



Published in final edited form as:

Multimodal Brain Image Anal (2011). 2011 ; 7012: 1–9. doi:10.1007/978-3-642-24446-9_1.

Accounting for Random Regressors: A Unified Approach to Multi-modality Imaging

Xue Yang¹, Carolyn B. Lauzon¹, Ciprian Crainiceanu², Brian Caffo², Susan M. Resnick³, and Bennett A. Landman¹

Xue Yang: Xue.Yang@vanderbilt.edu; Carolyn B. Lauzon: Carolyn.Lauzon@vanderbilt.edu; Ciprian Crainiceanu: ccrainic@jhsp.edu; Brian Caffo: bcaffo@jhsp.edu; Susan M. Resnick: resnicks@grc.nia.nih.gov; Bennett A. Landman: Bennett.Landman@vanderbilt.edu

¹Electrical Engineering, Vanderbilt University, Nashville TN, 37235 USA

²Department of Biostatistics, Johns Hopkins University, Baltimore MD, 21205 USA

³National Institute on Aging, National Institutes of Health, Baltimore MD, 21224

Abstract

Massively univariate regression and inference in the form of statistical parametric mapping have transformed the way in which multi-dimensional imaging data are studied. In functional and structural neuroimaging, the *de facto* standard “design matrix”-based general linear regression model and its multi-level cousins have enabled investigation of the biological basis of the human brain. With modern study designs, it is possible to acquire multiple three-dimensional assessments of the same individuals — e.g., structural, functional and quantitative magnetic resonance imaging alongside functional and ligand binding maps with positron emission tomography. Current statistical methods assume that the regressors are non-random. For more realistic multi-parametric assessment (e.g., voxel-wise modeling), distributional consideration of all observations is appropriate (e.g., Model II regression). Herein, we describe a unified regression and inference approach using the design matrix paradigm which accounts for both random and non-random imaging regressors.

Keywords

Model II regression; Inference; Statistical parametric mapping; Biological parametric mapping; model fitting

1 Introduction

The strong relationship between structure and biological function holds true from the macroscopic scale of multi-cellular organisms to the nano scale of biomacromolecules. Experience informs the clinical researcher that such structure-function relationships must also exist in the brain and, when discovered and quantified, will be powerful informers for early disease detection, prevention, and our overall understanding of the brain. Brain imaging modalities, such as positron emission tomography (PET) and magnetic resonance imaging (MRI), are primary methods for investigating brain structure and function. Quantification of the structure function relationship using imaging data, however, has been

challenging owing to the high-dimensional nature of the data and issues of multiple comparisons.

Statistical Parametric Mapping (SPM) enables exploration of relational hypotheses without *a priori* assumptions of regions of interest (ROIs) where the correlations would occur [1, 2]. SPM was limited to single modality regression with imaging data represented only in the regressand until extensions (e.g., Biological Parametric Mapping, BPM) were developed to enable multi-modality regression, allowing for imaging data to use considered for both regressors and regressand [3, 4]. These multi-modal methods rely on the traditional ordinary least squares approach in which regressors are exactly known (i.e., conditional inference). Although this assumption may be reasonable in SPM, where scalar regressors are likely to have significantly less variance than the regressand imaging data, such an assumption is clearly violated when both regressors and regressand are observations from imaging data. With BPM inference is not inverse consistent; interchanging the regressors and regressand images would yield different estimates of relationships. The inconsistent inverse behavior of BPM is a result of violated mathematical assumptions, not underlying biological truths. A researcher is seeking to uncover the two way structure-function relationship and a mathematical technique that optimizes an inverse consistent mapping rather than a unidirectional mapping would bring estimates closer to modeling these underlying physical truths.

Regression analysis accounting for errors in regressors would greatly improve the credibility of the truth model whilst reasonably considering the randomness of the imaging modality. Statistical methods accounting for random regressors have been developed and are collectively known as Model II regression [5, 6]. Surprisingly, Model II regression has not been generalized for the massively univariate case. To more accurately reflect clinical imaging data, herein we develop a general model that accounts for both random regressors and non-random regressors for use in the context of BPM and multi-modality image regression.

2 Theory

Our aim is to explain the observed intensity from one imaging modality, y , with a set of regressors, \mathbf{x} , of which at least one member is observed intensity from another imaging modality. We begin with a typical general linear model (GLM) and reformulate it to explicitly reflect the clinical imaging case of both random and non-random regressors. To begin, GLM is formulated as,

$$\mathbf{y} = \sum_{l=1}^d \mathbf{x}^{(l)} \beta^{(l)} + \boldsymbol{\varepsilon} \quad (1)$$

where d is the total number of regressors, $\boldsymbol{\varepsilon}$ is a parameterization of observational error in \mathbf{y} , and $\boldsymbol{\beta}$ is a vector of the fitted coefficients. Let $\sigma_{\mathbf{x}^{(l)}}$ and σ_y represent the common variance of each element about its truth in $\mathbf{x}^{(l)}$ and \mathbf{y} respectively. Then the $\mathbf{x}^{(l)}$ can be divided into two disjoint sets, fixed regressors whose values are considered to be exactly known, \mathbf{x}_f , s.t. ($\sigma_{\mathbf{x}_f} \ll \sigma_y$), and random regressors, \mathbf{x}_r , s.t. ($\sigma_{\mathbf{x}_r}$ not $\ll \sigma_y$). In BPM, all regressors are treated as

fixed regressors. By inclusion of random regressors, that Model-II diverges from BPM. Owing to their larger σ values, the observed \mathbf{x}_r are only therefore noisy estimates of the truth, \mathbf{x}_{rT} . Eq. 1 then becomes,

$$\mathbf{y} = \sum_{j=1}^q \mathbf{x}_{rT}^{(j)} \beta_r^{(j)} + \sum_{k=1}^m \mathbf{x}_f^{(k)} \beta_f^{(k)} + \varepsilon \quad (2)$$

where q and m represent the total number of random and fixed regressors respectively. The challenge here, that is not present in Model-I, is to estimate β_r and β_f given that \mathbf{x}_{rT} in Eq. 2 is not observed, but rather the noisy \mathbf{x}_r . There are many possible methods for solving for β , here we choose an approach that follows the ordinary least squares example. We seek a solution for the parameters, β , that maximize the log-likelihood of the model given the observed data (maximize $\ln P(\mathbf{y}, \mathbf{x}_r | \beta, \Sigma, \mathbf{y}_T, \mathbf{x}_{rT})$).

Let $i = 1, 2, \dots, n$ index the observations in \mathbf{y} and in each of the $\mathbf{x}_r^{(j)}$ (e.g. n = the number of subjects and i indexes the subject number). Let \mathbf{z}_i be a vector representing the i 'th set of observational errors in \mathbf{y} and in each $\mathbf{x}_r^{(j)}$. We can assume the errors follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix, Σ [7].

$$\mathbf{z}_i = \begin{bmatrix} y - y_T & x_r^{(1)} - x_{rT}^{(1)} & \dots & x_r^{(q)} - x_{rT}^{(q)} \end{bmatrix} \sim \Phi(\mathbf{0}, \Sigma) \quad (3)$$

Note that the observational errors, \mathbf{z} , are errors across subjects and do not condition errors across an image. Given that each $\mathbf{x}_r^{(j)}$ vector is observed from a unique experimental technique, it is reasonable to assume that the columns of \mathbf{z}_i are independent. We further assume that \mathbf{z}_i is independent as each subject is independent. Under these assumptions (normal and i.i.d.), the log-likelihood of the observed data, given the model in Eq. 2 is,

$$\ln P(\mathbf{y}, \mathbf{x}_r | \beta, \Sigma, \mathbf{y}_T, \mathbf{x}_{rT}) = \sum_{i=1}^n \ln \left(2\pi^{-n/2} \left(\det(\Sigma)^{-1/2} \right) \right) - \frac{1}{2} \sum_{i=1}^n (\mathbf{z}_i \Sigma^{-1} \mathbf{z}_i') \quad (4)$$

Maximizing the log likelihood, Eq. 4, is equivalent to minimizing $s = \sum_{i=1}^n (\mathbf{z}_i \Sigma^{-1} \mathbf{z}_i')$. With the assumption of independent observations (independent subjects, i , and independent experiments, $\mathbf{x}_r^{(j)}$), the covariance matrix, Σ , is diagonal with entries $\sigma_y^2 \quad \sigma_{x_r^{(1)}}^2 \quad \dots \quad \sigma_{x_r^{(q)}}^2$. Hence, s can be re-expressed as,

$$s = \sum_{i=1}^n \left(\sigma_y^{-2} (y_i - y_{T_i})^2 + \sum_{j=1}^q \sigma_{x_r^{(j)}}^{-2} (x_{r_i}^{(j)} - x_{rT_i}^{(j)})^2 \right) \quad (5)$$

where $x_{r_i}^{(j)}$ represents the i 'th element of $\mathbf{x}_r^{(j)}$. Eq. 5 is minimized when its partial derivatives w.r.t to each dependent variable is zero. We first solve for \mathbf{x}_{rT} at the minimum of Eq.5 by differentiating s with respect to $\mathbf{x}_{rT}^{(j)}$ and setting the result to $\mathbf{0}$ gives q total equations, one for

each $\mathbf{x}_{rT}^{(j)}$. With q equations and q unknowns, $\mathbf{x}_{rT}^{(j)}$ can be solved in terms of the other parameters. For a given term indexed by h ,

$$\mathbf{x}_{rT}^{(h)} = \frac{\beta_r^{(h)} \sigma_{x_r}^{2(h)} \left(\mathbf{y} + \mathbf{x}_r^{(h)} \beta_r^{(h)} - \sum_{j=1}^q \mathbf{x}_r^{(j)} \beta_r^{(j)} - \sum_{k=1}^m \mathbf{x}_k^{(j)} \beta_r^{(j)} \right) + \mathbf{x}_r^{(h)} \left(\sum_{j=1}^q \left(\beta_r^{(j)} \sigma_{x_r}^2 \right) - \beta_r^{(h)} \sigma_{x_r}^2 + \sigma_y^2 \right)}{\sigma_y^2 + \sum_{j=1}^q \beta_r^{(j)2} \sigma_{x_r}^2} \quad (6)$$

Substituting Eq. 6 for $\mathbf{x}_{rT}^{(j)}$ in Eq. 5, s becomes,

$$s = \frac{1}{\sigma_y^2} \sum_{i=1}^n \frac{\left(y_i - \left(\sum_{j=1}^q x_{r_i}^{(j)} \beta_r^{(j)} + \sum_{k=1}^m x_{f_i}^{(k)} \beta_f^{(k)} \right) \right)^2}{\sigma_y^2 + \sum_{j=1}^q \beta_r^{(j)2} \sigma_{x_r}^2} \quad (7)$$

Eq. 7 is now independent of the unknown \mathbf{x}_{rT} and provides an intuitive form as the model error in the numerator is balanced by the individual variances in the denominator and mirrors the more readily available multivariate case with non-random \mathbf{x} and the univariate case that accounts for a single random \mathbf{x} [8]. Eq. 7 is a function of two unknowns, $\boldsymbol{\beta}$ and σ . Note that the true σ does not factor into s , only the ratio of variance matters. If these ratios are known, then $\boldsymbol{\beta}$ is the only unknown; and the $\boldsymbol{\beta}$ that maximizes Eq. 4 can be solved by taking the partial derivative of s with respect to $\boldsymbol{\beta}$ and setting the result to 0. The partial derivative equations are nonlinear w.r.t. $\boldsymbol{\beta}$, and solving for the closed form solution of $\boldsymbol{\beta}$ is involved, so we employ numeric optimization methods.

In the Model II approach, the variance ratio needs to be known in order to minimize s by solving $\boldsymbol{\beta}$. The restriction arises because the number of unknown parameters is larger than the number of equations [9]. If we add the further assumption that the ratio of the overall variance across subjects is proportional to the ratio of the image noise variance, then we can estimate the model variance ratio by estimating the ratio of image noise for each modality.

Inference on $\boldsymbol{\beta}$

The maximum likelihood estimate of $\boldsymbol{\beta}$, $\boldsymbol{\beta}_{\text{est}}$, with a true value $\boldsymbol{\beta}_T$, is asymptotically normally distributed as $\boldsymbol{\beta}_{\text{est}} - \boldsymbol{\beta}_T \sim \Phi(0, n\mathbf{I}^{-1})$ where \mathbf{I} is the Fisher information matrix ($\mathbf{I}_{w,g} = \mathbf{E}[-\nabla_{\boldsymbol{\beta}} \ln P_i(y_i, \mathbf{x}_{r_i} \boldsymbol{\beta}, \Sigma, y_T, \mathbf{x}_{rT_i})]$), with $w = 1, 2, \dots, q+m$, $g = 1, 2, \dots, q+m$, P_i is the probability function for data i . Noting the distribution

$y_i - \left(\sum_{j=1}^q x_{r_i}^{(j)} \beta_{\text{est},r}^{(j)} + \sum_{k=1}^m x_{r_i}^{(k)} \beta_{\text{est},f}^{(k)} \right) \sim N \left(0, \sigma_y^2 + \sum_{j=1}^q \beta_{\text{est},r}^{(j)2} \sigma_{x_r}^2 \right)$, the Fisher information can be estimated from prior ratio and dataset. Allowing \mathbf{c} to represent the contrast vector and $\boldsymbol{\beta}_T = \boldsymbol{\beta}_{\text{null}}$ to be the null hypothesis, then it follows that the test-statistic, $\mathbf{c}'(\boldsymbol{\beta}_{\text{est}} - \boldsymbol{\beta}_{\text{null}}) / \sqrt{\mathbf{c}'(n\hat{\mathbf{I}})^{-1}\mathbf{c}}$ is t-distributed with $n-(q+m)$ degrees of freedom. This is an asymptotic t-test as all estimated parameters and Fisher information are asymptotically valid.

3 Methods and Results

3.1 Single voxel simulations with known true variance ratios

Model II regression is implemented using the Nelder–Mead method to find the optimized solution of β . For each of the following 3 scenarios, a simulated voxel with 50 observations (i.e., subjects) was studied using a model with one random regressor, one fixed regressor, and a single constant: $\mathbf{y}_T = \mathbf{x}_{rT}\beta_r + \mathbf{x}_f\beta_f + \beta_1$. In each of 500 Monte Carlo trials: β_T were chosen randomly from the uniform distribution (0–2); errors were added to y_{T_i} and x_{rT_i} from a normal distribution with fixed standard deviations, and the true variance ratio was assumed known. Model II and OLS performance were evaluated with the relative root mean squared error in β ($r\beta_{RMS}$).

- A. **Model II vs OLS response to $\sigma_x:\sigma_y$ ratios** (Figure 1A). Simulations were performed varying $\sigma_x:\sigma_y$. Model II regression performs equally well as OLS with small $\sigma_x:\sigma_y$, but becomes advantageous as more relative error is introduced into \mathbf{x}_r observations. The improvement is observed specifically on β_r , whereas the constants not associated with random regressors, β_1 and β_f , remain with equal accuracy in estimation between the two models.
- B. **Model II vs OLS response to number of random x regressors** (Figure 1B). The above model was altered by including up to 4 additional random regressors with randomized coefficients. Model II has smaller errors in the β_r estimates than OLS, however Model II becomes less advantageous with increases in the number of random x regressors. Note, the number of observations was not increased to compensate for the increased model complexity so less data per regressor is available with more regressors.
- C. **Model II sensitivity to the estimated ratio** (Figure 1C). To assess the response of Model II to estimated ratios that deviate from the truth, the assumed true ratio of variance was altered between 1/10th and 10 times its true value. Under the cases simulated here, Model II is insensitive to the ratio estimate for the range (0.5–2) and relatively insensitive over the range (.1–3).

At extremely incorrect ratio values, the β_r estimate rapidly loses accuracy. Based on this analysis we can apply Model II regression using estimated error ratio, with reasonable confidence in the methods' tolerance to mis-estimation of variance ratios.

3.2 Volumetric Imaging Simulation

Model II regression is incorporated as a regression method choice in the BPM toolbox for the SPM software using Matlab (Mathworks, Natick, MA). We simulated images of two modalities and regressed one modality on the other modality. The true regressor images are simulated from smoothed gray matter density images of 20 participants in the normal aging study of the BLSA neuroimaging project consisting of 79*95*69 voxels with 0.94*0.94*1.5 mm resolution [10]. The observed regressor images are simulated by adding zero mean Gaussian noise. The true regressand image intensity is “1.5*true regressor images – constant” inside the caudate region and equals a different constant everywhere else inside the brain mask. The observed regressand images were generated by adding zero mean

Gaussian noise to the true regressand images. The standard deviation of the noise is selected to make SNR=15. (SNR is defined as the mean signal divided by the standard deviation of noise). The observed regressor images and observed regressand images are used for simulation and the ratio of noise standard deviation is used as the ratio of observation error in Model II regression.

Fig. 2 presents the results for 20 simulated subject brains, each represented by a pairs of images. The statistical significance map is shown according to uncorrected p-value with $p < 0.001$ and 5 voxels extent threshold to exclude noise. Type I error and type II error are calculated at uncorrected p-value < 0.001 inside (for true-positive, TP, and false-negative, FN) and outside a caudate mask (for false-positive, FP, and true-negative, TN). In this simple model, both OLS and Model II regression control type I error as expected. Meanwhile, Model II regression improves true positive rate as compared to OLS regression.

3.3 Empirical demonstration of Model II regression

Image-on-image regression offers a direct opportunity to study associations between differing spatially located factors. As an illustrative example, consider potential correlations between GM tissue density (a structural measure) and PET signal (a measure of functional activity). A first model would associate tissue presence with greater functional signal. An analysis of modulating factors for this relationship (such as disease condition, intervention, or task) could reveal anatomical correlates of functional reorganization and shed light on the applicability of the structural-functional hypothesis.

Following this approach, we perform regression analysis of the relationship between anatomical MRI gray matter images (GM, as classified by SPM5) and functional PET images. PET and GM data were collected on a total of $n = 46$ observations (23 subjects imaged twice). Regression was performed in both directions in order to quantify both structure \rightarrow function and function \rightarrow structure relationships. The regression model used 1 random regressor and a single constant. For the Model II ratio, noise ratio is estimated following the method in [11]. The raw data and the resulting regression lines for a single voxel comparison are displayed. Fig 3 shows that Model II is symmetric, i.e., the mapping PET \rightarrow GM is the inverse of the mapping GM \rightarrow PET. The corresponding estimated variances for Model II are also smaller than the corresponding estimated variances in OLS forward regression and OLS inverse regression.

4 Discussion

Properly accounting for error is essential for valid parameter estimation and statistical inference. Herein, we have demonstrated that a full consideration of observation variability is feasible within the confines of a design matrix paradigm. Furthermore, we can readily consider simultaneous treatment of parameters with measurement error (\mathbf{x}_r) alongside traditionally defined fixed parameters (\mathbf{x}_f). Our formulation of “random observations” remains within the context of a “fixed effects” model as the β_r are deterministic parameters, as opposed to the classic “random effects” model where parameters are stochastic. These two approaches are complementary and could be combined for an appropriate experimental

framework. Extension of the Model II concepts to time series, hierarchical and complex model designs would be a fascinating area for continuing research.

We have observed substantial improvements in model fit and increased statistical power using Model II regression as opposed to OLS (Fig. 1A). While robust to model complexity (Fig. 1B) and prior estimation of observation variability (Fig. 1C), the improvements were not universal. When the OLS is appropriate (i.e., $\sigma_x \ll \sigma_y$), there was a slight increase in observed error (Fig. 1A); however, as the relative variance in \mathbf{x} increased, the OLS assumption of fixed regressors becomes increasingly violated and notable bias and increased variance could be observed in the OLS estimates. Examination of the inlays in Fig. 1 reveals a slight increase in error for the fixed parameters, β_f and β_1 , but a lack of appreciable bias.

The generalized Model II regression model can be used to analyze any complicated relationship by applying Taylor series and expand design matrix. Our presentation of Model II regression is inverse consistent, provides a logical framework for exploring relationships in multi-modal image analysis, and can help model relative uncertainty in imaging methods. Other error models may be more appropriate for specific imaging modalities and warrant further consideration.

Our presentation of the empirical analysis is preliminary and ongoing as is highlighted by the single voxel presentation. Interpreting the ratio of model variances is subject of active consideration as must consider the potential impact of both the imaging variability and model fit error in multiple dimensions. As discussed, we currently approximate this combined quantity as proportional to the imaging variability alone. Relaxing this assumption would greatly aid in generalization of this approach. These methods are available in open source as plug-ins for the SPM package.

Acknowledgments

This project was supported by NIH N01-AG-4-0012. This work described herein has not been submitted elsewhere for publication.

References

1. Friston K, Frith C, Liddle P, Dolan R, Lammertsma A, Frackowiak R. The relationship between global and local changes in PET scans. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism*. 1990; 10:458.
2. Friston K, Frith C, Liddle P, Frackowiak R. Comparing functional (PET) images: the assessment of significant change. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism*. 1991; 11:690.
3. Casanova R, Srikanth R, Baer A, Laurienti P, Burdette J, Hayasaka S, Flowers L, Wood F, Maldjian J. Biological parametric mapping: a statistical toolbox for multimodality brain image analysis. *NeuroImage*. 2007; 34:137–143. [PubMed: 17070709]
4. Oakes TR, Fox AS, Johnstone T, Chung MK, Kalin N, Davidson RJ. Integrating VBM into the General Linear Model with voxelwise anatomical covariates. *Neuroimage*. 2007; 34:500–508. [PubMed: 17113790]
5. York D. Least-squares fitting of a straight line. *Canadian Jour Physcs*. 1966; 44:1079–1086.
6. Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: But which regression? *Clinical and Experimental Pharmacology and Physiology*. 2010; 37:692–699. [PubMed: 20337658]

7. Friston K, Holmes A, Worsley K, Poline J, Frith C, Frackowiak R. Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*. 1994; 2:189–210.
8. Press, W. *Numerical recipes: the art of scientific computing*. Cambridge Univ Pr; 2007.
9. Carroll R, Ruppert D. *The Use and Misuse of Orthogonal Regression in Linear Errors-in-Variables Models*. *The American Statistician*. 1996; 50
10. Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, Bryan RN, Zonderman AB. One-year age changes in MRI brain volumes in older adults. *Cerebral Cortex*. 2000; 10:464. [PubMed: 10847596]
11. Rajan J, Poot D, Juntu J, Sijbers J. Noise measurement from magnitude MRI using local estimates of variance and skewness. *Physics in Medicine and Biology*. 2010; 55:N441. [PubMed: 20679694]

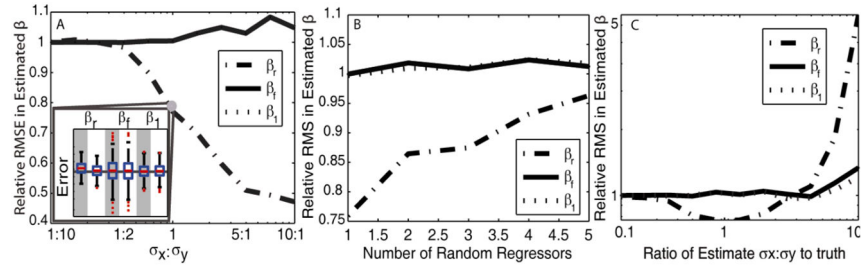
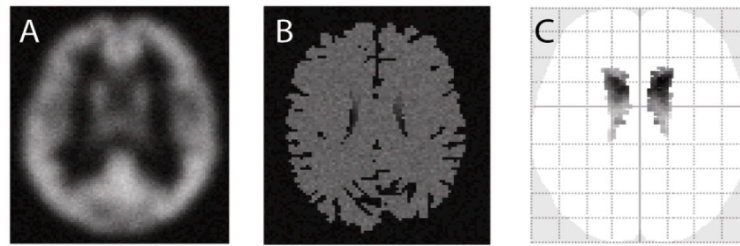


Fig. 1.

The relative RMSEs of Model II to OLS for each estimated coefficient (β_r , β_f , β_1) are plotted as a function of the ratio of the true standard deviations, $\sigma_x:\sigma_y$ (A), the number of random regressors, $x_r^{(s)}$ (B) and the accuracy of the ratio estimate (C). With increasing $\sigma_x:\sigma_y$ ratios, model II regression has increased relative accuracy in β_r estimates compared to OLS with increasing $\sigma_x:\sigma_y$ ratios. Comparison of the β_f , β_r , and β_1 error distributions formed is shown explicitly in the inlay for the point $\sigma_x:\sigma_y=1$ in (A). The gray column shows the OLS error and the white column shows the Model II error, the horizontal line is where the error is zero. In (C), for one unit σ_y , the estimated ratio μ_x was allowed to deviate from the ideal case, $\mu_x/\sigma_x = 1$. The common point shared in (A, B, C) is located in (B) at ‘Number of Random Regressors’ = 1, and (C) at ‘Ratio of Estimate to Truth’ = 1.



	Outside ROI		Inside ROI	
	FP	TN	FN	TP
Model II	0.005%	99.995%	12.2%	87.8%
OLS	0.096%	99.904%	16.6%	83.4%

Figure 2.

(A) is one regressor image, (B) is one regressand image in Model II BPM simulation. (C) is one dimension of the glass brain which is the projected statistical significance. The glass brain is shown according to uncorrected p-value $p < 0.001$ and 5 voxels extent threshold to exclude noise. The table is calculated with $p < 0.001$.

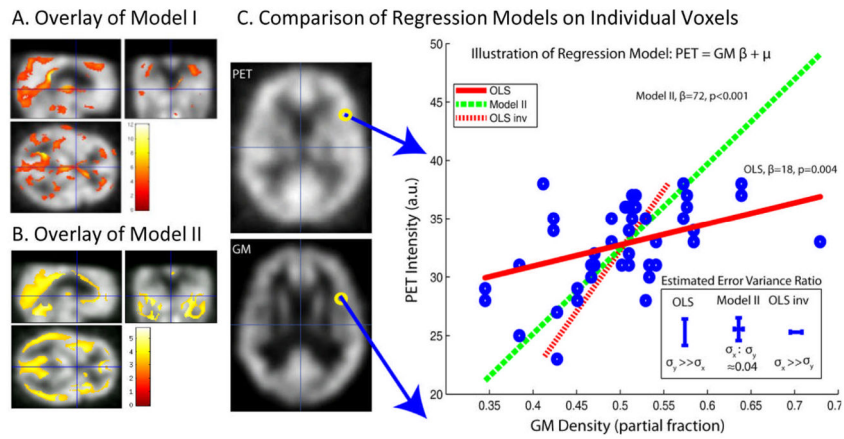


Figure 3. Model II and Model I (OLS) multi-modality regression analysis. Model I (A) and Model II (B) lead to distinct patterns of significant differences ($p<0.001$, uncorrected) when applied to identical empirical datasets and association models. Inspection of single voxel: PET vs Grey Matter MRI (GM) illustrates the reasons for the different findings (C). The GLM model used for the forward mapping is $\mathbf{y} = \mathbf{x}_r\beta_r + \beta_1$, where \mathbf{y} represents PET image intensity and \mathbf{x}_r represents GM normalized image intensity. On the left-hand side of (C), example images of PET and GM are shown, along with the location of the single example voxel whose regression analysis is displayed in the right-hand plot. The individual data points (blue circles) were fit using OLS (red lines) and Model II regression (green dashed line). The inverse mapping for OLS (red dash) is unique from the forward mapping (red full). The Model II mapping was found to be reversible and can be represented by the same line. Resulting error bars and corresponding $\sigma_x:\sigma_y$ value estimates are compared between OLS and Model II in the lower right-hand insert.