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

Hosting Laboratory: CERMEP - Imagrie du Vivant, Lyon, France
2024-2025

Abstract

[TODO]

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

Contents

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 (TAC) 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), volume of distribution (V_T), and Binding Potential (BP).

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:

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

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

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

$$C_T(t)\ =\ C_F(t)\ +\ C_B(t)\ +\ C_P(t). \tag{1.3}$$

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 (MR_{glu}) is calculated as

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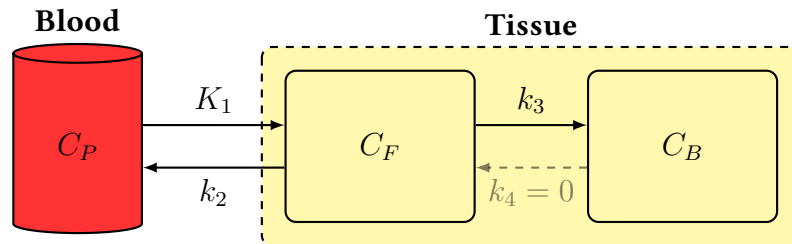

Figure 1.1: Schematic of the two-tissue compartment model (2TCM)

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

For $[^{11}C]$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

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

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. In practice, an average curve is derived from a representative cohort and then temporally aligned and scaled to the subject using one or more calibration points (e.g., early blood samples or image-derived peaks). PBIF can reduce invasiveness and acquisition burden, but it may introduce bias if the cohort is not well matched to the subject or if calibration is suboptimal.

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 the loss and mixing of signal that occurs when structures are small relative to the system's spatial resolution, causing activity to "spill out" of small objects and "spill in" from neighbors. In PET, this leads to systematic underestimation of arterial activity and contamination from surrounding tissue, especially for vessels like the ICA that have similar size as the effective resolution of PET camera (both around 5mm). In Figure 1.2, this effect is illustrated by simulating the effect of PET's point spread function (PSF) as well as noise on hot disk. Correcting PVE is essential for accurate IDIF estimation and downstream kinetic modeling.
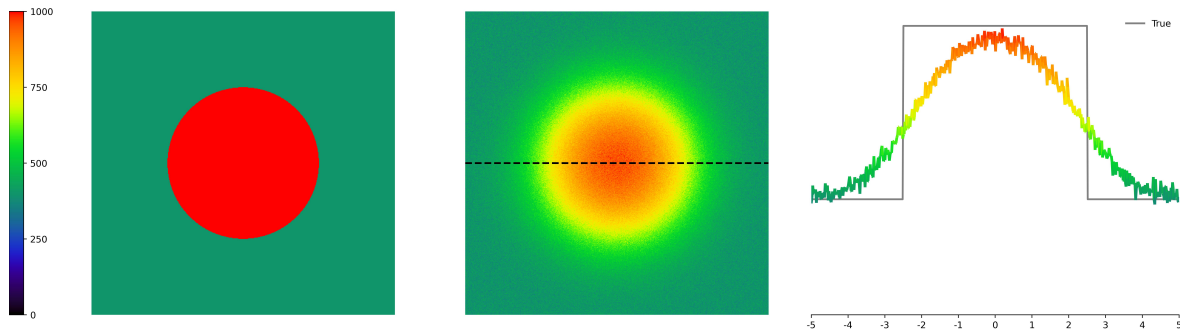

Figure 1.2: Partial Volume Effect shown on a uniform hot disk: left, the ground-truth disk; middle: the result of PET point spread function blurring and noise; right: intensity profile along the horizontal dashed-line compared to the ground truth (gray line), showing the spill-in and spill-out effects

1.5 Background

Many methods have utilized the first few frames of the dynamic PET, when the tracer has not yet reached tissue and predominantly resides in the arteries, to obtain a mask of the carotids [4]. These implementations range from fully manual approaches [5], to semi-automatic methods that use a few seed voxels for region growing or to initialize morphological operations [6, 7], and to fully automatic pipelines based on deep learning such as custom Convolutional Neural Networks (CNNs) [8, 9] and U-Net architectures [10].

However, even with sophisticated approaches, ICA segmentation directly from PET suffers from strong PVE. With the emergence of hybrid PET/MRI systems, 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 studies have

leveraged this to achieve accurate ICA segmentation using manual or semi-automatic procedures, typically relying on high-intensity thresholding and seeded or unseeded region growing [5, 11, 12, 13].

Unlike T1-weighted MRI, TOF-MRA is not always included in research and clinical protocols. Meanwhile, T1-weighted scans are routinely acquired for anatomical localization and image registration, among other purposes, which makes them a practical source for arterial segmentation when TOF-MRA is not part of the protocol. In this context, Rahman et al. [14] proposed a deep learning model that infers the ICA directly from T1-weighted images.

As noted above, even with a high-resolution anatomical arterial mask, PVE remains and must be corrected. Partial Volume Correction (PVC) methods estimate spill-in and spill-out coefficients between the ICA and surrounding tissue. Recovery Coefficients (RC) are commonly used, obtained by scanning cylindrical phantoms of various diameters and matching the ICA diameter to the closest phantom [6, 15]. The Geometric Transfer Matrix (GTM) method generalizes this by estimating regional spillover based on volume geometry and an explicit point spread function (PSF) model [16]. However, these linear mixing models do not account for additional effects such as time-varying noise or motion, and they typically treat the problem as a fixed linear unmixing without propagating uncertainty, which can limit recovery.

More advanced approaches include model-based matrix factorization that jointly estimates the input function and tissue activity [17]; non–deep-learning strategies that infer the blood curve without explicit vessel segmentation—such as factor analysis / cluster-IDIF, population-based input functions, and optimization-derived input functions [18, 15, 19]; and deep learning [20].

In this work, we propose a fully automatic pipeline for extracting the ICA mask from TOF-MRA. We then employ a Bayesian framework that couples GTM-based mixing with priors on the input function and tissue kinetics, explicitly modeling noise to improve IDIF estimation [21].

Materials and Methods

2.1 Dataset Description

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

Both datasets used the same imaging protocol. Using a Siemens Healthineers Biograph mMR simultaneous MR–PET system, an arterial time-of-flight MR angiography (TOF-MRA) was acquired in axial orientation with a voxel size of 0.3×0.3×0.7 mm, and a T1-weighted MRI was acquired in axial orientation with isotropic 1 mm voxels, during the 90-minute dynamic PET. Raw PET data were rebinned into 24 time frames (variable-length frames: 8×15 s, 3×60 s, 5×120 s, 1×300 s, 7×600 s) sinograms for dynamic reconstruction. Reconstruction yielded a voxel size of 1.04×1.04×2.08 mm^3 in a matrix of 344×344×127 voxels.

In the $[^{18}F]$FDG study, whole-blood and plasma AIFs were measured from 26 manually collected arterial samples (timing: every 5 s for the first minute, every 15 s until the second minute, and at 3, 5, 10, 20, 30, 45, 60, 75, 80, 85, and 90 minutes post-injection) and counted in a gamma counter. In the $[^{11}C]$Yohimbine study, 25 arterial blood samples were manually collected (timing: every 5 s for the first minute, every 10 s until the second minute, and at 5, 10, 30, 45, 60, and 90 minutes post-injection). The blood samples were counted in a gamma counter, then centrifuged to separate plasma, and the plasma activity was measured to compute the plasma fraction as the ratio of plasma to whole-blood activity at each time point.

2.2 Pre-processing

For both studies, the T1-weighted image was registered to the average PET image and to the TOF-MRA using the NiftyReg program, an affine registration method [22]. The two resulting affine transformation matrices were then composed to register the TOF-MRA directly to the PET space. Even though patients in the $[^{18}F]$FDG study were unconscious and PET and MRI were acquired simultaneously in the same session, we performed this step out of

caution to eliminate the possibility of misregistration. The T1-weighted image served as an intermediary because it shares anatomical features with both modalities, whereas directly registering TOF-MRA to PET is impractical due to PET's lower spatial resolution and the limited axial field-of-view of TOF-MRA. Finally, the brain in the T1-weighted image was segmented into regions of interest (ROIs) using the Hammersmith brain atlas [23] to obtain regional masks, and these masks were applied to the dynamic PET to extract regional TACs.

2.3   Carotid Segmentation

Figure 2.1 summarizes the carotid segmentation pipeline. Because vessels appear hyperintense in TOF-MRA, a high-intensity threshold can extract arterial structures. However, lesions—which were common in the comatose cohort—and venous structures can also appear hyperintense, and may be inadvertently selected by thresholding. To exclude these, a cuboid volume of interest (VOI) was defined in a reference space to cover the anatomical region where the ICAs are most likely to appear (Step I). The threshold was set to the 1st percentile of all non-zero voxels in the image and applied only within the VOI (Step II). A 3D connected-components filter was then applied, and the two largest connected components were retained as the left and right internal carotid arteries (Step III).

The reference image was the standard MNI152 atlas, padded by 50% in the inferior direction (negative z in voxel space) because the original atlas excludes the subcranial region relevant for the ICAs. The TOF-MRA was registered to this reference using an affine transformation, and the resulting matrix was used to map the VOI into TOF-MRA space before applying the thresholding and connectivity steps.
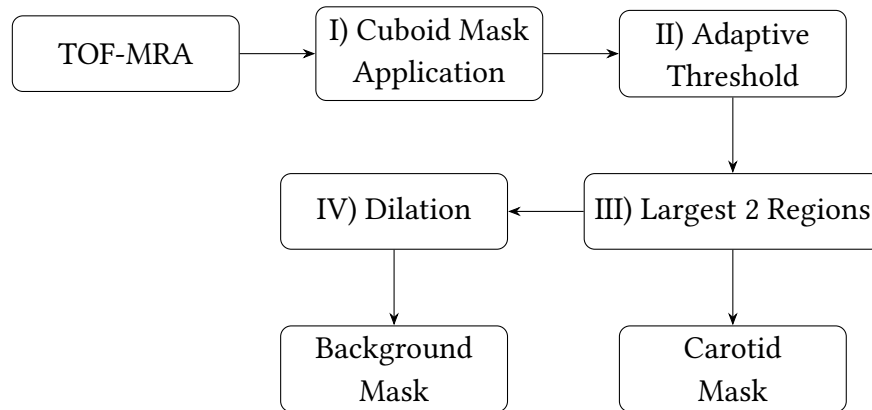

Figure 2.1: Carotid and background mask segmentation pipeline

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 [16].

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).

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

where $\omega_{n \to m}$ 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

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

with χ_n and χ_m denoting the binary masks of regions n and m, respectively, h the system's point spread function, and Ω_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 solved to recover the arterial and background TACs. However, GTM and other classical PVC methods often fail to fully recover the lost signal because they simplify the problem and ignore time-varying noise characteristics. As a result, they may propagate or even amplify noise. Time-varying noise arises from several factors, including the small diameter of the ICAs, very short early frames when the IF changes rapidly, and the exponential decay of the tracer, all of which reduce gamma counts detection by the PET camera and increase variance.

2.4.2 Bayesian Geometric Transfer Matrix

Modelling

For each subject, T_{IF} is modeled as a linear combination of a population mean and principal components obtained by applying principal component analysis (PCA) to AIFs from the cohort. Specifically, for each subject, a subset randomly selected subjects—excluding the subject under study—is used to build a PCA basis. After zero-centering the training AIFs, PCA yields principal axes $\phi_i(t)$ with explained variances λ_i. We define scaled components $v_i(t) = \sqrt{\lambda_i}\,\phi_i(t)$, and write

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

where $\mu_{IF}(t)$ is the population mean and the coefficients satisfy $\theta_i \sim \mathcal{N}(0,1)$ by construction. Here, p denotes the number of retained components.

Spectral analysis (SA) models tissue activity as the convolution of the input function with an impulse response written as a sum of exponentials [24]. This makes it more flexible and generalizable as it does not assume a particular compartment model and is convenient for describing background signal. Therefore, the background TAC is

$$T_{bg}(t) = \sum_{i=1}^{s} \alpha_i \left(T_{IF} \circledast e^{-\beta_i t}\right)(t), \tag{2.4}$$

where $\circledast$ denotes convolution, α_i is the amplitude, and β_i is the decay rate of the ith spectral component.

Noise

Accurate noise modeling is challenging because multiple sources contribute to the variance. Some effects (e.g. scatter, randoms, and motion) are partially corrected during reconstruction post-processing. The main contributor is low-count statistics, an inherent limitation of PET due to detector performance and overall camera sensitivity, which is particularly problematic for IDIF because the ICAs are small, early frames are very short during the bolus passage, and radioactive decay reduces detected events, increasing variance across frames.

While raw PET counts are Poisson-distributed, post-reconstruction noise is typically approximated as Gaussian [**TODO**]. To account for time-varying noise, we use per-frame weights to normalize variance across frames:

$$\omega_i = \frac{\Delta t_i}{c_i}\, \exp\left(-\frac{t_i\,\ln 2}{T_{1/2}}\right), \tag{2.5}$$

where Δt_i is the frame duration, c_i is the net true counts for frame i, and $T_{1/2}$ is the tracer half-life. Because c_i was unavailable, we substitute $C_T(t_i)$, the total reconstructed activity concentration at the frame mid-time.

The time-varying noise level is summarized by a weighted average variance:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \omega_i \, \sigma_i^2. \tag{2.6}$$

Estimation

Let $\mathcal{D}$ denote the observed PET TAC, and let $\Theta = (\theta_1, \ldots, \theta_p, \alpha_1, \beta_1, \ldots, \alpha_s, \beta_s)$ be the unknown model parameters, with an unknown Gaussian noise variance σ^2. We aim to estimate the joint posterior $p(\Theta, \sigma^2 \mid \mathcal{D})$.

By Bayes' rule,

$$p(\Theta, \sigma^2 \mid \mathcal{D}) \propto p(\mathcal{D} \mid \Theta, \sigma^2) \, \pi(\Theta) \, \pi(\sigma^2), \tag{2.7}$$

where $p(\mathcal{D} \mid \Theta, \sigma^2)$ is the likelihood and $\pi(\cdot)$ are the priors. We can therefore sample the posterior once the likelihood and priors are obtained.

The observed data are linked to the latent TACs through GTM. For convenience, define

$$\mathcal{T}(t) = \mathcal{G}(t; \Theta) = \mathcal{G}\left(\begin{bmatrix} T_{IF}(t; \theta_1, \ldots, \theta_p) \\ T_{tissue}(t; \alpha_1, \beta_1, \ldots) \end{bmatrix} \right), \tag{2.8}$$

where $\mathcal{G}$ is the GTM operator and $\mathcal{T}$ is the predicted activity (after PVE). Assuming Gaussian noise with frame weights ω_i, the likelihood is

$$p(\mathcal{D} \mid \Theta, \sigma^2) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\omega_i \left(\mathcal{D}(t_i) - \mathcal{T}(t_i) \right)^2}{2\sigma^2} \right). \tag{2.9}$$

We refer to this framework as the Bayesian Geometric Transfer Matrix (Bayesian GTM or BGTM).

The prior factorizes as

$$\pi(\Theta) = \prod_{i=1}^{p} \pi(\theta_i) \prod_{j=1}^{s} \pi(\alpha_j) \, \pi(\beta_j). \tag{2.10}$$

As noted above, $\theta_i \sim \mathcal{N}(0, 1)$. For the spectral parameters, we adopt broad uniform priors,

$$\alpha_i, \beta_i \sim \mathcal{U}(u_{\min}, u_{\max}). \tag{2.11}$$

For the noise variance, we use an inverse-gamma prior, $\pi(\sigma^2) \sim \Gamma^{-1}(a_0, b_0)$, conjugate to the Gaussian likelihood. We center this prior around an empirical weighted mean squared deviation between the observed carotid TAC and the mean AIF:

$$\mathbb{E}[X] = \frac{1}{N} \sum_{i=1}^{N} \omega_i \left(T_c(t_i) - \mu_{IF}(t_i) \right)^2. \tag{2.12}$$

Given a chosen coefficient of variation (CV), we set

$$a_0 = 2 + \frac{1}{\text{CV}^2}, \qquad b_0 = (a_0 - 1) \, \mathbb{E}[X]. \tag{2.13}$$

Sampling

Direct evaluation of the posterior **??** is intractable, so we employ Monte Carlo sampling. We use a Markov chain Monte Carlo scheme with a Metropolis-within-Gibbs strategy.

Gibbs updates draw each parameter from its univariate conditional posterior $p(\Theta_i \mid \mathcal{D}, \Theta_{-i}, \sigma^2)$ and the noise variance from $p(\sigma^2 \mid \mathcal{D}, \Theta)$. Metropolis–Hastings proposals explore Θ: at iteration m, propose $\Theta'_i = \Theta_i^{(m)} + \varepsilon_i b$, where ε_i is the step size and b is a Brownian perturbation. Accept with probability

$$\min\left(1, \ \frac{p(\Theta'_i \mid \mathcal{D}, \Theta_{-i}, \sigma^2)}{p(\Theta_i^{(m)} \mid \mathcal{D}, \Theta_{-i}, \sigma^2)}\right). \tag{2.14}$$

Because the inverse-gamma prior is conjugate, the conditional posterior for σ^2 is also inverse-gamma,

$$p(\sigma^2 \mid \mathcal{D}, \Theta) \sim \Gamma^{-1}(a, b), \tag{2.15}$$

with

$$a = a_0 + \frac{N}{2}, \qquad b = b_0 + \frac{1}{2}\sum_{i=1}^{N} \omega_i \left(\mathcal{D}(t_i) - \mathcal{T}(t_i)\right)^2. \tag{2.16}$$

Thus, σ^2 can be sampled directly without Metropolis–Hastings.

To improve mixing, each parameter k has its own proposal step size ϵ_k. During burn-in, step sizes are adapted every $L = 50$ Metropolis–Hastings updates by an arbitrary acceptance rate of 0.5: if the empirical rate exceeds 0.5, increase $\epsilon_k \leftarrow 1.1\,\epsilon_k$; otherwise decrease $\epsilon_k \leftarrow 0.9\,\epsilon_k$. After burn-in, step sizes are fixed and subsequent samples are retained.

Inference

After sufficient sampling, we use a modified maximum a posteriori (MAP) estimate by averaging the parameters of the top 0.1% of samples ranked by posterior probability and reporting this average as the solution.

2.5   Plasma Fraction Correction

By construction, IDIFs extracted from PET represent whole-blood activity. For [^{18}F]FDG, plasma metabolites are generally negligible over the scan, so the whole-blood AIF was used without plasma or metabolite correction [25]. For [^{11}C]Yohimbine, kinetic modeling requires the plasma parent activity, so whole-blood IDIFs were converted to plasma and corrected for metabolites using

$$C_{P,\text{parent}}(t) = \frac{C_{WB,\text{IDIF}}(t)}{R_{bp}} \times \text{PPF}(t), \tag{2.17}$$

where $R_{bp} = 0.661$ is the whole-blood–to–plasma ratio and $\mathrm{PPF}(t)$ is the plasma parent fraction modeled as a 3 exponential curve: [26]

$$\mathrm{PPF}(t) \;=\; 0.30\, e^{-t/0.88} \;+\; 0.09\, e^{-t/11.49} \;+\; 0.61\, e^{-t/65.77}. \tag{2.18}$$

Because V_T estimation uses total plasma parent activity, no additional free-fraction correction was applied.

2.6 Evaluation

2.6.1 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.

$$\mathrm{cAUC}(t) = \int_0^t IF(\tau)\, d\tau, \tag{2.19}$$

where IF is the input function.

2.6.2 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. Quantification was performed using graphical methods implemented in TPCCLIB [27]. The Patlak plot was used for $[^{18}\mathrm{F}]$FDG [3], and the Logan plot for $[^{11}\mathrm{C}]$Yohimbine [2].

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

$$\text{Average MAPE}(\mathrm{MR_{glu}}) = \frac{100}{N} \sum_{i=1}^{N} \left(\frac{1}{N_{\mathrm{ROI}}} \sum_{j=1}^{N_{\mathrm{ROI}}} \left| \frac{\mathrm{MR}_{\mathrm{glu},ij}^{\mathrm{IDIF}} - \mathrm{MR}_{\mathrm{glu},ij}^{\mathrm{AIF}}}{\mathrm{MR}_{\mathrm{glu},ij}^{\mathrm{AIF}}} \right| \right), \tag{2.20}$$

where N is the number of subjects.

Similarly, for $[^{11}\mathrm{C}]$Yohimbine, the volume of distribution (V_T) was calculated and the dataset error was computed as

$$\mathrm{MAPE}(V_T) = \frac{100}{N} \sum_{i=1}^{N} \left(\frac{1}{N_{\mathrm{ROI}}} \sum_{j=1}^{N_{\mathrm{ROI}}} \left| \frac{V_{T,ij}^{\mathrm{IDIF}} - V_{T,ij}^{\mathrm{AIF}}}{V_{T,ij}^{\mathrm{AIF}}} \right| \right). \tag{2.21}$$

Additionally, linear least-squares regression was performed between the regional $\mathrm{MR_{glu}}$ and V_T obtained using AIF and IDIF for each subject. The coefficient of determination (R^2) and regression slope (S) were computed per subject. The mean absolute percentage errors of these metrics across the dataset are

$$\mathrm{MAPE}(R^2) = \frac{100}{N} \sum_{i=1}^{N} \left| R_i^2 - 1 \right| \tag{2.22}$$

and

$$\mathrm{MAPE}(S) = \frac{100}{N} \sum_{i=1}^{N} \left| S_i - 1 \right|, \tag{2.23}$$

where R_i^2 and S_i denote the coefficient of determination and slope for subject i, respectively.

2.7 Simulation

PET simulations are integral to validating statistical and analytical PET methods because they provide ground truth that may be unavailable or affected by human or instrumental error in experimental data. In the previous section, we treated the AIF as the ground truth and evaluated methods against it. However, because AIF measurement and analysis involve multiple manual steps and devices operated by different personnel, errors can occur.

Therefore, as a complementary validation, we conducted realistic simulations and evaluated the method on these data as well. PET-SORTEO is a Monte Carlo PET simulation platform [28, 29] whose accuracy has been validated previously [30] and can closely approximate real acquisitions. This fidelity strongly depends on accurate descriptions of both the physical anatomy and tracer kinetics.

As input, PET-SORTEO requires the description and geometry of the PET machine, a 3D emission phantom segmenting the tissues into emitting regions, a 3D attenuation phantom describing tissue attenuation coefficients, and TACs for each emitting region. For this study, we simulated the $[^{18}\mathrm{F}]$FDG dataset using the Biograph mMR model to mirror the experimental datasets, and we derived numerical phantoms and TACs from the real MR and PET data.

For each subject, first, the brain atlas subregions were merged into larger regions: white matter (WM), gray matter (GM), gray nuclei (GN), cerebellar gray matter (CGM), and cerebellar white matter (CWM). An algorithm originally developed for pseudo-CT generation from T1-weighted MRI for bone attenuation correction was used to create masks for the extracerebral regions[31]. Values above 300 were labeled as bone (BONE) and values between -300 and 300 after excluding the brain as extracerebral soft tissues (SOFT) which includes the scalp, neck muscles, glands, eyes, etc. The leftover space between the skull and the brain was labeled as cerebrospinal fluid (CSF). Finally, the arterial mask (ICA) was obtained from the TOF-MRA segmentation (Section 2.3). These nine regions were combined to form the emission phantom.

The attenuation phantom was generated from the pseudo-CT by segmenting SOFT and BONE and assigning their respective attenuation coefficients.

The emission phantom was then registered to PET space to extract subject-specific TACs. These TACs were corrected for PVE using GTM and subsequently fitted with a 2TCM to denoise and obtain "ideal" TACs. The AIF used for fitting was the subject's measured AIF, itself fitted to an exponential model to remove noise [32].

Several practical considerations were applied. Our SOFT definition is broad; while adequate for simulation, it violates homogeneity assumptions in compartment modeling, so the unfitted TAC was retained for SOFT. The pseudo-CT algorithm was permissive for BONE, leading to overestimated BONE TACs due to spill-in; empirically, BONE activity was scaled by 0.5.

Extracting the correct activity of the CSF with this approach is not feasible since it's a very thin region which is heavily affected by PVE. Thus, CSF TAC was considered 0.2 of the fitted WM TAC.
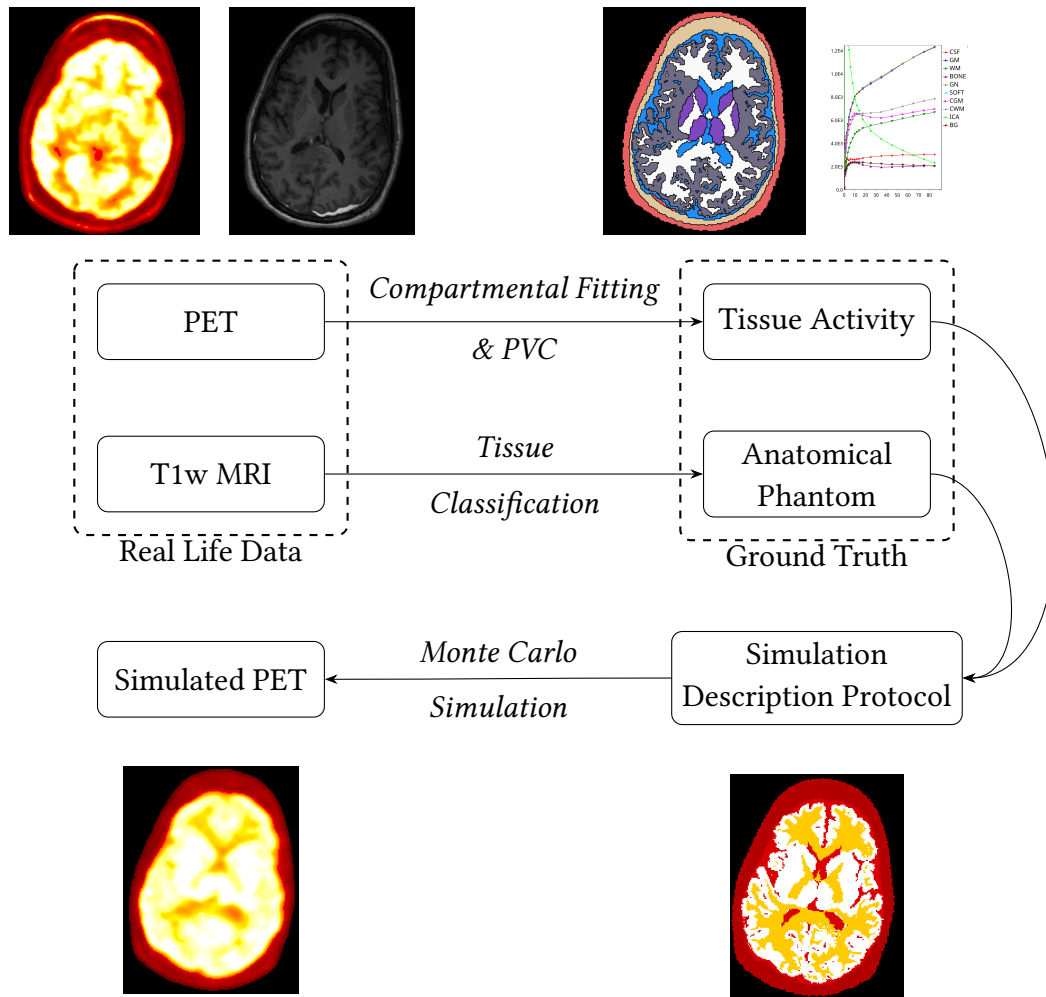

Figure 2.2

Results

3.1 Carotid Segmentation from TOF-MRA

As illustrated in Figure 3.2, the cuboid mask plays a crucial role in carotid segmentation. Because no ground-truth segmentation is available, results were evaluated by visual inspection. An example 3D rendering of a segmented ICA is shown in Figure 3.1.
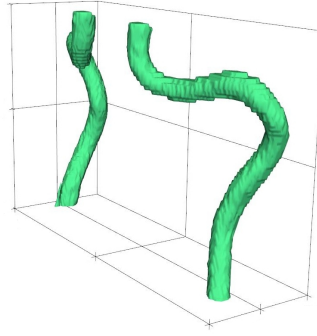

Figure 3.1: Three-dimensional visualization of the segmented internal carotid arteries.
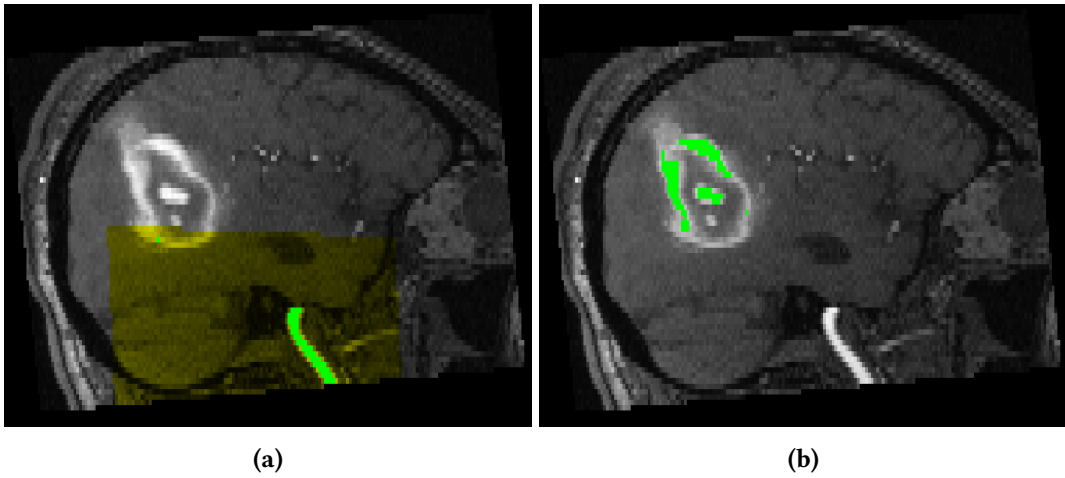

(a)	(b)

Figure 3.2: Comparison of carotid segmentation (green) with (a) and without (b) a cuboid mask (yellow).

3.2   Hyperparameter Tuning

The number of principal components p was set to 3 based on the cumulative explained-variance curve in Figure 3.3. With three components, the cumulative explained variance exceeds 90%. To assess the minimal population size needed to build the PCA model in the $[^{18}\text{F}]$FDG dataset, the analysis was repeated ten times with different random seeds for $N \in \{5, 10, 15, 20, 30, 40, 50, 60\}$, and median quantification errors were compared. As shown in Figure 3.4, performance plateaued at $N = 20$, which was therefore adopted. For the $[^{11}\text{C}]$Yohimbine dataset, the maximum feasible $N = 6$ was used given the cohort size of seven subjects.

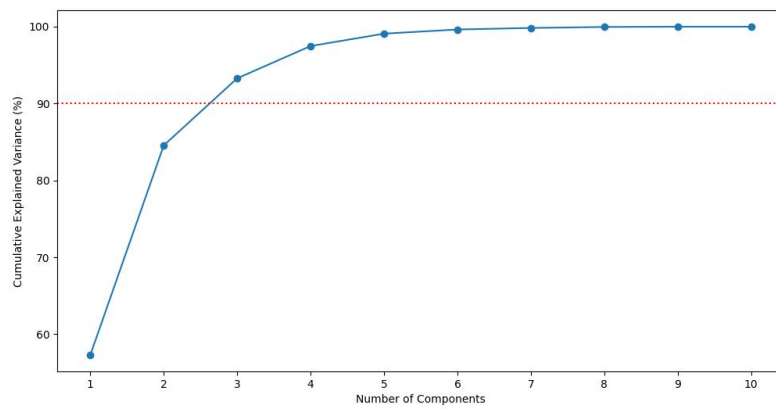**Figure 3.3:** Cumulative explained variance as a function of the number of PCA components.
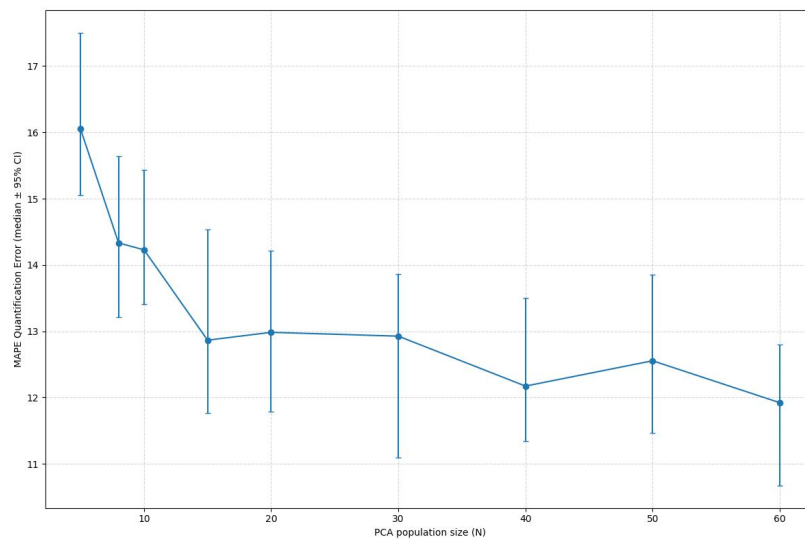

Figure 3.4: Median quantification error (MAPE) with 95% confidence intervals versus PCA population size.

In classical spectral analysis, the spectral range is chosen according to the isotope half-life (e.g., $10^{-4}\,\text{s}^{-1}$ to $1\,\text{s}^{-1}$ for ^{18}F) and discretized into $s = 100$ or $s = 1000$ fixed decay

rates β, while amplitudes α are fitted to the TAC [**TODO**]. Peaks in the resulting spectrum correspond to dominant kinetic components of the TAC. Estimating hundreds of amplitudes with MCMC would be impractical, so a compact basis is preferred in which both decay rates and amplitudes are learned. To select the basis size, we assumed the background kinetics could be approximated by the impulse response of either a 1TCM or a 2TCM, which entail one or two exponential basis functions, respectively. With two basis functions, the sampler frequently converged to nearly identical decay rates or pushed one rate to infinity, indicating a negligible contribution. Accordingly, a single exponential basis function was used for both datasets. As discussed in Section **??**, the PCA weights were assigned the prior $\theta_i \sim \mathcal{N}(0, 1)$. The spectral parameters were given weakly informative uniform priors $\alpha, \beta \sim \mathcal{U}(10^{-5}, 10^{-2})$ due to the lack of stronger prior knowledge.

Each Markov chain was run for 230,000 iterations with a 30,000-iteration burn-in to ensure convergence.

3.3 Simulation

Of the 59 [^{18}F]FDG subjects, 24 were excluded because large tumors or post-surgical skull openings caused the tissue-classification algorithm to fail. The remaining 35 subjects were used to build numerical phantoms and to generate TACs for the simulation protocol. The PET-SORTEO platform generated realistic PET in sinogram format based on these protocols. Sinograms were reconstructed with the same software and settings as the experimental data to ensure consistency.

The simulations were intended for quality control of the pipeline rather than replication of individual scans, and comparisons to real data were used only as sanity checks. Figure 3.5 shows sagittal slices from the second and last frames for a representative subject, including the real PET, the simulation input, and the simulation output. Less detail is visible in the neck region because we did not model the full arterial tree, the venous system, or glands. The early frame (top row) roughly coincides with the AIF peak and the activity in ICA is clearly visible. In the late frame, when the tissues reach steady state, the brain activity seems realistic and we see an adequete contrast between regions such as WM and GM.

Figure 3.6 compares a partial-volume−corrected TAC from the simulation output with its simulation input. As expected, there are small differences due to noise in larger regions such as GM and the cerebellum, and there is a clear bias in the ICA because it is strongly affected by partial-volume effects.

3.4 IDIF

Because BGTM builds on GTM and incorporates a population AIF prior, performance was compared with GTM-derived IDIF and with a PBIF defined as the mean AIF used in the PCA (μ_{IF}).
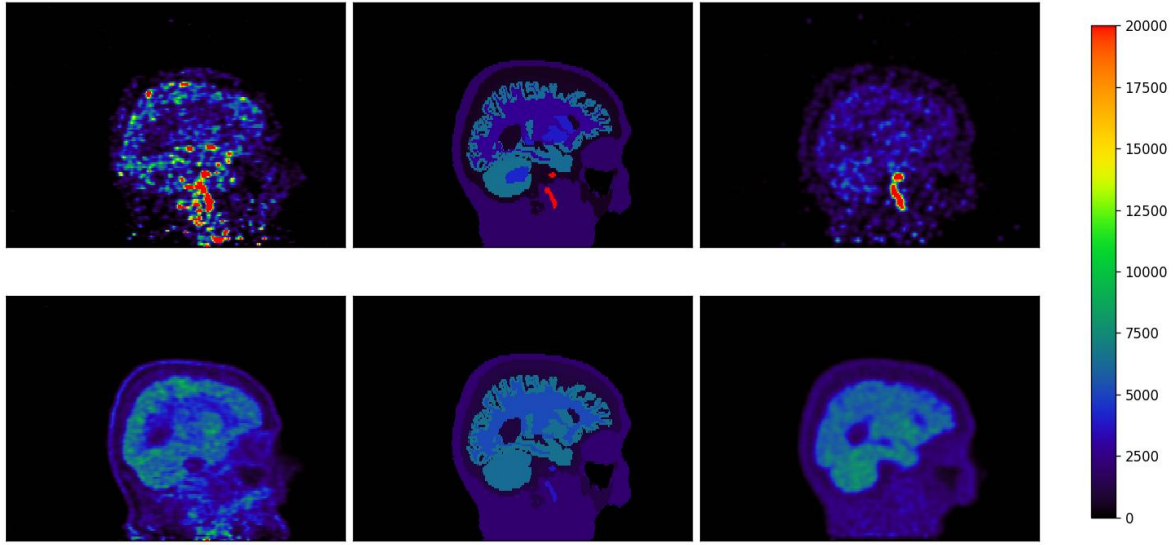
Figure 3.5: Sagittal slices from the second frame (top) and last frame (bottom) for real PET (left), simulation input (middle), and simulation output (right)

3.4.1 $[^{18}\text{F}]$**FDG Dataset**

The reported statistics correspond to the experiment yielding the median validation performance from the fine-tuning stage. Figure 3.7 illustrates best- and worst-case subjects. In the better-performing subject, BGTM outperforms GTM and PBIF (MR_{glu} MAPE of 1.2% versus 26% and 1.6%, respectively), whereas in the poorer-performing subject BGTM underperforms relative to GTM and PBIF (MR_{glu} MAPE of 36% versus 2.2% and 4.75%, respectively).

The average cAUC MAE across subjects was 13,024 for BGTM and 15,709 and 14,630 for GTM and PBIF, respectively (Figure 3.8b). In quantification, BGTM achieved lower errors than GTM and PBIF. Specifically, the average MR_{glu} MAPE was 13%, 24%, and 17% for BGTM, GTM, and PBIF, respectively (Figure 3.8a), and the average MR_{glu} MAE was 1.04, 1.98, and 1.40. The MAPE for the coefficient of determination (R^2) and the regression slope (S) was 1.75% and 11.2% for BGTM, versus 4.1% and 17% for GTM and 1.5% and 15.75% for PBIF, respectively.

Paired t-tests were used to compare BGTM with GTM and PBIF on quantification metrics (Table 3.1). Relative to GTM, BGTM showed significantly lower MR_{glu} MAE ($t = -5.934, p = 1.754 \times 10^{-7}$), MR_{glu} MAPE ($t = -5.459, p = 1.042 \times 10^{-6}$), R^2 absolute percentage error (APE) ($t = -2.762, p = 7.677 \times 10^{-3}$), and slope APE ($t = -4.030, p = 1.644 \times 10^{-4}$). Relative to PBIF, BGTM also yielded significantly lower MR_{glu} MAE ($t = -2.974, p = 4.275 \times 10^{-3}$), MR_{glu} MAPE ($t = -2.415, p = 1.890 \times 10^{-2}$), and slope APE ($t = -3.036, p = 3.583 \times 10^{-3}$), while the difference in R^2 APE was not significant ($t = 1.680, p = 9.832 \times 10^{-2}$).
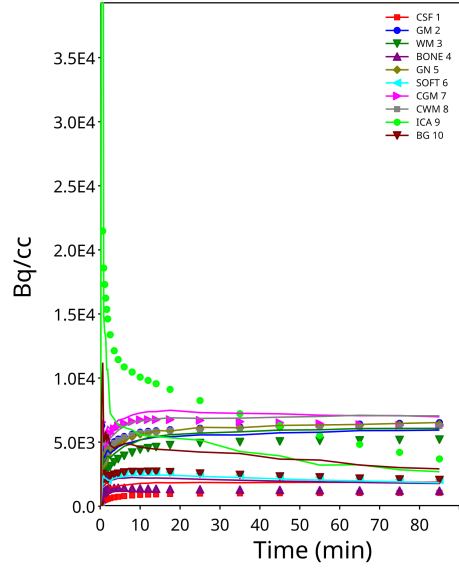
Figure 3.6: TAC Comparison of the simulation input (dots) and the partial-volume–corrected output (lines) of an example subject

3.4.2 $[^{11}C]$**Yohimbine Dataset**

Performance on the $[^{11}C]$Yohimbine dataset was lower than on $[^{18}F]$FDG. Figure 3.9 compares best- and worst-performing subjects. Across subjects, the average cAUC MAE was 55,499 for BGTM and 96,362 and 37,613 for GTM and PBIF, respectively (Figure 3.10b). By Logan analysis, the average V_T MAPE was 75.9%, 166.5%, and 29.5% for BGTM, GTM, and PBIF, respectively (Figure 3.10a).

3.4.3 **Simulated Dataset**

As a quality-control check on the simulations, the Direct IDIF (no PVC) from the experimental data was compared with the Direct IDIF from the simulated data. The simulated dataset showed a higher cAUC MAE than the experimental dataset (85,138 versus 34,929) and a higher AUC APE (46% versus 22%). Figure 3.13 shows the temporal error for one subject, where the simulated Direct IDIF underestimates activity relative to the experimental Direct IDIF.

This underestimation propagates to other IDIF methods and increases bias. Figure 3.11 shows best- and worst-case subjects. BGTM provided only a slight improvement in cAUC MAE relative to GTM (53,997 versus 57,991), and it performed worse than PBIF, which achieved an average cAUC MAE of 23,072 (Figure 3.12b). In absolute quantification, BGTM underperformed both comparators, with an average MR_{glu} MAPE of 48% versus 35.6% for GTM and 18.5% for PBIF. Paired t-tests showed no improvement over GTM for cAUC MAE ($t = -1.353$, $p = 0.185$) and statistically worse performance than PBIF for cAUC MAE ($t = 9.013$, $p = 1.56 \times 10^{-10}$), RMSE ($t = 8.805$, $p = 2.73 \times 10^{-10}$), and AUC APE ($t = 8.317$, $p = 1.04 \times 10^{-9}$).

Figure 3.7: Comparison of the IFs (a,d), cumulative AUC curves (b,e), and MR_{glu} regression lines (c,f) for one better-performing (top) and one poorer-performing (bottom) subject in the $[^{18}F]FDG$ dataset.

(a) MR_{glu} MAPE boxplot.   **(b)** cAUC MAE boxplot.
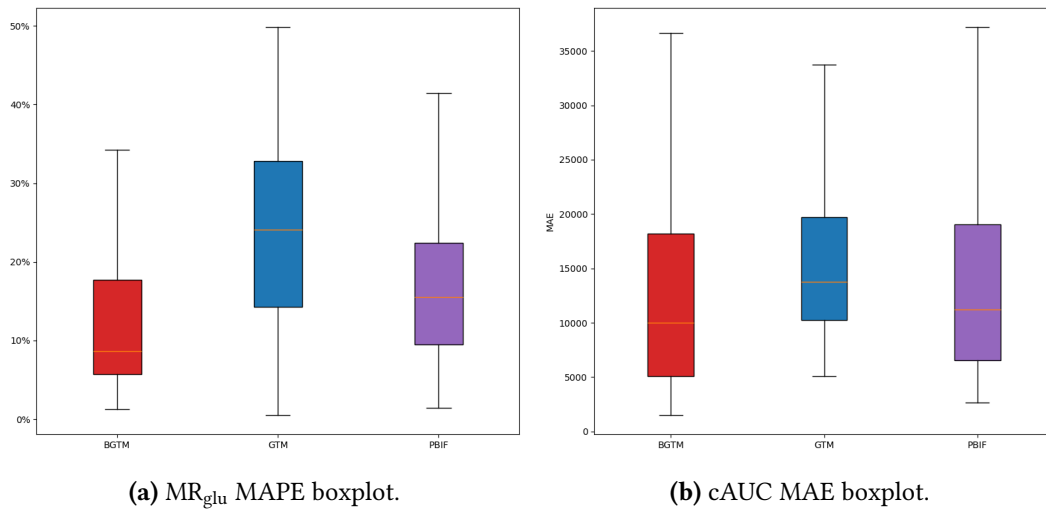
Figure 3.8: Boxplots of curve and quantification errors for the $[^{18}F]FDG$ dataset.
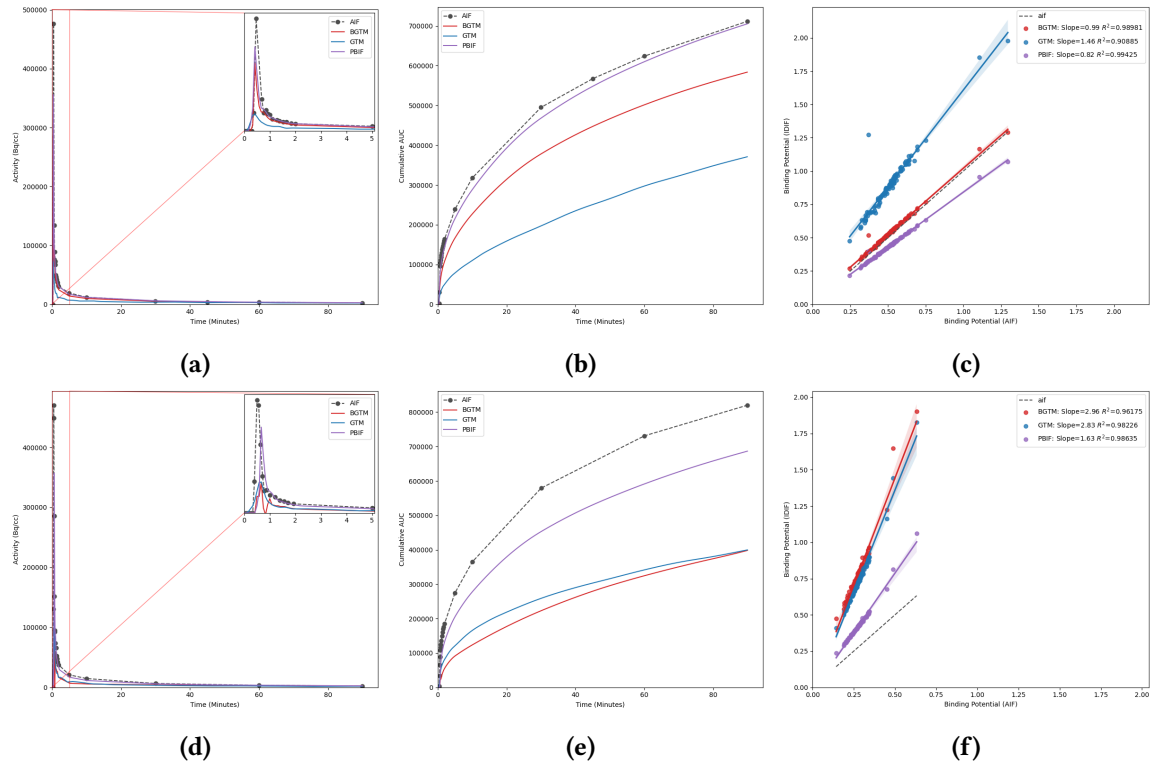
Figure 3.9: Comparison of the IFs (a,d), cumulative AUC curves (b,e), and V_T regression lines (c,f) for one better-performing (top) and one poorer-performing (bottom) subject in the $[^{11}C]$Yohimbine dataset.

(a) V_T MAPE boxplot.   **(b)** cAUC MAE boxplot.
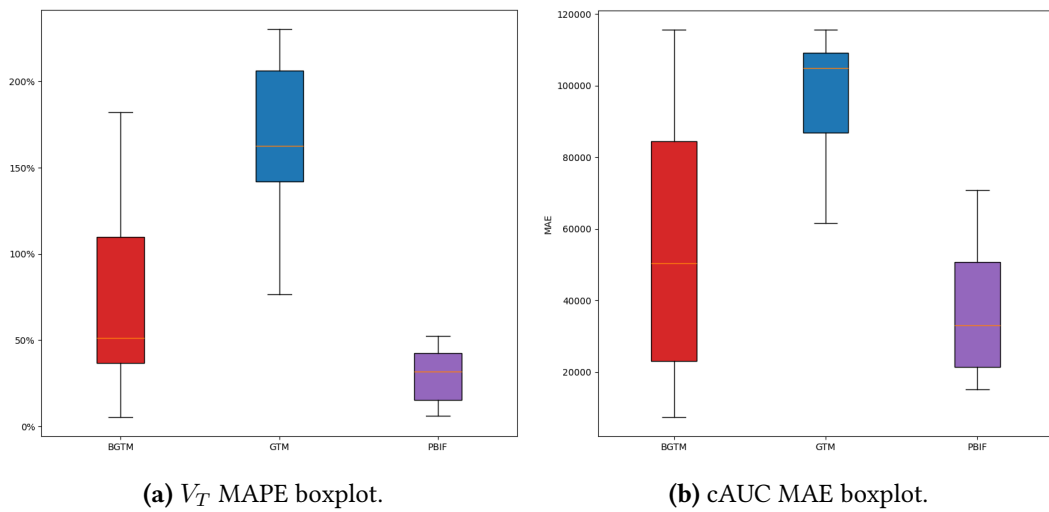
Figure 3.10: Boxplots of curve and quantification errors for the $[^{11}C]$Yohimbine dataset.
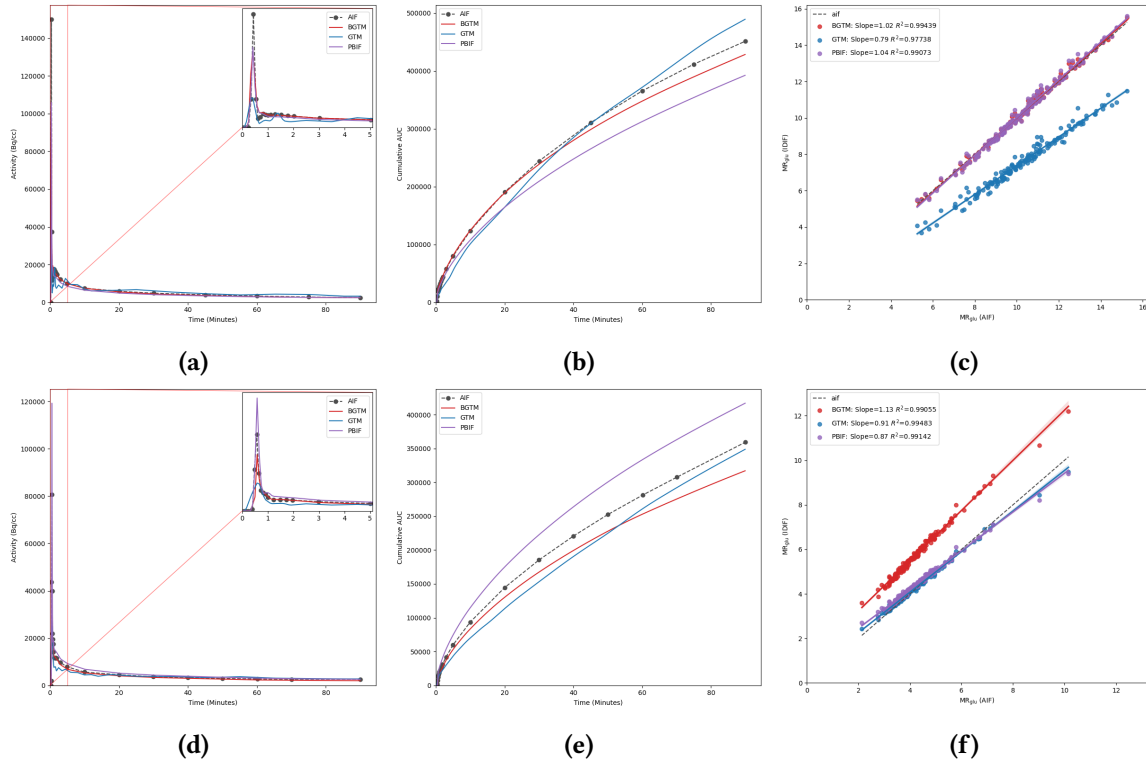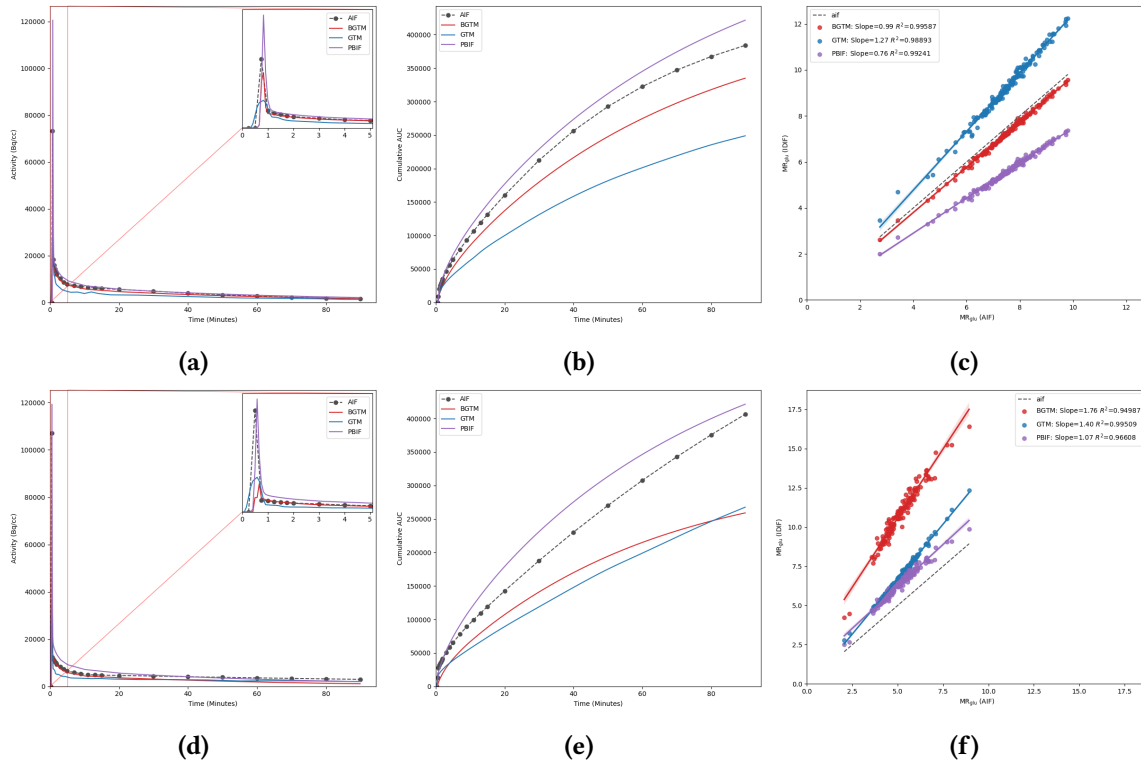
Figure 3.11: Comparison of the IFs (a,d), cumulative AUC curves (b,e), and MR_{glu} regression lines (c,f) for one better-performing (top) and one poorer-performing (bottom) simulated subject.


(a) MR_{glu} MAPE boxplot.

(b) cAUC MAE boxplot.

Figure 3.12: Boxplots of curve and quantification errors for the simulated dataset.

Figure 3.13: Relative temporal bias (%) of Direct IDIF in one experimental subject versus its simulated counterpart.

[^{18}F]FDG Dataset

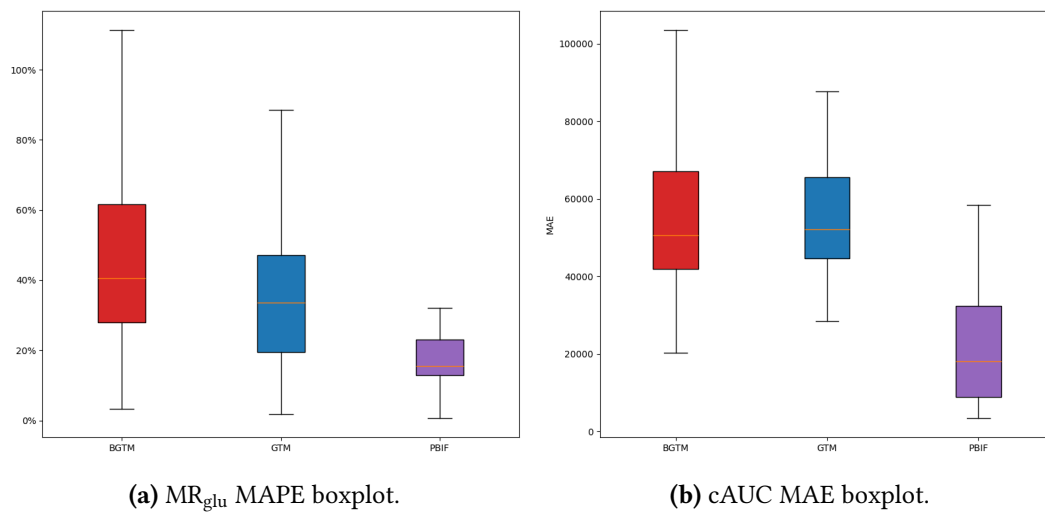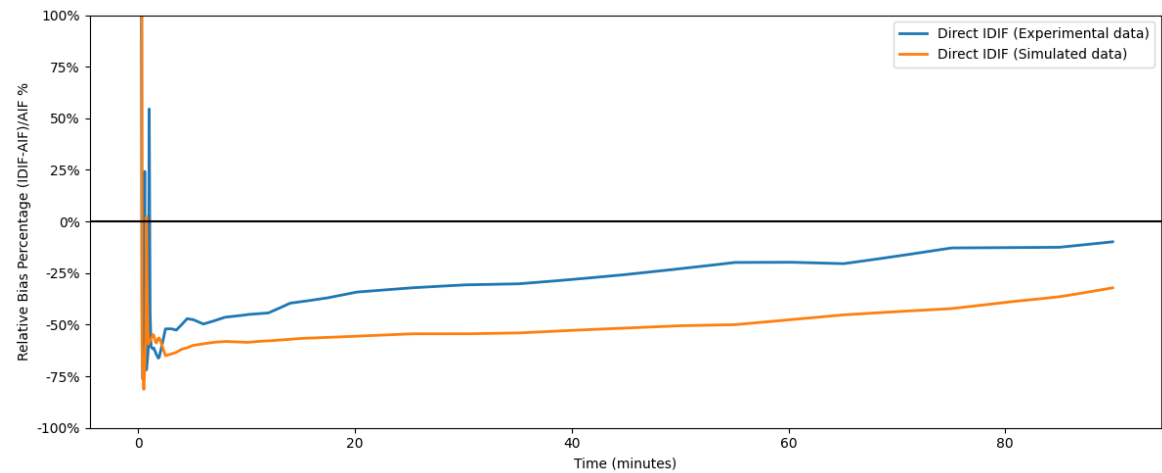| Metric	BGTM		GTM		PBIF		BGTM vs GTM		BGTM vs PBIF	
	μ	σ	μ	σ	μ	σ	t	p	t	p
IF cAUC MAE	13,024	10,410	15,709	7,884	14,630	10,268	-2.388	2.023×10^{-2} *	-1.278	2.065×10^{-1}
IF cAUC RMSE	20,293	16,078	23,380	11,979	23,334	16,803	-1.907	6.151×10^{-2} †	-1.439	1.555×10^{-1}
IF AUC APE (%)	11.93	10.76	11.57	11.50	14.34	11.08	0.297	7.674×10^{-1}	-1.862	6.767×10^{-2} †
MR$_{\text{glu}}$ MAPE (%)	13.19	10.57	24.36	13.68	17.17	12.41	-5.459	1.042×10^{-6} ***	-2.415	1.890×10^{-2} *
MR$_{\text{glu}}$ MAE	1.04	0.92	1.98	1.33	1.40	1.10	-5.934	1.754×10^{-7} ***	-2.974	4.275×10^{-3} **
R^2 APE (%)	1.76	4.07	4.13	8.90	1.50	3.53	-2.762	7.677×10^{-3} **	1.680	9.832×10^{-2}
Slope APE (%)	11.28	9.37	17.05	10.58	15.76	11.41	-4.030	1.644×10^{-4} ***	-3.036	3.583×10^{-3} **

Table 3.1: Comprehensive curve and quantification metrics for the [^{18}F]FDG dataset. Values are means (μ) and standard deviations (σ) across subjects. Paired two-sided t-tests compare GTM or PBIF against BGTM for each metric. Significance codes: * p<0.05, ** p<0.01, *** p<0.001, † p<0.10 (trend).

[¹¹C]Yohimbine Dataset

Metric	BGTM		GTM		PBIF		BGTM vs GTM		BGTM vs PBIF	
	μ	σ	μ	σ	μ	σ	t	p	t	p
IF cAUC MAE	55,499	40,271	96,362	20,343	37,613	20,456	-3.404	1.443×10^{-2}*	0.959	3.748×10^{-1}
IF cAUC RMSE	83,282	60,130	147,032	28,318	52,313	27,845	-3.789	9.083×10^{-3}**	1.196	2.768×10^{-1}
IF AUC APE (%)	26.14	19.02	52.75	3.53	17.74	14.78	-3.644	1.078×10^{-2}*	0.872	4.168×10^{-1}
V_T MAPE (%)	75.91	62.09	166.48	52.91	29.47	17.73	-3.628	1.100×10^{-2}*	2.588	4.133×10^{-2}
V_T MAE	0.324	0.211	0.812	0.360	0.132	0.072	-3.477	1.318×10^{-2}*	3.042	2.275×10^{-2}
R^2 APE (%)	2.04	1.83	2.27	3.09	2.06	2.40	-0.148	8.870×10^{-1}	-0.040	9.692×10^{-1}
Slope APE (%)	81.18	68.33	161.80	73.59	35.06	24.22	-3.167	1.939×10^{-2}*	2.409	5.263×10^{-2}

Table 3.2: Comprehensive curve and quantification metrics for the [¹¹C]Yohimbine dataset. Values are means (μ) and standard deviations (σ) across subjects. Paired two-sided t-tests compare GTM or PBIF against BGTM for each metric. Significance codes: * $p<0.05$, ** $p<0.01$, *** $p<0.001$, † $p<0.10$ (trend).

Simulated Dataset

Metric	BGTM		GTM		PBIF		BGTM vs GTM		BGTM vs PBIF	
	μ	σ	μ	σ	μ	σ	t	p	t	p
IF cAUC MAE	53,997	21,653	57,991	20,621	23,072	19,136	-1.353	1.849×10^{-1}	9.013	1.556×10^{-10}
IF cAUC RMSE	71,835	28,603	79,860	29,010	30,451	25,946	-2.105	$4.279 \times 10^{-2*}$	8.805	2.732×10^{-10}
IF AUC APE (%)	33.35	9.20	36.13	9.46	12.94	10.59	-1.612	1.162×10^{-1}	8.317	1.043×10^{-9}
MR$_{\text{glu}}$ MAPE (%)	48.14	27.09	35.60	22.39	18.56	11.58	2.678	1.133×10^{-2}	6.578	1.545×10^{-7}
MR$_{\text{glu}}$ MAE	4.62	2.13	3.49	2.21	1.82	1.08	2.463	1.901×10^{-2}	7.598	7.919×10^{-9}
R^2 APE (%)	0.35	0.88	0.65	0.88	0.32	0.59	-1.636	1.111×10^{-1}	0.609	5.466×10^{-1}
Slope APE (%)	45.58	23.70	39.97	23.74	16.86	11.37	1.297	2.035×10^{-1}	6.812	7.750×10^{-8}

Table 3.3: Comprehensive curve and quantification metrics for the simulated dataset. Values are means (μ) and standard deviations (σ) across subjects. Paired two-sided t-tests compare GTM or PBIF against BGTM for each metric. Significance codes: [*] p<0.05, [**] p<0.01, [***] p<0.001, [†] p<0.10 (trend).

Discussion

In this study, we sought to replace arterial sampling by estimating an image-derived input function (IDIF) from dynamic brain PET. We used a Bayesian Markov chain Monte Carlo (MCMC) framework that we refer to as Bayesian GTM (BGTM). BGTM models partial-volume effects and time-varying noise while jointly estimating the input function and the surrounding tissue activities.

Accurate IDIF requires a reliable segmentation of an artery. We proposed a fast and fully automatic yet simple segmentation algorithm that extracts the internal carotid arteries from the TOF-MRA image. We placed a simple cuboid mask over the expected ICA location, which allowed high-intensity thresholding to avoid non-arterial hyperintense tissues such as lesion. Because no ground-truth labels were available, evaluation relied on visual inspection. Future work should include expert annotations to enable statistical metrics such as the Dice coefficient. TOF-MRA is not routinely included in many imaging protocols, which limits the immediate applicability of it.

Spectral analysis was an appropriate choice for modeling the background tissue because it does not force a fixed kinetic model and is much more flexible. A fixed one- or two-tissue model can be mismatched when regional kinetics do not follow those forms. The input-function model used a PCA prior, which captured most of the variability with a few coefficients. However, it still relies on an sufficiently large AIF population. A parametric alternative such as the exponential model of Feng, Huang, and Wang [32] could reduce this dependence given tracer-specific informative priors on those parameters are available.

We treated noise as time-varying and used frame-wise weights to variance across time frames. Future work could estimate a separate noise level per frame rather than summarizing the variance with a single statistic.

Bayesian inference rests on the likelihood and the priors. Thus the quality and informativeness of the priors can drastically effect the results. By nature, in the PCA-based input function, the coefficients have a strong informative prior of $\theta_i \sim \mathcal{N}(0, 1)$. If a parametric input function model such as Feng were used instead, parameters would need tracer-specific priors that encode the underlying kinetics. One way to achieve this is to fit the parametric model to a population of AIFs and use the empirical parameter distributions as priors. For the spectral model of background kinetics, we lacked informative prior knowledge and used wide uniform priors. More informative priors could improve accuracy and robustness.

In the [^{18}F]FDG study, a larger PCA population was available and the subjects were

comatose which likely reduced inter-subject variability due to low glucose uptake and also eliminated the existence of motion artifacts. These factors allowed for achieving a high performance by BGTM which was not replicated in the [^{11}C]Yohimbine study. This cohort was small and consisted of healthy young adults with greater inter-subject variability. In addition, for this tracer, plasma parent correction was required and was applied with a population-based curve, which does not capture subject-specific metabolism. These factors likely contributed to the weaker results.

We ran Monte Carlo PET simulations as a complementary evaluation step with special advantage of examination with controlled conditions and known ground truth. The simulation protocol drew anatomy and kinetics from real scans to keep inputs realistic. The goal was quality control of the pipeline rather than replication of individual scans, and comparisons to real data were used only as sanity checks. As discussed in Section **??**, regional activities were broadly in line with the real data, and Figure 3.5 shows that early and late frames look consistent with the intended kinetics. However, Figure 3.6 shows substantial bias in the Direct IDIF for simulated data compared to the experimental data. Naturally, this bias in the propagated in to GTM and therefor BGTM, causing significantly worse performance compared to the experimental [^{18}F]FDG dataset.

Several factors could explain this bias. First, as for reasons pointed at in Section **??**, contrary to other regions, the TTAC of the SOFT region was taken raw from the experimental scan rather than fitting to a compartment model. This caused the activity in the SOFT region to follow a different kinetic profile than the rest of the regions. And since the background tissue lay inside the SOFT region, it has effected its signal and therefor effected the ICA signal by spill-in. A better approach would set plausible amplitudes and decay rates in an SA model for SOFT, generate its TACs from those parameters. This also has the added advantage of being able to compare BGTM's estimated spectral parameters with the known ground truth. Second, we resampled PET-derived TTACs and the fitted AIF ICA curve to mid-frame times to match the experimental framing, which can lose temporal detail. Third, the AIF was fitted with the Feng model using TPCCLIB, and some fits showed unrealistic peaks or fast recirculation, which can propagate bias to all IDIF methods.

To our knowledge, whole-brain Monte Carlo simulation with an explicit carotid input for IDIF validation has not been reported. If the present simulation framework is refined, it could support systematic studies across key factors. These factors include input-function shape (peak height, sharp versus broad peak, slow versus fast decay), background activity level, isotope, frame durations, and scanner geometry. Studying these dimensions would clarify when IDIF methods are reliable, highlight failure modes, and guide improvements to the BGTM implementation for adapting to different studies.

Conclusion

[TODO]

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] Tao Feng et al. "Image-derived and arterial blood sampled input functions for quantitative PET imaging of the angiotensin II subtype 1 receptor in the kidney". In: *Medical physics* 42.11 (2015), pp. 6736–6744.

[6] Praveen Dassanayake et al. "caliPER: a software for blood-free parametric Patlak mapping using PET/MRI input function". In: *Neuroimage* 256 (2022), p. 119261.

[7] Mark B Vestergaard et al. "Validation of kinetic modeling of [15O] H2O PET using an image derived input function on hybrid PET/MRI". In: *Neuroimage* 233 (2021), p. 117950.

[8] Matteo Ferrante et al. "Physically informed deep neural networks for metabolite-corrected plasma input function estimation in dynamic PET imaging". In: *Computer methods and programs in biomedicine* 256 (2024), p. 108375.

[9] Junyu Chen et al. "Deep learning-derived arterial input function". In: *arXiv preprint arXiv:2505.24166* (2025).

[10] Rugved Chavan et al. "An end-to-end deep learning pipeline to derive blood input with partial volume corrections for automated parametric brain PET mapping". In: *Biomedical Physics & Engineering Express* 10.5 (2024), p. 055028.

[11] Lalith KS Sundar et al. "Towards quantitative [18F] FDG-PET/MRI of the brain: automated MR-driven calculation of an image-derived input function for the non-invasive determination of cerebral glucose metabolic rates". In: *Journal of Cerebral Blood Flow & Metabolism* 39.8 (2019), pp. 1516–1530.

[12] Hasan Sari et al. "Estimation of an image derived input function with MR-defined carotid arteries in FDG-PET human studies using a novel partial volume correction method". In: *Journal of Cerebral Blood Flow & Metabolism* 37.4 (2017), pp. 1398–1409.

[13] Thies H Jochimsen et al. "Fully automated calculation of image-derived input function in simultaneous PET/MRI in a sheep model". In: *EJNMMI physics* 3 (2016), pp. 1–17.

[14] Syed Taskeen Rahman et al. *Deep Learning Based Internal Carotid Artery Segmentation of MR Images for PET Image Derived Input Function Estimation.* 2024.

[15] Chul Hyoung Lyoo et al. "Image-derived input function derived from a supervised clustering algorithm: methodology and validation in a clinical protocol using [11C](R)-rolipram". In: *PLoS One* 9.2 (2014), e89101.

[16] 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.

[17] Yu-Hua Dean Fang et al. "Image quantification for TSPO PET with a novel image-derived input function method". In: *Diagnostics* 12.5 (2022), p. 1161.

[18] Urban Simoncic and Paolo Zanotti-Fregonara. "Image-derived input function with factor analysis and a-priori information". In: *Nuclear medicine communications* 36.2 (2015), pp. 187–193.

[19] André H Dias et al. "Clinical validation of a population-based input function for 20-min dynamic whole-body 18F-FDG multiparametric PET imaging". In: *EJNMMI physics* 9.1 (2022), p. 60.

[20] Keisuke Kawauchi et al. "A convolutional neural network-based system to estimate the arterial plasma radioactivity curve in 18F-FDG dynamic brain PET study". In: *Nuclear Medicine Communications* 44.11 (2023), pp. 1029–1037.

[21] 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.

[22] Marc Modat et al. "Global image registration using a symmetric block-matching approach". In: *Journal of medical imaging* 1.2 (2014), pp. 024003–024003.

[23] 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.

[24] Vincent J Cunningham and Terry Jones. "Spectral analysis of dynamic PET studies". In: *Journal of Cerebral Blood Flow & Metabolism* 13.1 (1993), pp. 15–23.

[25] Sanjiv S Gambhir et al. "Simple noninvasive quantification method for measuring myocardial glucose utilization in humans employing positron emission tomography and fluorine-18 deoxyglucose". In: *Journal of Nuclear Medicine* 30.3 (1989), pp. 359–366.

[26]   Chloé Laurencin et al. "Modeling [11C] yohimbine PET human brain kinetics with test-retest reliability, competition sensitivity studies and search for a suitable reference region". In: *NeuroImage* 240 (2021), p. 118328.

[27]   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`.

[28]   Anthonin Reilhac et al. "PET-SORTEO: A Monte Carlo-based simulator with high count rate capabilities". In: *IEEE Transactions on Nuclear Science* 51.1 (2004), pp. 46–52.

[29]   A Reilhac et al. "Validation and application of PET-SORTEO for the geometry of the Siemens mMR scanner". In: *PSMR Conference. Cologne.* 2016.

[30]   Anthonin Reilhac et al. "PET-SORTEO: validation and development of database of simulated PET volumes". In: *IEEE Transactions on Nuclear Science* 52.5 (2005), pp. 1321–1328.

[31]   Inès Mérida. "Development and application of multi-atlas attenuation correction for brain imaging with simultaneous PET-MR". 2017LYSE1116. PhD thesis. 2017. URL: `http://www.theses.fr/2017LYSE1116/document`.

[32]   Dagan Feng, Sung-Cheng Huang, and Xinmin Wang. "Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography". In: *International journal of bio-medical computing* 32.2 (1993), pp. 95–110.