difficulty with this approach is that the anatomically defined cortical regions may not be functionally appropriate. The functional measurements do more often than not vary just as much within a region as between regions. To average distinctively different voxel TACs might alter the output in unexpected ways and significant areas within a region might not be discovered. Functionally interesting patterns might also be overlooked (Poldrack, 2007).

2.3.2. Smoothing to improve signal-to-noise ratio

A common way of improving the noise properties on a voxel level is to spatially filter the data within each time frame with a Gaussian filter, i.e. convolution with a Gaussian-shaped kernel. The true or desired signal component of a measured data set is usually changing rather smoothly, while the noise often is seen as rapid random changes both in space and in time. If this assumption is true, a smoothing of the data can improve the signal to noise ratio and sensitivity, by removing the high frequencies in the data. The motivation for spatially smoothing brain data is that close regions in the brain often exhibit similar characteristics and brain functions are assumed to be clustered in the brain. Another reason for spatially smoothing the data can be to account for subject variability in a group analysis at voxel level. To register subjects to a common space is a difficult process as there are high variability in the structure of brains, so voxels with similar function will quite possibly not be exactly aligned.

The spatial smoothing does have some drawbacks, with the most important one being edge artefacts. Several functional regions of the brain have rather abrupt borders. Especially at the border of the whole brain, the change between the highly functional cerebral cortex and cerebrospinal fluid with no neural functionality is rapid. Cerebral cortex also borders to white matter, which has a low concentration of neurons. Another problematic area is in the folds of the cortex, where close regions will lack direct physiological connections. If the filter smooth across these borders, the resulting signal will be inaccurate (Hagler et al., 2006).

It is clear that spatial smoothing also reduces the spatial resolution. In addition, signal peaks can be displaced. One situation when this can occur is when there are two neighboring regions with high amounts of signal. A smoothing might then result in a displaced peak between these regions instead of one peak for each region. In the presence of high noise levels such as in PET data, smoothing might not effectively remove high frequencies. If the signal of a data point is very large due to noise, a Gaussian smoothing will spread this signal to neighboring data points instead of filtering it out (Reimold et al., 2006; Sacchet and Knutson, 2013).

2.4. Magnetic resonance imaging

Magnetic resonance imaging uses magnetic resonance of protons in the brain to get highly detailed structural 3D images of the brain. All protons have a weak magnetic moment. The subject is placed in a scanner with a large magnetic coil surrounding the subject. In the strong magnetic field, many of the protons in the body gets aligned. A varying electromagnetic field is briefly produced with a special frequency called the resonance frequency, which makes the protons spin. The radiofrequency is turned off and the protons aligns again with the magnetic field. In this relaxation, the protons themselves produce radio frequency signal which is measured by a receiver coil. By using gradient coils, the strength of the magnetic field can be varied to different locations, which also makes the released radio frequency vary predictably dependent on location. The measured frequencies are collected in k-space, which is in frequency domain. This holds the raw data of a

MR session. The 3D volume can then be reconstructed by an inverse Fourier transform.

Protons in different tissues have different relaxation rates and will therefore get different intensities in the reconstructed image. By varying different parameters that have different effects on the relaxation time of different tissue types, the intensity distribution will differ. The two most common intensity distributions are T_1 weighted and T_2 weighted images. The MR scans for this report are T_1 weighted. In T_1 weighted images fat appears brighter than water, which has the result that white matter is brighter than gray matter. In T_2 weighted images, the opposite is true.

The process of MR imaging is fast and, therefore, several volumes can be recovered for noise reduction. It is also harmless for the patient. (Hanson, 2009). Structural MRI has a good contrast between different important structures of the brain, such as between white matter and gray matter. As it images structural information of the brain rather than functional information it is superior to PET when it comes to segmenting the cerebral cortex and other regions. Futhermore, it has a superior resolution and significantly lower noise level than PET.

3. Related works

Many approaches have been developed to improve the low signal-to-noise ratio in PET data. Basic methods focuses on the processing step of spatially smoothing the time frames after the reconstruction procedure and previous to the kinetic analysis of the time frames. One of the problems with filtering by convolution with a Gaussian kernel, i.e. Gaussian filtering, is the introduction of edge artifacts between regions of high signal and regions of low signal.

General edge preserving low-pass filters exist and some have been evaluated for use with PET data, like the bilateral filter which is a locally adaptive smoothing filter with edge-preserving and noise-reducing qualities (Hofheinz1 et al., 2011). Wavelet denoising has been proposed by Lin et al. (2001), which uses wavelet transform that transforms the data into a multiscale representation where noise ideally can be more easily removed and the issue with peak shifts present in Gaussian filtering can be reduced. Another such method by Turkheimer et al. (2008) uses structural information from MRI together with PET data together with wavelet transform. A spatial filtering procedure can also be incorporated in the reconstruction procedure, such as anisotropic diffusion for use of anatomical priors (Chan et al., 2009). With the use of a surface representation of the cortical layer obtained from structural MRI, as in this thesis, the problem with edge artifacts is not as critical, because the smoothing is preformed only across the cortical surface where neighboring regions will be similar in signal strength and the properties of regions can be assumed to vary rather smoothly in comparison to the boundaries between different tissue types. The signal will however still be affected by noise.

An issue with these methods is that they concentrate on the spatial resolution at each time frame without regarding the other time frames simultaneously. If the difference in activity between two regions and the noise level are of the same order of magnitude, such a method could risk to introduce an unwanted bias. Some methods have been developed to regularize both across time and space. One such method for post-reconstruction filtering is based on spatio-temporal anisotropic diffusion without the use of a certain kinetic model (Tauber et al., 2011).

A focus of this thesis is the regularization of the model MRTM2 in the kinetic analysis by using a Bayesian framework. Earlier methods using MRTM or similar models without the second step of MRTM2 includes regularizing on pre-computed estimates from smoothed data (Zhou et al., 2002) and regularization by truncated singular value decomposition, boundary constraints and Tikhonov regularization (Buchert et al., 2003). A Bayesian approach to kinetic analysis using wavelets has been developed by Turkheimer et al. (2006) and a data-driven approach with the use of Sparse Bayesian Learning has been developed by Peng et al. (2008). Some methods, such as Global-Two-Stage filtering, use a population-based prior for the kinetic analysis that assumes that voxels within regions have similar charasteris-

tics (Tomasi et al., 2011). Another method using this a-piori assumption working on multiple scales has been developed by Rizzo et al. (2012). The regions are segmented from structural MRI data. However, the requirement of a a-priori knowledge of the functionality of such regions is an issue with these methods, especially within the cortical layer. Another approach uses prior voxel-by-voxel information from other brains (Fang et al., 2012), which can be problematic because the registration between brains is rarely perfect. Some recent papers have proposed direct parameter estimations on the raw PET data, i.e. combining the reconstruction process and the kinetic analysis (Wang and Qi, 2012). The idea of incorporating the whole data analysis process into one model is tempting, but the actual performance of these models have not yet been compared to other more complicated models. All these models have been validated on volume data.

The conventional approach to the statistical comparison of parameters between brains is to independently fit a general linear model at each vertex or voxel. This is the approach used in FreeSurfer and that will be evaluated in this thesis. The amount of independent tests considered at once will be high, which inflates the risk of finding significant areas where there actually are none. This is often handled by techniques to correct for multiple comparisons which threshold the already estimated p-values, often with the result of a lack of power. However, other models have been developed within the Bayesian framework that model the between-brain data of all vertices or voxels simultaneously, with a regularization on sparsity and coherence. Two such model are the MVB scheme(Friston et al., 2008) and the relevance voxel machine(Sabuncu and Van Leemput, 2013). The last model was validated by using FreeSurfer's cortical surface representation and between-brain mapping which yielded reasonably good results.

4. Materials and methods

In this chapter the specific models developed and tools used and assessed in this thesis will be described. In section 4.1 the mathematical framework leading up to the Multilinear Reference Tissue Model (MRTM2) will be described, including the one-tissue compartment model, model linearization and the use of a reference tissue. In section 4.2, the study used in this thesis will be described and some initial results from the study will be presented. In section 4.3, the software package FreeSurfer will be described together with its tools and specific capabilities. In section 4.4, the pipeline for a surface-based MRTM2 using FreeSurfer will be described. In section 4.5, the incorporation of MRTM2 into a Bayesian framework will be explained which includes the development of the regularized MRTM2 and further attempts to move MRTM2 into a fully Bayesian framework. Lastly, in section 4.6, some statistical tools for evaluating the repeatability of a model will be described.

4.1. Kinetic analysis with a one-tissue compartment model

The goal of a PET data analysis is to measure the density of a certain substance within a tissue of interest. In the case of this thesis, the tissue of interest is the neural tissue of the brain and we want to measure the concentration of serotonin transporters. This can be measured for a smaller tissue region as a voxel or a larger region as the frontal cortex or the whole cerebral cortex. The measured quantity is called binding potential (BP) and is the ratio between the available transporters and a dissociation constant. The dissociation constant describes how tightly the injected radioactive tracer binds to the serotonin transporters. More than one scan session has to be conducted to directly measure the concentration of available transporters, but this is usually not necessary as the dissociation constant often can be considered the same across subjects. BP can be estimated from a PET time activity curve by setting up a mathematical model that describes the kinetic process of the tracer's path from the blood to the neural tissue and back to the blood.

The circulatory system of the human body continuously supply the brain with oxygen and other vital substances with blood as the transporting liquid. A substance enters and leaves the brain tissue through a semipermeable blood-brain barrier. The process of blood flow can be modeled by compartment models, where one compartment corresponds to a physiologically separate section of the injected tracer. The tracer concentration within each compartment is assumed to be homogeneous with rapid transfer between compartments. Neuronal activity, such as serotonin transmission, has been shown to be closely coupled with the blood flow. Therefore, a

one-tissue compartment model is usually sufficient. This has been shown to be the case for the tracer used in this thesis, DASB (Ginovart et al., 2001). Other functional activities might require more compartments, e.g. metabolism is modeled by a two-tissue compartment model because the radiotracer stays a while in mitochondria to be metabolised (Maguire, 2007).

The one-tissue compartment model is based on the Fick principle. This principle says that the change in concentration of a substance in brain tissue is given by the difference of substance concentration in arterial blood and in venous blood multiplied by the rate that the substance is able to move through the blood-brain barrier. If this rate is assumed to be constant through the whole PET procedure, The Fick princple can be expressed as

$$\frac{dC_t}{dt} = K_1 C_a - K_1 C_v, \tag{4.1}$$

where K_1 is the rate constant for transfer from blood to tissue (ml g^{-1} min^{-1}), C_t is the radioactivity concentration in tissue, C_a is the radioactivity concentration in arterial blood and C_v is the radioactivity concentration in venous blood (Bq/ml). The decay corrected radioactive counts from PET data is assumed to be equivalent to the substance concentration in the tissue. The ratio between C_t and C_v is called a distribution volume V_d(ml ml). Inserting this ratio in equation 4.1 we get the differential equation

$$\frac{dC_t}{dt} = K_1 C_a - \frac{K_1}{V_d} C_t. \tag{4.2}$$

The ratio K_1/V_d is treated as one parameter called the clearance rate constant k_2. Assuming that 1 ml of tissue corresponds to 1 g, k_2 has the unit min^{-1}. This model is shown in figure 4.1.1, together with the (Maguire, 2007).

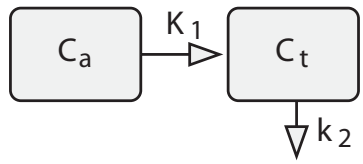

Figure 4.1.1. One-tissue compartment model.

When estimating the parameters of the model, the process is assumed to be in a constant state, without a change in transfer or clearance rate.

4.1.1. Linearization of the one-tissue compartment model

To ease computation of the parameter estimates, compartment models are often linearized so they can be solved by linear least-squares fitting. Least squares is much faster to solve than a direct kinetic analysis of a compartment model. The

linearization is done by simply integrating the differential equation (Van den Hoff, 2007). Integrating the one-tissue compartment model in equation 4.2 between zero and a variable time yields

$$C_t(T) = K_1 \int_0^T C_a(t)dt - k_2 \int_0^T C_t(t)dt. \tag{4.3}$$

This relationship will be linear at every time point starting from T=0 for a tracer with one-tissue compartment kinetics. The parameters of the model can then be estimated by setting up a system of equation, with one equation for each time frame of a TAC, and solving it with a multilinear least squares method. However, this method assumes that the predictors are noiseless and is thus bad at handling noisy predictors. $\int_0^T C_t(t)dt$ will contain high levels of noise for smaller regions and at voxel-level. Neither $\int_0^T C_a(t)dt$ will be noise free.

4.1.2. Modeling with a reference tissue

To continuously draw blood samples during a PET scan is a difficult and invasive procedure. A common way to avoid this is to instead use a reference tissue. For SERT analysis, the reference tissue should ideally have no serotonin transporter, so the radioactive signal only comes from nondisplaceable tracer, i.e. free and non-specifically bound tracer. Furthermore, the reference tissue should ideally have similar free and non-specific properties and similar density to the tissue of interest. A reference tissue model can then compare the distribution volumes of the reference region and the region of interest, resulting in the binding potential on the form

$$BP_{nd} = \frac{V_t - V_{nd}}{V_{nd}} = \frac{V_t}{V_{nd}} - 1, \tag{4.4}$$

where V_t is the total distribution volume of the tissue and V_{nd} is the distribution volume of nondisplaceable tracer. The amount of nondisplaceable tracer is assumed to be the same in the reference region and the region of interest. With the reference tissue only containing nondisplaceable tracer, its total distribution volume will be equal to the nondisplaceable distribution volume. The reference tissue model is illustrated in figure 4.1.2 with a prime sign denoting reference tissue (Lammertsma, 2007).

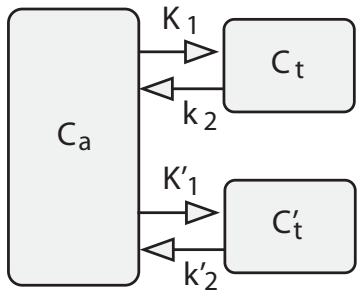

Figure 4.1.2. Reference tissue model, reference tissue with prime sign.

As the distribution volume $V_t = K_1/k_2$ the binding potential will be

$$BP_{nd} = \frac{K_1 k_2'}{K_1' k_2} - 1. \tag{4.5}$$

One method that uses reference tissue together with the one-tissue compartment model is the non-linear simplified reference tissue model SRTM. There are no regions in the brain that completely lack serotonin transporters, but white matter in both cerebellum and cerebrum and cerebellar cortex have very low concentration of transporters. Cerebellar cortex is usually chosen as it has similar properties to e.g. the cerebral cortex (Kish et al., 2005).

An example of the TACs included in a reference tissue model at a voxel or vertex level is shown in figure 4.1.3.

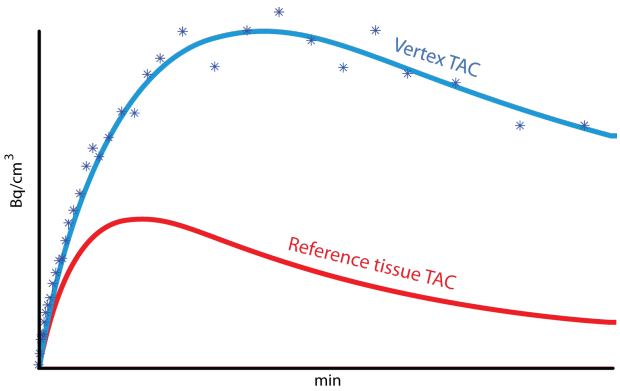

Figure 4.1.3. Reference tissue TAC.

4.1.3. The multilinear reference tissue method MRTM2

The multilinear reference tissue method combines the use of a reference tissue with linearization. Mathematically, the linearization of a reference tissue model can be obtained by solving equation 4.3 for $\int_0^T C_a(t)$ for the target tissue, which yields

$$\int_0^T C_a(t)dt = \frac{k_2}{K_1}\int_0^T C_t(t)dt + \frac{1}{K_1}C_t(T), \tag{4.6}$$

and for the reference tissue, which yields

$$\int_0^T C_a(t)dt = \frac{k_2'}{K_1'}\int_0^T C_t'(t)dt + \frac{1}{K_1'}C_t'(T). \tag{4.7}$$

The blood concentration term C_a can then easily be eliminated by combining equations 4.6 and 4.7. The resulting equation is then rearranged to make C_t the dependent variable, which yields

$$C(T) = \frac{K_1}{K_1'} k_2' \int_0^T C'(t)dt - k_2 \int_0^T C(t)dt + \frac{K_1}{K_1'} C'(T). \tag{4.8}$$

By dropping the subscript of the tissue concentrations in equation 4.8 and defining the relative radioligand delivery R_1 as K_1/K_1' the following equation is obtained:

$$C(T) = R_1 k_2' \int_0^T C'(t)dt - k_2 \int_0^T C(t)dt + R_1 C'(T). \tag{4.9}$$

This is the multilinear reference tissue model (MRTM), which can be solved as a three-parameter multilinear least-squares problem for the whole time activity curve at once for a voxel's or an averaged ROI time activity curve. The ordinary least squares assumes that the independent variables of the model are noise free. The time activity concentration curve $C(T)$ at each voxel will be a noisy measurement and the integral of it is present as an independent variable. However, the integral of a noisy time curve will at least be less noisy than the curve itself.

Another problem with this model is that the reference tissue's clearance rate k_2' will be estimated as a different value at each voxel. The reference tissue's time curve C' should be the same at each voxel and therefore should its clearance rate also be constant. To deal with this problem MRTM is only used for a preliminary analysis of the data with the only output being a fixed value of k_2'. In this analysis only high binding regions are used as target tissue. The equation is then rearranged to the two-parameter multilinear reference tissue model MRTM2 on the form

$$C(T) = R_1 k_2' \left(\int_0^T C'(t)dt + \frac{1}{k_2'} C'(T) \right) - k_2 \int_0^T C(t)dt. \tag{4.10}$$

The binding potential can then be obtained by the relationship

$$BP_{nd} = R_1 k_2'/k_2 + 1. \tag{4.11}$$

For improved readability, BP_{nd} will be referred to as simply BP throughout this work. The whole multilinear model can be formulated in matrix form for a voxel n with K time steps as

$$\mathbf{c}_n = \mathbf{X}_n \mathbf{w}_n + \boldsymbol{\epsilon}_n, \tag{4.12}$$

where

$$\mathbf{c}_n = \begin{bmatrix} C_n(T_1) \\ \vdots \\ C_n(T_K) \end{bmatrix},$$

$$\mathbf{X}_n = [\mathbf{x}_{1,n}, \mathbf{x}_{2,n}] = \begin{bmatrix} \int_0^{T_1} C'(t)\mathrm{d}t + \frac{1}{k_2'} C'(T_1) & -\int_0^{T_1} C_n(t)\mathrm{d}t \\ \vdots & \vdots \\ \int_0^{T_K} C'(t)\mathrm{d}t + \frac{1}{k_2'} C'(T_K) & -\int_0^{T_K} C_n(t)\mathrm{d}t \end{bmatrix},$$

$$\mathbf{w}_n = \begin{bmatrix} R_{1,n} k_2' \\ k_{2,n} \end{bmatrix},$$

and $\boldsymbol{\epsilon}_n$ contains noise. The unknown parameters in $\mathbf{w}_n$ have to be estimated with $w_{1,n} \geq 0$ and $w_{2,n} \geq 0$.

MRTM2 has been shown to be a useful method for estimating BP in larger regions with a good control of the noise levels of PET data with [^{11}C]DASB. The model does however still contain noisy predictors. In the paper of Ichise et al. (2003) that introduced the method, it was shown that the two-step procedure effectively reduces the bias and variability of regional estimates in a simulation study and is stable for high-binding areas and mid-binding areas, such as cerebral cortex, in a study with data from the ECAT 47 scanner with a spatial resolution full width half maximum of 9.3 mm. It did however have problems in regions with lower binding, such as white matter. This is perhaps not surprising, as cerebral white matter has been shown to have almost as low binding as cerebellum (Kish et al., 2005). In the simulation study it was shown that the accuracy is comparable to that of the non-linear method SRTM with the same two-step procedure and direct kinetic analysis on the one-tissue compartment model with radioactivity data from blood. At the same time, it is more computationally efficient to compute.

In a test-retest study by Kim et al. (2006) the method has been shown to have high reliability in most regions when comparing the regional BP estimate from an average ROI TAC or when averaging voxel level BP estimates over a region. The reliability measure was ICC(3,1) and the scanner used was GE Advance with a reconstructed voxel size of 6 mm. A slight negative bias was present at all reported regions, perhaps due to the short time between test and retest scan of 1 hour. The effect of noise at individual voxels were however not reported in detail, but seems to have been stable.

With a higher scan resolution, such as the 1.2 mm resolution of the scans used in this thesis, MRTM2 is unstable at voxel level. This is conventionally handled by reducing the spatial resolution of the data by spatially smoothing each volume with a Gaussian filter previous to MRTM2 estimations.

4.2. Data set - A sex-hormone fluctuation study

The examined data set of 61 healthy female volunteers is from an ongoing project at NRU, headed by Vibe G. Frøkjær (MD,PhD). The project examines the neuropsychobiological effects of pharmacologically introduced sex-hormone fluctuations. It has previously been shown in epidemiological studies that there is an increased vulnerability to neuropsychiatric disorders directly after a pregnancy and in the menopausal transition period (Munk-Olsen et al., 2006; Freeman et al., 2006). Both these events are characterized by sex-hormone fluctuations in terms of a rapid

decline in sex-hormone production from late pregnancy to post partum and of a period of increased fluctuation in ovarian sex-hormone production that ultimately ceases when menopause is reached. Because sex hormone levels has been shown to be linked to the serotonergic transmitter system (Frøkjær et al., 2010) and the serotonin transporter function as a key regulater of the serotonergic signaling is linked to depression and many other neuropsychiatric disorders, the project includes imaging of the serotonin transporter. The hypothesis of the study was that sex hormone fluctuation provokes depressive symptoms and that the emergence of depressive symptoms would be coupled to a change in this marker of serotonergic signaling. Thus one of the goals is to address if one of the pathways by which sex-hormone fluctuations provokes mood changes could be over serotonin signaling. The study was approved for NRU by the Regional Ethics Committee in Copenhagen and data acquisition was finalized by the end of 2012.

Two PET scans were acquired for each subject: a baseline and a follow-up scan. Between the two scans, 31 subjects were given a GnRH-agonist treatment to obtain a transient stimulation and a subsequent downregulation of ovarian sex-hormone production, while the remaining 30 were given placebo injections of saline. The study design was doubleblinded. Furthermore, Hamilton rating scale for depression was used to characterize the change in the subjects' depressive symptoms from baseline. The Hamilton score is determined by a semistructured interview where the interviewing doctor follows a questionnaire to score different symptoms, such as anxiety, feelings of guilt, insomnia and agitation. The scores are then weighted together to give an overall score; a lower score means less depressive symptoms and a higher score more depressive symptoms. A score of 0-7 is considered normal and a score of 20 or higher indicates moderate to severe depression. One Hamilton score was assessed at baseline and one at follow-up.

The follow-up scan was acquired as close as possible to the same time in the menstrual cycle as the baseline scan for the placebo group. The baseline scan was acquired at cycle day 5-8, GnRHa intervention was started on cycle day 22 and the follow-up scan was acquired 14-19 days after intervention. For most subjects the scans were acquired one cycle apart, i.e. roughly one month apart ($N = 50, 32.6 \pm 3.4$ days), while some were 2 cycles apart ($N = 8, 64.5 \pm 7.5$ days), 3 cycles apart (92 days) or 4 cycles apart (133 and 122 days). The follow-up scan in the placebo group was on average acquired slightly later in the cycle than the baseline scan. The subjects were screened for variables that could alter the serotonin transmitter system, such as past or present psychiatric or neurological disorders and substance abuse. As none of these variables were present, they were regarded as healthy. Further demographics is shown in table 4.2.1. The age, body mass index and Hamilton score at baseline are fairly similar between the groups. Additionally, no difference between groups in baseline sex hormones was observed.

Table 4.2.1. Demographics of the 61 subjects enrolled in the study (mean $\pm$ SD).

	Active group, $N = 31$	Placebo group, $N = 30$
Age (years)	23.3 ± 3.3	25.2 ± 5.9
Body mass index (kg/m^2)	23.2 ± 2.3	23.4 ± 3.9
Hamilton score baseline	1.2 ± 1.5	1.6 ± 2.2

The PET scans and Hamilton scores at baseline and follow-up will be examined in this thesis to compare the results of surface-based analysis with FreeSurfer to the initial results of the study obtained on a regional level of analysis. This will serve as a basis to discuss the capabilities, advantages and issues with the tools of FreeSurfer and an overall surface-based approach. Moreover, 24 subjects from the placebo group will be used in a pseudo test-retest analysis to conclude on the repeatability of the different modeling approaches. The 24 selected subjects are the placebo subjects whom were scanned with one month in between baseline and follow-up (32.9 ± 3.1 days).

4.2.1. Image acquisition

The list-mode data from the HRRT PET scanner was reconstructed to 38 dynamic time frames by the iterative ordinary Poisson ordered-subset expectation maximization (OP-OSEM) 3D method with resolution modeling (Hong et al., 2007; Comtat et al., 2008; Sureau et al., 2008). The 38 frames had the sequence 6×5 s, 10×15 s, 4×30 s, 5×2 min, 5×5 min and 8×10 min. The total acquisition time was 90 min. The mean intravenously injected dose of [^{11}C]DASB was 577 ± 43 MBq for the active group and 591 ± 11 MBq for the placebo group. The reconstructed voxel size was $1.2188 \times 1.2188 \times 1.2188$ mm. The radioactive decay was accounted for and the dynamic frames were motion corrected. The motion between frames was estimated by the automatic image registration routines from Woods et al. (1998).

Each subject was also scanned by a 3T Verio Siemens Medical Inc scanner to acquire a structural T1-weighted magnetic resonance scan, which was corrected for field inhomogenities bias field intensities previous to the analysis in this report.

4.2.2. Radiotracer [^{11}C]DASB

The radioactive tracer [^{11}C]DASB is used in this study to image the serotonin transporter binding by PET. The tracer was introduced by Houle et al. (2000) and is regarded as one of the superior tracers for this purpose. [^{11}C]DASB binds with high selectivity to the serotonin transporter and has been shown to correlate well with regional post-mortem receptor densities and distribution pattern. Furthermore, it has been shown to give reliable results when used with MRTM2 (Ichise et al., 2003).

4.2.3. Initial results for the sex-hormone fluctuation study

The study was successful in inducing an initial rise followed by a downregulation of sex-hormones by the time of the follow-up scan in the treated group. It did also provoke a rise in Hamilton score on average in this group, as seen in figure 4.2.1.

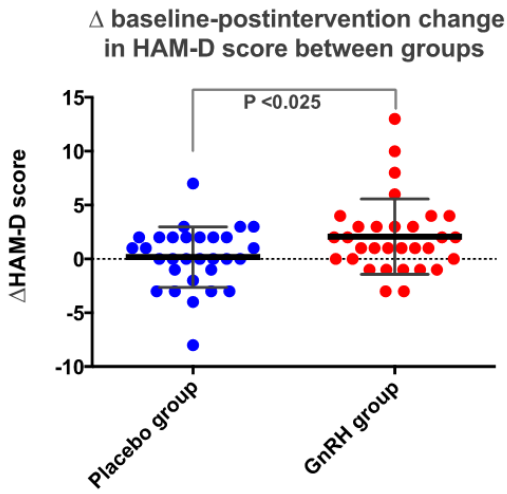

Figure 4.2.1. Difference in Hamilton score from baseline to follow-up in the placebo and active group.

The PET data was automatically segmented into cortical and subcortical regions and averaged time activity curves were obtained for these regions. This automatic segmentation is based on the segmentation by Svarer et al. (2005). The segmentation method used was the one in the software Statistical Parametric Mapping (SPM5) by the Wellcome Trust Centre for Neuroimaging, London, which uses structural information from MRI data. Regional BP estimates were by modeling the regional averaged TACs by MRTM2 in PMOD version 3.0 by PMOD Technologies Ltd.

The data was analyzed with the general linear model

$$\Delta\text{Hamilton} = \beta_0 + \beta_1 \text{ group} + \beta_2 \ \Delta\text{BP} + \beta_3 \text{ group } \Delta\text{BP}, \tag{4.13}$$

where group is a categorical variable for group belonging (treated or placebo), $Delta$Hamilton is the change in Hamilton score between baseline and follow-up and ΔBP is the change in binding potential between baseline and follow-up. This model is set up to measure how significant the difference in slope is between the treated group and the placebo group, i.e. the interaction between group and ΔBP. The global result on BP from the total neocortical region is shown in figure 4.2.2.

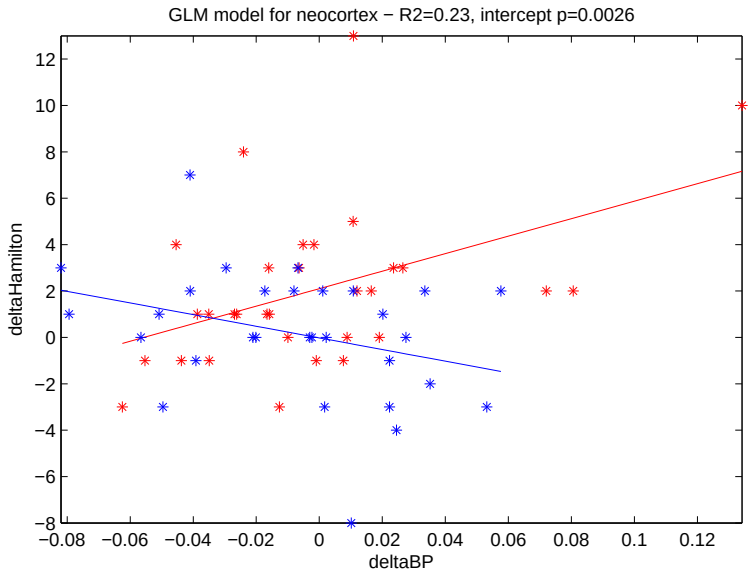

Figure 4.2.2. Change in Hamilton score plotted against change in global neocortical BP. Red shows active group and blue shows placebo group. The lines corresponds to the fit of the general linear model.

The significance of the difference in slope within neocortex is consistent with the hypothesis that an acute change in sex-hormonal levels is coupled with the serotonergic signaling system in the brain. However, two remarks could be made: the slope of the placebo group is slightly negatively correlated to the change in BP although not significantly, and there is one subject who drives the slope of the treated group. The negative slope of the placebo group might have to do with the time of acquisition of the follow-up scan, which was on average slightly later in the menstrual cycle than the baseline scan. That one subject who is driving the active slope does make the conclusions that can be drawn from the study less certain. Nevertheless, it is consistent with previous studies where there are only a limited number of women in a population that develop neuropsychiatric disorders after pregnancy or in the menopausal transition period. The development of the other active subjects do follow the same trend as the driving subject, but to a lower degree. With a higher number of subjects within the study, the statistical power would of course have been greater. It is however rarely plausible in a neurobiological study, due to the cost and complexity of conducting PET studies and in the case of this study the ethics in pharmacologically challenging healthy women. 61 subjects is a rather high number for these kind of studies. The significance in some other regions is shown in figure 4.2.2.

Table 4.2.2. Significance of difference in slope for some brain regions.

Region	p-value
Neocortex	0.003
Prefrontal	0.03
Pallidostriatum	0.21
Midbrain	0.12

The difference in slope between the active group and the placebo group is significant in the whole neocortex and prefrontal cortex, but not in subcortical regions. Thus, the effect seems to be a cortical phenomenon which motivates a further exploratory analysis of the cortical layer with a surface-based analysis.

4.3. Parametric imaging with FreeSurfer

FreeSurfer is an open-source software package for automatic analysis of the human brain developed by the Athinoula A. Martinos Center for Biomedical Imaging at Massachusetts General Hospital. It has been developed especially for taking the highly-folded structure of cerebral cortex into account in statistical analysis of structural or functional measurements. FreeSurfer version 5.1 for Linux-based operating systems was used in this thesis.

FreeSurfer automatically segments the borders of the cortex, the inner boundary to white matter and the outer boundary to cerebrospinal fluid, and constructs surfaces at these borders so statistics can be computed in-between these two surfaces. The segmentation is done on a T1-weighted MR image. Non-brain tissue is first removed in the MR image using a hybrid watershed/surface deformation procedure (Segonne et al., 2004) and the image's intensity is normalized (Sled et al., 1998). The boundary between white matter and cortical gray matter is segmented and a surface tessellation of this segmentation is performed for each hemisphere (Dale et al., 1999). The surface of a hemisphere is inflated by a procedure that is driven by the convexity or concavity of the surface. A spring term smooths the surface while a second term preserves a desirable amount of the original metric properties (Fischl et al., 1999a). Topological defects will be identified by iteratively mapping the inflated surface to a sphere, where the defects will make the mapping non-continuous, and will be corrected by locally retessellating the original surface (Fischl et al., 2001; Segonne et al., 2007). After topological correction, a final inflated surface and a final spherical surface are produced.

The original surface is then deformed following intensity gradients to place the outer and inner border at the location where the greatest shift in intensity constitutes the optimal change in tissue type (Dale et al., 1999; Fischl and Dale, 2000). This produces the outer and inner surface, that are shown together with the inflated surface in figure 4.3.1. As seen in this figure, the inflated surface can be used for a better visualization of measurements in deep sulcus structures. The measurement in the figure in red to green is the measurement of convexity acquired from the inflation procedure.

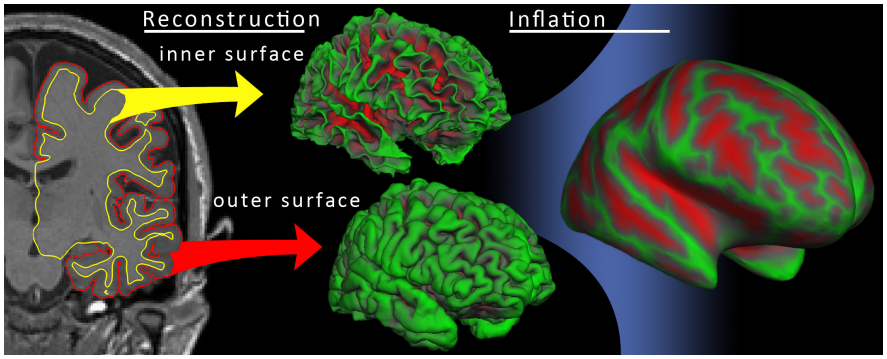

Figure 4.3.1. Important outputs from the cortical surface reconstruction procedure: The inner and outer surfaces and the inflated surface.

As the inner, outer, inflated and spherical surfaces all originates from one original surface, they will have vertex correspondence. The vertex correspondence of a small part of an inner and outer surface is shown in figure 4.3.2, in a simplified 2D setting. For functional measurements, where the functional image has been registered to the MR image, the value for a vertex is chosen as the value of the voxel coinciding with the middle of the inner-outer vertex pair, i.e. the middle of the blue lines in the figure. A vertex value could also be averaged from several voxels coinciding with the vertex-pair line. However, regarding segmentation issues and varying layer thickness the middle voxel is a good choice.

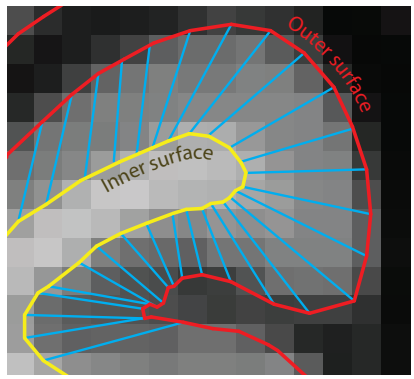

Figure 4.3.2. Vertex correspondence between inner and outer surface. The functional data for a vertex is taken from the voxel at the middle of this line.

To perform a vertex- or voxelwise analysis between brains, the individual brains have to be registered to a common space. This is a difficult procedure due to the high structural variability between brains. As mentioned before, the highly folded cortical layer is a particular challenge. This is one of the main issue to motivate the use of FreeSurfer, where the surface-based representation makes it possible to use detailed structural information in a mapping procedure. In FreeSurfer, the mapping is done in spherical space, as seen in figure 4.3.3. An average subject has

already been pre-computed in FreeSurfer based on several brains. Each subject of the analysis can thus be directly registered to this average subject. The registration procedure uses a measurement of convexity and is performed across multiple scales. The measurement of convexity, seen as red to green in the figure, is obtained in the inflation procedure. Convexity reflects an average property of an area and is therefore less noise-prone than e.g. mean curvature. This also means that it emphasizes the consistent folding patterns over more variable patterns. It has been shown that this method results in a greater accuracy than e.g. non-linear inter-subject mapping in voxel space to an MNI atlas, as both topological structures and sulci patterns of the cortex are specifically taken into account (Fischl et al., 1999b).

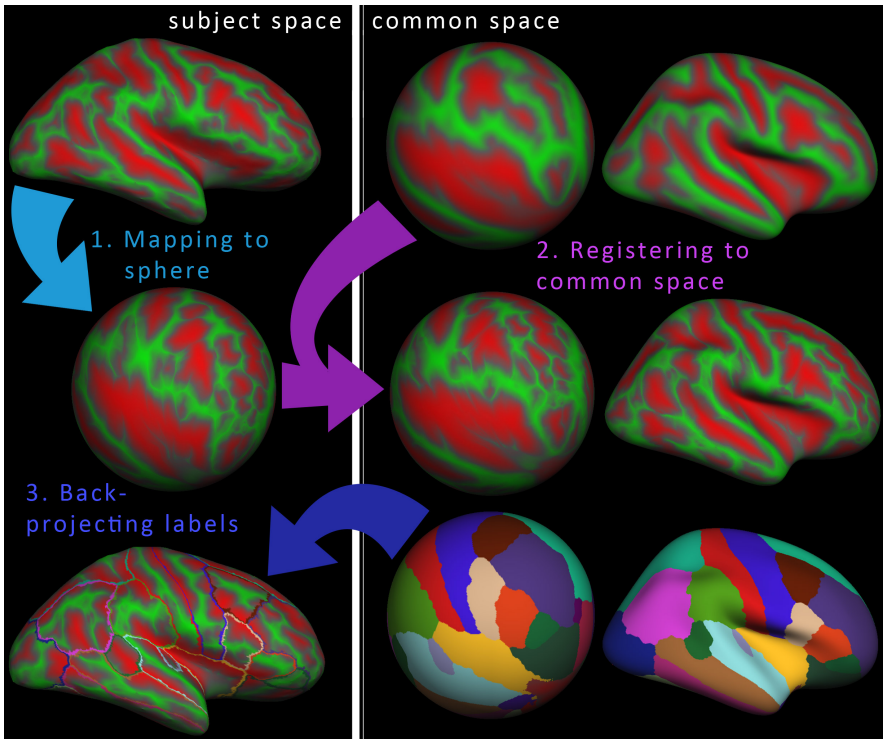

Figure 4.3.3. Spherical non-linear registration to common space.

The average subject has also been labeled with the Desikan-Killany Atlas for standard structural regions, such as superior frontal gyrus. These labels can easily be back-projected to the space of each subject (Fischl et al., 2004).

Another motivation for using FreeSurfer is the issue of spatially smoothing the data. As described in section 2.3.2, a Gaussian filtering of the data in voxel-space will result in a mixture of signal between the cerebral cortex and surrounding tissue with less or no neuronal activity. In addition, a greater level of smoothing might be desired in a between-brain analysis because the registration of brains to a common space is often far from perfect. A filter for this purpose typically has a full width at half maximum (FWHM) of 5 to 10 mm. Because the thickness of the cortical layer ranges between 1 and 5 mm, it is quite clear that smoothing neighboring voxels

with a filter of such an extent will result in severe edge artifacts. In FreeSurfer the smoothing can instead be done in the direction of the surface, illustrated in figure 4.3.4. This results in a mixing of only close cortical, which can be assumed to have similar amounts of activation. The smoothing process in FreeSurfer is an iterative process that yields an approximate FWHM (Hagler et al., 2006).

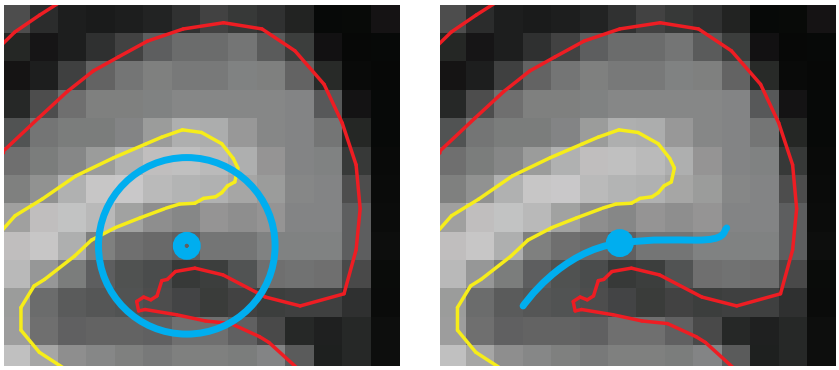

Figure 4.3.4. 2D illustration of the principle difference in smoothing across voxels and across a surface.

4.3.1. Boundary-based registration of multimodal data

An algorithm for multimodel co-registration has been developed for FreeSurfer, called boundary-based registration (BBR), that make use of the reconstructed surfaces of the cortical layer. It is used to rigidly transform a volume with functional data, such as PET, to the structural MR volume. It initializes the registration procedure by performing a registration in SPM or FSL. In the case of this thesis the method of SPM5 is used as initialization. The method of SPM tries to maximize the normalized mutual information to make the joint histogram of the two volumes as compact as possible. This method usually works well when two intra-subject volumes are close from the beginning, but do have a risk of ending up in a local minimum if the noise level is high in local regions, the intensity distribution is very different between the two volumes or some parts are missing in one volume.

In contrast to the method in SPM, BBR treats the reference and input volume differently and can thus make the whole registration procedure more robust. The input volume can be of any modality with some tissue contrast, while the reference volume is represented by the inner and outer cortical surface. The cortex boundaries are used because the cortex is highly folded and a good registration to it is assumed to also result in a good registration of the whole volume. The cost function involves the local intensity differences within inner-outer vertex pairs. The intensity of a vertex is computed from the intensities of the input volume's voxels which for the moment corresponds to the vertex and a small distance inwards towards the center of the brain. The optimization of the cost function is either a minimization or maximization problem, depending on if white matter is lighter or darker than gray matter in the input volume. The optimization problem is solved by a search

function and gradient descent in different steps (Greve and Fischl, 2009). In the case of PET data, a good input volume is the mean volume of the time frames. The early and late time frames are excluded as they are more error prone. An example of a registered mean PET volume together with the inner cortical surface is shown in figure 4.3.5.

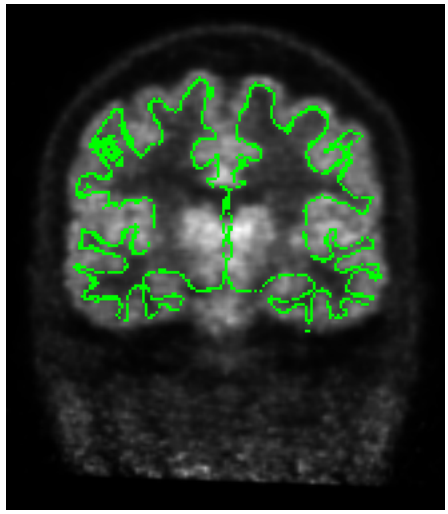

Figure 4.3.5. FreeSurfer registration.

4.3.2. Group analysis with the general linear model

The general linear model(GLM) is the most commonly used model for testing the significance of different hypotheses from a data set. It assumes that the errors follow a Gaussian distribution. In FreeSurfer the testing is done independently at each single vertex, a univariate test, without including informations from other vertices. Information from neighbouring vertices will of course be included in the model when the surface has been smoothed. The GLM is defined as

$$\mathbf{y}_{(s\times v)} = \mathbf{X}_{(s\times p)}\mathbf{B}_{(p\times v)} + \mathbf{U}_{(s\times v)}, \tag{4.14}$$

where $\mathbf{y}$ is the dependent variable matrix, $\mathbf{X}$ is the design matrix containing information on each subject, $\mathbf{B}$ is the parameter matrix and $\mathbf{U}$ contains the residual error. s is the number of subjects, v is the number of vertices and p is the number of regression parameters. The number of parameters depend on the independent variables: for each combination of levels in categorical variables there will be one intercept and one parameter for each continuous variable. E.g. for two categorical variables with two levels and two continuous variables, there would be 12 parameters. The parameters are estimated with

$$\mathbf{B} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \tag{4.15}$$

The residual error is then computed as

$$\mathbf{U} = \mathbf{y} - \mathbf{X}\mathbf{B}. \tag{4.16}$$

To measure the significance of different combinations of parameters, a contrast matrix $\mathbf{C}_{(j\times p)}$ is used to compute a F-value

$$\mathbf{F} = \frac{\mathbf{G}^{\mathrm{T}}\left(\mathbf{C}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{C}^{\mathrm{T}}\right)^{-1}\mathbf{G}}{j\mathrm{var}(\mathbf{U})}, \tag{4.17}$$

which is used in an F-test to get a p-value. A low p-value means a high significance of the parameter combination. If the contrast matrix only contains one row ($j = 1$), then $F = t^2$ where t is used instead in a two-tailed t-test (Smith, 2013).

As described in section 4.2.3, the initial results of the study that will be examined in this thesis showed one subject that has a high influence on the result of a GLM on the global neocortical region. It can therefore be interesting to see what influences it has on a vertex level. One method to measure the influence of one data point on a GLM analysis is Cook's distance (Cook, 1977). It measures the effect on an estimated parameter when deleting a data point by combining the studentized residual distance and leverage into one measurement. Cook's distance can be defined as

$$D_i = \frac{r_i^2 h_i}{p(1-h_i)^2\hat{\sigma}^2} \tag{4.18}$$

where r_i is the residual of data point i, h_i is the leverage of data point i, p is the number of parameters of the model and $\hat{\sigma}^2$ is the estimated variance of the fit. The leverage is the i:th diagonal elements of the hat matrix $\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$. A Cook's distance greater than 1 is a good indication of a influential data point that can be worth looking into. Another proposed value is a Cook's distance greater than $4/n$ where n is the number of subject, which would equal 0.07 for the sex-hormone study.

4.3.3. Correction for multiple comparisons

In a statistical analysis, two types of errors are important: type I and type II errors. A type I error is when a test shows a statistical significance where there is no true significance: a false positive. A type II error is the opposite: a false negative. An often used threshold for p-values is 0.05, which corresponds to a 95% confidence interval. This means that there is a 95% chance that the tested condition is significant. In brain imaging it is common to test the statistical significance independently at each vertex. To use a 0.05 threshold for so many simultaneous tests might result in a high number of false positives, which inflates the risk of finding significant areas where there actually are none. Therefore, several different approaches have been developed to control the probability of false positives. This is referred to as correcting for multiple comparisons. Early methods, such as Bonferroni correction, focuses only on the probability of false positives with the result of an increased risk of false negatives. It is important that the method does not introduce too many false negatives, as this will reduce the statistical power of the analysis and actual significant areas might be overlooked. Two methods for multiple comparisons that have

been implemented in FreeSurfer will be assessed in this thesis.

One simple and often used approach is to find a p-value threshold with a procedure that controls the false discovery rate (FDR). The false discovery rate is the proportion of false positives among the results over a chosen threshold in a data set, whether the data set is the whole brain or only a region of the brain. This method was introduced to have a better control over the probability of false negatives than Bonferroni correction. First, the vertex p-values are ordered from smallest to largest as

$$p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(V)}. \tag{4.19}$$

The threshold is then chosen to be the p-value with the largest i for which

$$p_{(i)} \leq \frac{i}{V}\frac{q}{c(V)}, \tag{4.20}$$

where V is the total number of vertices and q is a value from 0 to 1 corresponding to the maximum tolerated FDR. It can be set to a conventional value as 0.05, but could also be larger in some cases. $c(V)$ is a constant which is chosen depending on assumptions about the distributions of p-values across vertices. A generally suitable formula is $c(V) = \sum_{i=1}^{V} 1/i$. A problem with this approach is that it becomes more conservative as correlations between the tests increases (Genovese et al., 2002). With the often high amount of smoothing in a between-brains analysis, this method might lack the needed statistical power. Nevertheless, it is an often used method in these settings as it is easy to implement and understand.

A more advanced approach is the clusterwise correction by Monte Carlo simulation. This method assumes that significant areas in the brain will be clustered and focuses on finding sufficiently large and significant clusters in a p-value map. The output will be a number of clusters with clusterwise p-values. It starts by generating random maps of the same size as the p-value map and with independent normally distributed noise. The residual smoothness level of the p-map is estimated and the simulated maps are smoothed to this level. At each simulation the simulated map is thresholded to a user-defined p-value threshold and the maximum cluster size is recorded. After many map generation, a distribution of cluster sizes is obtained. The original p-map is then thresholded at the same p-value and the size of the obtained clusters is compared to the distribution, which results in a clusterwise p-value that is regarded as corrected for multiple comparisons as the distribution will show how often a cluster of that size can appear by chance. To get a reliable estimation from the Monte Carlo simulation, several thousand generated maps have to be produced which can take several hours. In FreeSurfer this process have been pre-computed for different levels of smoothing and can be transformed to a new case by the use of z-statistics, so for every new case it is just a matter of looking up the cluster-sizes and their corresponding p-values in a table (Hagler et al., 2006).

4.4. Surface-based MRTM2 pipeline by using FreeSurfer

To get more stable binding potential estimates when using MRTM2 for parametric imaging, the data can be spatially smoothed with a Gaussian filter previous to the modeling of the data. As described in section 4.3 the spatial smoothing in voxel space results in partial volume effects by mixing in signal or lack of signal from white matter and cerebrospinal fluid. It also mixes signal in cortical folds between spatially close areas with no direct functional connections. This is the motivation for implementing a surface-based method for MRTM2. The smoothing can thus be done across the surface between vertices that are functionally closer. A pipeline was developed for a full across-brain analysis using the FreeSurfer software together with version 8.1 of the numerical computing environment MATLAB developed by MathWorks. The pipeline was inspired by the yet unpublished work of Douglas N Greve at the Athinoula A. Martinos Center for Biomedical Imaging at Massachusetts General Hospital (Personal communication, spring 2013). The pipeline is shown in figure 4.4.1.

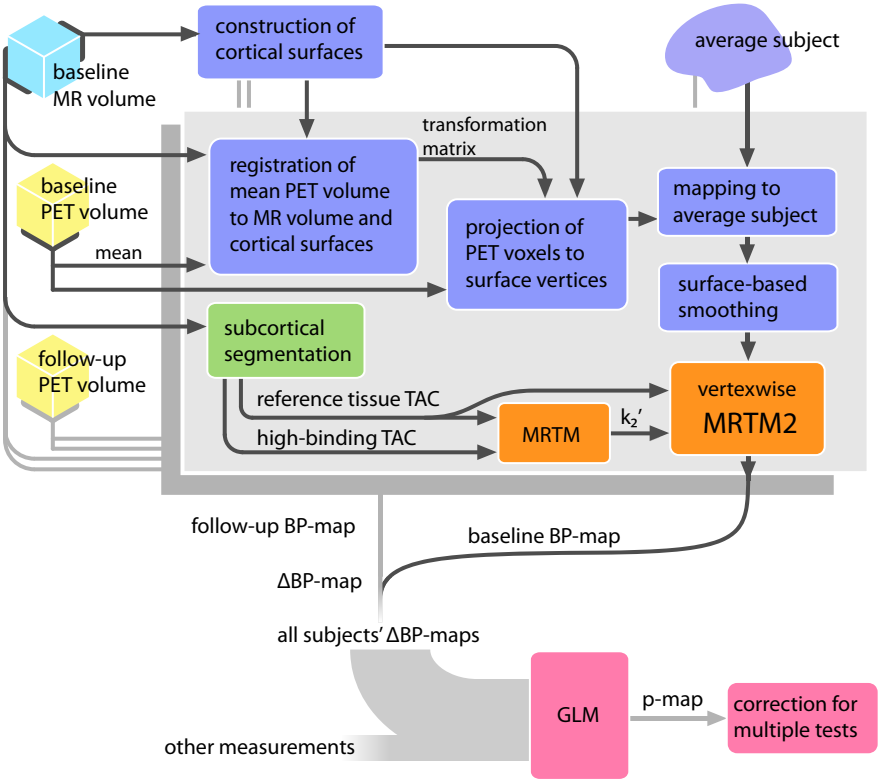

Figure 4.4.1. Pipeline of surface-based MRTM2.

The lilac boxes represents steps performed in FreeSurfer. The inner, outer, spherical and inflated surfaces are obtained from the baseline MR volume for each subject, as described in the previous section. Both the baseline scan and the follow-up

scan are registered to the baseline MR volume and surfaces, to obtain baseline-follow-up vertex correspondence. The mean PET volume of time frames 7 to 32 is used for the registration. Earlier and later time frames are excluded due to their higher level of noise. The registration is done by the boundary-based registration method with initialization by SPM5's co-registration method. The rigid transformation matrix obtained by this registration is then used to transform all the time frames. They are then projected onto the cortical surface and mapped to the average subject where they are smoothed prior to vertexwise modeling by MRTM2. A high level of smoothing is used to account for the imperfections in the mapping to the average subject, typically 10 mm.

To obtain the regional averaged reference tissue TAC and high-binding tissue TAC for the modeling with MRTM, a subcortical segmentation is needed. The subcortical segmentation in FreeSurfer could be used for this purpose. However, it is not as good as the cortical segmentation which is FreeSurfer's specialty. Therefore, the subcortical segmentation that was used for the previous results of the sex hormone study is used also here, described in section 4.2.3. In addition, the surface-based results can then be better compared to previous results. The subcortical regions used can be seen in figure 4.4.2. The reference region in orange is cerebellar cortex, which is used without vermis. The high-binding regions are thalamus in green, caudate in blue and putamen in pink.

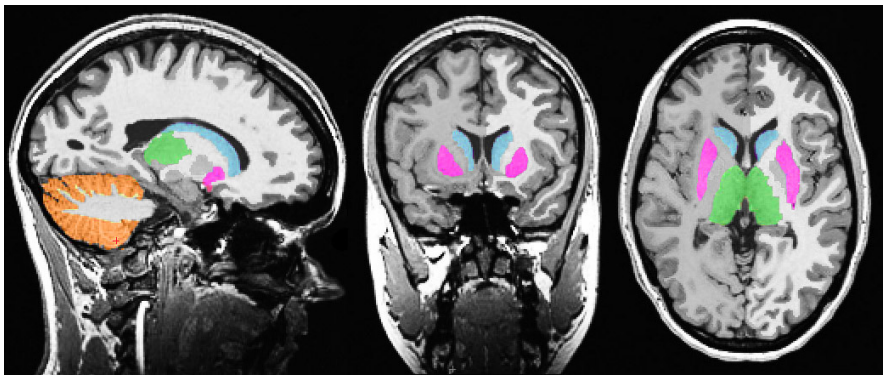

Figure 4.4.2. High-binding regions and reference region.

Subsequently, the estimation of a constant reference tissue clearance rate k_2' is done by MRTM. The input for the vertexwise MRTM2 is the spatially smoothed time frames, the regionally averaged reference tissue TAC and the reference tissue clearance rate k_2'. The output of the MRTM2 is a map of BP values. The baseline BP-map is subtracted from the follow-up BP-map to obtain the ΔBP-map. The ΔBP-maps of all subjects are collected and the between-brain statistical analysis can start.

The general linear model is fitted independently at each vertex and generates a map of p-values. This p-map can then be corrected for multiple comparison.

The MATLAB implementation of MRTM2 can be seen in appendix A.

4.5. Incorporating MRTM2 into a Bayesian framework

When pre-smoothing the data previous to parametric imaging, it is assumed that vertices close to each other will have similar characteristics. This assumption makes it possible to lend information from neighboring vertices to improve the signal to noise ratio at each vertex position. The spatial smoothing is done at each frame, so the assumption is more specifically that the radioactivity at a certain time point should be similar at close vertices. However, the distribution between blood vessels and neuronal tissue might differ between neighboring vertices and thus their contribution to the overall radioactivity. Furthermore and more importantly, with a high noise level such as in PET data the noise might not get effectively reduced by Gaussian filtering. A high noise peak at one vertex will be reduced in height by filtering but it will also spread to neighboring vertices. If the peak is high enough, this might result in a larger area having severe errors in the estimated parameters.

A more reasonable approach might be to instead assume that the parameters themselves should be similar at neighboring vertices. With the parameter estimated by MRTM2, this means that the rate of delivery and the rate of clearance of radioligand should vary smoothly across the tissue. To be able to use this assumption in the modeling procedure, it has to be directly incorporated into the model. A pre-processing step as pre-smoothing is not an option. This is also a more sound approach, to incorporate the assumptions you make on the data in a mathematical model rather than pre-processing the data. Of course, the dynamic frames are already modeled from the raw list-mode data of a PET acquisition procedure. As mentioned in chapter 3 some recent attempts have been made to combine this and estimate the parameters directly on list-mode data. However, many more aspects of the data have to be incorporated in the model for such an approach to work and it can become difficult to handle, especially when interested in the highly-folded cortical layer.

In a Bayesian framework, the assumption that parameters should vary smoothly across the cortical layer can be built into a model as prior knowledge by using Bayes' rule presented in equation 4.22. As presented in section 4.1.3, MRTM2 can be written in matrix form for one vertex n as

$$\mathbf{c}_n = \mathbf{X}_n \mathbf{w}_n + \boldsymbol{\epsilon}_n, \tag{4.21}$$

where the data consists of the response vector $\mathbf{c}_n$, the predictors in matrix $\mathbf{X}_n$ and the parameters in the parameter vector $\mathbf{w}_n$. For a whole PET scan there are N vertices, K timesteps and $M = 2$ parameters at each vertex. The whole problem can be written as

$$\mathcal{X} \in \mathbb{R}^{K \times M \times N} \text{ with slices } \mathbf{X}_1, \ldots, \mathbf{X}_N,$$

$$\mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_N] \in \mathbb{R}^{K \times N}$$

and

$$\mathbf{W} = [\mathbf{w}_{n=1}, \ldots, \mathbf{w}_{n=N}] = [\mathbf{w}_{m=1}, \mathbf{w}_{m=2}]^{\mathrm{T}} \in \mathbb{R}^{M \times N}.$$

Bayes' rule can be written as

$$p(\mathbf{W}|\mathbf{C},\mathcal{X}) = \frac{p(\mathbf{C}|\mathbf{W},\mathcal{X})\,p(\mathbf{W})}{p(\mathbf{C},\mathcal{X})}. \tag{4.22}$$

The denominator will be constant and can thus be removed from the calculations. We then have that the likelihood function $p(\mathbf{C}|\mathbf{W},\mathcal{X})$ times the prior probability $p(\mathbf{W})$ equals the posterior probability $p(\mathbf{W}|\mathbf{C},\mathcal{X})$.

The likelihood is simply the model of MRTM2 in equation 4.21 collected for all vertices, with the assumption of equal and independently distributed noise. The joint probability is then the product of the probability of each event, i.e.

$$p(\mathbf{C}|\mathcal{X},\mathbf{W},\beta) = \prod_{n=1}^{N} \mathcal{N}(\mathbf{c}_n|\mathbf{X}_n\mathbf{w}_n,\ \beta^{-1}\mathbf{I}) \tag{4.23}$$

with the hyperparameter β introduced as the precision term, the inverse of the variance. To find the mean of this distribution is equivalent to performing multilinear least squares at each vertex independently.

4.5.1. Regularized MRTM2

The prior probability should reflect the made assumption that parameters of neighboring vertices should be similar in value. This is the prior assumption that we want to incorporate into the model. An easy way is to start with the squared distance between two parameters of neighboring vertices, i.e.

$$(w_{n=1} - w_{n=2})^2 \tag{4.24}$$

This is the distance that we want to regularize. The parameter differences of all neighboring vertices can be collected in a matrix $\mathbf{\Gamma}$, with as many columns as the number of vertices and as many rows as the number of pairs of neighboring vertices. The row corresponding to one vertex pair will contain zeros everywhere except for the column corresponding to the first vertex which will contain 1 and the column corresponding to the second vertex which will contain -1. The squared distances for all vertices can then be summarized as $\mathbf{w}_m\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}\mathbf{w}_m^{\mathrm{T}}$ for parameter m. A hyperparameter α is introduced to govern the amount of regularization. A prior probability can then be written as

$$p(\mathbf{W}|\alpha) = \prod_{m=1}^{2} \mathcal{N}(\mathbf{w}_m|\mathbf{0},\ (\alpha\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma})^{-1}) \tag{4.25}$$

for the two independently distributed parameters of MRTM2. To make further calculations easier, the likelihood and prior are rewritten for a common parameter vector $\mathbf{w} = [\mathbf{w}_{m=1}^{\mathrm{T}}, \mathbf{w}_{m=2}^{\mathrm{T}}]$ of length NM. Setting

$$\mathbf{G} = \begin{bmatrix} \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma} \end{bmatrix} \tag{4.26}$$

the prior probability becomes

$$p(\mathbf{w}|\alpha) = \mathcal{N}_{NM}(\mathbf{w}|0, (\alpha\mathbf{G})^{-1}), \tag{4.27}$$

To be able to rewrite the likelihood function we vectorize $\mathbf{C}$ as

$$\mathbf{c} = [\mathbf{c}_1^{\mathrm{T}}, \ldots, \mathbf{c}_N^{\mathrm{T}}]^{\mathrm{T}} \tag{4.28}$$

and introduce a matrix equivalent to the tensor $\mathcal{X}$ on the form

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{x}_{1,2} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_{2,1} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{x}_{2,2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{x}_{N,1} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{x}_{N,2} \end{bmatrix}, \tag{4.29}$$

where the first half of the matrix includes the predictor vectors belonging the first parameter and the second half includes the predictor vectors belonging to the second parameter. The likelihood can then be rewritten as

$$p(\mathbf{c}|\mathbf{X}, \mathbf{w}, \beta) = \mathcal{N}_{NK}(\mathbf{c}|\mathbf{X}\mathbf{w},\ \beta^{-1}\mathbf{I}). \tag{4.30}$$

We now assume that the hyperparameters β and λ are known so we can regard them as constants. The posterior probability can then be obtained as

$$p(\mathbf{w}|\mathbf{c}, \mathbf{X}) = \mathcal{N}(\mathbf{w}|\mathbf{m}, \mathbf{S}^{-1}) \tag{4.31}$$

with

$$\mathbf{m} = \beta\mathbf{S}^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{c} \tag{4.32}$$

$$\mathbf{S} = \beta\mathbf{X}^{\mathrm{T}}\mathbf{X} + \alpha\mathbf{G}. \tag{4.33}$$

This is a conditional distribution of $\mathbf{w}$ given $\mathbf{c}$ and can be derived from the marginal distribution of $\mathbf{w}$ as in equation 4.27 and the conditional distribution of $\mathbf{c}$ given $\mathbf{w}$ as in equation 4.30. This has been derived for two general Gaussian distributions by Bishop (2006). The maximum a posteriori probability (MAP) estimate is then on the form

$$\mathbf{w}_{MAP} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda\mathbf{G}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{c}. \tag{4.34}$$

where $\lambda = \alpha/\beta$. This approach can be called the regularized MRTM2. The hyperparameter λ can be set by the user to reach a desired level of regularization, much as setting the desired level of pre-smoothing in the conventional approach. It could be interesting to note that if λ is set to zero, the equation does become equivalent to a least squares fit at each vertex independently.

To directly implement this equation would be computationally demanding. The matrices $\mathbf{X}$ and $\mathbf{\Gamma}$ are sparse, but will still have a dimension of $NK \times 2N$ and the number of vertex pairs $\times N$, respectively. A brain hemisphere contains above 100 0000 vertices with each vertex has on average 6 neighbors. However, some the structural characteristics of the involved products can be observed:

- The matrix $\mathbf{\Gamma}^{\mathrm{T}}\mathbf{\Gamma}$ will have the dimensions $N \times N$. The diagonal of this matrix will contain the total number of neighbors of each vertex. If the first indexed vertex has 5 neighbors, the first value of its column will be 5. The rest of the elements of the column will be zero, except for the rows corresponding to its neighbors which will contain -1.
- The matrix $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ will have the structure

$$\begin{bmatrix} \mathbf{x}_{1,1}^{\mathrm{T}}\mathbf{x}_{1,1} & \dots & \mathbf{0} & \mathbf{x}_{1,1}^{\mathrm{T}}\mathbf{x}_{1,2} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{x}_{N,1}^{\mathrm{T}}\mathbf{x}_{N,1} & \mathbf{0} & \dots & \mathbf{x}_{N,1}^{\mathrm{T}}\mathbf{x}_{N,2} \\ \mathbf{x}_{1,1}^{\mathrm{T}}\mathbf{x}_{1,2} & \dots & \mathbf{0} & \mathbf{x}_{1,2}^{\mathrm{T}}\mathbf{x}_{1,2} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{x}_{N,1}^{\mathrm{T}}\mathbf{x}_{N,2} & \mathbf{0} & \dots & \mathbf{x}_{N,2}^{\mathrm{T}}\mathbf{x}_{N,2} \end{bmatrix}. \tag{4.35}$$

- $\mathbf{X}^{\mathrm{T}}\mathbf{c}$ will have the structure

$$\begin{bmatrix} \mathbf{x}_{1,1}^{\mathrm{T}}\mathbf{c}_1 & \dots & \mathbf{x}_{N,1}^{\mathrm{T}}\mathbf{c}_N, & \mathbf{x}_{1,2}^{\mathrm{T}}\mathbf{c}_1 & \dots & \mathbf{x}_{N,2}^{\mathrm{T}}\mathbf{c}_N \end{bmatrix}^{\mathrm{T}}. \tag{4.36}$$

- The problem that has to be solved will consist of a sparse inverse of dimension $N \times N$ multiplied by a vector with the dimension $N \times 1$. This can be efficiently solved by MATLABs backslash operator and sparse matrix capabilities. The whole problem can be solved in under 2 seconds on NRUs servers with Xeon 5690 3.47GHz CPUs.

The implementation in MATLAB can be seen in appendix B.

4.5.2. A move towards a fully Bayesian approach

For the regularized MRTM2, the hyperparameters are considered known and can be chosen by the user. However, the optimal parameters are of course not known. For a fully Bayesian treatment of the data, the optimal hyperparameters should be estimated from the data. Such a procedure would also relieve the user from making a subjective decision without having much to go on. When treating the hyperparameters as unknown, the full posterior probability $p(\mathbf{w}|\mathbf{c}, \mathbf{X})$ will become an integration over all possible values of the hyperparameters. A problem arises when using the prior developed for regularized MRTM2, because the prior is not proper. As seen in figure 4.5.1, the prior has an infinite distribution. This might not be a problem when regarding the problem as a user-defined regularization problem, but it becomes problematic when we need to integrate over all possible values of the hyperparameters to obtain the posterior probability of $\mathbf{w}$ given the data.

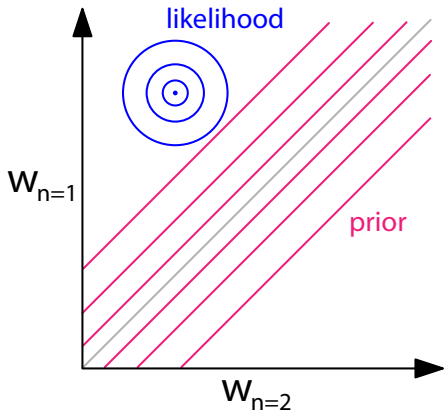

Figure 4.5.1. Improper prior used in regularized MRTM2.

To make the prior distribution proper, a bound has to be set on the distribution. This can be done by introducing a second hyperparameter, with the relationship

$$\alpha_1(w_{n=1} - w_{n=2})^2 + \alpha_2(w_{n=1} + w_{n=2})^2. \tag{4.37}$$

This will lead to a proper prior distribution with zero mean, shown in figure 4.5.2 for one parameter pair.

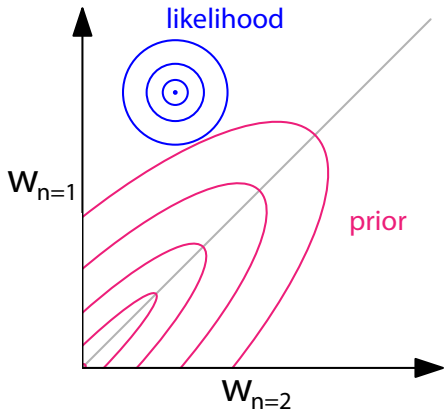

Figure 4.5.2. A proper prior.

To make the prior more adaptive, the mean of the distribution could also be estimated instead of assuming that it would be zero. To not inflate the number of hyperparameters needed to be estimated, one mean for each parameter type could be introduced. This yields the prior distribution

$$p(\mathbf{w}|\boldsymbol{\theta}) = \mathcal{N}_{NM}(\mathbf{w}|\boldsymbol{\mu}, (\alpha_1\mathbf{G} + \alpha_2\mathbf{F})^{-1}), \tag{4.38}$$

with $\mathbf{F}$ being similar in structure to $\mathbf{G}$, but with 1 instead of -1 at the relevant off-diagonal elements. $\boldsymbol{\mu}$ will necessarily be a vector of the same length as $\mathbf{w}$, but will

be μ_1 for all values of the first parameter type and μ_2 for all values of the second parameter type. For an easier notation, all hyperparameters of the prior probability is collected in a hyperparameter vector $\boldsymbol{\theta}$. The posterior is then on the form

$$p(\mathbf{w}|\mathbf{c}) = \int_{\boldsymbol{\theta},\beta} p(\mathbf{c}|\mathbf{w},\beta)p(\mathbf{w}|\boldsymbol{\theta})p(\boldsymbol{\theta},\beta|\mathbf{c})\,\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\beta. \tag{4.39}$$

For a more readable notation the dependencies on $\mathbf{X}$ have been omitted. This integration is difficult to solve, but an often used approximation framework called the evidence approximation treats the distribution $p(\boldsymbol{\theta},\beta|\mathbf{c})$ as sharply peaked around some estimated values $\boldsymbol{\theta}^*$ and β^* (Bishop, 2006). The posterior distribution of $\mathbf{w}$ can then be written on the familiar form

$$p(\mathbf{w}|\mathbf{c}) \approx p(\mathbf{w}|\mathbf{c},\boldsymbol{\theta}^*,\beta^*) = p(\mathbf{c}|\mathbf{w},\beta^*)p(\mathbf{w}|\boldsymbol{\theta}^*). \tag{4.40}$$

The posterior distribution for $\boldsymbol{\theta}$ and β is given by

$$p(\boldsymbol{\theta},\beta|\mathbf{c}) \propto p(\mathbf{c}|\boldsymbol{\theta},\beta)p(\boldsymbol{\theta})p(\beta). \tag{4.41}$$

If the prior distributions of the hyperparameters can be assumed to be relatively flat, the values $\boldsymbol{\theta}^*$ and β^* can thus be obtained by maximizing the marginal likelihood $p(\mathbf{c}|\boldsymbol{\theta},\beta)$.

When having a Gaussian likelihood function as in equation 4.30 and a Gaussian prior distribution as in equation 4.27, the marginal distribution of $\mathbf{c}$ can be derived as

$$p(\mathbf{c}|\boldsymbol{\theta},\beta) = \mathcal{N}(\mathbf{c}|\mathbf{X}\boldsymbol{\mu},\beta^{-1}\mathbf{I} + \mathbf{X}(\alpha_1\mathbf{G} + \alpha_2\mathbf{F})^{-1}\mathbf{X}^{\mathrm{T}}), \tag{4.42}$$

which has been showed by Bishop (2006). This probability will thus be maximized to obtain the optimal hyperparameters. For an easier maximization process, the logarithm of this distribution will be maximized. To maximize the logarithm of a function is equivalent to directly maximizing the function, but makes the process easier. With such a large data set, the probability for the optimal solution will can be so small that an underflow problem can occur if the logarithm is not used. Thus, the logarithm of the function is

$$\begin{aligned}\ln p(\mathbf{c}) = -\frac{NK}{2}\ln 2\pi - \frac{1}{2}\ln\|\beta^{-1}\mathbf{I} + \mathbf{X}\mathbf{H}^{-1}\mathbf{X}^{\mathrm{T}}\| - \\ \frac{1}{2}(\mathbf{c} - \mathbf{X}\boldsymbol{\mu})^{\mathrm{T}}\left(\beta^{-1}\mathbf{I} + \mathbf{X}\mathbf{H}^{-1}\mathbf{X}^{\mathrm{T}}\right)^{-1}(\mathbf{c} - \mathbf{X}\boldsymbol{\mu}).\end{aligned} \tag{4.43}$$

where $\mathbf{H} = \alpha_1\mathbf{G} + \alpha_2\mathbf{F}$. With the use of the matrix determinant lemma for an invertible square matrix on the form

$$\|\beta^{-1}\mathbf{I} + \mathbf{X}\mathbf{H}^{-1}\mathbf{X}^{\mathrm{T}}\| = \frac{\|\beta^{-1}\mathbf{I}\|\|\beta\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{H}\|}{\|\mathbf{H}\|}, \tag{4.44}$$

the Woodbury matrix identity

$$(\beta^{-1}\mathbf{I} + \mathbf{X}\mathbf{H}^{-1}\mathbf{X}^{\mathrm{T}})^{-1} = \beta\mathbf{I} - \beta\mathbf{X}(\beta\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{H})^{-1}\mathbf{X}^{\mathrm{T}}\beta \tag{4.45}$$

and some rearranging the final equation can be written as

$$\ln p(\mathbf{c}) = \frac{1}{2}(-NK\ln 2\pi + NK\ln\beta - \ln\|\mathbf{U}\| + \ln\|\mathbf{H}\| - \beta(\mathbf{c}-\mathbf{X}\boldsymbol{\mu})(\mathbf{c}-\mathbf{X}\boldsymbol{\mu})^{\mathrm{T}} + \beta^2((\mathbf{c}-\mathbf{X}\boldsymbol{\mu})\mathbf{X})^{\mathrm{T}}\mathbf{U}^{-1}((\mathbf{c}-\mathbf{X}\boldsymbol{\mu})\mathbf{X})), \tag{4.46}$$

with $\mathbf{U} = \beta\mathbf{X}^{\mathrm{T}}\mathbf{X} + \mathbf{H}$. This function has to be maximized. Partial first order derivatives of this function can be derived for each hyperparameter without too much trouble and used in e.g. a Gradient descent algorithm. However, they will involve the computation of several traces of large systems of equations with the size $\mathbf{M}^{-1}_{2N\times 2N}\mathbf{N}_{2N\times 2N}$, where N is the number of vertices. The number of vertices in one hemisphere is above 100 000, so even if the system of equations will be sparse the computation time will be vast.

A simpler approach could be to use a non-derivative optimization function, which usually needs many more iterations but works well for a small number of parameters. In this thesis, the MATLAB function `fminsearch` is used to directly optimize on the cost function $-\ln p(\mathbf{c})$. `fminsearch` uses the Nelder-Mead Simplex Method (Lagarias et al., 1998). The heaviest computational task at each iteration will be to compute the logarithm of the determinants of the huge matrices (size $2N \times 2N$). The determinants will be too large to be represented as a standard double-precision floating-point number. Fortunately, the logarithm of such a determinant will be possible to represent and can be computed without explicitly computing the determinant. Futhermore, the matrices will be sparse and symmetrical positive definite. The process starts by performing the LDL-decomposition of the matrix, which is the most efficient decomposition for large sparse symmetrical matrices, in comparison to e.g. Cholesky decomposition. The variant of the decomposition for sparse matrices is on the form

$$\mathbf{P}^{\mathrm{T}}\mathbf{M}\mathbf{P} = \mathbf{L}^{\mathrm{T}}\mathbf{D}\mathbf{L} = (\mathbf{L}\mathbf{D}^{\frac{1}{2}})^{\mathrm{T}}\mathbf{L}\mathbf{D}^{\frac{1}{2}} \tag{4.47}$$

where $\mathbf{P}$ is a permutation matrix, $\mathbf{L}$ is a lower unitriangular matrix and $\mathbf{D}$ is a diagonal matrix. It is related to the original matrix's logarithm as

$$\ln\|\mathbf{M}\| = 2\sum \ln \mathrm{diag}(\mathbf{L}\mathbf{D}^{\frac{1}{2}}) \tag{4.48}$$

This decomposition can also be used for a fast computation of $\mathbf{U}^{-1}((\mathbf{c}-\mathbf{X}\boldsymbol{\mu})\mathbf{X})$, with the decomposition already obtained for the computation of the log determinant. The solution is computed with MATLAB's backslash operator as

```
P * (L' \ (D \ (L \ (P' * (c - X*mu)))))
```

One issue with most optimization algorithms is the possibility to end up in a local minimum and thus not find the true optimal values. Therefore, it is important

to choose a plausible starting point for the algorithm. Another issue is the issue of what data to optimize the parameters on, e.g. to obtain different hyperparameters for different hemispheres of a brain or different subjects in a study might make them less comparable. However, the relationship between number of samples and computation time is non-linear. More samples from e.g. different subjects in one optimization scheme makes the computation time higher than optimizing each subject separately.

The MATLAB implementation of the MRTM2 function with `fminsearch` can be seen in appendix C and the cost function can be seen in appendix D.

4.6. Repeatability Analysis

The comparison of different mathematical models and the assessment of the quality of a certain model's outcome can be done in various ways. Three related standard terms can be used: accuracy, trueness and precision (JCGM, 2012). For a model to be valuable it should be accurate, meaning a high trueness and a high precision. A low accuracy can be due to poor trueness although it is precise, or to low precision although it is on average true. To measure trueness is obviously more difficult in PET modeling than to measure precision. The true concentrations are not known. Post-mortem analyses can give an insight into the serotonin concentrations, but are typically carried out on a small portion of the brain to say something about the larger region. How much the concentration can vary on smaller regions such as voxels or vertices is hard to say. This is also the problem in doing simulated studies, as the trueness of the simulation will be hard to assess. The precision of a study is easier to assess and can be done on a test-retest study, but also for these studies it is an issue with not knowing in very much detail how the binding potential can vary spatially but perhaps more importantly over time. The precision of model on a data set that has been acquired over a short period of time and with the same instruments to keep conditions constant is often called the repeatability of the model. In this thesis, the repeatability of the surface-based vertexwise MRTM2 and the regularized MRTM2 will be evaluated by the use of the psuedo-test-retest data set. As this data set is part of a placebo group, placebo effects might be present. Furthermore, the time between the test and retest scan was around one month which might result in a true inaccuracy between the scans. However, this could also be the case for sessions closer together. The models will be compared on a vertex-level and on a regional level where the vertex-wise BP will be averaged on a larger region. In the regional comparison the models will also be compared to MRTM2 modeling on regional averaged TACs, both from FreeSurfer regions and from the results of the initial PMOD analysis. Apart from exploratory analyses, two statistical measurements will be used.

A simple procentual measurement of bias will be used on the form

$$\frac{\text{retest} - \text{test}}{\text{test}}. \tag{4.49}$$

On the regional analysis, the average between subjects will serve as a measurement

of bias and the standard deviation will serve as a measurement of variability of a model. On the vertex analysis, the measurement is less ideal because of the presence of negative values. Therefore, the denominator will be changed to its absolute value.

Furthermore, the reliability of the models will be assessed by the intraclass correlation coefficient (ICC). This coefficient is often used when measurements have been obtained from different observers measuring the same quantity. There are several types of ICC whereof two will be used in this thesis. They employ a two-way random effects model to compare the within-sample variance to the between-sample variance. ICC(2,1) with absolute agreement will be used for a within-subject analysis of repeatability between vertices of the test and retest scan. It is on the form

$$ICC(2,1) = \frac{BMS - EMS}{BMS + (k-1)EMS + k(JMS - EMS)/n}, \tag{4.50}$$

where n is the number of vertices, k is the number of scans, BMS is the between-vertex mean square, EMS is the residual mean square and JMS is the between-scans mean square. For the regional analysis, the ICC(3,1) for consistency will be used used to measure the repeatability between subjects. It is on the form

$$ICC(3,1) = \frac{BMS - EMS}{BMS + (k-1)EMS}, \tag{4.51}$$

where BMS is the between-subject mean square (Shrout and Fleiss, 1979; Raemaekers et al., 2007).

5. Results

For visualization of the results, the results of the first subject in the study will serve as an example to be able to show some of the advantages and disadvantages of the different approaches. The subject, with CIMBI number 50584, is from the placebo group and had 36 days between the baseline and the follow-up session. It is a rather average subject so the conclusions that can be drawn from this subject can be generalized to the whole population. Although, it should be mentioned that it is slightly more well-behaved than the average, meaning less negative and overly large BP estimates. For the presented results, the automatic surface segmentation of each subject in the study was carefully inspected and manually edited where deemed necessary. In section 5.1, the results of the surface-based vertexwise MRTM2 will be compared to the conventional voxel-based approach. In section 5.2, the results of the vertexwise GLM with and without correction for multiple comparisons will be shown and the ability of the between-brain registration of FreeSurfer will be validated. In section 5.3, the results of the edited data set will be compared to results of the original non-edited data set for a brief discussion on FreeSurfer's automatic segmentation algorithm and the need for manual editing. In section 5.4, the regularized MRTM2 will be compared to vertexwise MRTM2. First, the results of an exploratory analysis will be presented on a few subjects, followed by measurements on all subjects. Furthermore, averaged mean BP for some regions will be compared to the initial results of the study which includes MRTM2 estimations on regional averaged TACs. Finally, a few results of the attempt to treat the data in a fully Bayesian way will be shown in section 5.5.

5.1. Validation of surface-based vertexwise MRTM2

The mean PET volume of each subject and session was registered to the corresponding MR volume at baseline by the Boundary-Based Registration method of FreeSurfer. An example of the registration is shown in figure 5.1.1 for three slices of the baseline scan of the first subject. By closer visual inspection of some subjects, it proved to be a robust method. The cost function gave a value in the close proximity of 0.3 for each scan.

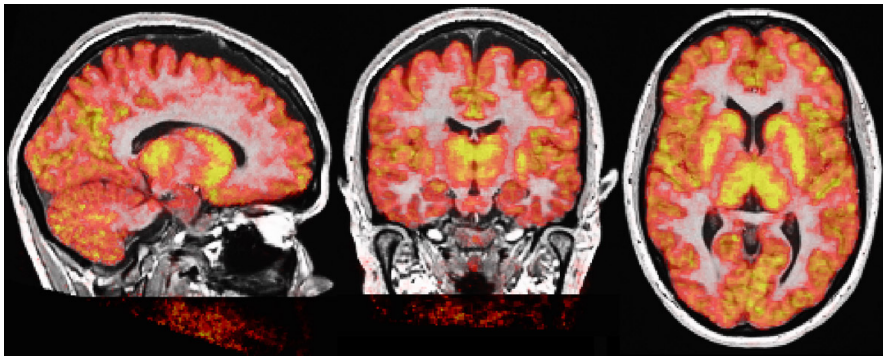

Figure 5.1.1. Mean PET volume of time frame 7 to 32 registered to MR volume by BBR (first subject at baseline).

Figure 5.1.2 shows the BP-maps estimated on raw data and smoothed data with a FWHM 5 mm volume-based Gaussian kernel. It is clear that MRTM2 is highly unstable on raw data, e.g. with high values at unexpected places such as the skull and white matter. However, it should be noted that rather few values are estimated as negative within the cortical layer. In contrast, the smoothed BP-map is more stable and as expected white matter voxels are generally lower than cortical voxels which in turn are lower than voxels in the high-binding subcortical regions. However, it is clear that many outer cortical voxels have been affected by CSF signal. Another thing to note is that voxels with low or no signal have been estimated as negative, which is incorrect.

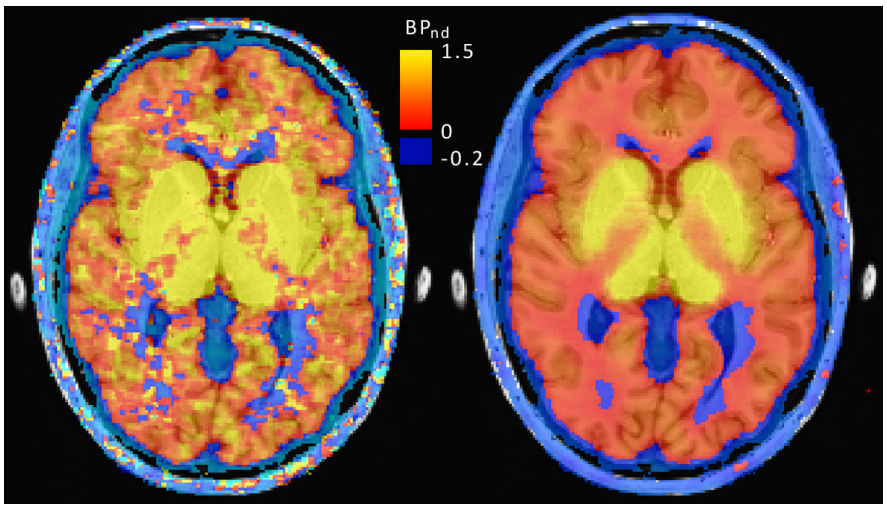

Figure 5.1.2. BP-maps estimated by MRTM2 on raw data and data smoothed by a volume-based Gaussian filter, FWHM 5 mm, overlayed on the MR volume.

To compare to the surface-based smoothing, the raw data was also modeled by MRTM2 on the cortical surface (figure 5.1.3). The result seems rather noisy and many vertices have been estimated as negative.

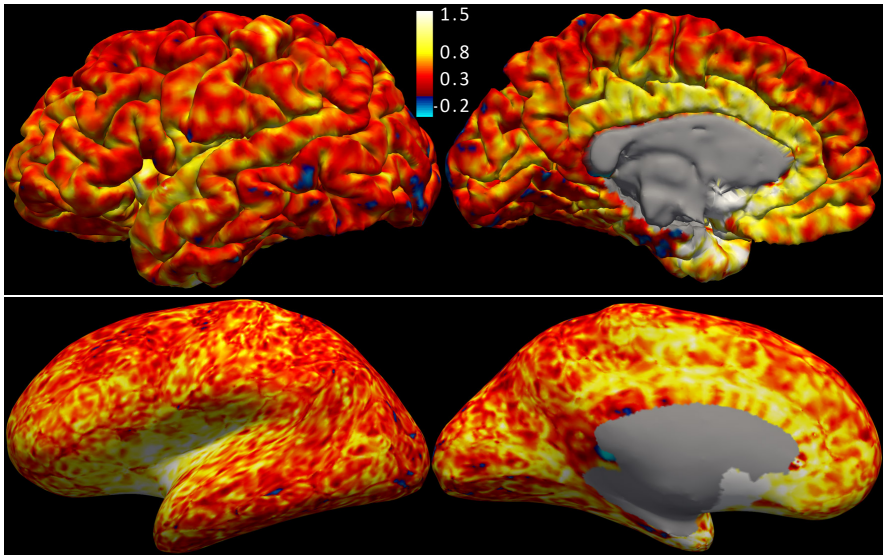

Figure 5.1.3. Raw BP-map on the outer surface and the inflated surface. Lateral and medial view.

The volume-based smoothed BP-map was projected onto the cortical surface (figure 5.1.4). By projecting it onto the cortical surface, it becomes even clearer how the volume-based smoothing affects the cortical region. The BP in gyrus regions are systematically lowered compared to the raw BP-map.

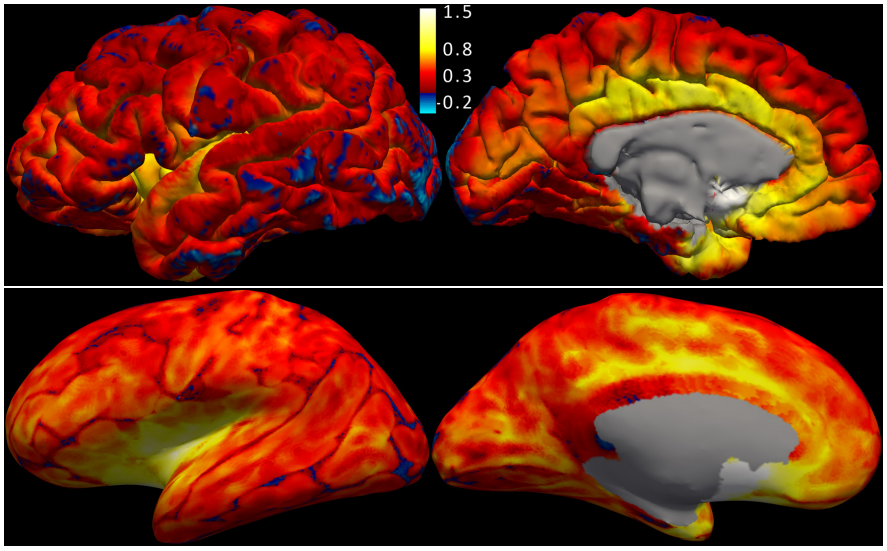

Figure 5.1.4. BP-map (FWHM 5 mm) estimated on volume-based smoothed data on the outer surface and the inflated surface. Lateral and medial view.

The BP-map estimated on data smoothed across the surface is shown in figure 5.1.5. Comparing this to the volume-based smoothing, it can be seen that the volume-

based smoothing results in lowered BP in deep folds and a mirror effect at the walls of the folds. Unsurprisingly, the volume-based smoothing mixes signal between different tissue types and between the walls of a fold. When comparing surface-based smoothing to the estimation on raw data, it can be seen that surface-based smoothing results in fewer negative values and a less noisy appearance but also a loss in resolution.

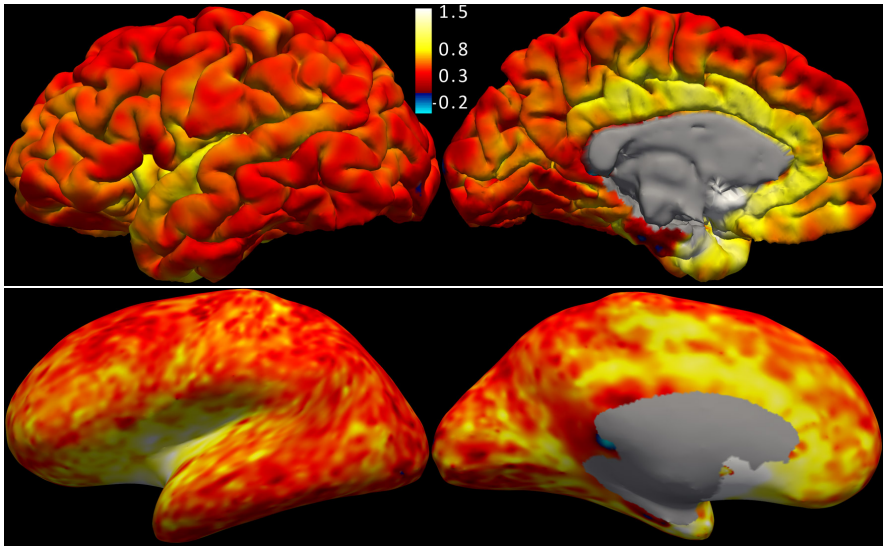

Figure 5.1.5. BP-map estimated on data smoothed across the cortical surface (approximately FWHM 5 mm). Visualized on the outer surface and the inflated surface. Lateral and medial view.

In table 5.1.1, some statistics are shown comparing the three BP-maps on the cortical surface. The main findings are that MRTM2 on raw data results in vastly too high and low BP values, while the volume-based smoothing results in more reasonable values but with an average lowering of BP and many negative values. In contrast, surface-based smoothing results in few negative values. However, the minimum BP is lower than the one obtained through volume-based smoothing.

Table 5.1.1. Comparison between BP estimations on raw data, data smoothed on voxel level and data smoothed on surface (first subject at baseline). Implausible estimates refers to the number of vertices that have BP estimates that are either negative or above 10.

BP estimations	mean $\pm$ SD	max	min	implausible estimates
On raw data	0.51 ± 0.33	20.11	-10.72	1028
FWHM 5 mm - in volume	0.36 ± 0.24	2.66	-0.39	3254
FWHM 5 mm - on surface	0.49 ± 0.24	3.64	-0.48	161

Two rather large areas still had negative BP after 5 mm surface-based smoothing.

One is seen on the medial view in figure 5.1.5 and the other one was in the back of the occipital lobe at the same height as the first one. To see if there are any segmentation issues involved within these areas, figure 5.1.6 shows the two areas on a slice of the MR volume. The inner area is a typical segmentation issue for FreeSurfer and was present in almost all subjects, which resulted in a presence of negative BP. In the areas where the two hemispheres are connected to each other, the MR intensity between different tissue types becomes blurred. Thus, it becomes difficult to correctly segment these areas. The other area, in the back of the brain, is correctly segmented. This is a part of the occipital lobe which includes regions with low concentration of serotonin transporters.

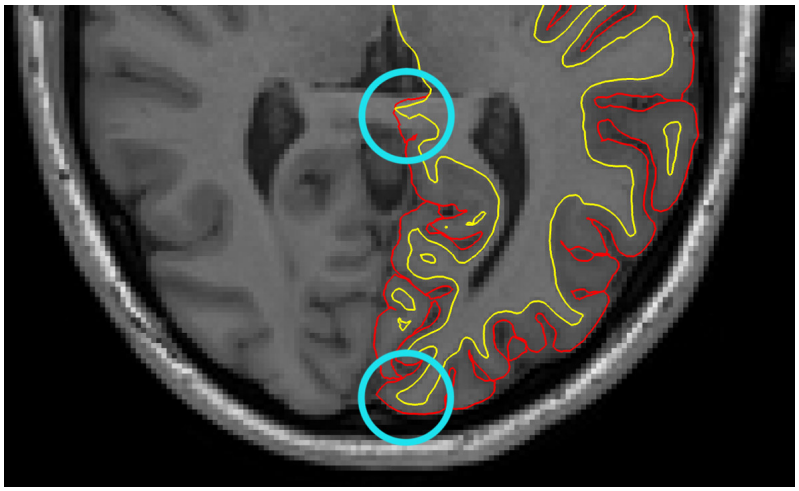

Figure 5.1.6. Cortical segmentation at areas that still had negative BP after surface-based smoothing of 5 mm FWHM in the back of the brain.

5.2. Between-brain analysis in FreeSurfer

For the between-brain analysis, the subjects were mapped to the average subject and the baseline and follow-up scan were smoothed with FWHM 10 mm in common space before the modeling with MRTM2. The resulting ΔBP-map for the first subject is shown in figure 5.2.1. It can be seen that the level of smoothing is rather high. Also 5 mm and 15 mm smoothing were tried, with less significant result in the between-brain analysis, which could suggest that 10 mm is the required level of smoothing to account for the variability between brains.

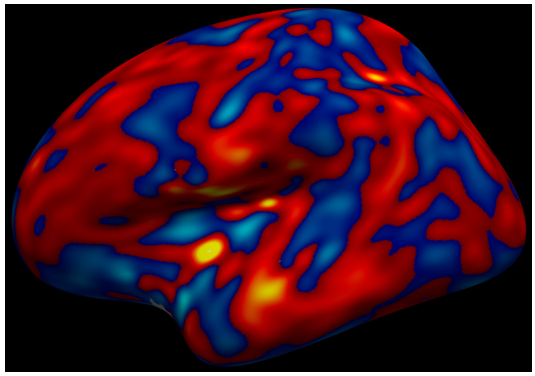

Figure 5.2.1. ΔBP-map for first subject in common space with 10 mm smoothing.

5.2.1. General linear model

The general linear model in equation 4.13 was fitted at each vertex independently to find the local significance of the difference in slope between the treated group and the placebo group. The uncorrected p-maps for both hemispheres thresholded at $p = 0.05$ are shown on the outer surface in figure 5.2.2 and on the inflated surface in figure 5.2.3.

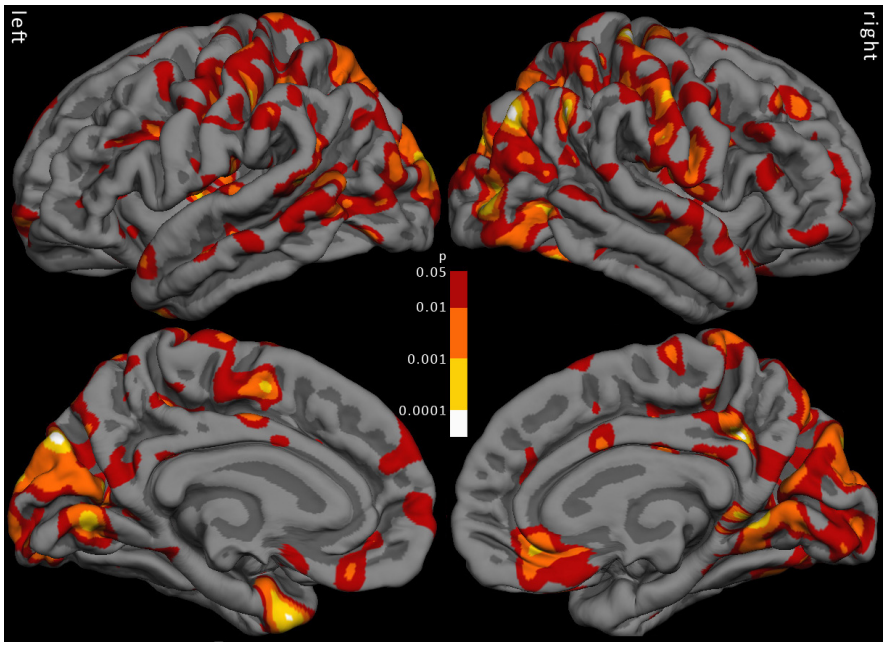

Figure 5.2.2. Uncorrected p-maps for both hemispheres with the threshold $p = 0.05$. Lateral and medial view.

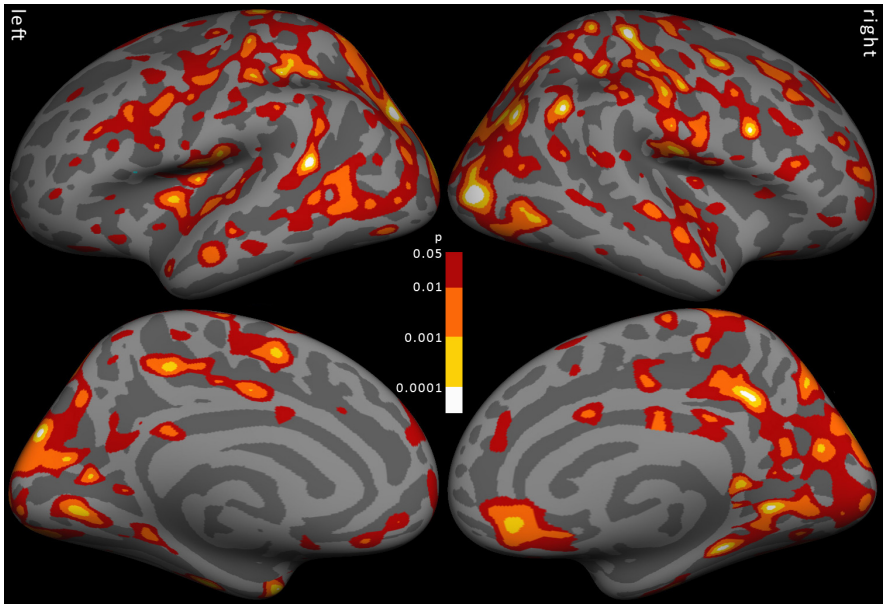

Figure 5.2.3. p-values hamilton.

The most significant areas seems to be in the occipital and parietal lobe. Unfortunately, these regions are not of interest in the sex-hormone study. These are also the regions that show most coherence between the two hemispheres. Coherency between the hemispheres is important because functional differences between them are not expected. The frontal cortex, a region of interest in the analysis, did show a significant effect in the initial regional analysis but seems virtually unsignificant in this more local analysis.

The p-map was corrected for multiple comparisons by FDR with a q-value of 0.05, shown in figure 5.2.4, and with clusterwise correction with a vertexwise threshold at $p = 0.01$ and a clusterwise threshold of 0.05, shown in figure 5.2.5. The FDR procedure leaves very few and small areas as significant. The areas do not have any coherency between the hemispheres. As the p-map is smooth, this shows the apparent conservatism of the FDR procedure when it comes to correlated data. It could of course also mean that the GLM does not have any strong significant results, but in any case, FDR does not in this case yield any useful results.

In contrast, the clusterwise correction does result in some interesting clusters. Especially the highly significant cluster in lower parietal lobe shows some interesting coherent characteristics between the hemispheres. Other significant clusters with some coherency are located at the top of the postcentral gyrus and at the medial end of occipital lobe. However, these do not seem to be exactly at the same location in both the hemispheres. Unfortunately, these areas is not interesting in the study at hand, but do show that the clusterwise correction can be a valuable method for finding patterns on the cortical surface. However, the uniform clusters lacks detailed information and the their boundaries seem rather arbitrary. This could make it hard in many cases to draw conclusions from such an analysis alone.

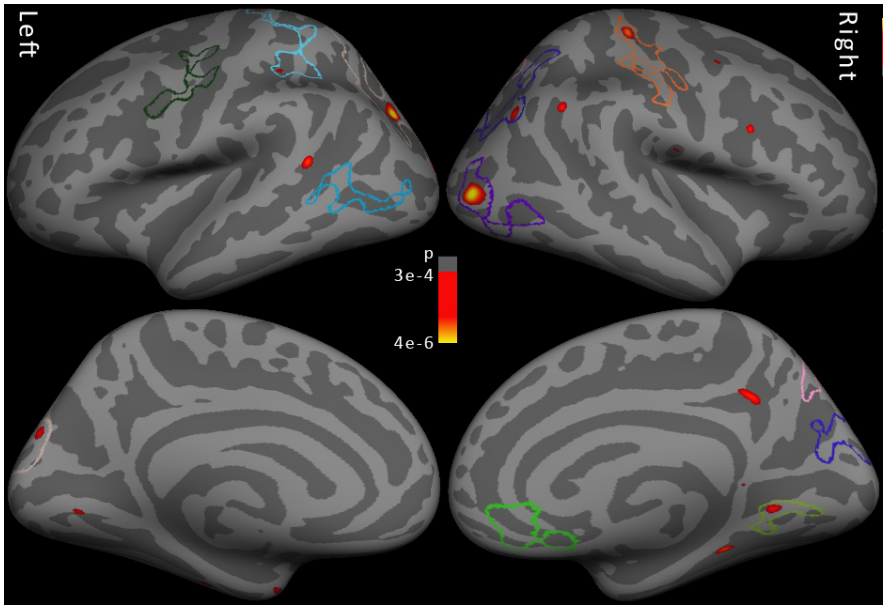

Figure 5.2.4. False Discovery Rate, q=0.05. Contours of clusters from the clusterwise correction are overlayed on the surface. Left and right hemispheres in lateral and medial view.

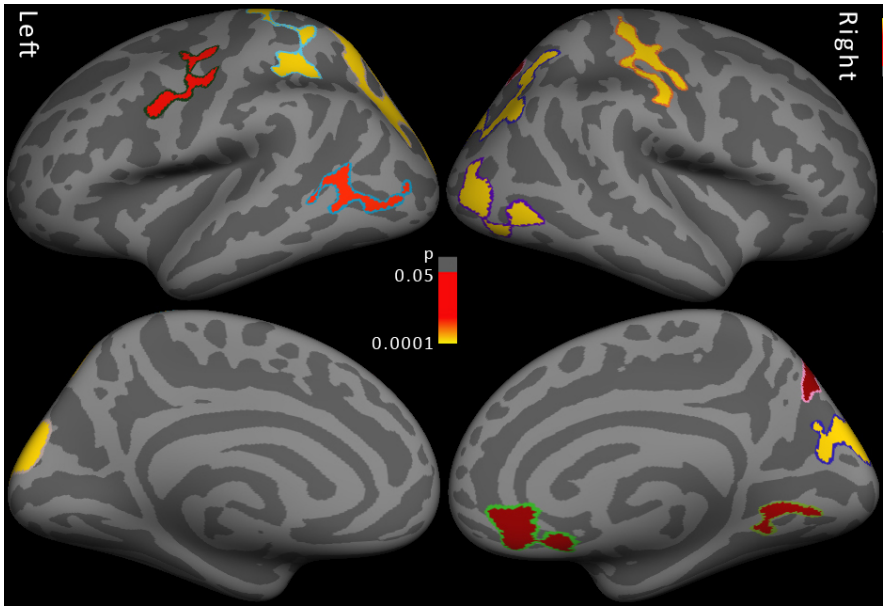

Figure 5.2.5. Cluster analysis for a vertexwise p-value threshold of 0.01 and a clusterwise threshold of 0.05. Left and right hemispheres in lateral and medial view.

In the initial analysis described in section 4.2.3, one subject was present that showed

a great influence on the outcome of the model on regional averaged TACs of neocortex. For that test, the subject had a Cook's distance of 0.27. The visualization capabilities of FreeSurfer makes it easy to visual this subjects' Cook's distance across the surface. This is shown in figure 5.2.6. Compared to other subjects, this subject does have a very large Cook's distance in many regions of the cerebral cortex. However, when comparing to the significance map in figure 5.2.3 it can be noted that the subject does not have a very large Cook's distance at the significant areas. This suggests that these areas can be more certain in their significance.

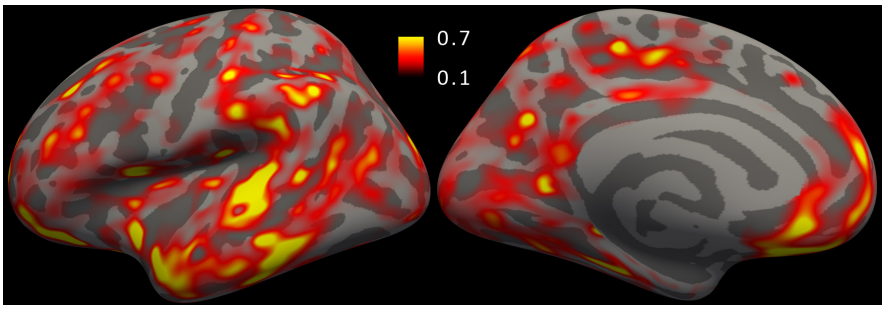

Figure 5.2.6. Cook's distance for driving subject on global neocortical level on left hemisphere.

5.2.2. Problems with registration to common space

As described in section 4.3, individual brains are mapped to the average subject by fitting patterns of convexity. It is therefore interesting to examine these patterns and to see how well they are mapped to the average subject. In figure 5.2.7, 8 subjects convexity patterns are shown mapped to the average subject in lateral view. The first brain shows the average convexity patterns that they are mapped to. It can be seen that the main gyrus structures of the brain are well mapped to the average subject, such as pre- and postcentral gyrus and superior and medial temporal gyrus. However, many areas seems to have less clear patterns that are more difficult to map, such as in the frontal lobe and between occipital lobe in the back of the brain and surrounding regions. Futhermore, many smaller structures do not seem to be well-mapped.

In figure 5.2.8, the convexity patterns are shown on the medial view. Mapping problems can be seen in the upper region in the back of the brain. Many smaller patterns are also not well-mapped. It seems clear that the high level of smoothing for the between-brain analysis is necessary in the present general linear model framework, despite other issues with spatial smoothing. It should be noted that the FreeSurfer between-brain registration algorithm has been validated and compared to other registration methods with good results.

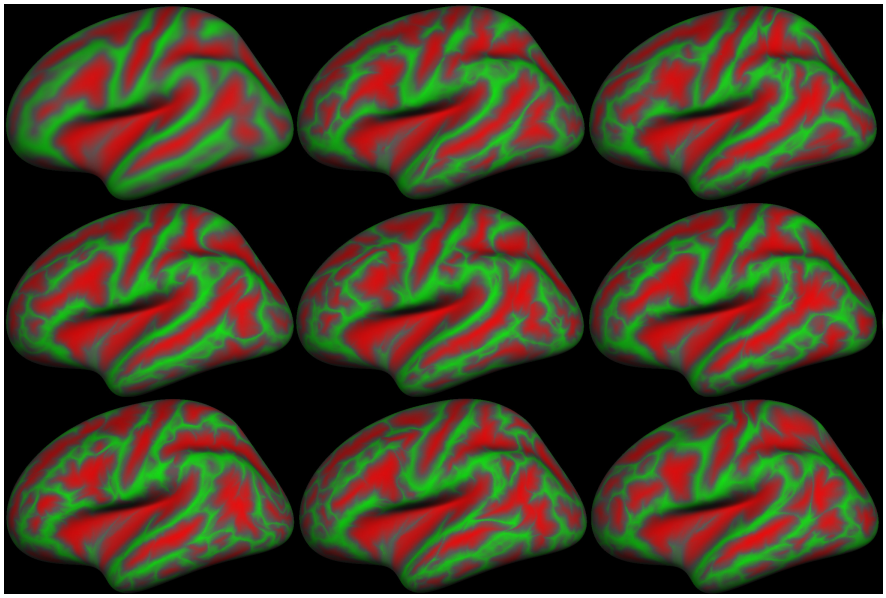

Figure 5.2.7. Convexity patterns on the average subject, left hemisphere lateral view. On the first brain the average convexity pattern is shown that is used as the reference when mapping individual brains to the average surface. The convexity patterns of the 8 first subjects of the sex-hormone study is shown mapped to the average subject.

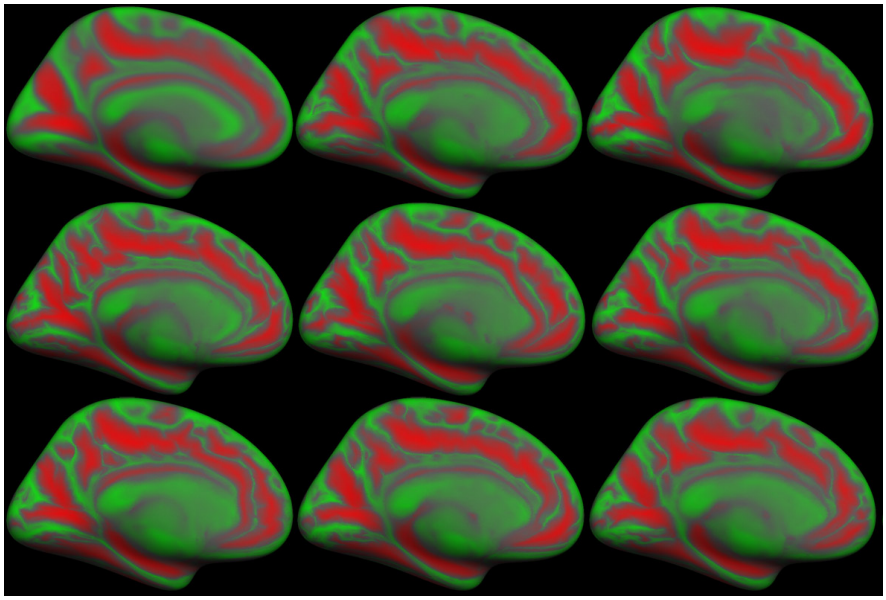

Figure 5.2.8. Convexity patterns on the average subject, left hemisphere medial view. The convexity patterns of the 8 first subjects of the sex-hormone study is shown mapped to the average subject. First subject shows the average convexity patterns that are mapped to

To further investigate the performance of the mapping to average space and to generalize the findings of visually comparing 8 subjects convexity patterns, the standard deviation was computed for the mapped convexity patterns of all subjects in the sex-hormone study, shown in figure 5.2.9. It can be seen that the most severe differences are in sulcus regions in the transition from occipital lobe to temporal lobe, and in the parietal lobe. Also middle frontal gyrus is affected.

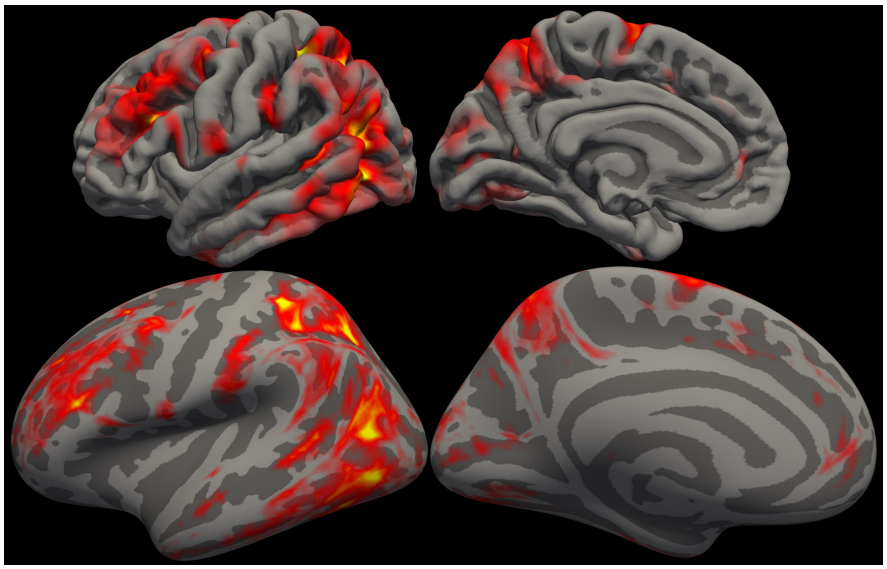

Figure 5.2.9. Standard deviation of the mapped convexity patterns, $n = 61$. Outer surface and inflated surface of left hemisphere, lateral and medial view

To examine the results of these registration issues on a group analysis, the convexity standard deviation map was overlayed on the p-map in figure 5.2.3. The continuity between the measurements are quite striking. The p-values in frontal cortex follows the line of the standard deviation convexity area at its immediate proximity. In occipital to temporal cortex a p-value cluster goes right between two areas with high standard deviation of convexity. Furthermore, a highly significant p-value area in parietal cortex also goes between two areas of high standard deviation of convexity. If this is due to the worse between-brain mapping in these areas or have other structural explanations is hard to say. Notice also that the ΔBP-maps in the GLM have been smoothed after brain mapping, which might be why the p-value clusters seem to extend slightly into affected areas.

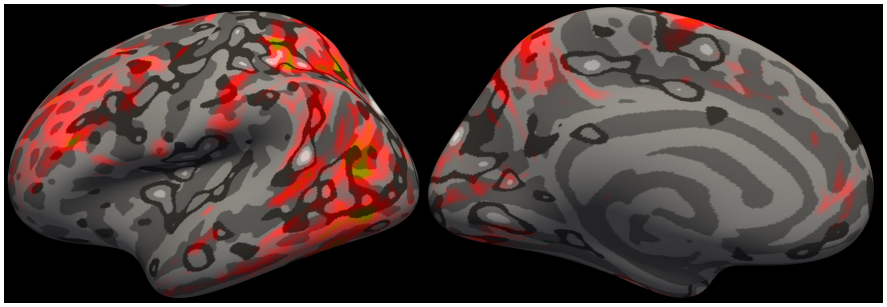

Figure 5.2.10. Standard deviation of the mapped convexity patterns overlayed on the p-map of the general linear model performed in this thesis, thresholded at p=0.05. Outer surface and inflated surface, lateral and medial view.

5.3. Editing of surface segmentation

The cortical segmentation of FreeSurfer was checked and manually edited where deemed necessary. The segmentation was checked by two persons, whereof one was a medical doctor with experience of neuroimaging data. The segmentation was carefully edited and reviewed to be able to assess the effect of manual smoothing. The automatic segmentation of FreeSurfer works well, but as any automatic segmentation it does result in quite a few errors in a surface reconstruction procedure on average. Only two subjects were judged to not need manually editing and most subjects were quite significantly edited. Some areas were systematically segmented with errors, such as the upper central medial area of the brain, lower frontal cortex and at the very back of occipital cortex. Most issues regarded the outer surface. An example of a typical notch before and after editing at a upper central medial area can be seen in figure 5.3.1, together with a more rare problem with the inner surface.

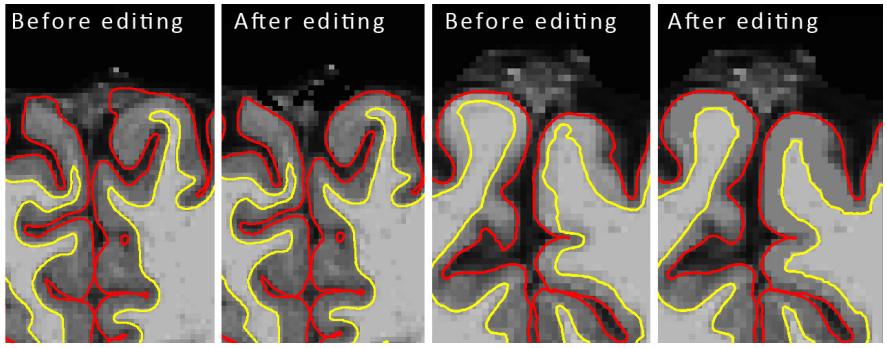

Figure 5.3.1. Editing segmentations.

Many subjects were edited, but the question is to what extent segmentation issues actually affects a statistical analysis. Obviously, most segmentation issues will be at different areas for different subjects and with the 61 subjects in the study, the study can be seen as rather robust even if one or two subject would have a bad

segmentation at one vertex. In figure 5.3.2, negative correlations and paired t-tests of differences in volume between the edited and non-edited data set can be seen for each vertex. Most areas are well correlated and had no significant statistical difference. There are however some areas with significant difference, most notably in the areas which we could see had systematic errors.

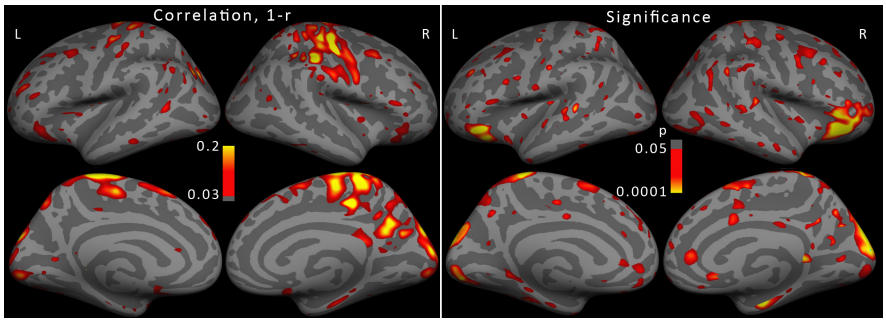

Figure 5.3.2. Negative correlation of cortical volume between edited and non-edited data sets, and significance of the difference in cortical volume.

The areas most affected by segmentation issues were not significant in the general linear model, so conclusions about the effect of segmentation errors on a statistical group analysis are hard to draw from this study. A small change in p-value did however change the result of the clusterwise analysis on the left hemisphere, as seen in figure 5.3.3. The change in p-value was quite small, so this result might say more about the cluster analysis and p-value thresholding than about the effect of segmentation issues. Nevertheless, segmentation issues will give a higher level of noise and uncertainty that under some unfortunate coincidences might effect the outcome of a statistical analysis.

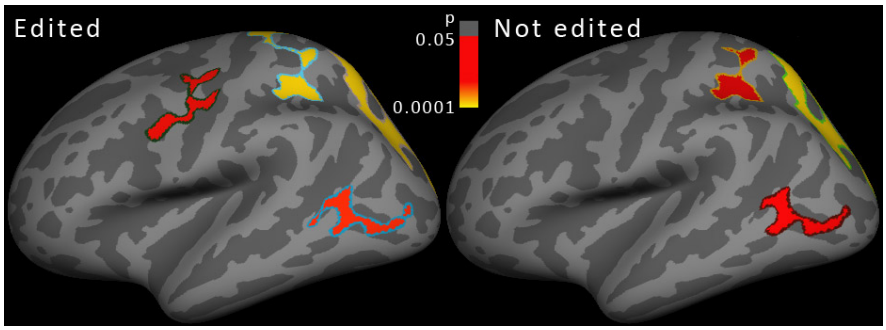

Figure 5.3.3. Difference between edited data set and non-edited data set on cluster analysis.

5.4. Regularized MRTM2 compared to MRTM2 with pre-smoothing

The results of regularized MRTM2 compared to the vertexwise MRTM2 with a pre-smoothing step will first be presented by visualization and an exploratory analysis on a few subjects. A repeatability analysis on vertex level will then draw conclusions on the whole test-retest data set, followed by a regional analysis.

5.4.1. Visualization and exploratory analysis

The results will first be presented and visualized on the first subject of the sex-hormone study. Three levels of smoothing and regularization will be compared. To find a level of regularization that is directly comparable to a level of smoothing is difficult due to the quite different characteristics of the two models. However, with three levels the progress from a low level of regularization to a high level of regularization can be shown. The range of smoothing and regularization has been chosen to be as comparable as possible when taking all aspects into account. A visualization of the result on the baseline scan is shown in figure 5.4.1.

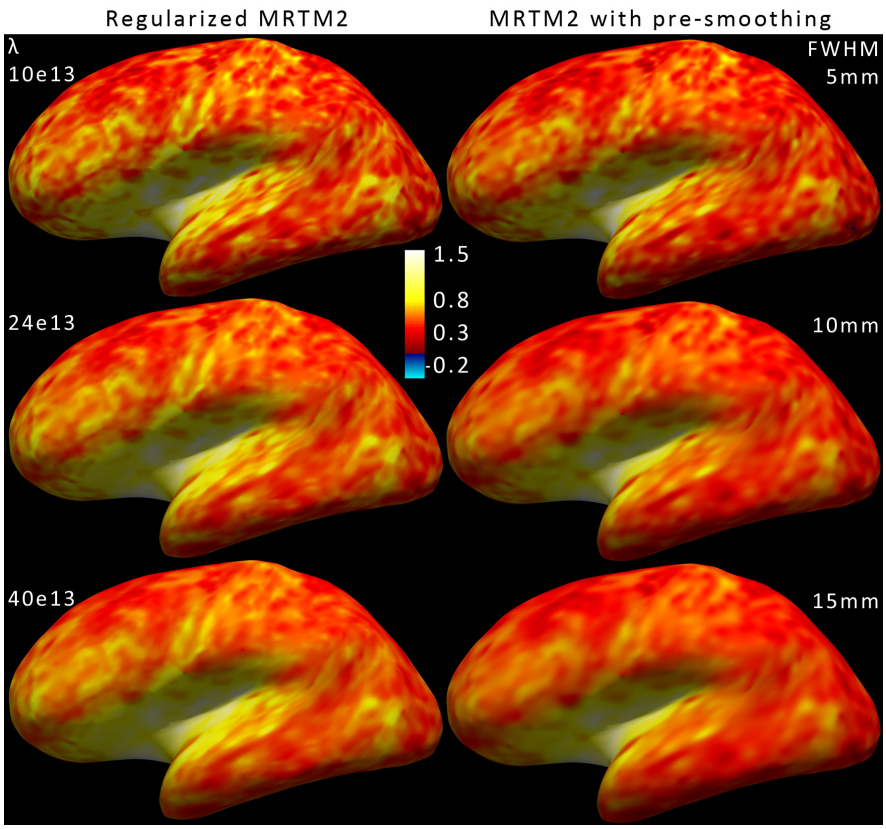

Figure 5.4.1. baseline BP for different levels of smoothing and regularization. Left hemisphere, lateral view.

Regularized MRTM2 seems to retain a higher resolution than the pre-smoothed vertexwise MRTM2, even at the highest regularization level. It is sharper and more detailed, without appearing noisy. The 15 mm smoothed BP-map is highly blurred and also the 10 mm is quite blurry. The 5 mm smoothed map contains a small negative area in occipital cortex, while the map with the lowest level of regularization is able to produce a positive result in that area. The result of the left hemisphere in medial view is shown in figure 5.4.2.

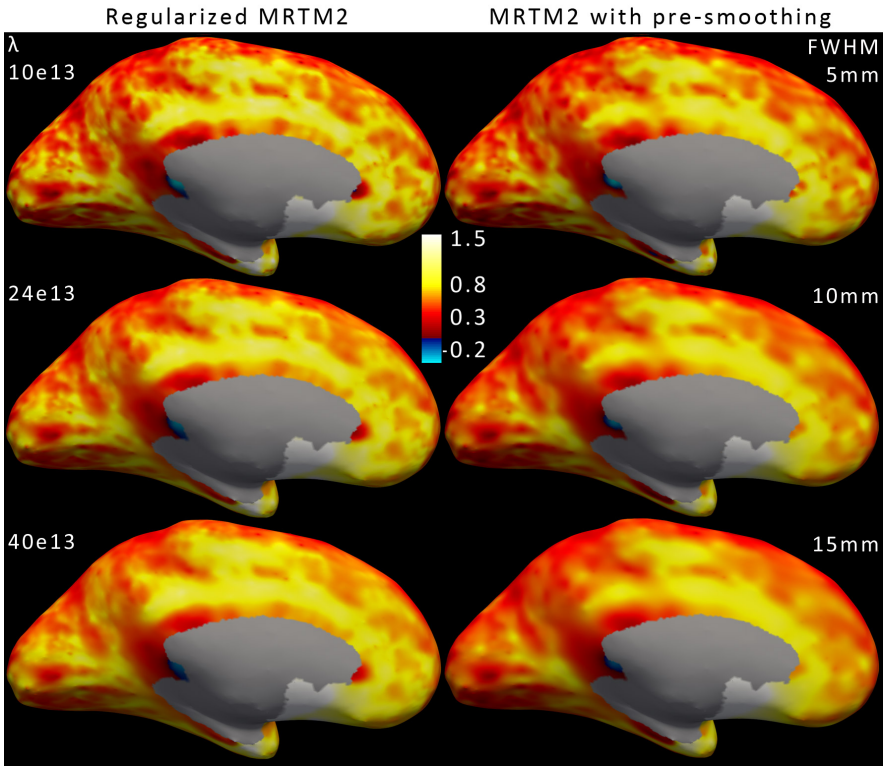

Figure 5.4.2. baseline BP for different levels of smoothing and regularization - Medial view.

The same conclusions can be drawn from this view as for the lateral view. However, a negative area close to the medial wall is larger for all regularization levels than for the smoothed maps. As shown earlier, this area is often segmented too far into other tissue types. Although the area is larger in the regularized maps, it is moved closer to the medial wall which might be reasonable. This move could be due to the fact that it is regularizing on neighboring vertices but not vertices within the medial wall. This means that vertices close to the wall will have less neighbors to regularize to. Another characteristic of the regularization noticeable in this figure is that a higher regularization yields a slightly higher signal on average.

Some statistics on the first subject is shown in figure 5.4.1. Both models works fairly well on the baseline scan but the regularized MRTM2 is superior on the follow-up scan. The number of vertices outside the plausible range is lower for

the regularized case and it's range is shorter, even though it retains a higher resolution. It should be noted that even BP values of 3 is quite high for the cortical layer. Furthermore, the standard deviation is comparable between the two models and the mean value becomes slightly higher for higher levels of regularization, but stayes the same for different levels of smoothing.

Table 5.4.1. Comparison between BP estimations with vertex-by-vertex MRTM2 on pre-smoothed data and BP estimations with regularized MRTM2. Implausible estimates refers to the number of vertices that have BP estimates that are either negative or above 10.

MRTM2 with pre-smoothing	mean $\pm$ SD	max	min	not in range
baseline				
FWHM 5 mm	0.49 ± 0.24	3.64	-0.48	161
FWHM 10 mm	0.49 ± 0.23	3.33	-0.36	86
FWHM 15 mm	0.49 ± 0.22	3.15	-0.26	70
follow-up				
FWHM 5 mm	0.50 ± 0.31	38.81	-0.49	507
FWHM 10 mm	0.50 ± 0.24	8.37	-0.43	205
FWHM 15 mm	0.49 ± 0.23	5.38	-0.38	123
Regularized MRTM2	mean $\pm$ SD	max	min	not in range
baseline				
λ =10e13	0.54 ± 0.25	3.45	-0.42	131
λ =24e13	0.56 ± 0.23	3.00	-0.30	84
λ =40e13	0.57 ± 0.21	2.71	-0.23	55
follow-up				
λ =10e13	0.55 ± 0.25	3.49	-0.44	194
λ =24e13	0.57 ± 0.22	2.62	-0.35	108
λ =40e13	0.58 ± 0.21	2.33	-0.28	71

ΔBP-maps for different levels of smoothing and regularization are shown in figure 5.4.3 on the left hemisphere in lateral view. This is the baseline BP-map subtracted from the follow-up BP-map. The average value is comparable between the equivalent level of regularization and smoothing. However, some noticeable differences can be observed. The general structures of the signal are obtained through different levels of regularization, with only a lowering of the average signal and a more coherent result with higher levels of regularization. In contrast, the smoothed BP-maps get noticeably smoothed out with higher level of smoothing, spreading strong signals to neighboring areas. Furthermore, the some rather large differences between the smoothed BP-maps and the regularized BP-maps can be observed in insula and in the front of temporal cortex.

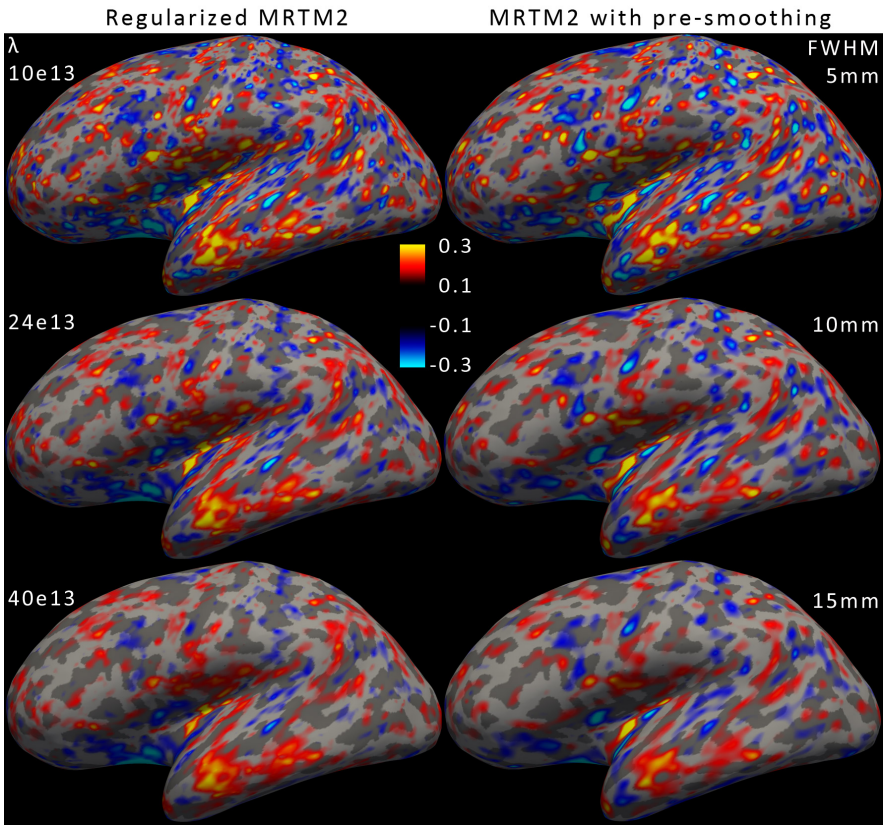

Figure 5.4.3. Δ BP for different levels of smoothing and regularization. Left hemisphere, lateral view.

ΔBP-maps for different levels of smoothing and regularization are shown for the left hemisphere in medial view in figure 5.4.4 in order for completeness. The same conclusions can be drawn from this as for the lateral view. The signal is different especially in frontal cortex, where a big blob of signal appears in the pre-smoothed maps in comparison to only very low signal in the regularized maps.

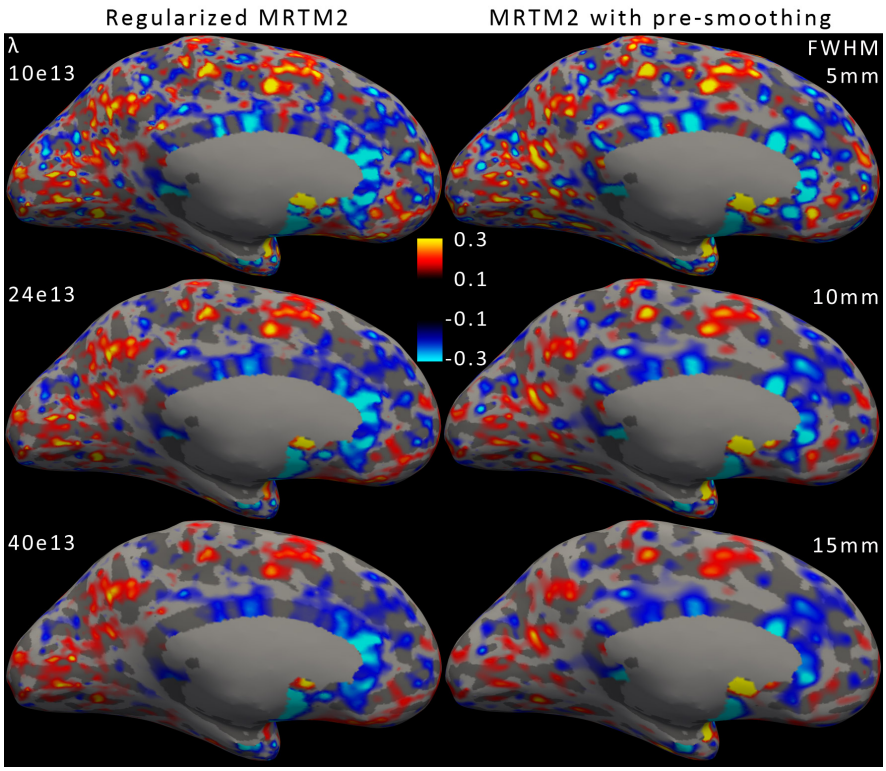

Figure 5.4.4. ΔBP for different levels of smoothing and regularization. Left heimsphere, medial view.

The regularized MRTM2 regularizes the estimated parameters k_2 and R_1 to be similar to the parameters of neighboring vertices. To visualize the effect this regularization has on the parameters, a scatter plot on vertex pairs for each parameter is shown in figure 5.4.5 for the mid-level of regularization. It is contrasted with the estimations on raw data and on smooth data with 10 mm FWHM.

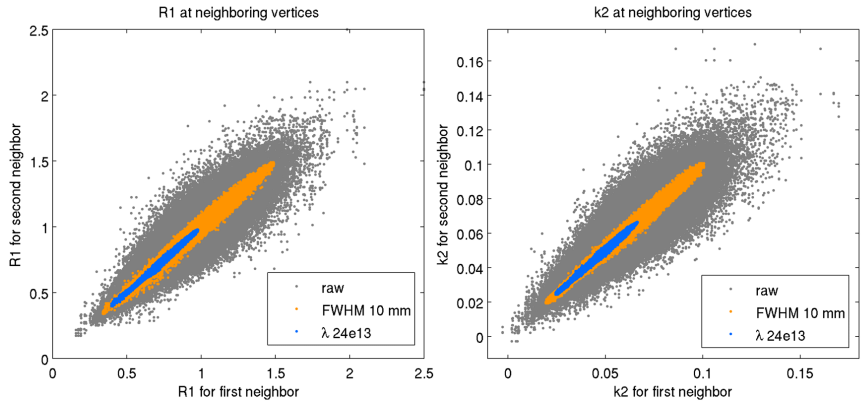

Figure 5.4.5. Parameter estimates for neighboring vertices comparing the λ24e3 regularization level to 10 mm smoothing level and estimations on raw data.

This figure shows that the smoothing have a similar effect on the distribution as in the regularized case, bringing neighboring vertices closer together, but to a lesser degree. The regularization is high and more importantly seems to draw the mean of the distribution closer to zero compared to the other distributions. This is somewhat surprising, but could be the reason why the model seems to result in a slightly higher average BP value with increased level of regularization. However, only the contour of the distributions can be seen in the figure, the mean of the distribution might still be similar across the distributions. Importantly, the high regularization do not result in a lowering of the variance of the BP estimations compared to a high smoothing level.

The parameter k_2 is plotted against the parameter R_2 in figure 5.4.6. The various lines spreading out from the origo shows combinations of k_2 and R_1 that results in a certain BP value. All vertices above the line of BP equal to zero will be negative. It is clear that the regularization approach treats the parameters differently than the pre-smoothing approach, with the result of a more dense distribution. This is expected as it is regularizing directly on the parameters. It seems to lower the mean of the distribution, but again only the border of the distribution is shown. It does however alter at least the range of R_1 in a way that is not consistent with the literature, e.g. in the paper by Kim et al. (2006) the regional average of R_1 was for many regions around 1, while 1 actually is outside of the range of this distribution. However, this does not alter the range of BP in a similar way and the variance of BP is not significantly lowered compared to the pre-smoothing approach.

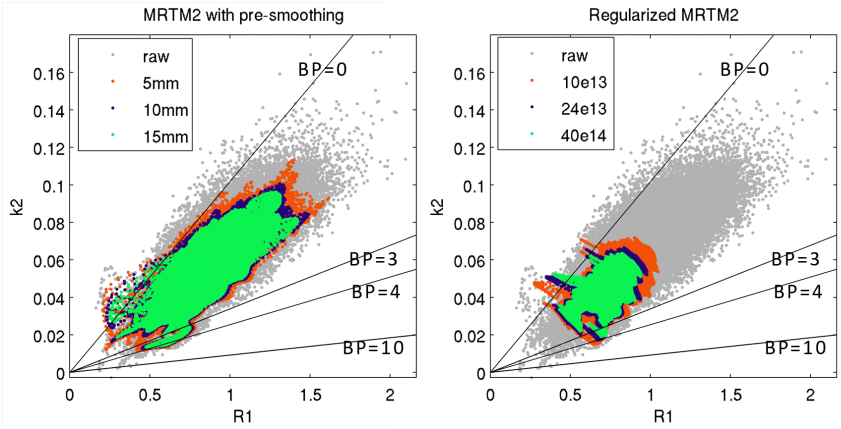

Figure 5.4.6. k_2 estimates against R_1 estimates for MRTM2 with pre-smoothing and regularized MRTM2. Black lines show different BP levels. All vertices above the BP$=0$ line will have negative BP.

In figure 5.4.7, scatter plots between baseline and followup BP for the 5 first subjects are shown for the different levels of regularization and smoothing. This clearly shows that the regularized MRTM2 handles extreme noise better than the pre-smoothed approach. As described earlier, the smoothing of data with high levels of noise might not be able to effectively filter out the noise, but might instead spread out the a noisy value of a data point to neighboring vertices. This does in fact seem to be the case. In addition, the regularization results in a better consistency between

hemispheres, with an equal range of values and a more similar shape.

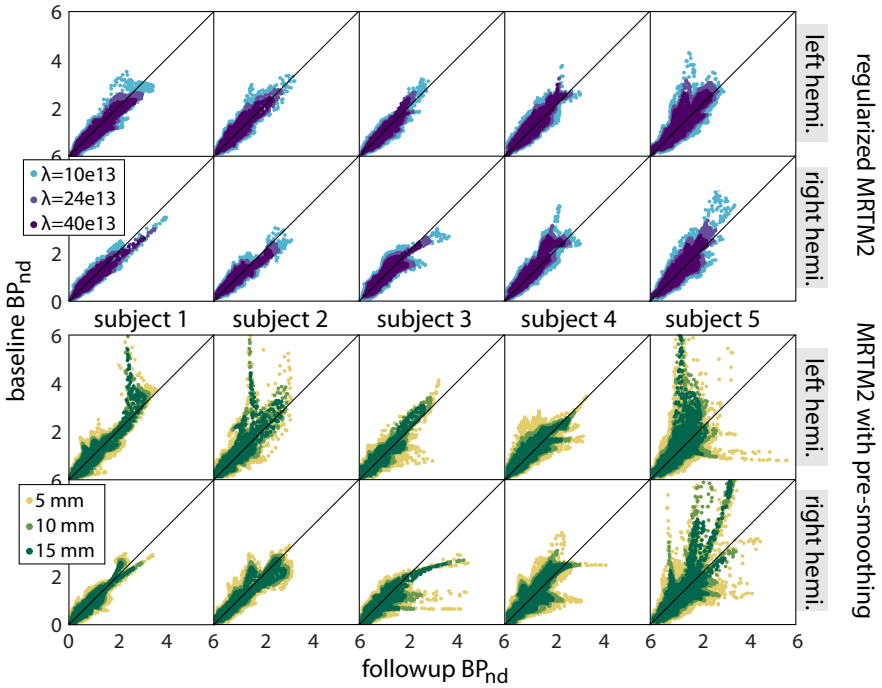

Figure 5.4.7. Scatter plot between baseline and followup BP for the first 5 subjects with different levels of smoothing. Right and left hemisphere are shown separately.

In figure 5.4.8, the scatter plot has been zoomed in and FreeSurfer regions have been colored. Only the highest level of smoothing and regularization is shown. Mainly regions with a high average BP is involved in the extremely high values in the scatter plots from pre-smoothed data, such as insular. Also here it is clear that the regularized MRTM2 handles extreme noise better and yields a more coherent result.

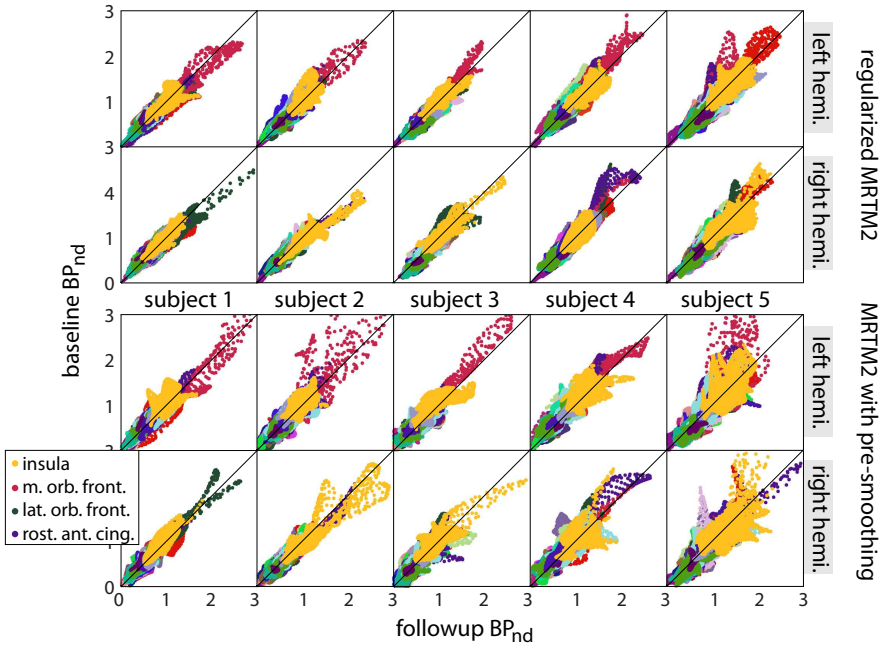

Figure 5.4.8. Zoom-in on a scatterplot between baseline and followup BP for the first 5 subjects and the highest level of smoothing and regularization. FreeSurfer regions have been colored. Right and left hemisphere are shown separately.

In figure 5.4.9, the scatter plot is shown for values around zero. Especially isthmus cingulate is a problematic regions for both methods, but mainly for the pre-smoothing approach. Isthmus cingulate is the region close to the medial wall that was previously seen to give negative results in the visualization of the first subject. Some other regions are also problematic for the pre-smoothing approach, but not for the regularized MRTM2.

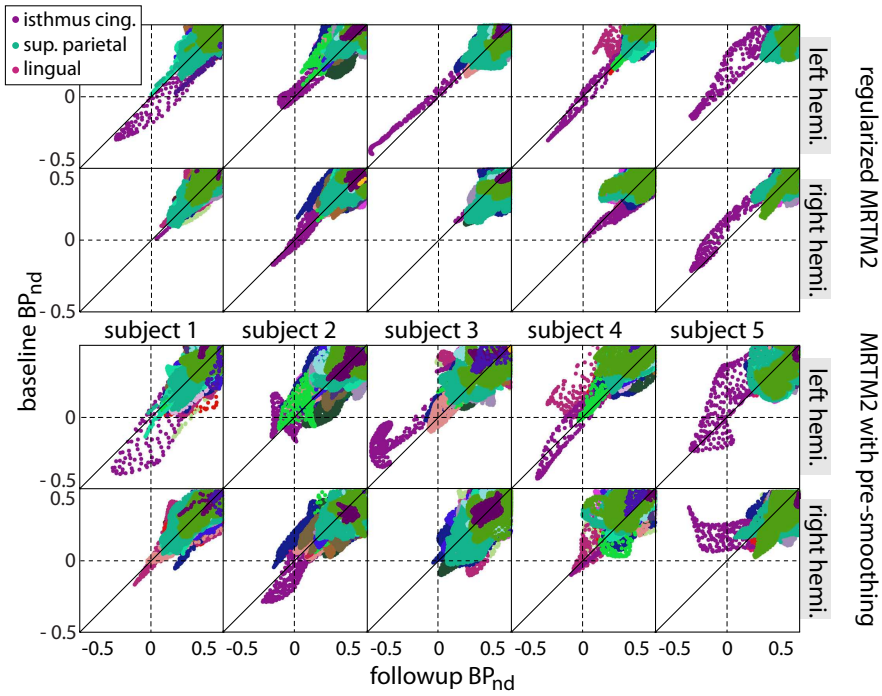

Figure 5.4.9. Zoom-in on a scatterplot to the proximity of 0 between baseline and followup BP for the first 5 subjects and the highest level of smoothing and regularization. FreeSurfer regions have been colored. Right and left hemisphere are shown separately.

5.4.2. Repeatability analysis on vertex level for all subjects

To be able to draw conclusions for all subjects in the test-retest data set some measurements have been performed on vertex level for each subject and level of smoothing and regularization. These measurements have been computed for the whole brain. In figure 5.4.10, the number of vertices that have negative BP or a BP above 10 have been plotted for each subject with smoothing levels on the x-axis and regularization levels on the y-axis. On average, an increased level of regularization and smoothing results in less vertices that are not in the range. Regularized MRTM2 outperforms MRTM2 with pre-smoothing for almost all subjects. Also when comparing the highest level of smoothing with the mid-level of regularization, the regularization approach performs better on average.

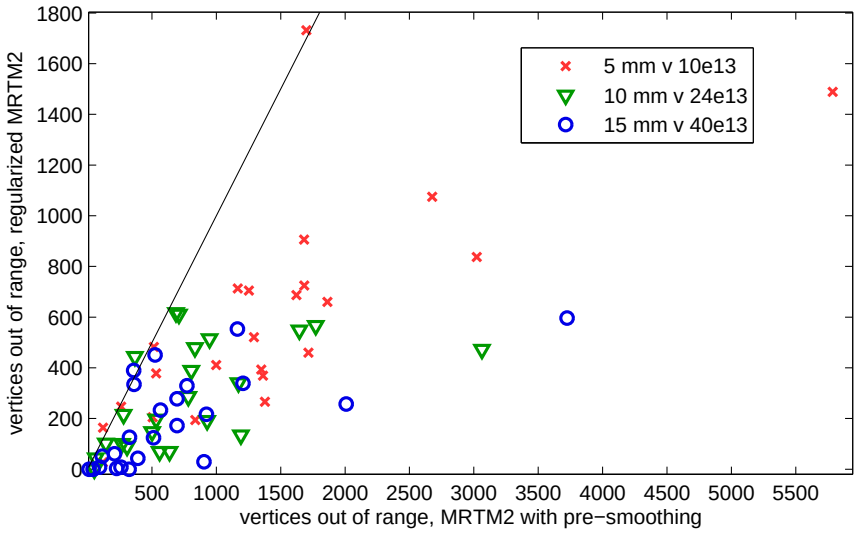

Figure 5.4.10. Number of vertices for each subject that have negative BP or a BP above 10. Levels of smoothing plotted against levels of regularization.

An inappropriately high level of regularization or smoothing could result in a good agreement between test and retest scan simply because of a very low variance within each scan. Therefore, the variance of all vertices for each subject has been calculated, which is shown in figure 5.4.11. The figure shows that the variance on average goes down both for increasing level of smoothing and regularization. Furthermore, the variance is slightly higher on average for the two higher levels of regularization compared to the two higher levels of smoothing, which shows that a high agreement between the scans would not come from a very low variance.

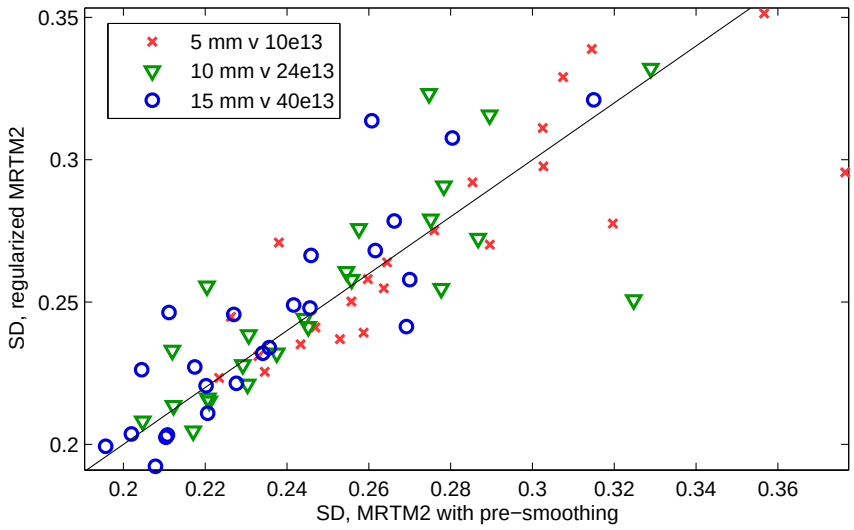

Figure 5.4.11. Standard deviation of all vertices of a subject. Level of smoothing plotted against level of regularization.

In figure 5.4.12, an estimation of the procentual bias between test and retest is shown for each subject. The absolute value of the bias was taken to be able to more clearly see which method gives a higher or lower bias. The measurement does not give any conclusive results, with an equal amount of bias on average. However, this measurement is not very suitable for a vertex-by-vertex estimation as the vertices can vary greatly and a procentual measurement can thus become misleading. An interesting observation can be made though, that the different levels of smoothing does not result in any significant change in bias while a change in regularization does change the bias, although not always in the same direction. Nevertheless, some systematic patterns seems to be present. A certain amount of bias is also expected with a month between the test and retest session, so it is hard to say which level of bias is appropriate.

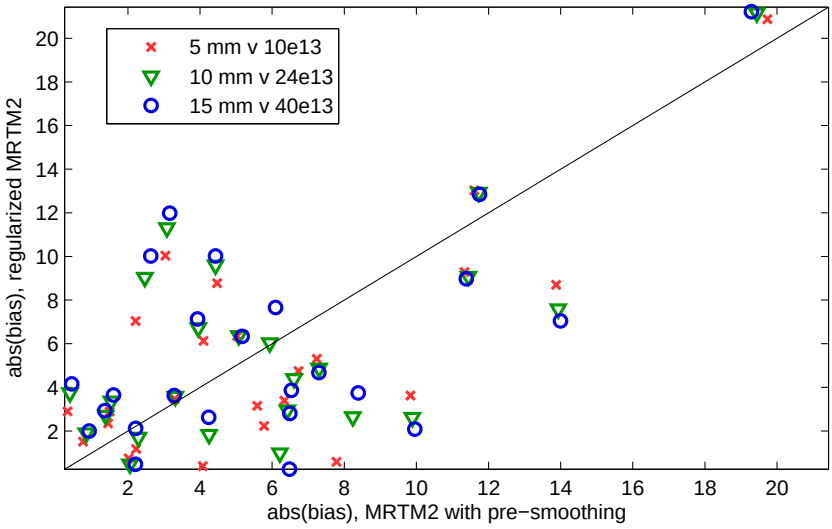

Figure 5.4.12. Absolute value of the procentual bias between test and retest scan for each subject. Levels of smoothing plotted against levels of regularization.

In figure 5.4.13, the ICC(2,1) is shown, measured on all vertices in the brain within each subject, with levels of smoothing plotted against levels of regularization. Only 5 subjects obtain a higher ICC with the smoothing approach, the rest have a higher ICC with the regularization approach. With lower levels of smoothing and regularization this trend is even more pronounced. This shows that the regularization MRTM2 is a repeatable method with good vertex-level precision.

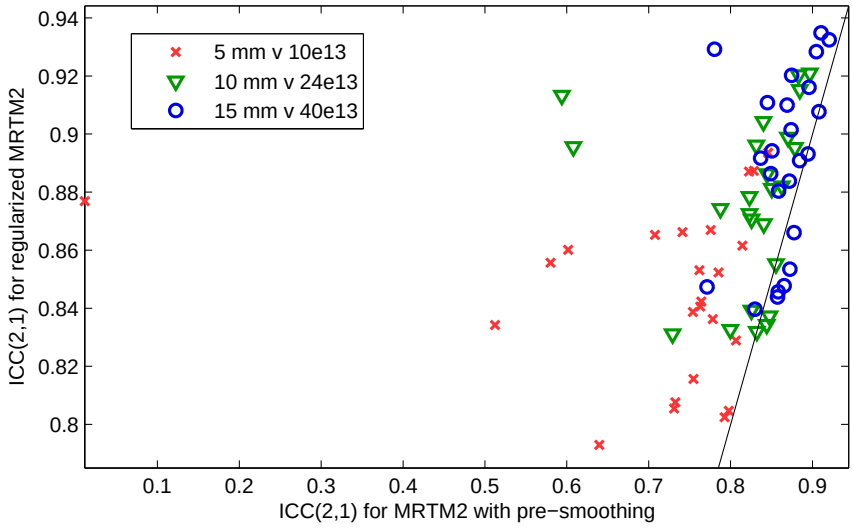

Figure 5.4.13. ICC(2,1) between test and retest scan for each subject. Levels of smoothing plotted against levels of regularization.

5.5. Regional repeatability analysis

The repeatability was also assessed on a regional level to be able to compare to the methods used in the initial analysis of the sex-hormone study. Six regions were chosen for this analysis on the basis of finding FreeSurfer regions that could be compared to regions used in the initial analysis with PMOD. The regions, except for neocortex, can be seen in figure 5.5.1. Brown represents the frontal cortex, lilac represents parietal cortex, green represents occipital cortex, red represents the superior temporal gyrus and yellow represents insula.

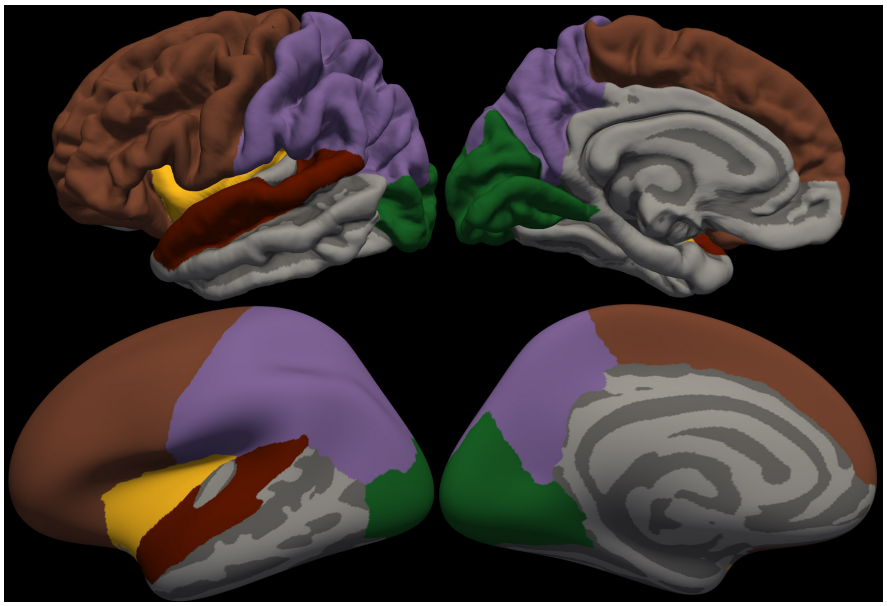

Figure 5.5.1. Regions used in regional analysis.

The ICC(2,1) was performed on the vertices of these regions, which can be seen in figure 5.5.2. It gives similar results as ICC(2,1) on all brain vertices, except for superior temporal gyrus where the ICC is equal for regularization and smoothing on average. However, the results of the other regions shows that the global results can be generalized to smaller regions.

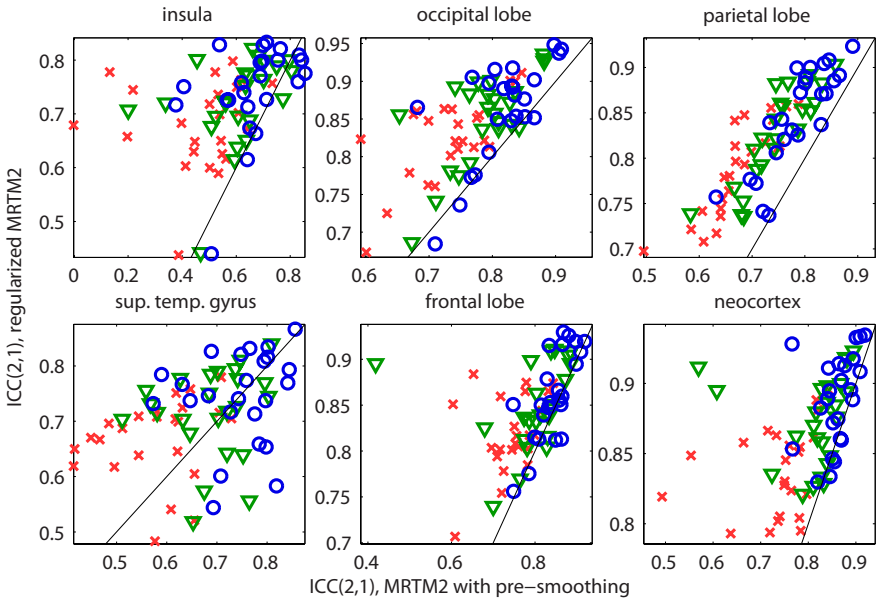

Figure 5.5.2. Vertexwise ICC(2,1) for different regions.

Regional data was obtained from the vertex-based approaches by averaging BP of all vertices within a region, to be able to compare the smoothing and regularization approach to results from averaged regional TACs. The ICC(3,1), measuring consistency across subjects on regional data, is shown in table 5.5.1. Both increased level of regularization and smoothing yielded a slightly lower but not significantly lower ICC. The ICC(3,1) is comparable across methods.

Table 5.5.1. ICC(3,1) on regional data measuring the consistency across subjects for two levels of regularization and smoothing compared to MRTM2 estimates on averaged regional TACs for FreeSurfer regions and in PMOD.

Region	$\lambda 10^{13}$	$\lambda 40^{13}$	5 mm	15 mm	ROI	PMOD
Insula	0.92	0.92	0.92	0.92	0.91	0.92
Occipital	0.93	0.91	0.92	0.91	0.91	0.91
Parietal	0.80	0.78	0.82	0.82	0.81	0.79
Sup.temporal	0.88	0.89	0.90	0.89	0.88	0.89
Frontal	0.87	0.86	0.88	0.88	0.87	0.86
Neocortex	0.91	0.89	0.91	0.90	0.89	0.89

In table 5.5.2 the results of mean procentual bias across subjects is shown, obtained from regional data. For most regions the regularization resulted in a slightly higher mean bias, but still showed a reasonably low bias.

Table 5.5.2. Mean of estimated procentual bias on regional data across subjects for two levels of regularization and smoothing compared to MRTM2 estimates on averaged regional TACs for FreeSurfer regions and in PMOD.

Region	$\lambda 10^{13}$	$\lambda 40^{13}$	5 mm	15 mm	ROI	PMOD
Insula	-0.53	-0.82	1.30	-0.22	-0.49	-0.47
Occipital	0.82	0.81	-0.20	-0.26	-0.26	-0.07
Parietal	1.90	1.78	1.30	1.23	1.33	1.34
Sup.temporal	0.31	0.30	-0.51	-0.61	-0.63	-0.90
Frontal	1.97	1.81	1.04	1.00	1.00	1.24
Neocortex	0.84	0.78	0.23	0.08	0.08	0.20

Also the variability, measured as the standard deviation of the procentual bias, showed comparable results across methods, as seen in table 5.5.3.

Table 5.5.3. Variability measured as the standard deviation of the procentual bias on regional data across subjects for two levels of regularization and smoothing compared to MRTM2 estimates on averaged regional TACs for FreeSurfer regions and in PMOD.

Region	$\lambda 10^{13}$	$\lambda 40^{13}$	5 mm	15 mm	ROI	PMOD
Insula	5.39	5.21	6.01	5.95	6.20	6.12
Occipital	8.65	9.68	8.55	8.67	8.85	8.58
Parietal	10.59	10.78	11.26	11.37	11.70	17.19
Sup.temporal	7.00	6.84	6.83	6.95	7.40	7.07
Frontal	8.82	9.14	8.88	9.00	9.27	9.67
Neocortex	6.71	7.23	7.10	7.33	7.73	8.28

The general linear model used for the sex-hormone study was fitted on neocortex for the different types of regional data, showed in figure 5.5.3.

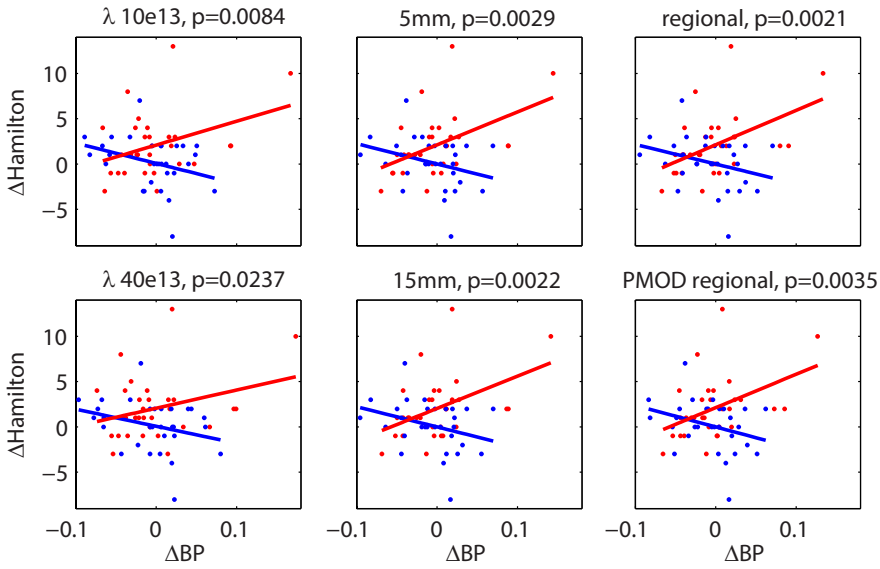

Figure 5.5.3. GLM fitted on neocortex for different types of regional data.

The regularized MRTM2 gave the least significant result. It can be seen that ΔBP for some subjects are rather different from the other methods.

5.6. Fully Bayesian approach

The developed model for a fully Bayesian treatment of the data was tried. The process with `fminsearch` was run for several different starting points, to be able to assess the models behavior and robustness against ending up in a local minima. The different starting points can be seen in table 5.6.1. A reasonable starting value for β was chosen as

$$\frac{1}{\text{var}(\mathbf{c})} = 3.36e-9.$$

A reasonable starting value for α_1 was chosen to a level that would be equivalent to a λ of 24e13 in the regularized MRTM2 approach, and a starting value for α_2 was chosen as a value considerably less than this. μ_1 and μ_2 got starting values by averaging the results of a MRTM2 on raw data. The starting points were then varied to different degrees, including only estimating on β, α_1 and α_2.

Table 5.6.1. Starting points for fminsearch

Starting points nr.	1	2	3	4	5
β	3.36e-9	3.36e-9	3.36e-9	1	3.36e-9
α_1	1e6	1e6	1e3	1	1e10
α_2	1e2	1e2	1e5	1	1e7
μ_1	0.09	-	0.09	1	0.09
μ_2	0.06	-	0.06	1	0.06

The first, second and third search went fastest with around 300 iterations and took around 1 hour to complete. The first search resulted in the lowest cost, although not significantly lower than the second and third search, while the last two searches resulted in a higher cost. The third and last search took a bit longer than 2 hours to complete. The resulting optimized hyperparameters can be seen in table 5.6.2.

Table 5.6.2. Resulting optimized hyperparameters and range of estimated BP

Starting points nr.	1	2	3	4	5
β	8.67e-9	8.67e-9	8.67e-9	8.64e-9	4.49e-9
α_1	1.30e4	1.30e4	1.30e4	3.02	1.83e10
α_2	3.38e-3	1.35e-4	1.40e-7	1.9	3.78e-4
μ_1	0.09	-	3.06	0.01	0.01
μ_2	0.05	-	1.57	0.16	3.34e-10
min BP	6.49	6.49	6.49	15.74	0.71
max BP	-0.58	-0.58	-0.58	-10.64	0.68

Unfortunately, none of the results gave a reasonable regularization of BP. In particular, α_2 was estimated so low that it did not significantly contribute to the result and subsequently did μ_1 and μ_2 not contribute significantly either. The varying of the starting points also showed the presence of local minima.

The obtained result for the first set of starting points is showed in figure 5.6.1, a result only slightly more regularized than estimating directly on the raw data.

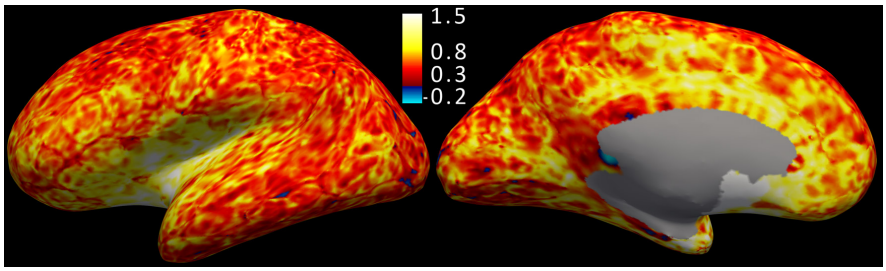

Figure 5.6.1. Result from fminsearch.

6. Discussion

The pipeline for a surface-based application of MRTM2 proved successful. The cortical segmentation enabled the smoothing of the data on the cortical layer and thus the structure of cerebral cortex could more accurately be taken into account compared to volume-based smoothing that was shown to introduce severe mixing of particularly non-cortical signal. However, many areas surrounding cerebral cortex were estimated to have negative BP, which is inaccurate. This is probably due to the use of a reference region with a very low but still present concentrations of serotonin transporters. Also some cortical areas were estimated to have negative BP without any segmentation issues involved and with a surface-based smoothing of FWHM 5 mm. One investigated area was in the occipital lobe, which does contain areas with a low concentration of serotonin transporters. It could be that some areas do have a natural occurring signal that is lower than the signal in cerebellum. This is problematic for the reference tissue approach and could also mean that other areas are estimated as lower than they truly are. This might not be a problem when equally lowered values are compared to each other, but still shows a significant flaw in the reference tissue approach. It also raises a question about how to treat negative BP. In this thesis it has been treated as a indicator of the presence of noise as they do stand out in the visualization of BP-maps. This treatment might not be totally accurate.

A problem with using a surface-based approach is the dependence on a well- segmented cortical representation. The FreeSurfer segmentation algorithm did result in some segmentation issues in most segmented subjects, even though it is a rather robust method. These segmentation errors are both time consuming to correct and sometimes hard to find. In this thesis, each segmentation was carefully reviewed multiple times by two persons, including a medical doctor with experience of neuroimaging data, and manually edited where deemed appropriate. Still, some errors where not found, even systematic errors like the shown error in isthmus cingulate. In this thesis, the lengthy editing process did not give any significant results in the study at hand in the statistical group analysis. However, it did change the outcome of the clusterwise correction for multiple comparisons due to a small change in p-value. This probably says more about the problem with thresholding p-maps than about the need for manual editing. It should be noted though, that the areas shown to be most affected by systematic segmentation errors where not part of the significant areas of the statistical group analysis, which means that conclusions about the need for manual editing are difficult to draw from this study. A paired t-test of the difference of volume between the edited and non-edited data set did show significance in some areas. However, many subjects had large errors in these areas, so it might be enough to look through the segmentation in search for these fewer and more easily noticeable large errors. It is of course always recommended to quickly look through a segmentation for large errors. In general, the FreeSurfer algorithm does a reasonably good job at segmenting the cortical layer.

The between-brain analysis of FreeSurfer involved some issues. The mapping of individual brains to a common space showed some registration issues that seemed to have quite significant effects on the outcome of the general linear model and had the magnitude to possibly change interpretations of the performed statistical analysis. These registration issues were worse in some particular areas, such as middle frontal gyrus, parietal cortex and in the sulcus between occipital cortex and other regions. The non-linear procedure of registering highly variable structures of individual brains to a common space will be bound to have registration issues. Often, the registration irregularities are accounted for by smoothing with a rather large Gaussian filter, which also was done in this thesis. However, the presented findings suggests that such a filtering procedure does not accurately account for these irregularities. Another related problem is the correction for multiple comparisons. There are many more methods for correcting for multiple comparison, but at least the two methods assessed in this thesis did both have their shortcomings, with a lack of power contrasted with a lack of detail, and a general problem of thresholding at a certain value. The whole procedure of smoothing the data equally in every direction to account for the problem with imperfect vertex-by-vertex correspondence between individual brain and then performing a test at each vertex independently and then correct for the many performed independent tests seems like a problematic pipeline. The need of developing a more sophisticated but still efficient procedure seems evident.

The developed regularized MRTM2 seems to handle high levels of noise better than the pre-smoothed vertexwise MRTM2, while retaining a higher resolution and detail, even at high amounts of regularization. It resulted in a higher repeatability between scans on a vertex level both at the whole cerebral cortex and at different regions of the cortical layer, without lowering the within scan variance. It showed better coherence between hemispheres and did not significantly depart from the other modeling approaches on a regional level. It did however slightly but consistently increase the average BP with increased regularization. One issue that would need further attention is the effect on the range of the parameters R_1 and k_2, which has not been fully examined in this thesis. Another issue is that it changed some ΔBP values in the general linear model so it resulted in a less significant result. It is hard to say what should be the true values, as the regularized MRTM2 treats the data in quit a different way. But this could suggest that some bias is introduced in the modeling and would need further examination.

The attempt to treat the data with MRTM2 in a fully Bayesian approach with optimization of the hyperparameters was not successful. The resulting optimal values did not result in an adequate regularization of the BP estimations and the risk of getting caught in a local minimum in the optimization procedure was also clearly present. However, the final optimization procedure together with the use of LDL-decomposition to solve the determinant of huge matrices present in the cost function was computationally rather efficient. Clearly more prior knowledge needs to be incorporated into the model. A possible prior knowledge is the bounds on BP. The presence of too high values and negative values are known to be implausible and this knowledge could thus be incorporated into the model. The problem with that approach is that the binding potential is not directly a part of the model.

However, the bounds on BP can be translated to a relationship between the two parameters in the model. As seen in figure 6.0.1, a negative BP is equivalent to the first parameter being smaller than the second parameter. A BP of above e.g. 10 is equivalent to the first parameter being larger than 11 times the second parameter. These two relationships create a bounded area of appropriate values. A change of coordinate system makes it possible to only regularize one parameter which could make it easier to incorporate this information into the model.

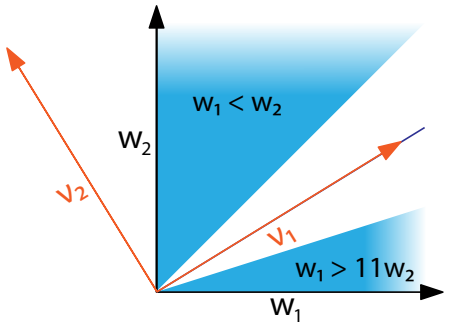

Figure 6.0.1. Bounds related to BP limits and change of coordinate system.

Furthermore, the parameter in the new coordinate system it could be limited by a function such as in figure 6.0.2.

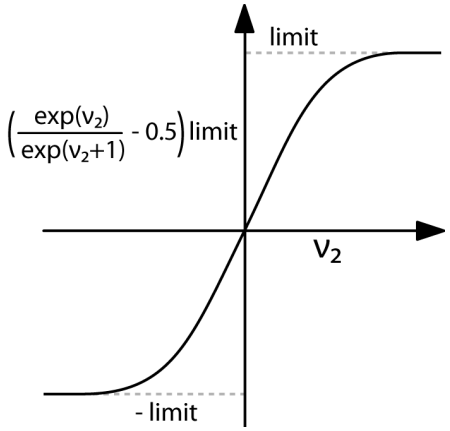

Figure 6.0.2. Possible way to limit the range.

7. Conclusions and Outlook

The aim of this project was to improve the filtering and modeling of PET data, to better handle the high noise level present in such data. It should lessen the artifacts of conventional volume-based spatial filtering with a Gaussian kernel. The focus was on the cerebral cortex due to its highly-folded and thin structure, which makes it particularly unsuited for the conventional approach to filtering.

For this purpose a pipeline for a surface-based approach to the modeling with the multilinear reference tissue method MRTM2 was developed. The surface representation of the cerebral cortex was obtained by the use of the software package FreeSurfer. The cortical representation enabled the smoothing of the data on the cortical layer and thus the structure of the cerebral cortex could more accurately be taken into account in the filtering procedure. In contrast with the conventional volume-based approach it proved successful with less edge artifacts. To assess the developed models, a baseline-follow-up study was used that images the sertonin transporter. Some problems where seen both in the volume-based approach and the surface-based approach with the presence of estimations of negative binding potential, which is improbable. This is probably due to the reference tissue approach with the use of cerebellum as a reference tissue, which does have some concentration of serotonin transporters.

Furthermore, the statistical tools for between-brain analysis of FreeSurfer was evaluated with the surface-based MRTM2. A group analysis was performed between the subjects of the study, with some apparent issues. The mapping of individual brains to a common space showed some registration issues that seemed to have quite significant effects on the outcome of the statistical group analysis and had the magnitude to possibly change interpretations of the performed analysis. These registration issues were worse in some particular areas, such as middle frontal gyrus, parietal cortex and in the sulcus between occipital cortex and other regions. Conventionally the discontinuity from a brain mapping to common space is accounted for by a large Gaussian filter previous to modeling in common space. However, the presented findings suggests that such a filtering procedure does not accurately account for the registration irregularities. Another related problem is the correction for multiple comparisons. Two methods were assessed with some problems observed such as a lack of power for one method contrasted with a lack of details for the other method, and a general problem of thresholding at a certain value.

A Bayesian framework was used to directly incorporate the data filtering into the mathematical model. The model was based on MRTM2 and assumed that close regions have similar parameters and regularized the data on this assumption. It was shown to handle high levels of noise better than the vertexwise surface-based approach with pre-smoothing, while at the same time retaining high resolution and detail, even at high amounts of regularization. In addition, it resulted in a higher re-

peatability on a vertex level between baseline and follow-up scan within a subject, both on the whole cerebral cortex and different regions of the cortical layer, without lowering the within scan variance. This was measured by the intraclass correlation coefficient with absolute consitency ICC(2,1). It showed better coherence between hemispheres and did not significantly depart from the other modeling approaches on a regional level. It did however slightly but consistently increase the average BP with increased regularization. When used on a general linear model of the study in the report, it showed less significance than the other methods, including estimation on regional averaged TACs and the vetexwise surface-based approach which all yielded rather similar significant results. It was not clear why it performed worse, if it was due to a higher specificity or just introduced bias.

Futhermore, an attempt was made to treat the data in a fully Bayesian approach by further developing the previous model and to including optimization over the hyperparameters of the model. The resulting optimal hyperparameters did not result in an adequate regularization of the BP estimations and the risk of getting caught in a local minimum in the optimization procedure was also clearly present. However, the final developed optimization procedure together with the use of LDL-decomposition to solve the determinant of huge matrices present in the cost function was computationally rather efficient. Clearly more prior knowledge needs to be incorporated into the model.

7.1. Outlook

The regularized MRTM2 using a Bayesian framework showed some nice results compared to the surface-based pre-smoothed model. However, it also showed some characteristics that could be further examined, e.g. the fact that the average binding potential went slightly up with increased regularization. To be able to use the model in a fully Bayesian approach, it needs further development to result in useful optimized results, which could include the incorporation of the prior knowledge of bounds on the binding potential.

The group analysis that was carried out showed some problems with registration irregularities when brains were mapped to the common space and issues with the need of having to correct for multiple comparison. A more sophisticated model for group analysis could be developed with the Bayesian framework that do not assume vertex-to-vertex correspondence and that can adapt to different levels of coherence between the subjects, for example by using the measure of convexity currently used for mapping between brains in FreeSurfer.

Bibliography

Bailey, D. L., Townsend, D. W., Valk, P. E., and Maisey, M. N., editors (2005). *Positron Emission Tomography - Basic Science*. Springer-Verlag.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Buchert, R., Wilke, F., van den Hoff, J., and Mester, J. (2003). Improved Statistical Power of the Multilinear Reference Tissue Approach to the Quantifiquation of Neuroreceptor Ligand Binding by Regularization. *Journal of Cerebral Blood Flow and Metabolism*, 23:612–620.

Bue Klein, A. (2010). *Brain-Derived Neurotrophic Factor (BDNF): Interactions with the serotonergic system and its potential as a biomarker in neurological and neuropsychiatric diseases*. PhD thesis, Faculty of health science, University of Copenhagen.

Chan, C., Fulton, R., Feng, D., and Meikle, S. (2009). Regularized image reconstruction with an anatomically adaptive prior for positron emission tomography. *Phys. Med. Biol.*, 54:7379–7400.

Charnay, Y. and Léger, L. (2010). Brain serotonergic circuitries. *Dialogues in clinical neuroscience*, 12:471–487.

Comtat, C., Sureau, F. C., Sibomana, M., Hong, I. K., Sjöholm, N., and Trébossen, R. (2008). Image based resolution modeling for the HRRT OSEM reconstructions software. *IEEE Nuclear Science Symp. Conf. Rec.*, pages 4120–4123.

Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, 19(1):15–18.

Dale, A., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179 – 194.

Daws, L. C. and Gould, G. G. (2011). Ontogeny and regulation of the serotonin transporter: Providing insights into human disorders. *Pharmacology & Therapeutics*, 131:61–79.

Fang, Y.-H. D., El Fakhri, G., Becker, J. A., and Alpert, N. (2012). Parametric imaging with Bayesian prior: A validation study with 11-Altropane PET. *Neuroimaging*, 61:131–138.

Fischl, B. and Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20):11050–11055.

Fischl, B., Liu, A., and Dale, A. M. (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Medical Imaging*, 20(1):70–80.

Fischl, B., Sereno, M. I., and Dale, A. (1999a). Cortical surface-based analysis: Ii: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195 – 207.

Fischl, B., Sereno, M. I., Tootell, R. B., and Dale, A. M. (1999b). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., and Dale, A. M. (2004). Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, 14(1):11–22.

Freeman, E. W., Sammel, M. D., Lin, H., and Nelson, D. B. (2006). Associations of hormones and menopausal status with depressed mood in women with no history of depression. *Arch Gen Psychiatry*, 63(4):375–82.

Friston, K., Chu, C., Miranda, J. M., Hulme, O., Rees, G., Penny, W., and Ashburner, J. (2008). Bayesian decoding of brain images. *NeuroImage*, 39:181 – 205.

Frøkjær, V. G., Erritzoe, D., Juul, A., Årup Nielsen, F., Holst, K., Svarer, C., Madsen, J., Paulson, O. B., and Knudsen, G. M. (2010). Endogenous plasma estradiol in healthy men is positively correlated with cerebral cortical serotonin 2A receptor binding. *Psychoneuroendocrinology*, 35:1311—1320.

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of Statistical Maps in Functional Neuroimaging Using False Discovery Rate. *NeuroImage*, 15:870–878.

Ginovart, N., Wilson, A. A., Meyer, J. H., Hussey, D., and Houle, S. (2001). Positron Emission Tomography Quantification of [^{11}C]-DASB Binding to the Human Serotonin Transporter: Modeling Strategies. *Journal of Cerebral Blood Flow and Metabolism*, 21:1342–1353.

Greve, D. N. and Fischl, B. (2009). Accurate and robust image alignment using boundary-based registration. *NeuroImage*, 48:63–72.

Hagler, D. J., Pinar Saygin, A., and Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fmri data. *NeuroImage*, doi:10.1016/j.neuroimage.2006.07.036.

Hanson, L. G. (2009). Introduction to Magnetic Resonance Imaging Techniques. Technical report, Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre.

Hofheinz1, F., Langner, J., Beuthien-Baumann, B., Oehme, L., Steinbach, J., Kotzerke, J., and van den Hoff, J. (2011). Suitability of bilateral filtering for edge-preserving noise reduction in PET. *EJNMMI Research*, 23(1).

Hong, I. K., Chung, S. T., Kim, H. K., Kim, Y. B., Son, Y. D., and Cho, Z. H. (2007). Ultra fast symmetry and SIMD-based projection-backprojection (SSP) algorithm for 3-D PET image reconstruction. *IEEE Trans. Med. Imag.*, 26:789–803.

Houle, S., Ginovart, N., Hussey, D., Meyer, J. H., and Wilson, A. A. (2000). Imaging the serotonin transporter with positron emission tomography: initial human studies with [^{11}C]DAPP and [^{11}C]DASB. *European Journal of Nuclear Medicine and Molecular Imaging*, 27:1719–1722.

Ichise, M., Liow, J.-S., Lu, J.-Q., Takano, A., Model, K., Toyama, H., Suhara, T., Suzuki, K., Innis, R. B., and Carson, R. E. (2003). Linearized Reference Tissue Parametric Imaging Methods: Application to [^{11}C]DASB Positron Emission Tomography Studies of the Serotonin Transporter in Human Brain. *Journal of Cerebral Blood Flow and Metabolism*, 23:1096–1112.

Innis, R. B. (2007). In vivo PET imaging understood from the perspective of in vitro receptor binding. In Maguire, R. P. and Leenders, K. L., editors, *PET pharmacokinetic course manual*, chapter 7. International society of cerebral blood flow and metabolism.

JCGM (2012). *International vocabulary of metrology - Basic and general concepts and associated terms*. 3 edition.

Kim, J. S., Ichise, M., Sangare, J., and Innis, R. B. (2006). PET Imaging of Serotonin Transporters with [^{11}C]DASB: Test-Retest Reproducibility Using a Multilinear Reference Tissue Parametric Imaging Method. *Journal of Nuclear Medicine*, 47(2):208–214.

Kish, S. J., Furukawa, Y., Chang, L.-J., Tong, J., Ginovart, N., Wilson, A., Houle, S., and Meyer, J. H. (2005). Regional distribution of serotonin transporter protein in postmortem human brain - Is the cerebellum a SERT-free brain region? *Nuclear Medicine and Biology*, 32:123–128.

Lagarias, J., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization*, 9(1):112–147.

Lammertsma, A. A. (2007). Receptor kinetics - modelling and practical approach. In Maguire, R. P. and Leenders, K. L., editors, *PET pharmacokinetic course manual*, chapter 8. International society of cerebral blood flow and metabolism.

Lin, J., Laine, F., and Bergmann, S. (2001). Improving PET-based physiological quantification through methods of wavelet denoising. *IEEE Trans. Biomed. Eng.*, 48:202–212.

Maguire, R. P. (2007). Cerebral blood flow - single-tissue-compartment model. In Maguire, R. P. and Leenders, K. L., editors, *PET pharmacokinetic course manual*, chapter 4. International society of cerebral blood flow and metabolism.

Munk-Olsen, T., Munk Laursen, T., Bøcker Pedersen, C., Mors, O., and Bo Mortensen, P. (2006). New Parents and Mental Disorders: A Population-Based Register Study. *JAMA*, 296(21):2582–2589.

Peng, J.-Y., Aston, J. A. D., Gunn, R. N., Liou, C.-Y., and Ashburner, J. (2008). Dynamic Positron Emission Tomography Data-Driven Analysis Using Sparse Bayesian Learning. *IEEE Transactions on Medical Imaging*, 27(9):1356–1369.

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*.

Raemaekers, M., Vink, M., Zandbelt, B., van Wezel, R., Kahn, R., and Ramsey, N. (2007). Test–retest reliability of fMRI activation during prosaccades and antisaccades. *NeuroImage*, 36:532–542.

Reimold, M., Slifstein, M., Heinz, A., Müller-Schauenburg, W., and Bares, R. (2006). Effect of spatial smoothing on t-maps: arguments for going back from t-maps to masked contrast images. *Journal of Cerebral Blood Flow and Metabolism*, 26:751–759.

Rizzo, G., Turkheimer, F., Keihaninejad, S., Bose, S., Hammers, A., and Bertoldo, A. (2012). Multi-Scale hierarchical generation of PET parametric maps: Application and testing on a [11C]DPN study. *NeuroImage*, 59:2485–2493.

Sabuncu, M. R. and Van Leemput, K. (2013). The Relevance Voxel Machine (RVoxM): A Self-tuning Bayesian Model for Informative Image-based Prediction. *IEEE Trans Med Imaging*, [E-publication ahead of print].

Sacchet, M. D. and Knutson, B. (2013). Spatial smoothing systematically biases the localization of reward-related brain activity. *NeuroImage*, 66:270–277.

Saulin, A., Savli, M., and Lanzenberger, R. (2012). Serotonin and molecular neuroimaging in humans using PET. *Amino Acids*, 42:2039–2057.

Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., and Fischl, B. (2004). A hybrid approach to the skull stripping problem in mri. *NeuroImage*, 22(3):1060 – 1075.

Segonne, F., Pacheco, J., and Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Trans Med Imaging*, 26:518–529.

Shrout, P. E. and Fleiss, J. L. (1979). Intra Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86(2):420–428.

Sled, J. G., Zijdenbos, A. P., and Evans, A. C. (1998). A non-parametric method for automatic correction of intensity non-uniformity in mri data. *IEEE Transactions on Medical Imaging*, 17:87 – 97.

Smith, J. (august 2013). FreeSurferWiki - mri-glmfit. http://surfer.nmr.mgh.harvard.edu/fswiki/glm_fit.

Sureau, F. C., Reader, A. J., Comtat, C., Leroy, C., Ribeiro, M.-J., Buvat, I., and Trébossen, R. (2008). Impact of Image-Space Resolution Modeling for Studies with the High-Resolution Research Tomograph. *The Journal of Nuclear Medicine*, 49:1000–1008.

Svarer, C., Madsen, K., Hasselbalch, S. G., Pinborg, L. H., Haugbøl, S., Frøkjær, V. G., Holm, S., Paulson, O. B., and Knudsen, G. M. (2005). MR-based automatic delineation of volumes of interest in human brain PET images using probability maps. *NeuroImage*, 24:969–979.

Tauber, C., Stute, S., Chau, M., Spiteri, P., Chalon, S., Guilloteau, D., and Buvat, I. (2011). Spatio-temporal diffusion of dynamic PET images. *Phys. Med. Biol.*, 56:6583–6596.

Tomasi, G., Bertoldo, A., Cobelli, C., Pavese, N., Tai, Y. F., Hammers, A., and Turkheimer, F. E. (2011). Global-two-stage filtering of clinical PET parametric maps: Application to [11C]-(R)-PK11195. *Neuroimaging*, 55:942–953.

Turkheimer, F. E., Aston, J. A. D., Asselin, M.-C., and Hinz, R. (2006). Multi-resolution Bayesian regression in PET dynamic studies using wavelets. *NeuroImage*, 32:111–121.

Turkheimer, F. E., Boussion, N., Anderson, A. N., Pavese, N., Piccini, P., and Visvikis, D. (2008). PET Image Denoising Using a Synergistic Multiresolution Analysis of Structural (MRI/CT) and Functional Datasets. *J Nucl Med*, 49:657–666.

Van den Hoff, J. (2007). Linearisations. In Maguire, R. P. and Leenders, K. L., editors, *PET pharmacokinetic course manual*, chapter 6. International society of cerebral blood flow and metabolism.

Wang, G. and Qi, J. (2012). An Optimization Transfer Algorithm for Nonlinear Paramteric Image Reconstruction From Dynamic PET Data. *IEEE Transactions on medical imaging*, 31:1977–1987.

Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., , and Mazziotta, J. C. (1998). Automated image registration: I. General methods and intrasubject, intramodality validation. *J Comput Assist Tomogr*, 22(1):139–152.

Zhou, Y., Endres, C., Brasic, J. R., and Huang Sung-Cheng, D. F. W. (2002). Linear regression with spatial constraint to generate parametric images of lingand-receptor dynamic PET studies with simiplified reference tissue model. *Neuroimaging*, 62:774–781.

A. MATLAB function for vertex-by-vertex MRTM2

```
function [BP,R1,k2] = MRTM2(C,refC,refk2,timestep)
% R1 and BP_nd estimated with MRTM2.
% INPUT
% C : Matrix of tissue tracer concentrations, col: time, row: sample
% refC : Vector of reference tissue tracer concentrations
% refk2 : Tracer clearance rate constant of reference tissue
% timestep : vector of time steps between frames
% OUTPUT
% BP : Binding potential
% R1 : Relative tracer delivery

nSamp = length(C);
nTime = length(timestep);
BP = zeros(1,nSamp);
R1 = zeros(1,nSamp);
k2 = zeros(1,nSamp);

% Cumulative trapezoidal numerical integration
cumintC = cumtrapz(timestep,C,1);
cumintRef = cumtrapz(timestep,refC);

disp('start')
% MRTM2
for v=1:nSamp
    if ~mod(v,1000)
        fprintf('\b. ');
    end
    if cumintC(nTime,v)>0
        X = [cumintRef + (1/refk2)*refC, cumintC(:,v)];
        param = X\C(:,v);
        BP(v) = -((param(1)/param(2)) + 1);
        k2(v) = -param(2);
        R1(v) = param(1)/refk2;
    else
        BP(v) = NaN;
        k2(v) = NaN;
        R1(v) = NaN;
    end
end
```

B. MATLAB function for regularized MRTM2

```
function [BP,R1,k2] = rMRTM2(C,refC,refk2,timestep,nbrs,lambda)
%Regularized MRTM2. N : number of vertices,
%                   K : number of time steps.
%----Input-------------------------
% C : Matrix of tissue tracer concentrations, size(C) = [N,K]
% refC : Reference tissue TAC, size(refC)=[1,K]
% refk2 : Reference tissue clearance rate constant,
%   estimated by MRTM
% timestep : Time steps between frames,size(timestep) = [1,K]
% nbrs : Each row contains the indices to a vertex's
%   neighbors, NaN at remaining elements if not max number
%   of neighbors. size(nbrs) = [N,(max number of neighbors)]
% lambda : Regularization parameter, e.g. 10e13.
%----Output------------------------
% BP : Binding potential
% R1 : Relative tracer delivery
% k2 : Tracer clearance rate constant

N = size(C,1);
cumintC = cumtrapz(timestep,C,2);
cumintRef = cumtrapz(timestep,refC,2);
X1 = repmat(cumintRef + (1/refk2)*refC,N,1);
X2 = -cumintC;

%Computes regularization matrix G
ind = ~isnan(nbrs);
vert = repmat((1:N)',1,size(nbrs,2));
diagG = sum(ind,2);
diagG = [diagG; diagG];
offdiagG = -ones(sum(diagG),1);
gm = [1:2*N, vert(ind)', vert(ind)'+N];
gn = [1:2*N, nbrs(ind)', nbrs(ind)'+N];
G = sparse(gm, gn, [diagG; offdiagG], 2*N, 2*N);

%Computes X'*X and X'*c
diagX2X1 = repmat(sum(X2.*X1,2),2,1);
X = [X1;X2];
diagXX = sum(X.^2,2);
dm = [1:2*N, 1:2*N];
dn = [1:2*N, N+1:2*N, 1:N];
XX = sparse(dm, dn, [diagXX; diagX2X1], 2*N, 2*N);
Xc = sum([C;C].*X,2);
```

```
%Computes W
S = XX + lambda*G;
W = S\Xc;
R1 = W(1:N)/refk2;
k2 = W(N+1:end);

%Computes BP
BP = R1*refk2./k2 - 1;
```

C. MATLAB function for Bayesian MRTM2 using fminsearch

```
function [BP,R1,k2,fval,exitflag,output,hp] = ...
    bMRTM2(C,refC,refk2,timestep,nbrs,hp0)
%Bayesian MRTM2 using fminsearch. N : number of vertices,
%                                 K : number of time steps.
%----Input-------------------------
% C : Matrix of tissue tracer concentrations, size(C) = [N,K]
% refC : Vector of reference tissue tracer concentrations, size(refC)=[1,K]
% refk2 : Tracer clearance rate constant of reference tissue
% timestep : Vector of time steps between frames, size(timestep) = [1,K]
% nbrs : Each row contains the indices to a vertex's neighbors, NaN at
%        remaining elements if not max number of neighbors.
%        size(nbrs) = [N,(max number of neighbors)]
% hp0 :  Starting point for hyperparameters,
%  i.e. [beta,alpha1,alpha2,mu1,mu2] = [1\var(C(:)),1e5,1e2,0.09,0.06]
%
%----Output-------------------------
% BP : Binding potential
% R1 : Relative tracer delivery
% k2 : Tracer clearance rate constant
% other outputs from the fminsearch procedure

N = size(C,1);
K = size(C,2);

cumintC = cumtrapz(timestep,C,2);
cumintRef = cumtrapz(timestep,refC,2);
X1 = repmat(cumintRef + (1/refk2)*refC,N,1);
X2 = -cumintC;

%Computes regularization matrices G and F
ind = ~isnan(nbrs);
vert = repmat((1:N)',1,size(nbrs,2));
diagG = sum(ind,2);
diagG = [diagG; diagG];
offdiagG = -ones(sum(diagG),1);
gm = [1:2*N, vert(ind)', vert(ind)'+N];
gn = [1:2*N, nbrs(ind)', nbrs(ind)'+N];
G = sparse(gm, gn, [diagG; offdiagG], 2*N, 2*N);
F = sparse(gm, gn, [diagG; -offdiagG], 2*N, 2*N);

%Computes X'*X and X'*c
```

```
diagX2X1 = repmat(sum(X2.*X1,2),2,1);
X = [X1;X2];
diagXX = sum(X.^2,2);
dm = [1:2*N, 1:2*N];
dn = [1:2*N, N+1:2*N, 1:N];
XX = sparse(dm, dn, [diagXX; diagX2X1], 2*N, 2*N);
Xc = sum([C;C].*X,2);

%Searches for optimal hyperparameters
options = optimset('PlotFcns',@optimplotx);
[hp,fval,exitflag,output] ...
    = fminsearch(@costfun,hp0,options,G,F,XX,C,X,N,K);

%Estimates MRTM2 parameters
if length(hp)==3
    hp(4) = 0;
    hp(5) = 0;
elseif length(hp)==4
    hp(5) = hp(4);
end

Hmu = sparse(N*2,1);
H = hp(2)*G + hp(3)*F;
Hmu(1:N) = sum(hp(4)*H(1:N,1:N),2);
Hmu(N+1:end) = sum(hp(5)*H(N+1:end,N+1:end),2);
S = (hp(1)*XX + H);
W = S\(hp(1)*Xc+Hmu);
W = [W(1:N),W(N+1:end)]';
R1 = W(1,:)/refk2;
k2 = W(2,:);

%Computes BP
BP =  W(1,:)./W(2,:) - 1;
```

D. Cost function for fminsearch

```
function cost = costfun(hp,G,F,XX,C,X,N,K)
%costfunction to be used with fminsearch and bMRTM2 to find optimal
%hyperparameters

%Checks if any hyperparameter is negative, if so return NaN
if sum(hp<0)>0
    cost = NaN;
    return
end

%Orders hyperparameters
beta = hp(1);
alpha1 = hp(2);
alpha2 = hp(3);
if length(hp)==3
    mu1 = 0;
    mu2 = 0;
elseif length(hp)==4
    mu1 = hp(4);
    mu2 = mu1;
else
    mu1 = hp(4);
    mu2 = hp(5);
end

%Computes determinant of H
H = alpha1*G + alpha2*F;
[L,D,~] = ldl(H);
diagLD = diag(L*D);
if sum(diagLD < 0) > 0
    logdetH = NaN;
else
    logdetH = 2*sum(log(sqrt(diagLD)));
end

%Computes determinant of S
S = beta*XX + H;
[L,D,P] = ldl(S);
diagLD = diag(L*D);
if sum(diagLD < 0) > 0
    logdetS = NaN;
else
    logdetS = 2*sum(log(sqrt(diagLD)));
end

%Computes cost function
```

```
CXmu = C - X(1:N,:)*mu1 - X(N+1:end,:)*mu2;
cXmu2 = sum(sum(CXmu.^2));
cXmuX = sum([CXmu;CXmu].*X,2);

cost = .5*( N*K*log(2*pi) - N*K*log(beta) + logdetS - logdetH + ...
    beta*cXmu2 - beta^2*cXmuX'*(P * (L' \ (D \ (L \ (P' * cXmuX))))) );
```