



UNIVERSITY OF CLAUDE BERNARD LYON 1  
Polytech Lyon

---

**Estimation of Image-Derived Arterial Input Function in  
Brain PET Imaging:  
Application to Modeling PET Dynamics of Glucose  
Metabolism in Patients with Impaired Consciousness**

---

By SEPAND ALI MADAD SOLTANI

Master 2 in Medical Device Engineering  
Internship Report

Supervised by INÉS MÉRIDA and NICOLAS COSTES

Academic Advisor: KEVIN TSE VE KOON

2024-2025



# Abstract

Positron Emission Tomography (PET) combined with kinetic modeling enables quantitative assessment of physiological processes through the use of an arterial input function (AIF). Traditional AIF acquisition via arterial sampling is invasive and challenging, particularly in comatose patients. This study aimed to enhance a non-invasive Bayesian Geometric Transfer Matrix (BGTM) method for estimating image-derived input function (IDIF) using Time-of-Flight Magnetic Resonance Angiography (TOF-MRA) acquired simultaneously with dynamic PET on a hybrid PET/MR scanner. The proposed method integrates TOF-MRA-guided carotid segmentation with a Bayesian framework incorporating population-based AIF priors to address partial volume effects. Evaluated on a cohort of 56 comatose patients, the refined segmentation pipeline, incorporating a cuboid mask to exclude non-carotid tissues, improved robustness in challenging cases. The BGTM approach demonstrated a 57% reduction in mean absolute error in the cumulative Area Under the Curve (cAUC) errors (14,202 vs. 33,764 vs,  $p < 0.0001$ ) and a significant reduction (14.1% vs. 33.0%,  $p < 0.0001$ ) in mean absolute percentage error (MAPE) for glucose metabolic rate ( $MR_{glu}$ ) quantification compared to conventional GTM method. However, high variability in  $MR_{glu}$  and cAUC errors (71% and 65% coefficient of variation, respectively) highlights inconsistencies in the method. These results position BGTM as a promising pathway toward reducing reliance on invasive arterial sampling, though addressing residual variability remains critical for future clinical applications.

**Keywords:** Dynamic FDG-PET, Image-Derived Input Function, Hybrid PET/MRI, Bayesian Framework

# Contents

<b>Abstract</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Positron Emission Tomography . . . . .	1
1.2 Kinetic Modeling . . . . .	2
1.3 Input Function . . . . .	3
1.3.1 Arterial Input Function . . . . .	3
1.3.2 Population-Based Input Function . . . . .	3
1.3.3 Image Derived Input Function . . . . .	3
1.4 Partial Volume Effect . . . . .	4
1.5 Background . . . . .	4
<b>2 Materials and Methods</b>	<b>6</b>
2.1 Dataset Description . . . . .	6
2.2 Pre-processing . . . . .	6
2.3 Carotid Segmentation . . . . .	7
2.4 Partial Volume Correction . . . . .	8
2.4.1 Geometric Transfer Matrix . . . . .	8
2.4.2 Bayesian Geometric Transfer Matrix . . . . .	9
2.5 Evaluation . . . . .	12
2.5.1 IF Curves . . . . .	12
2.5.2 Quantification . . . . .	13
2.6 Simulation . . . . .	14
<b>3 Results</b>	<b>16</b>
3.1 Carotid Segmentation from MRA-TOF . . . . .	16
3.2 Hyperparameter Tuning . . . . .	16
3.3 Simulation . . . . .	17
3.4 IDIF . . . . .	17
<b>4 Discussion</b>	<b>20</b>
<b>5 Conclusion</b>	<b>21</b>
<b>References</b>	<b>22</b>

# Introduction

## 1.1 Positron Emission Tomography

Positron Emission Tomography (PET) is an in vivo functional imaging technique widely used in clinical and research settings to monitor physiological and biochemical processes. In PET, a biologically active molecule is labeled with a positron-emitting radioisotope, serving as a radiotracer, and then injected into the body. As the radiotracer accumulates in target tissues, its radioactive decay produces positrons, which interact with electrons to emit pairs of gamma photons in nearly opposite directions. These photons are detected by the PET scanner, and image reconstruction algorithms generate a three-dimensional representation of the tracer distribution. This imaging modality allows for the investigation of metabolic changes, receptor binding, and other biochemical processes, providing invaluable information in oncology, neurology, cardiology, and other fields.

There are two main categories in PET image acquisition: static imaging and dynamic imaging. Static PET involves acquiring a single scan after the radiotracer injection. This single snapshot offers a powerful yet simplified view of tracer distribution. The common quantification metric in static imaging is the Standardized Uptake Value (SUV), which normalizes tissue uptake by the injected dose and weight of the subject, allowing for a semi-quantitative comparison of tracer accumulation across different tissues or over time [1]. Due to its simplicity, static PET is widely used in clinical settings; however, it also has limitations. Because it reflects only one time point, the SUV cannot capture the temporal dynamics of tracer uptake and clearance, and various physiological factors may influence its measurements, thereby reducing its accuracy.

Dynamic PET imaging provides a more comprehensive view of radiotracer kinetics by acquiring a series of images over a period ranging from a few minutes to more than an hour post-injection, depending on the tracer type. Instead of a single snapshot, dynamic imaging produces time-activity curves (TACs) that illustrate how tracer concentration in each tissue changes throughout the scanning period. This approach enables the measurement of physiological parameters such as the tracer rate of influx ( $K_i$ ) for radiotracers with irreversible uptake (e.g. [ $^{18}\text{F}$ ]FDG), volume of distribution ( $V_T$ ), and the rates of phosphorylation and dephosphorylation.

## 1.2 Kinetic Modeling

To quantify pharmacokinetic parameters, kinetic modeling is employed. Compartmental modeling is the most popular and is considered the most accurate approach in kinetic modeling. In compartmental modeling, the distribution and kinetics of a radiotracer are described by dividing the system into distinct compartments, each representing a pool of tracer that behaves uniformly. Interactions between compartments can be unidirectional or bidirectional, meaning the tracer may either move in and out or only enter a compartment. Various graphical models (e.g., the Logan [2] and Patlak [3] methods), as well as classical compartmental model fitting approaches, are used to analyze tracer kinetics.

Figure 1.1 shows the two-tissue compartment model (2TCM), also known as the three-compartment model, in series mode. This model comprises one tissue compartment for the free tracer,  $C_F(t)$ , and another for the receptor-bound tracer,  $C_B(t)$ , in addition to an external compartment representing the tracer concentration in the plasma or blood, denoted as the input function  $C_P(t)$ .

The tracer kinetics are governed by a series of first-order differential equations, in which the exchange rates between the compartments are considered constant:

$$\frac{dC_F(t)}{dt} = K_1 C_P(t) - (k_2 + k_3)C_F(t) + k_4 C_B(t), \quad (1.1)$$

$$\frac{dC_B(t)}{dt} = k_3 C_F(t) - k_4 C_B(t), \quad (1.2)$$

where  $K_1$ ,  $k_2$ ,  $k_3$ , and  $k_4$  are the constant rate parameters.

The total radiotracer tissular kinetic measured by PET (the PET data),  $C_T(t)$ , is given by

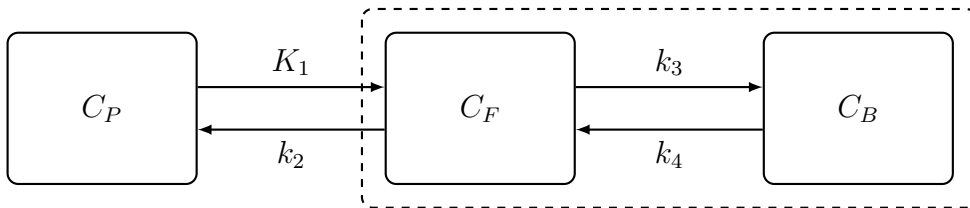
$$C_T(t) = C_F(t) + C_B(t) + C_P(t). \quad (1.3)$$

Thus to solve this system of equations and to estimate  $K_1$ ,  $k_2$ , and  $k_3$  parameters, we must fit the model using the measured PET TACs ( $C_T$ ) and the input function ( $C_P$ ).

For [ $^{18}\text{F}$ ]FDG quantification, the metabolic rate of glucose ( $\text{MR}_{\text{glu}}$ ) is calculated as

$$\text{MR}_{\text{glu}} (\mu\text{mol}/\text{min}/100\text{g}) = \frac{[C]}{LC} \cdot \frac{K_1 \times k_3}{k_2 + k_3}. \quad (1.4)$$

where  $[C]$  denotes the glucose concentration, and  $LC$  is the lumped constant.



**Figure 1.1:** Schematic of the two-tissue compartment model (2TCM)

For [ $^{11}\text{C}$ ]Yohimbine quantification, we utilize the Volume of Distribution ( $V_T$ ) which is the ratio of radiotracer concentration in the target tissue ( $C_T$ ) to the plasma ( $C_P$ ):

$$V_T = \frac{C_T}{C_P}$$

Using the Logan plot method this can be directly estimated from these two values or to be fitted to a compartment model. In the latter case,  $V_T$  can be calculated as

$$V_T = \frac{K_1}{k_2} \left(1 + \frac{k_3}{k_4}\right)$$

## 1.3 Input Function

### 1.3.1 Arterial Input Function

The arterial input function (AIF) is considered the gold standard for obtaining the input function. It is determined by inserting an arterial catheter into the patient and continuously drawing blood samples to measure the radiotracer concentration, thereby obtaining the blood activity curve used in the quantification model. However, this procedure is invasive and can cause discomfort, potentially discouraging patients from undergoing future examinations. Furthermore, this method is labor-intensive and requires trained personnel to manage both the subject and the measurement devices.

### 1.3.2 Population-Based Input Function

Population-Based Input Function (PBIF) is a method for replacing the subject specific AIF with the average AIFs of a population of other subjects.

### 1.3.3 Image Derived Input Function

The image-derived input function (IDIF) has been proposed as a non-invasive alternative for obtaining the input function. IDIF techniques typically involve identifying vascular structures or regions with high blood pool activity within the imaging field and extracting the input function directly from the PET images. For example, in whole-body, cardiac or small animal PET studies the aorta is visible in the in the Field of View (FOV) of the PET camera and is used as the source for IDIF. In brain PET imaging, the Internal Carotid Arteries (ICA) are the largest vessels present in the FOV and have a diameter of approximately 5 mm much smaller than of the aorta. Their smaller size makes IDIF more challenging due to Partial Volume Effects

## 1.4 Partial Volume Effect

Partial Volume Effect (PVE) is

## 1.5 Background

Many methods have utilized the first few frames of the dynamic PET which the tracer has not reached tissues yet and mainly exists in the arteries, to obtain the mask of the carotids [4]. These implementation vary from completely manual [TODO], semi-automatic by just providing a few seed voxels for a region growing algorithm [TODO] or for the center of a dilation, and completely automatic by using sophisticated deep learning methods such as custom Convolutional Neural Networks (CNN) [TODO] and U-NET [TODO].

However even with the sophisticated approaches, ICA segmentation directly from PET still suffer from the strong PVE present. With the emergence of hybrid PET/MRI machines, it has become feasible to acquire both functional and anatomical data simultaneously. MRI provides high-resolution soft tissue contrast, while PET captures metabolic activity. For instance, time-of-flight MR angiography (TOF-MRA) delivers excellent arterial contrast. Many methods have taken advantage of this fact to deliver simple yet very accurate segmentation of the ICA by manual or semi-automatic procedures usually involving high intensity thresholding or seeded or not-seeded region growing algorithms [TODO].

Unlike T1-weighted MRI, TOF-MRA is not usually available in studies thus TODO [TODO] proposed a deep learning model to extract the ICA from the T1-weighted images. [TODO] in the caliPER software enabled the semi-automatic extraction of the arteries from both TOF-MRA as well as T1-weighted images.

However, as previously mentioned, even with a high-resolution anatomical mask of the arteries, the issue of PVE still exists and must be accounted for. Many Partial Volume Correction (PVC) methods involve in deriving the spill-in and spill-out coefficient of the ICA and the surrounding tissue. Recovery Coefficient (RC) is the most popular method of estimating the spill-in and spill-out coefficient [TODO, TODO, TODO, TODO]. They are computed by scanning cylindrical phantoms of different diameters and matching the diameter of the ICA to them [TODO]. Geometric Transfer Matrix (GTM) is a more general method adapted to any type of volume by estimating spillage coefficients by based on the geometry of the VOI and modelling the PSF of the PET[TODO]. However, these methods do not account for the apparent noise in the PET and thus they fail to fully recover the lost signal due to noise.

More complex methods have been proposed such as model-based matrix factorization to estimate the input function and surrounding tissue activity [TODO] and deep learning approaches [TODO, TODO]

In this work, we propose a fully automatic pipeline for extracting the ICA mask from the TOF-MRA. We utilize a Bayesian framework to model the input function and the surrounding tissue based on their interaction according to GTM by incorporating prior knowledge



---

and also accounting for noise [5].

# Materials and Methods

## 2.1 Dataset Description

Two experimental dataset were available. In the first study, 59 acute comatose patients were included between 7 days and 30 days after the coma onset ( $46 \pm 16$  years old; 21 females). PET data were acquired in list mode during 90 min from the injection of an intravenous bolus of  $[^{18}\text{F}]\text{FDG}$ . The second study, included 7 healthy subjects ( $25 \pm 3$  years old; all male) with injection of an intravenous bolus of  $[^{11}\text{C}]\text{Yohimbine}$ .

Both dataset used the same exact following imaging protocol. Using a Siemens healthineers Biograph mMR simultaneous MR and PET machine, an arterial time-of-flight MR (TOF-MRA) in axial orientation, with a voxel size of  $0.3 \times 0.3 \times 0.7$  mm as well as a T1-weighted MRI in axial orientation, with isometric voxel size of 1mm were acquired during the 90 minute dynamic PET. Raw PET data were rebinned into 24-time frames (variable length frames :  $8 \times 15$  s,  $3 \times 60$  s,  $5 \times 120$  s,  $1 \times 300$  s,  $7 \times 600$  s) sinograms for dynamic reconstruction. Reconstruction yielded a voxel size of  $1.04 \times 1.04 \times 2.08$  mm<sup>3</sup> in a matrix of  $344 \times 344 \times 127$  voxels.

In the  $[^{18}\text{F}]\text{FDG}$  study, AIF in the whole blood and plasma were measured from 26 arterial blood samples manually collected (with the following timing: every 5 s for the first minute, every 15 s until second minute, and at times 3, 5, 10, 20, 30, 45, 60, 75, 80, 85 and 90 minutes post-injection) and counted with a gamma counter.

In the  $[^{11}\text{C}]\text{Yohimbine}$  study, 25 arterial blood samples were manually collected (with the following timing: every 5 s for the first minute, every 10 s until second minute, and at times 5, 10, 30, 45, 60, and 90 minutes post-injection). The blood samples were counted with a gamma counter and were centrifuged and the extracted plasma was counted again to calculate the plasma fraction [TODO correct this part]

## 2.2 Pre-processing

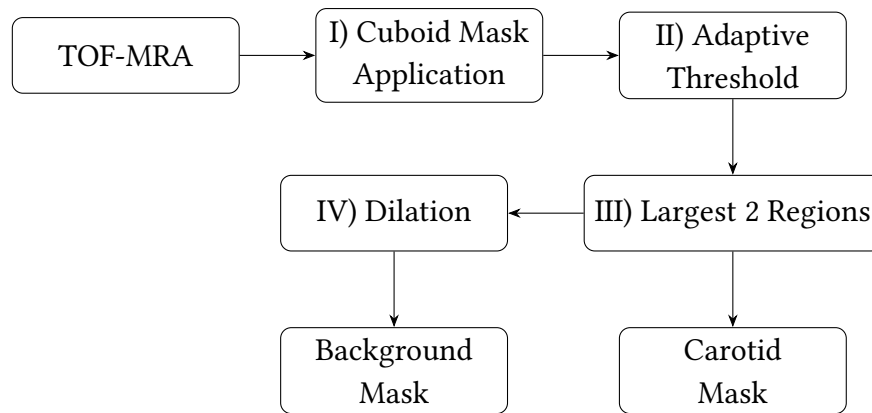
For both studies, the T1-weighted image was registered to both the average PET image and to the TOF-MRA image using a affine registration method. The two resulting affine transformation matrices were then combined to register the TOF-MRA directly to the PET. Even though in the  $[^{18}\text{F}]\text{FDG}$  study, the patients were completely unconscious and the PET

and MRI images were acquired simultaneously in the same session, this was still done out of an abundance of caution. The T1-weighted image was used as a medium since it has common spatial features with both the other two modalities and directly registering TOF-MRA to the PET would be impractical due to the low spatial resolution of the PET and the limited axial FOV of the TOF-MRA images. Lastly, the brain was segmented into regions of interest (ROI) based on the Hammersmith brain atlas [6].

## 2.3 Carotid Segmentation

Figure 2.1 shows the complete pipeline for the carotid segmentation. Since vessels appear as hypersignal in TOF-MRA, a high-intensity thresholding technique can be used to extract the arteries from the image. However, other tissues such as brain lesion—which were very common in the comatose dataset—and venous structures may also appear as hypersignal and can interfere with the selection. To exclude them, in a reference image, a cuboid volume of interest (VOI) was defined where the carotid arteries are most likely to appear (Step I). The threshold value was calculated as the 1% percentile of all non-zero voxels in the image and then it was applied to only to the voxels inside the VOI (Step II). A 3D connectivity constraint was employed to the remaining voxels to extract all fully connected volumes and among those, the top 2 biggest were consider as the left and right internal carotid artery and extracted(Step III). The threshold value was calculated as the 1% percentile of all non-zero voxels in the image and then it was applied to only to the voxels inside the VOI (Step II). A 3D connectivity constraint was employed to the remaining voxels to extract all fully connected volumes and among those, the top 2 biggest were consider as the left and right internal carotid artery and extracted(Step III).

The reference image used was the standard MNI152 atlas, which was padded by 50% more voxels in the inferior direction of the brain (negative Z in voxel-space) since the original atlas excludes the area below the brain which is of the interest for us. The TOF-MRA image was then registered to the reference image using an affine registration technique and the obtained affine matrix was used to register and then apply the VOI to the TOF-MRA.



**Figure 2.1:** Carotid and background mask segmentation pipeline

## 2.4 Partial Volume Correction

### 2.4.1 Geometric Transfer Matrix

As we discussed before in ??, direct extraction of the radioactivity in the arteries is not practical due to PVC. Geometric Transfer Matrix aims to account for this loss of signal by considering the observed TACs are linear combination of the true value and other effecting regions [7].

Here we define two regions, the carotid and the surrounding tissues (background). A mask for extracting the activities of the latter was obtained by dilating the carotid mask by 5 mm and subtracting the voxels corresponding to the carotid mask (Figure 2.1, Step IV).

$$\underbrace{\begin{bmatrix} T_c \\ T_{bg} \end{bmatrix}}_{\text{Observed}} = \underbrace{\begin{bmatrix} \omega_{c \rightarrow c} & \omega_{bg \rightarrow c} \\ \omega_{c \rightarrow bg} & \omega_{bg \rightarrow bg} \end{bmatrix}}_{\text{GTM}} \cdot \underbrace{\begin{bmatrix} T_{IF} \\ T_{tissue} \end{bmatrix}}_{\text{Unknown}}, \quad (2.1)$$

where  $\omega_{n \rightarrow m}$  are the spill-in and spill-over coefficient of region  $n$  onto region  $m$ , which is obtained by convolving the binary mask of region  $n$  with the system's point spread function and integrating the resulting intensity over region  $m$ , normalized by the total signal in region  $m$ . where

$$\omega_{n \rightarrow m} = \frac{\int_{\Omega_m} (h * \chi_n)(r) dr}{\int_{\Omega_m} (h * \chi_m)(r) dr}, \quad (2.2)$$

with  $\chi_n$  and  $\chi_m$  denoting the binary masks of regions  $n$  and  $m$ , respectively,  $h$  the system's point spread function, and  $\Omega_m$  the spatial domain of region  $m$ .

$T_c$  and  $T_{bg}$  are respectively the observed carotid and background TACs and  $T_{IF}$  and  $T_{tissue}$  are the real unknown TACs of the carotid (the input function) and the background tissue.

By inverting the GTM, this system of equations can be easily solved and the "true" TAC in the artery and background be computed. However, GTM along with other classical PVE methods usually fail to fully recover the lost signal as they are a simplification that does not consider the whole picture particularly accounting for the time-variant noise experienced in the signal. This causes these methods to not only fail to correct the noise but end up amplifying them because of their assumptions. The time-variant noise can be attributed to number of contributing factors, namely, the small size of the arteries, very short frame-times at the beginning of the scan where the changes in the input function are very rapid and rapid decay of the radiotracer which in all result in low-count statistic hence having higher noise.

### 2.4.2 Bayesian Geometric Transfer Matrix

#### Modelling

For each subject,  $T_{IF}$  is modeled as a linear combination of a population mean and its principal components derived by performing principal component analysis (PCA) on a set of arterial input functions (AIFs) collected from the population. Specifically, for each subject, a subset of 20 random subjects is selected from the dataset—excluding the subject under study—to construct the PCA model. The data are first zero-centered, and PCA is applied to extract the principal axes  $\phi_i(t)$  and their corresponding explained variances  $\lambda_i$ . Each axis is then scaled by  $\sqrt{\lambda_i}$  to yield components  $v_i(t)$  with distribution of  $\mathcal{N}(0, 1)$ . The input function is then modeled as:

$$T_{IF}(t) = \mu(t) + \sum_{i=1}^p \theta_i v_i(t), \quad (2.3)$$

with

$$v_i(t) = \sqrt{\lambda_i} \phi_i(t),$$

where  $p$  is the number of components,  $\mu(t)$  is the population mean AIF,  $\phi_i(t)$  are the principal axes obtained from PCA,  $\lambda_i$  are their explained variances, and  $v_i(t)$  are the scaled components. Due to this scaling, the weighting coefficients  $\theta_i$  will have a distribution of  $\sim \mathcal{N}(0, 1)$ .

Spectral Analysis (SA) has been proposed as a decomposition method for describing the tissue activity in dynamic PET. This method produces a spectrum of kinetic components by modelling the TACs as a convolution of the input function with a impulse response function [TODO].

The impulse response function is considered as sum of multiple exponentials. Therefore the background TAC will be

$$T_{bg} = \sum_{i=1}^s \alpha_i \cdot T_{IF} \otimes e^{-\beta_i t}.$$

Where  $\otimes$  denotes the convolution operator,  $\alpha_i$  is the coefficient, and  $\beta_i$  is the frequency of the  $i$ th spectrum.

#### Noise

Accurately modelling the noise is very challenging as there are numerous sources of them. Some sources of noise such as scattering effect, random events, and motion artifacts among others are corrected to some degree by the reconstruction program. However the biggest contributor of noise is "low-count statistics" which is caused by short time frames at the begin of the scan and by the exponential decay of the radioisotope towards the end of the scan apparent by the low dose of the tracer injection. This combined with the small size of the carotid introduces a massive challenge for recovering the lost signal.

Noise in raw PET counts is considered to be a Poisson distribution however after the reconstruction process it is widely assumed to be a Gaussian distribution [TODO]. To account for the previously mentioned time-variant noise, weights are used to normalize the noise along all frames to a single statistic. We define the weights as

$$\omega_i = \frac{\Delta t_i}{c_i} e^{\frac{-t_i \ln(2)}{T_{1/2}}}$$

where  $\omega_i$  is the weight,  $\Delta t_i$  is the frame duration, and  $c_i$  is the net-true counts detected by the PET camera at the  $i$ th frame and  $T_{1/2}$  is the half-life of the radioisotope. However,  $c_i$  was substituted here with  $C_T(t_i)$ , the total radioactivity concentration at the mid-time frame due to unavailability of these counts in our datasets.

The time-variant noise variances can be normalized as the weighted mean of the per frame variance.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N \omega_i \sigma_i^2.$$

## Estimation

Let us define  $\mathcal{D}$  as the observed PET TAC,  $\Theta = (\theta_1, \dots, \theta_p, \alpha_1, \beta_1, \dots, \alpha_s, \beta_s)$  as the set of unknown parameters describing the system, and consider an unknown Gaussian noise with variance of  $\sigma^2$  present in the system. We seek to estimate the joint probability  $p(\Theta, \sigma^2 | \mathcal{D})$  of the model parameters  $\Theta$  and noise level given the observed data  $\mathcal{D}$

According to the Bayes rule:

$$p(\Theta, \sigma^2 | \mathcal{D}) \propto p(\mathcal{D} | \Theta, \sigma^2) \cdot \pi(\Theta) \cdot \pi(\sigma^2)$$

where  $p(\Theta, \sigma^2 | \mathcal{D})$  is the posterior distribution,  $p(\mathcal{D} | \Theta, \sigma^2)$  is the likelihood, and  $\pi(\Theta)$  and  $\pi(\sigma^2)$  are the prior knowledge that we have on the set of parameters and the noise. Thus we can sample the posterior distribution by calculate the likelihood and having the priors.

We can derive the observed data from the GTM response of the model ???. For convenience, we define

$$\mathcal{T}(t) = \mathcal{G}(t; \Theta) = \mathcal{G} \left( \begin{bmatrix} T_{IF}(t; \theta_1, \dots, \theta_p) \\ T_{tissue}(t; \alpha_1, \beta_1, \dots) \end{bmatrix} \right)$$

where  $\mathcal{G}$  is the GTM model and  $\mathcal{T}$  is the perceived activity based on the GTM model (effected by PVE). Since we considered the noise to be Gaussian, the likelihood would be

$$p(\mathcal{D} | \Theta, \sigma^2) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{\omega_i (\mathcal{D}(t_i) - \mathcal{T}(t_i))^2}{2\sigma^2} \right).$$

Hence from this point forward, we call this method as the Bayesian Geometric Transfer Matrix (Bayesian GTM or BGTM).

The prior of parameters can be expanded as

$$\pi(\Theta) = \prod_{i=1}^p \pi(\theta_i) \prod_{j=1}^s \pi(\alpha_j) \pi(\beta_j).$$

As mentioned before, the PCA weighting coefficients have a normal distribution of  $\mathcal{N}(0, 1)$  which is also used as their prior knowledge. The prior for the frequency and amplitude in SA was chosen as a reasonable uniform distribution range.

$$\alpha_i, \beta_i \sim \mathcal{U}(u_{\min}, u_{\max})$$

The prior distribution for the noise variance,  $\pi(\sigma^2)$ , is specified as an inverse-gamma distribution with hyperparameters  $a_0$  and  $b_0$  since inverse-gamma is the prior conjugate for the variance of a normal distribution

$$\pi(\sigma^2) \sim \Gamma^{-1}(a_0, b_0)$$

In this work, the hyperparameters are selected so that  $\pi(\sigma^2)$  is centered around an empirical estimate of the TAC variance, computed as the weighted root mean squared error between the observed carotid TAC and the mean AIF.

$$\mathbb{E}[X] = \frac{1}{N} \sum_{i=1}^N (T_c(t_i) - \mu(t_i))^2$$

Thus by choosing an arbitrary Coefficient of Variation (CV), we will have

$$a_0 = 2 + \frac{1}{CV^2}, \quad b_0 = (a_0 - 1) \mathbb{E}[X]$$

## Sampling

Evaluating the posterior distribution ?? directly is not feasible due to its complexity and lack of a closed-form solution. Thus we utilize a Monte-Carlo sampler to estimate the posterior distribution by taking sufficient random samples. In this work, we utilize a Markov Chain Monte-Carlo with a hybrid Metropolis-within-Gibbs sampling method.

The Gibbs sampler allows for separately drawing sample for each parameter one by one based on its univariate posterior conditional distribution  $p(\Theta_i \mid \mathcal{D}, \Theta_{-i}, \sigma^2)$  for each  $\Theta_i$  where  $\Theta_{-i}$  is the parameter vector except  $\Theta_i$  and  $p(\sigma^2 \mid \mathcal{D}, \Theta)$  for the  $\sigma^2$ .

By iteratively doing this, Gibbs sampling results in the desired multivariate conditional posterior distribution  $p(\Theta, \sigma^2 \mid \mathcal{D})$ .

The exploration of the parameters  $\Theta$  is done using the Metropolis-Hasting algorithm. In this algorithm, a new parameter is proposed based on slightly changing the previous

sample. For example at the  $m$ th iteration in the chain, a new proposal for  $\Theta_i^{(m)}$  is given as  $\Theta'_i = \Theta_i^{(m)} + \varepsilon_i \cdot b$ , where  $\varepsilon_i$  is the scaling factor or step-size and  $b$  is a Brownian motion. The proposal is either accepted and is kept as the new sample or rejected and the parameter stays unchanged. The acceptance probability is given by

$$\min \left( 1, \frac{p(\Theta'_i | \mathcal{D}, \Theta_{-i}, \sigma^2)}{p(\Theta_i^{(m)} | \mathcal{D}, \Theta_{-i}, \sigma^2)} \right)$$

By choosing inverse-gamma for the noise prior which is conjugate prior, the posterior distribution is also a inverse-gamma distribution.

$$p(\sigma^2 | \mathcal{D}, \Theta) \sim \Gamma^{-1}(a, b)$$

where

$$a = a_0 + \frac{N}{2}, \quad b = b_0 + \sum_{i=0}^N \omega_i (\mathcal{D}(t_i) - \mathcal{T}(t_i))^2.$$

Since we have a closed-form expression for the posterior distribution of the noise variance, we can just simply draw samples without the Metropolis-Hasting acceptance/rejection method.

To ensure efficient exploration of the parameter space, each parameter  $k$  is assigned its own proposal step size  $\epsilon_k$ . During the burn-in phase, these step sizes are adapted in blocks of  $L = 50$  Metropolis-Hastings updates by monitoring the empirical acceptance rate for each parameter. If the acceptance rate for a parameter exceeds 0.5, its step size is increased by an arbitrary 10% ( $\epsilon_k \leftarrow 1.1 \epsilon_k$ ); otherwise, it is decreased by 10% ( $\epsilon_k \leftarrow 0.9 \epsilon_k$ ). This adaptive tuning continues until the end of the burn-in period, after which all step sizes are fixed and subsequent samples are retained.

## Inference

After sufficient sampling, we employ a modified Maximum a Posteriori (MAP) method by calculating the average parameters of the top 0.1% samples with the highest posterior probability and considering it as the solution.

## 2.5 Evaluation

### 2.5.1 IF Curves

The performance of the proposed IDIF estimation was first evaluated by computing the mean absolute error (MAE) between the cumulative area under the curve (cAUC) of the estimated IDIF and the *ground true* AIF. cAUC was considered to be a more suitable metric



since it provides an integrated measure of tracer exposure over time and is less sensitive to local fluctuations or noise in the curve compared to the directly comparing the TACs.

$$\text{cAUC}(t) = \int_0^t IF(\tau) d\tau, \quad (2.4)$$

where  $IF$  is the input function.

### 2.5.2 Quantification

However, because the cAUC error does not fully capture the impact of IDIF deviations on kinetic parameters, absolute quantification was also performed to evaluate the performance of the estimated IDIF against the gold standard AIF. This was achieved by utilizing an reversible and irreversible two-tissue compartment model (2TCM) via non-linear fitting respectively with the `fitk3` for the  $[^{18}\text{F}]$ FDGdataset and `fitk4` program for the  $[^{11}\text{C}]$ Yohimbinedataset from the TPCCLIB library developed at the Turku PET Centre [8]—applying the model fitting once with the IDIF and once with the AIF as the input function.

The brain atlas was applied to the PET image and regional TACs were acquired by averaging voxels over time. For  $[^{18}\text{F}]$ FDG, the regional ( $\text{MR}_{\text{glu}}$ ) was then calculated, The mean absolute percentage error (MAPE) of the  $\text{MR}_{\text{glu}}$  in each ROI was calculated and then averaged across the entire dataset:

$$\text{Average MAPE}(\text{MR}_{\text{glu}}) = \frac{100\%}{N} \sum_{i=1}^N \left( \frac{1}{N_{\text{ROI}}} \sum_{j=1}^{N_{\text{ROI}}} \left| \frac{\text{MR}_{\text{glu},ij}^{\text{IDIF}} - \text{MR}_{\text{glu},ij}^{\text{AIF}}}{\text{MR}_{\text{glu},ij}^{\text{AIF}}} \right| \right), \quad (2.5)$$

where  $N$  is the number of subjects.

Similarly, for  $[^{11}\text{C}]$ Yohimbine, the Volume of Distribution ( $V_T$ ) was calculated and the dataset error was calculated as

$$\text{MAPE}(V_T) = \frac{100}{N} \sum_{i=1}^N \left( \frac{1}{N_{\text{ROI}}} \sum_{j=1}^{N_{\text{ROI}}} \left| \frac{V_{T,ij}^{\text{IDIF}} - V_{T,ij}^{\text{AIF}}}{V_{T,ij}^{\text{AIF}}} \right| \right). \quad (2.6)$$

Additionally, linear least-squares regression was performed between the regional  $\text{MR}_{\text{glu}}$  and  $V_T$  obtained by quantification with using AIF and IDIF for each subject. The coefficient of determination ( $R^2$ ) and the regression slope ( $S$ ) were obtained for each subject. The mean absolute errors of these metrics across the dataset are given by

$$\text{MAE}(R^2) = \frac{1}{N} \sum_{i=1}^N |R_i^2 - 1| \quad (2.7)$$

and

$$\text{MAE}(S) = \frac{1}{N} \sum_{i=1}^N |S_i - 1|, \quad (2.8)$$

where  $N$  is the number of subjects and  $R_i^2$  and  $S_i$  denote the coefficient of determination and the regression slope for subject  $i$ , respectively.

## 2.6 Simulation

PET simulations are an integral part of the validation of statistical and analytical PET methods as they have the advantage of knowing the exact ground truth which is not always available or might be effected by human error or instrumental malfunction. In the previous section, we considered the AIF to be the gold standard and the ground truth and the evaluation was done by comparing the method to it. But since the AIF is taken and analysed manually and involves multiple devices and processes operated by different personnels (machine operators, nurses, and biochemists) at different steps, there is a large possibility for mistakes.

For this reason, as a complementary validation step to the experimental datasets, simulations were conducted and the method will be evaluated on them as well.

PET-SORTEO is a realistic Monte-Carlo PET simulation platform which its accuracy has been validated numerous times ?? and for that reason can be considered as close to a real PET acquisition. However, this accuracy heavily depends on describing the brain physical anatomy and the tracer's biochemical processes in the tissue.

Thus, as the input, PET-SORTEO requires the description and performance of a PET machine, a 3D emission numerical phantom segmenting the brain in to different emitting regions, a 3D attenuation numerical phantom describing the attenuation coefficient of the tissues, and lastly the Time Activity Curves in each emitting regions.

For this study, we seek to replicate the  $[^{18}\text{F}]$ FDG dataset again in the simulation. Thus we chose the same Biograph mMR scanner as the real dataset and we took inspirations from the real MR and PET data to create our numerical phantoms and the TACs.

Firstly, using the regional brain atlas of the subject, the sub-regions were merged to create larger regions of White Matter(WM), Gray Matter(GM), Gray Nuclei (GN), Cerebellar Gray Matter(CGM), and Cerebellar White Matter(CWM). An algorithm that was originally designed for creating pseudo CT from T1-weighted image for bone attenuation correction in PET reconstruction was used here to create masks for the bone ?. Values above 300 were considered as the bone (BONE) and values between -300 and 300 were considered as Extra-cerebral Soft Tissues (SOFT) — after subtracting the brain—which includes the scalp, neck muscles, glands, eyes, etc. The left over area between the skull and the brain was considered as the Cerebrospinal fluid (CSF). And lastly, from the segmentation of TOF-MRA ??, the mask of the artery (ICA) was obtained. And lastly, the 9 different regions were overlapped together to create the final emission phantom.

The attenuation phantom was again created from the pseudo-CT by segmenting the scan in to SOFT and BONE and filling them with their respective attenuation coefficients.

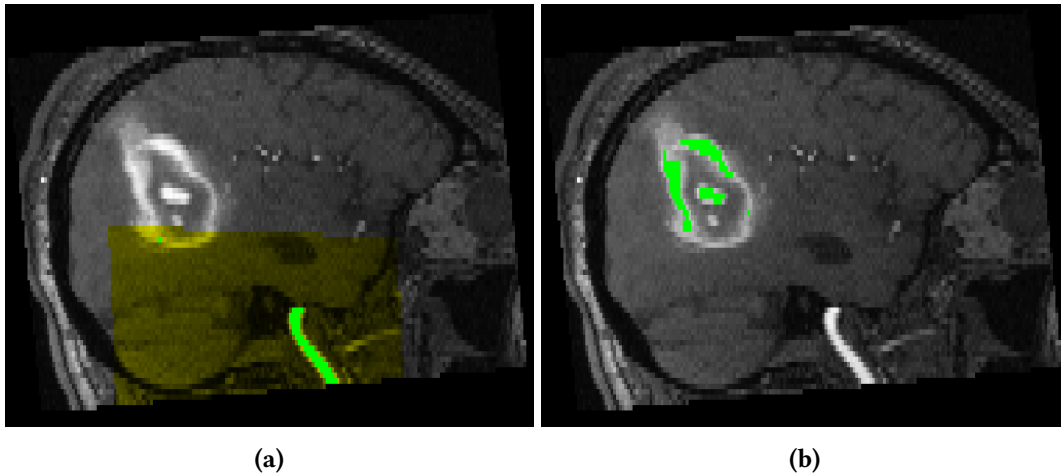
The resulting emission phantom was then registered to the PET space to extract TACs from the PET scan of the subject. The TACs were then corrected for PVE by GTM. The partial volume corrected TACs were then fitted to a (2TCM) to remove any possible noise and to achieve a so-called "perfect" TACs. The AIF used in the fitting was also the real AIF of the subject that was fitted to a exponential model for the same reason??. A few consideration were still taken in to account. Namely, our definition of SOFT tissue is very broad

and although it is sufficient for the purpose of simulation, it breaks the rule of homogeneity required for compartment modelling thus the raw unfitted TAC was kept for this specific region. The pseudo-CT generation algorithm was very permissive in its definition of the BONE thus the resulting TAC was higher than reality since signal from the brain and the scalp was leaked in to it. So it was empirically decided that the value of the BONE should be halved. The CSF region is very narrow at points therefor it is heavily effected by PVE so GTM fails to correct for it. So the CSF TAC was considered to be 0.2 of the fitted WM TAC.

# Results

## 3.1 Carotid Segmentation from MRA-TOF

As illustrated in Figure 3.1, the cuboid mask plays a crucial role in the carotid segmentation. Because no ground truth segmentation is available, visual inspection was used to evaluate the results.



**Figure 3.1:** Comparison of carotid segmentation (green) with (a) and without (b) a cuboid mask (yellow). In the absence of the cuboid mask, the segmentation algorithm fails to capture the carotid and instead incorrectly identifies the brain lesion

## 3.2 Hyperparameter Tuning

The number of principle components  $p$  was chosen as 3 by analysing the explained variance ratio plot in Figure ???. We see with only 3 components we get more than 90% explained variability. To prove the availability of the method, we seek to find the lowest possible number of population needed for the PCA in our  $[^{18}\text{F}]\text{FDG}$  dataset. To do this, the experiment was repeated 10 times each time with a different random seed for choosing population for  $N = \{5, 10, 20, 30, 40, 60\}$  and the resulting quantification errors were averaged. As apparent in Figure ??, we see a plateau in the performance after  $N = 10$  [TODO] and it was

chosen as the ideal population sample count. For the  $[^{11}\text{C}]$ Yohimbine dataset since it only has 7 subjects,  $N = 6$  was chosen.

In model fitting using SA, an appropriate spectral range is chosen based on the radioisotope half-life (e.g.  $10^{-4}\text{s}^{-1}$  to  $1\text{s}^{-1}$  for  $^{18}\text{F}$ ) and is then divided into  $s = 100$  or  $s = 1000$  fixed frequencies ( $\beta$ ) then their amplitudes ( $\alpha$ ) are fitted to the TAC. The peaks in the resulting spectrum are considered as the [TODO] frequencies of the TAC. However, estimating 100 or 1000 different amplitude parameters using MCMC would not be computationally feasible. Instead we must choose an appropriate number of spectral components and let the model to tune the frequencies and the amplitudes together.

To derive the number of basis functions, we assumed the activity in the surrounding tissue can either be modelled as 1TCM or 2TCM which their Impulse Response Function (IRF) have respectively one and two basis functions. Experiments with two basis functions showed that the model rarely results in two distinctive frequencies and would generally converge to similar values for both frequencies. Thus for both datasets one basis function was chosen.

As discussed in Section ??, the prior distribution used for the PCA weighting coefficients was  $\theta_i \sim \mathcal{N}(0, 1)$ . The prior for the SA parameters were considered a very uninformative uniform distribution  $\mathcal{U}(10^{-5}, 10^{-2})$  as we did not have concrete prior knowledge on these parameters.

### 3.3 Simulation

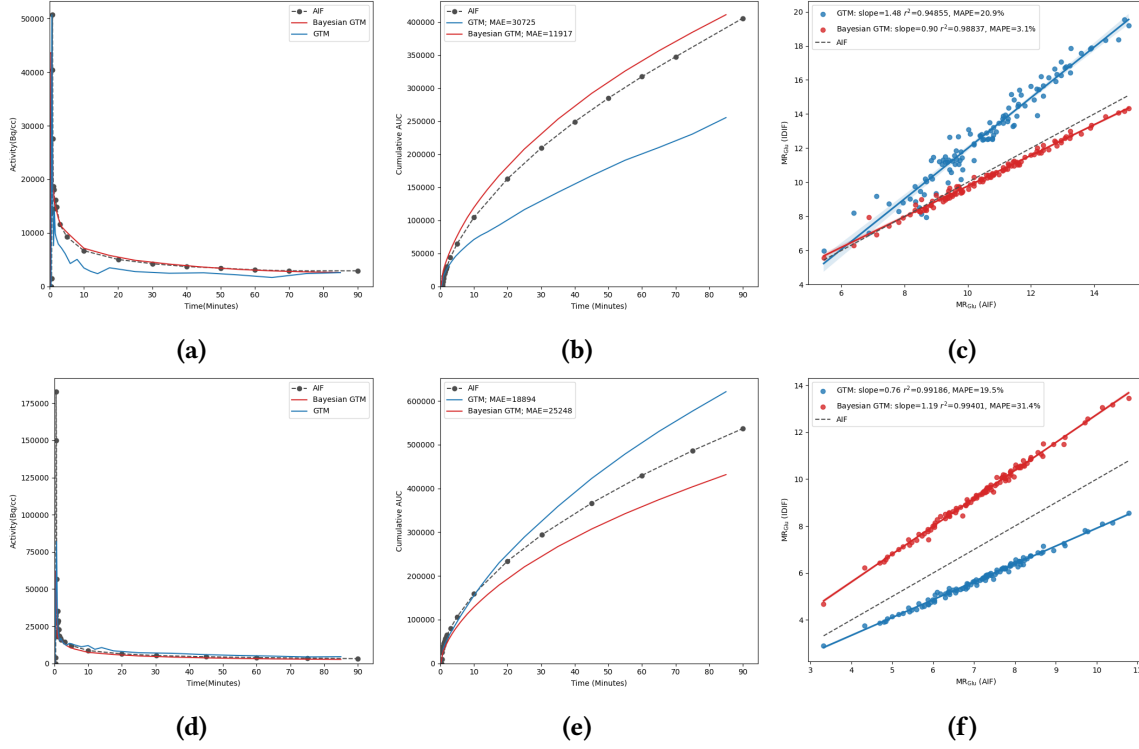
From the 59 subjects in the  $[^{18}\text{F}]$ FDG dataset, 24 subjects were excluded due to having very large brain tumors or opened skulls due to open brain surgery which caused the tissue classification algorithm to fail. The rest of the 35 subjects were used to create the numerical phantoms and to generate TACs for the simulations. Although the point of the simulation was not to replicate the same subject exactly but only use it as an inspiration for what real TACs look like, the results of the simulations were compared to the real scans as a quality control step. In Figure ?? we see the 2nd and last frame axial [TODO] view of the real PET, the simulation input, and the simulation output. As expected we see less details in the neck area (SOFT region) since we did not consider the complete arterial and venous structure as well as glands. But more importantly, we can clearly see the activity in the internal carotids. In the last frame we can also see the activity of the brain to be very realistic.

However, the

### 3.4 IDIF

Since BGTM relies on the GTM method and also depends on a population AIF. Thus the performance should be compared compared to the IDIF derived from GTM method and also the Population Based Input Function (PBIF) which is the mean AIF used in the PCA.

Figure 3.2 compares the two methods for one of the best- and worst-performing subjects. In the well-performing subject, BGTM significantly outperforms GTM ( $MR_{glu}$  MAPE of 3.1% vs. 20.9%); however, in the poorly performing case, BGTM falls short of GTM ( $MR_{glu}$  MAPE of 31.4% vs. 19.5%). The average mean absolute error of the cAUC curves across the dataset was 14,202 for BGTM and 33,764 for GTM (Figure 3.3b).

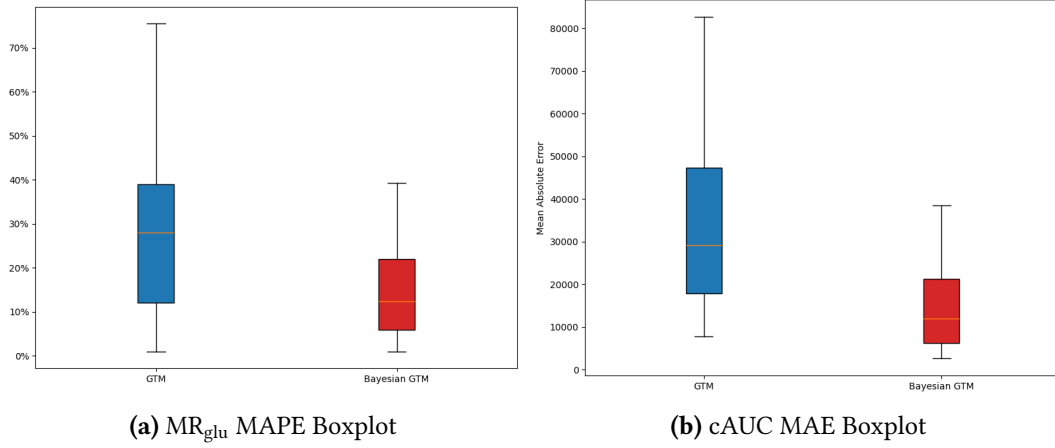


**Figure 3.2:** Comparison of the IFs (a,d), cumulative AUC curves (b,e), and  $MR_{glu}$  regression lines (c,f) for one of the best-(top row) and worst-performing (bottom row) subjects

ROI-based quantification was carried out using both IDIF methods, with BGTM yielding significantly better performance. Specifically, the BGTM and GTM methods achieved an average  $MR_{glu}$  mean absolute percentage error (MAPE) of 14.1% and 33%, respectively (Figure 3.3a), and an average  $MR_{glu}$  MAE of 1.42 and 3.5. In addition, the MAE for the coefficient of determination ( $R^2$ ) and the regression slope ( $S$ ) were 0.004 and 0.14 for BGTM, compared to 0.030 and 0.304 for GTM, respectively.

A paired t-test was conducted to compare the performance of BGTM and GTM across previously mentioned metrics. The results, summarized in Table 3.1, indicate that BGTM significantly outperforms GTM in cAUC MAE ( $t = 7.44$ ,  $p = 7.2 \times 10^{-10}$ ),  $MR_{glu}$  MAPE ( $t = 4.32$ ,  $p = 6.5 \times 10^{-5}$ ),  $MR_{glu}$  MAE ( $t = 4.41$ ,  $p = 4.8 \times 10^{-5}$ ), and  $MR_{glu}$  Slope MAE ( $t = 4.73$ ,  $p = 1.6 \times 10^{-5}$ ), demonstrating the effectiveness of the proposed method. However, no significant difference was observed in  $MR_{glu}$   $R^2$  MAE ( $t = 1.45$ ,  $p = 0.15$ ).

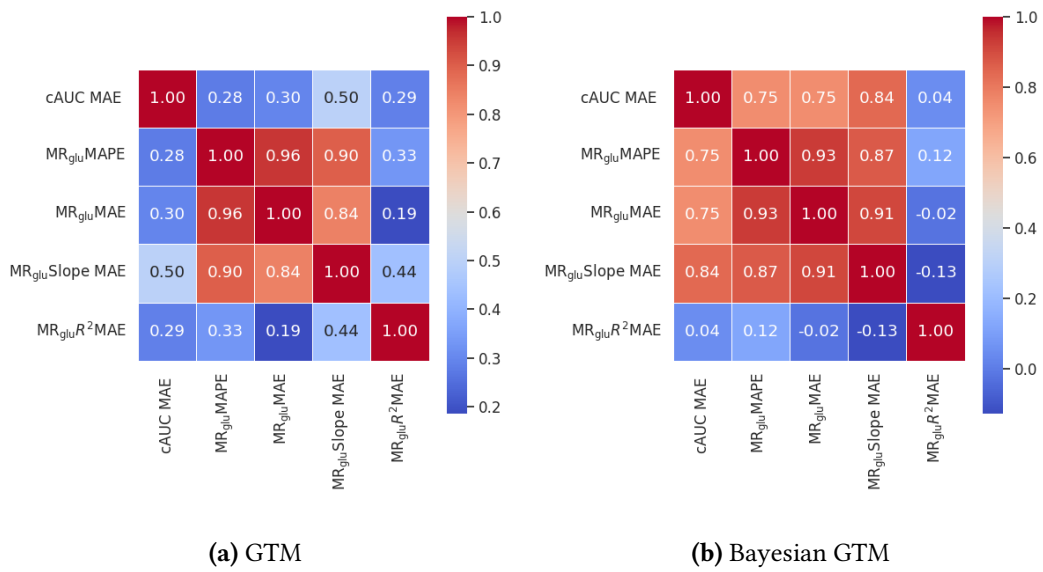
As illustrated in Figure 3.4, there is a strong correlation between the cAUC MAE and the quantification errors suggesting cAUC as a reliable intermediate metric.



**Figure 3.3:** Boxplot of curve and quantification errors

Metric	BGTM		GTM		Paired T-Test	
	$\mu$	$\sigma$	$\mu$	$\sigma$	T-Value	P-Value
IF cAUC MAE	14,202	9,190	33,764	21,212	7.44	$7.2 \times 10^{-10}$
MR <sub>glu</sub> MAPE (%)	14.1	10.1	33.0	31.5	4.32	$6.5 \times 10^{-5}$
MR <sub>glu</sub> MAE	1.42	1.07	3.50	3.38	4.41	$4.8 \times 10^{-5}$
MR <sub>glu</sub> $R^2$ MAE	0.004	0.006	0.030	0.132	1.45	$1.5 \times 10^{-1}$
MR <sub>glu</sub> Slope MAE	0.14	0.109	0.304	0.230	4.73	$1.6 \times 10^{-5}$

**Table 3.1:** Summary of performance metrics for BGTM and GTM methods and their paired t-test.



**Figure 3.4:** Correlation matrix of different metrics for Bayesian GTM and GTM methods

# Discussion

The Bayesian Geometric Transfer Matrix (BGTm) method demonstrated significant improvements in IDIF accuracy compared to classical GTM approaches through enhanced handling of partial volume effects. This advancement was achieved by integrating TOF-MRA-guided carotid segmentation with population-based AIF prior knowledge within a Bayesian framework.

The refined segmentation algorithm, particularly through implementation of the cuboid mask, proved crucial for processing challenging data from comatose patients who frequently present brain lesions (Figure 3.1). However, occasional over-conservatism in vessel selection suggests the region-growing and thresholding parameters can be optimized to accurately capture the complete carotid structure. The lack of reference or ground truth carotid mask prevented the evaluation of the segmentation. This also disabled us from accurately evaluating the proposed IDIF algorithm, as the calculated errors are accumulated with the segmentation error as well. With labeled data, deep learning approaches such as U-Net can also be explored to improve segmentation accuracy [9].

The cAUC error reduction (Figure 3.3b) confirms Bayesian priors effectively constrain AIF estimates to physiological ranges. The strong correlation between cAUC errors and quantification errors (Figure 3.4) validates cAUC as a robust intermediate metric for IDIF assessment, enabling rapid validation without requiring full kinetic modeling.

Quantitative analysis revealed strong agreement between BGTm and AIF-derived  $MR_{glu}$  values, with the Bayesian approach achieving a 57% reduction in absolute error compared to conventional GTM methods (Table 3.1). Despite this improvement, the observed MAPE results exhibited substantial dispersion (Coefficient of Variation = 71%). This inconsistency is particularly evident in outlier cases where BGTm underperformed (Figure 3.2). Although, further investigation is required to fully characterize the underlying causes, one plausible factor could be TOF-MRA/PET misregistration resulting in inaccurate TAC extraction.

Our analysis was limited to a comatose patient cohort from a single imaging center, which may limit generalizability to populations with normal cerebral blood flow. Future validation should incorporate multi-center studies involving both healthy subjects and different radiotracers to establish broader applicability.

Additionally, the current PCA implementation of randomly selecting 10 subjects from the population lacks practical viability for standardized implementation. Future iterations should only be limited to a predefined population.



# Conclusion

This study demonstrates that combining TOF-MRA-guided carotid segmentation with partial volume correction in a bayesian framework improves image derived input function estimation. While the method showed significant improvement in IDIF estimation accuracy, variability in accuracy underscores the challenges posed in artery segmentation as well as partial volume correction. Future validation must address limitations in the current PCA framework by establishing predefined reference cohorts and expanding to multi-center studies with diverse patient populations. By refining anatomical alignment protocols and ensuring consistent prior knowledge integration, this approach can enhance reliability for clinical translation, ultimately supporting safer, patient-friendly quantitative PET imaging.

# References

- [1] John W Keyes. “SUV: standard uptake or silly useless value?” In: *Journal of Nuclear Medicine* 36.10 (1995), pp. 1836–1839.
- [2] Jean Logan et al. “Graphical analysis of reversible radioligand binding from time—activity measurements applied to [N-11C-methyl]-(-)-cocaine PET studies in human subjects”. In: *Journal of Cerebral Blood Flow & Metabolism* 10.5 (1990), pp. 740–747.
- [3] Clifford S Patlak, Ronald G Blasberg, and Joseph D Fenstermacher. “Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data”. In: *Journal of Cerebral Blood Flow & Metabolism* 3.1 (1983), pp. 1–7.
- [4] Peter Young et al. “Image-derived input functions from dynamic 15O–water PET scans using penalised reconstruction”. In: *EJNMMI physics* 10.1 (2023), p. 15.
- [5] Zacharie Irace et al. “Bayesian partial volume correction for image derived input function”. In: *Journal of Cerebral Blood Flow and Metabolism*. Vol. 41. 1\_ SUPPL. SAGE PUBLICATIONS INC. 2021, pp. 229–229.
- [6] Alexander Hammers et al. “Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe”. en. In: *Hum. Brain Mapp.* 19.4 (Aug. 2003), pp. 224–247.
- [7] Olivier G Rousset, Yilong Ma, and Alan C Evans. “Correction for partial volume effects in PET: principle and validation”. In: *Journal of nuclear medicine* 39.5 (1998), pp. 904–911.
- [8] Vesa Oikonen et al. *TPCCLIB*. Version 0.6.20. Retrieved on 2nd February 2025. Turku PET Centre, University of Turku, 2018. URL: <https://gitlab.utu.fi/vesoik/tpcclib>.
- [9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18. Springer. 2015, pp. 234–241.