Sequential Monte Carlo and Compartmental Models

John Aston

Statslab, DPMMS, University of Cambridge

Warwick Vacation School, Sept 24, 2015

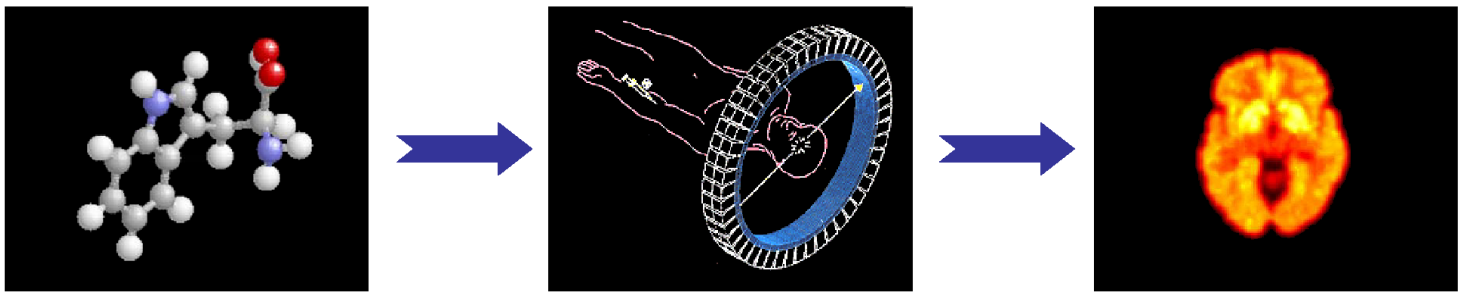
Acknowledgements

- Adam Johansen (Warwick)
- Yan Zhou (NUS)

Outline

Outline

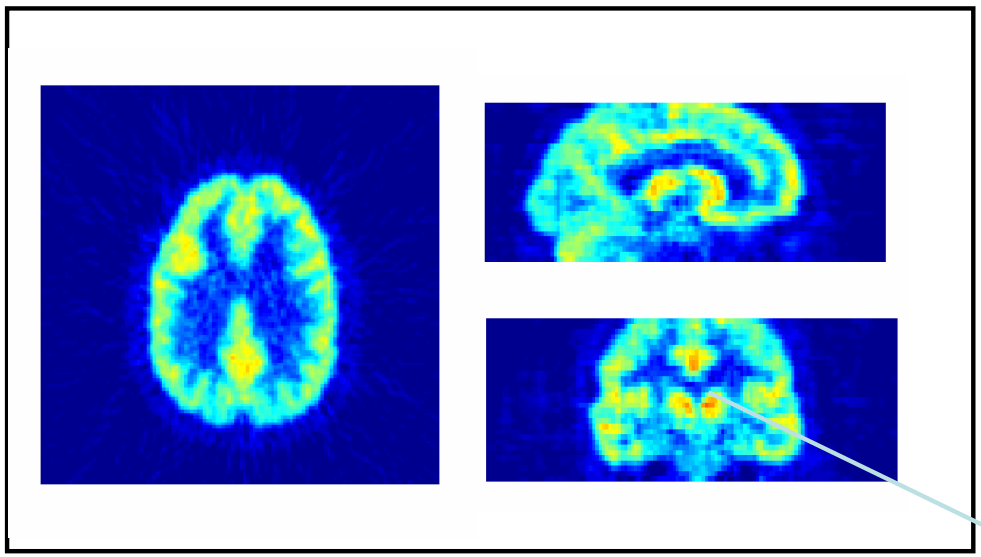
Positron Emission Tomography (PET)



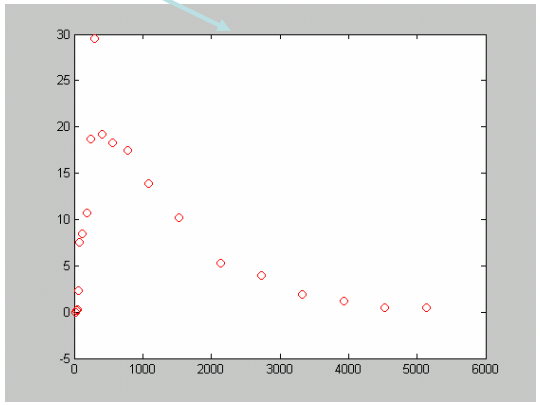
1. Radioactively label compounds
2. Scan Subjects to produce tomographic data
3. Reconstruct tomographic data to produce images

Neuroreceptor Imaging

SCH23390 - Dopamine Imaging



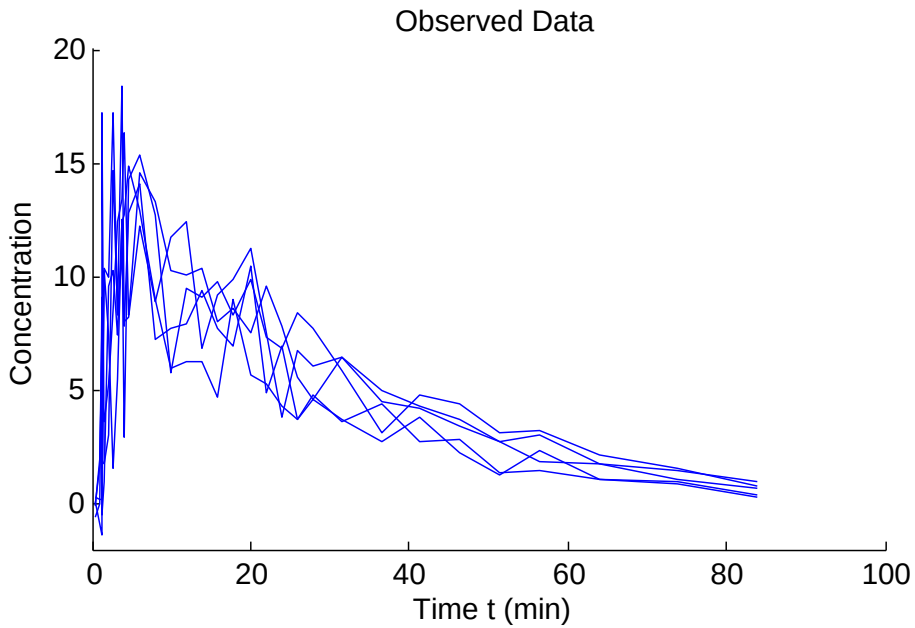
$\approx 10^7$ data points
in a single 4D scan



128x128x31 - Spatial Dim
20 - Temporal Dim

Neuroreceptor Imaging

Diprenorphine - Sample time courses

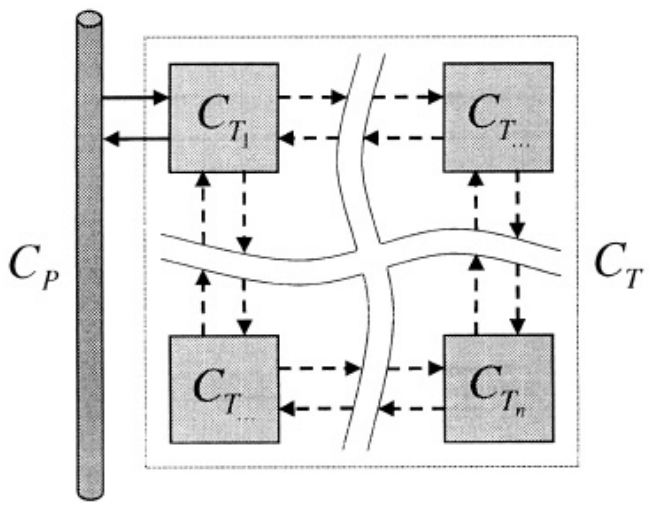


General compartmental model for PET tracers

Model equation

$$C_T(t) = C_P(t) * \sum_{i=1}^{n} \omega_i \exp(-\nu_i t)$$

- $C_T(t)$: Tissue time-activity function
- $C_P(t)$: Blood/Plasma input function
- n: Number of compartments
- ω_i: Weight (relative tracer volume) in compartment i
- ν_i: Diffusion rate of tracer in compartment i
- Both ω_i and ν_i must be positive.



PET compartmental analysis

Volume of distribution

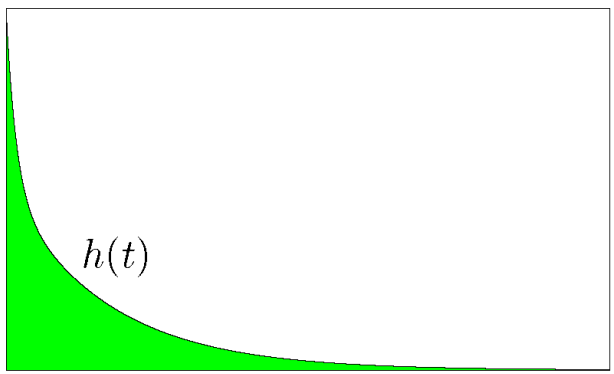
In PET analysis we are interested in the total volume of tracer present in the tissue (volume of distribution):

$$\mathrm{V}_T \equiv \int_0^\infty \sum_{i=1}^{n} \omega_i \exp(-\nu_i t) dt = \sum_{i=1}^{n} \frac{\omega_i}{\nu_i}$$

V_T is defined as the area under

$$h(t) = \sum_{i=1}^{n} \omega_i \exp(-\nu_i t)$$

(green area in the graph)



PET compartmental analysis

Standard "Statistics Text Book" Solution

$$C_T(t) = C_P(t) * \sum_{i=1}^{n} \omega_i \exp(-\nu_i t) + \epsilon$$

Set n equal to a known number (usually 1 or 2 or at most 3). Then you have known solutions to the general equation which can be solved using standard algorithms such as Non Linear Least Squares.

A few problems

- How to choose n
- Algorithms often not very stable
- Algorithms often often take a while to converge (and we have a $10^5 - 10^6$ spatial points to look at).

PET compartmental analysis

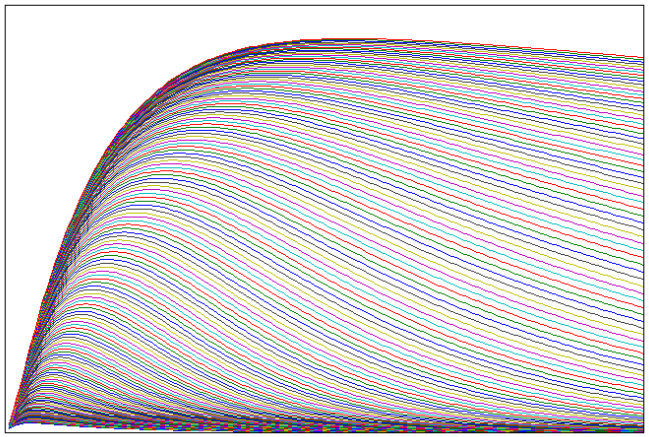
Linear problem formulation

- Define $y_j \equiv \frac{1}{t_j - t_{j-1}} \int_{t_{j-1}}^{t_j} C_T(t)dt$.
- Define $\Phi_{ji} \equiv \frac{1}{t_j - t_{j-1}} \int_{t_{j-1}}^{t_j} \int_0^t C_P(\tau) \exp(-\nu_i(t-\tau))d\tau dt$.

The problem can be reformulated as the linear equation

$$\mathbf{y} = \mathbf{\Phi}\boldsymbol{\omega} + \epsilon,$$

where each column of $\mathbf{\Phi}$ is a basis vector.



Plot of basis vectors

PET compartmental analysis

PET Spectral Analysis (Cunningham and Jones, 1993)

$$\min_{w_i \geq 0,\ 1 \leq i \leq n} \|\mathbf{y} - \mathbf{\Phi w}\|_2^2.$$

Solved using NNLS algorithm (Lawson and Hanson, 1974).

DEPICT (Gunn et al, 2002)

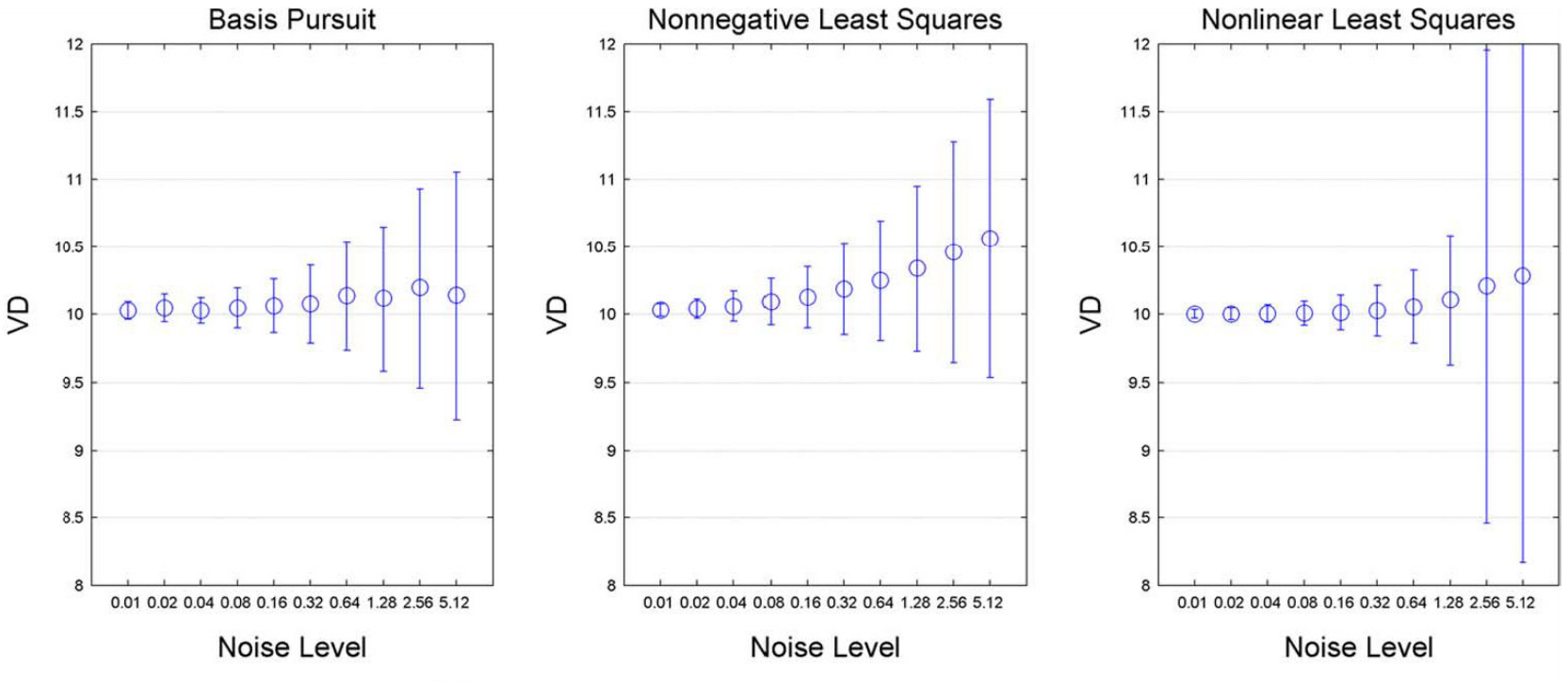
$$\min_{w} \left(\|\mathbf{y} - \mathbf{\Phi w}\|_2^2 + \lambda \|\mathbf{w}\|_1 \right).$$

Solved using basis pursuit algorithm (Chen, Donoho and Saunders, 1999).

Analysis of noise level simulated data

V_T mean comparison

NNLS and DEPICT have noise level dependent bias (and parameter dependent too).



Outline

Bayes Rule

If we know something about the parameters, then this can be updated when we see the data.

$$\textit{posterior}(\text{params}|\text{data}) \propto \textit{likelihood}(\text{data}|\text{params})\textit{prior}(\text{params})$$

or

$$p(\theta|y) = \frac{l(y|\theta)p(\theta)}{p(y)} \propto l(y|\theta)p(\theta)$$

and in PET we usually know something about the parameters (even if it is only a physiologically reasonable range)

▸ Basis Functions

Monte Carlo Methods

Idea is to take a sample of data and see what the parameters look like.
E.g.

$$\theta = E_p(f(X)) \quad X \sim p$$

which implies

$$\theta = \int f(x)p(x)dx$$

So we could sample a lot of X's (say $X_1, \ldots, X_n$ from p in an independent way) and then evaluate them to get an idea of the value of θ.

$$\hat{\theta} = \frac{1}{N}\sum_{i=1}^{N} f(X_i)$$

Importance Sampling

Problem 1 - we cant always sample from p.
Idea - Sample from something else q and then adjust.

$$\theta = \int \frac{f(x)p(x)}{q(x)} q(x)dx$$

which implies

$$\theta = E_q(\frac{f(x)p(x)}{q(x)})$$

So we could sample $X_1, \ldots, X_n$ from q in an independent way and then evaluate them to get an idea of the value of θ.

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} \frac{f(X_i)p(X_i)}{q(X_i)}$$

so long as $q(x) > 0$ whenever $f(x)p(x)$ is non-zero.

Importance Sampling

Problem 2 - we might not even know the normalising constant of p.
Idea - We can still use importance sampling.
Let

$$p(x) = \eta(x)/Z$$

where η is the unnormalised density and Z is the normalising constant (unknown).

$$\theta = \frac{1}{Z}\int \frac{f(x)\eta(x)}{q(x)} q(x)dx = \frac{1}{Z}\int f(x)w(x)q(x)dx$$

and

$$Z = \int w(x)q(x)$$

So again we could sample $X_1, \ldots, X_n$ from q in an independent way and then evaluate them both integrals above to idea of the value of θ.

Sequential Importance Sampling

Problem 3 - How do you choose a good importance density
Idea - Start somewhere and move slowly towards the right density
In the Bayesian context, start at the prior and move towards the posterior.

1. Define a sequence $p_k(\theta|y) \propto l(y|\theta)^{\gamma_b} p(\theta)$
2. Sample from $p(\theta) = p_0(\theta|y)$
3. Iterate:
   - 3.1 Move the sample at iteration $b-1$ to a sample at iteration b by using a Markov Kernel.
   - 3.2 Evaluate the Importance Samples and calculate their weights

Types of Markov Kernel could be e.g. random walks.

Resampling

Problem 4: Only a few weights dominate
Idea: Resample to get a new set.
We now have a set of samples (X_i^b, W_i^b) of our distribution of interest.
We can measure the variation in those (normalised) weights

$$ESS = \left(\sum_{i=1}^{N} (W_i^b)^2 \right)^{-1}$$

This takes a value between 1 and N. If it is too low, say $N/2$ or less, then we resample the samples to get a new set each with weight $1/N$.
This new set is then taken to the next iteration.

Sequential Monte Carlo Samplers

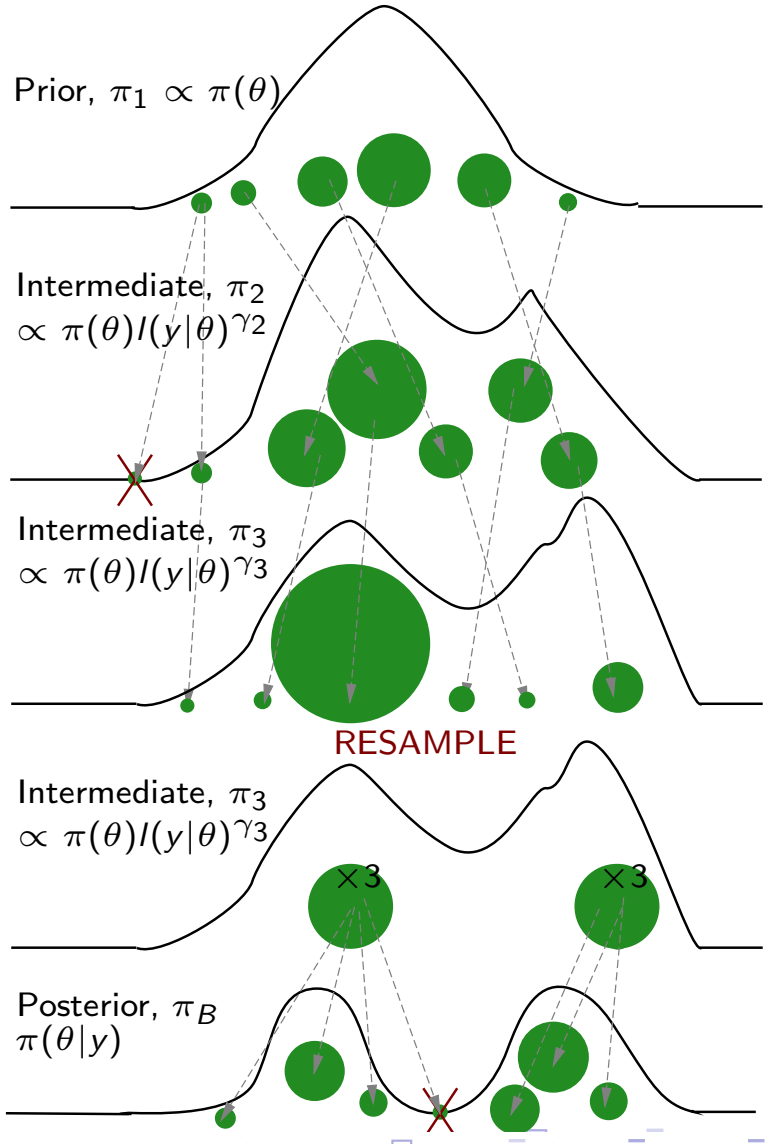
(Del Moral et al, 2006)

SMC Samplers

$$\pi_b \propto p(\theta) l(y|\theta)^{\gamma_b}$$

where

- $p(\theta)$ = prior of the model parameters
- $l(y|\theta)$ = likelihood
- $0 = \gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_B = 1$, a tempering schedule



Extensions with SMC - Identifiability

Take a similar approach to optimization using simulated annealing:

$$\pi^b \propto p(\theta) l(\theta|y)^{\gamma_b}$$

If γ_b is allowed to go above one, then if parameter distribution doesnt degenerate to point mass as γ_b increases, then parameter is unidentifiable.

Model Selection

For a given model M, the normalising constant, Z_B

$$Z_B = p(y|M).$$

It turns out that SMC samplers can be used to give unbiased estimates of ratios of normalising constants (Del Moral, 2004, Del Moral et al, 2006), via the reweighting of the particles

$$\frac{Z_b}{Z_{b-1}} \approx \frac{\widehat{Z_b}}{Z_{b-1}} = \sum_{i=1}^{N} W_{b-1}^{(i)} w_b(\theta_{b-1:b}^{(i)})$$

$$= \text{Normalising Constant for weights at iter } b = \bar{W}_b$$

where $W_{b-1}^{(i)}$ are the particle weights and $w_b(\theta_{b-1:b}^{(i)})$ are the unnormalised incremental weights (see Del Moral et al, 2006). Z_B, can thus be approximated as:

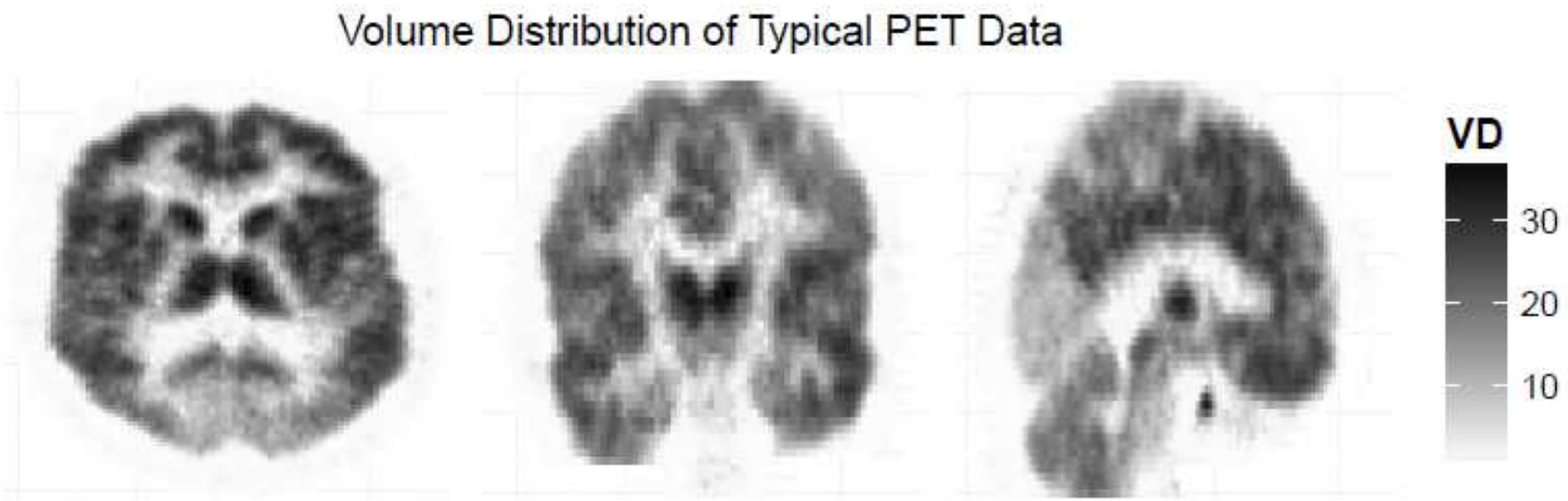
$$\widehat{Z_B} = \widehat{Z_1} \prod_{b=2}^{B} \bar{W}_b$$

Outline

SMC PET Implementation

- Subject-by-subject analysis of $[^{11}C]$-Diprenorphine PET Scans.
- Models chosen from 1, 2, and 3 - tissue compartmental models
- Gaussian / t-distribution Measurement Errors (again proportional to true signal).
- Prior specification used incorporate biological meaningful information (see Zhou et al, 2013)
- Parameter of interest is volume of distribution which can be derived from all models
- Allows option of model selection or averaging (model selection used in results)
- N=1000 particles, and approximately 180-200 intermediate distributions used (random quantity based on CESS).

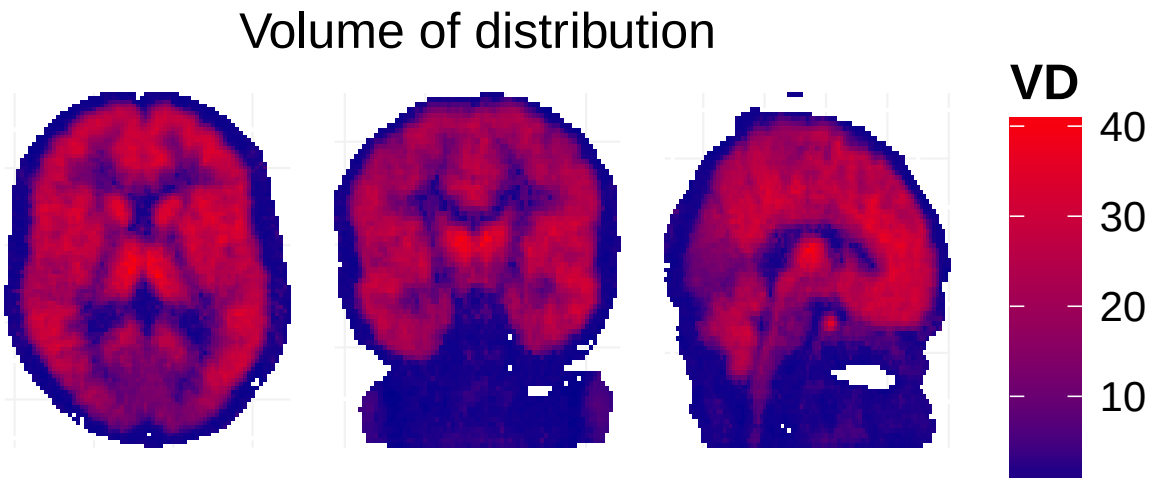
PET Results

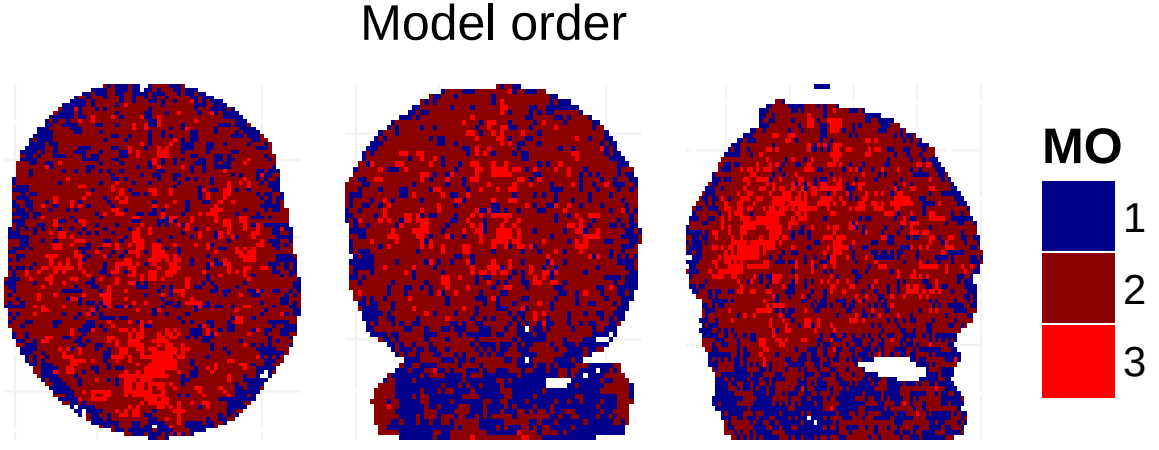


[^{11}C]-Diprenorphine imaging of the Opioid receptors. Healthy control part of study into epilepsy.

Final Estimates of V_T and Model Order

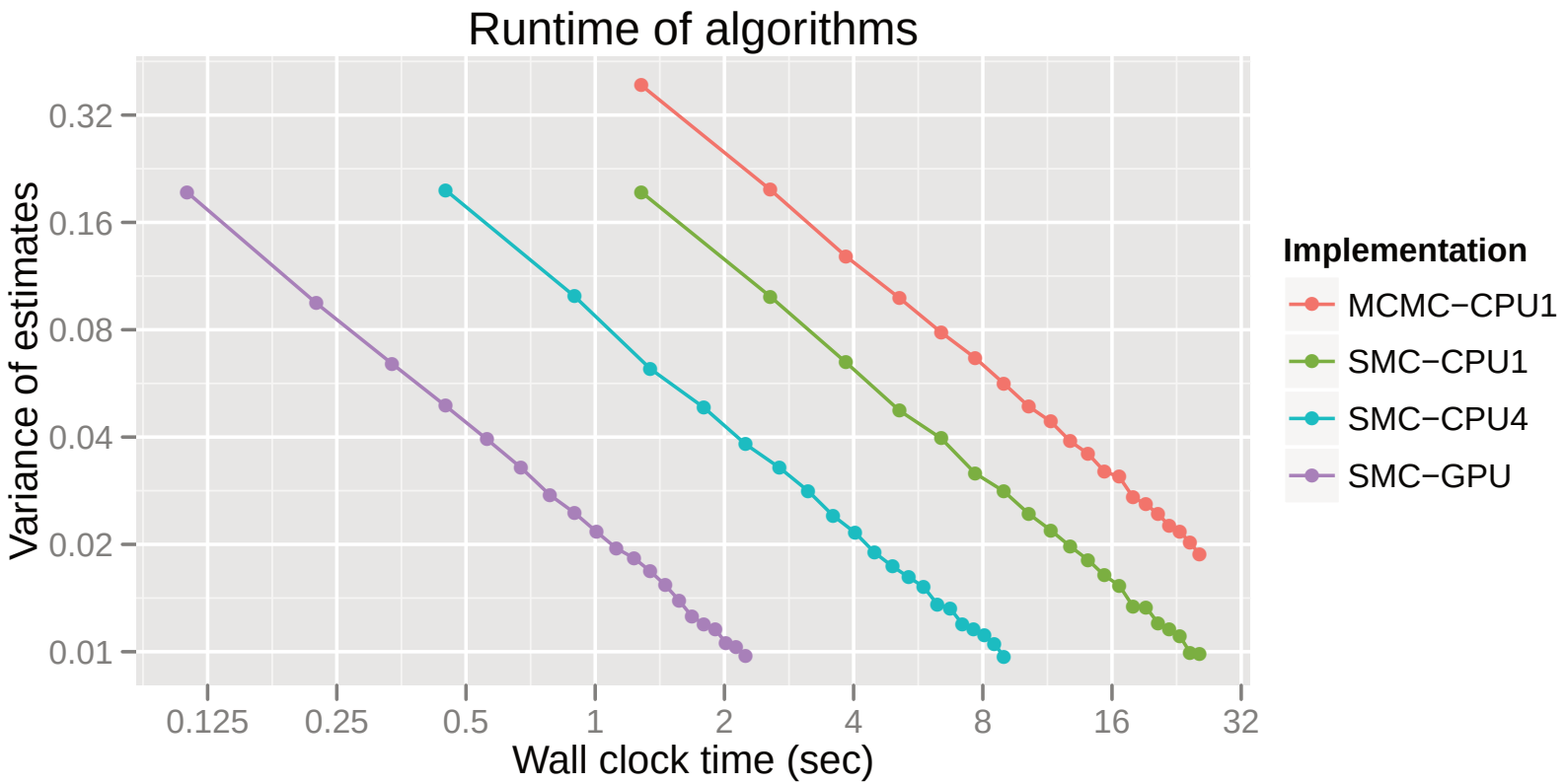
[^{11}C]-Diprenorphine





Computational Results

(Zhou et al, 2012)



Conclusions

- Compartmental Models are ubiquitous as biomedical models and naturally lend themselves to SMC analysis
- Compartmental Models give considerable insight and SMC can help parameter estimation and possibly even identifiability
- PET can interrogate many neurochemical systems and SMC is now computationally implementable even at the voxel level, allowing not only parameter estimation but also uncertainty measures.