ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 25 (2005) 159–168

Wavelet variance components in image space for spatiotemporal neuroimaging data

John A.D. Aston,[a,*] Roger N. Gunn,[b] Rainer Hinz,[c] and Federico E. Turkheimer[c,d]

[a]*Institute of Statistical Science, Academia Sinica, 128 Academia Road, Sec 2, Taipei 11529, Taiwan*
[b]*Translational Medicine and Technology, GlaxoSmithKline, UK*
[c]*Imanet Hammersmith, Hammersmith Hospital, London, UK*
[d]*Department of Neuropathology, Imperial College London, UK*

Received 12 July 2004; revised 20 October 2004; accepted 26 October 2004
Available online 5 January 2005

Neuroimaging studies place great emphasis on not only the estimation but also the standard error estimates of underlying parameters derived from a temporal model. This allows inferences to be made about the signal estimates and resulting conclusions to be drawn about the underlying data. It can often be advantageous to interrogate temporal models after spatial transformation of the data into the wavelet domain. Wavelet bases provide a multiresolution decomposition of the spatial data dimension and an ensuing reduction in spatial correlation. However, widespread acceptance of these wavelet techniques has been hampered by the limited ability to reconstruct both parametric and error estimates into the image domain after analysis of temporal models in the wavelet domain. This paper introduces a derivation and a fast implementation of a method for the calculation of the variance of the parametric images obtained from wavelet filters. The technique is proposed for a class of estimators that have been shown to be useful in neuroimaging studies. The techniques are demonstrated for both functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) data sets.

Keywords: Wavelets; Variance; Mean square error; fMRI; PET

Introduction

Wavelets are now increasingly used in data analysis of neuroimaging studies. For both positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), wavelets are used at different points in the chain of analysis (Bullmore et al., 2003; Desco et al., 2001; Ruttimann et al., 1998; Turkheimer et al., 1999, 2000a). Many such methods use the entire four-dimensional data set to perform this analysis, making use of temporal estimates of error to allow thresholding of wavelet components in the spatial wavelet domain (Müller et al., 2003; Ruttimann et al., 1998, 2000b). However, current methods only provide estimates of variance of the wavelet coefficients prior to thresholding and performing the inverse transform. In this work, we introduce a method for the estimation of the resulting error estimates of the parameters in image space through the analysis of the residuals to the model.

Wavelets and functional imaging

Wavelet methods are of use in neuroimaging due to the inherent problems of data analysis in the image domain both spatially and temporally. In this paper, the wavelet analysis considered will be conducted exclusively in the spatial domain. In functional imaging, data are spatially correlated due to the intrinsic way data are measured, and ignoring this correlation will lead to spurious results. However, the decorrelating properties and the sparse nature of the signal representation, due to the multiresolution analysis provided by wavelet transformation, lead to much more effective methods.

Statistical parametric mapping (SPM) (Ashburner et al., 1999) is the defacto standard in analysis of brain mapping studies. Wavelet methods offer a different type of analysis to the SPM method, as the output is a regularized parametric map as opposed to probability of null-hypothesis map (p-map). Effective comparison can be difficult without the error estimates being available from the wavelet analysis. Previous work (Van De Ville et al., 2003, 2004b) has looked at comparing wavelet methods and SPM methods especially in the area of Gaussian smoothing, which is used by SPM to improve signal to noise ratio. Indeed, it has been shown (Van De Ville et al., 2003) that for certain carefully chosen wavelets, a smoothing similar to SPM can be performed in the wavelet domain. In addition, a method of thresholding the wavelets to retain the underlying knowledge about the error estimates has also been proposed (Van De Ville et al., 2004b), as well as an estimate of the standard errors indirectly obtained using an absolute value wavelet (Van De Ville et al., 2004a).

* Corresponding author. Fax: +886 2 2783 1523.
E-mail address: jaston@stat.sinica.edu.tw (J.A.D. Aston).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.10.037

Wavelet analysis is also of great interest in dynamic PET studies where there can be quantitative measurements made through the use of calibration to measured radioactivities (Turkheimer et al., 1999). While hypothesis testing is still of interest, the estimation of the signal is of prime importance and methods that concentrate on signal estimation are highly applicable. Studies of this type include metabolism studies and neuroreceptor ligand studies, examples of which will be detailed below.

Estimation techniques

MSE evaluation of methods differs from those of the traditional SPM setting. The focus of SPM is to provide the minimum linear unbiased variance estimate (best linear unbiased, BLU) of the parameter of interest and perform hypothesis testing based on this estimate. MSE comprises two components, variance and squared bias. MSE estimators obtain a large reduction in the variance at the expense of introducing little bias into the estimate. Therefore, while on average the estimate of the parameter will be slightly different from the truth (biased), any one estimate of the parameter could be far closer to the truth (smaller variance), assuming the bias is small. As the study is only performed once, it can be argued that MSE is a natural choice for finding the parameter, and methods with smaller MSEs preferred.

Wavelet filters, through shrinkage thresholding, reduce the variance of the parameters by introducing a small amount of bias, and can be shown to be better estimators in terms of MSE than a traditional BLU estimator. However, it has not previously been possible to quantify the reduction in variance achieved in neuroimaging studies. Here, this problem is addressed with a fast procedure for the calculation of the variance of images after wavelet regularization.

The methods discussed here are primarily considered in four-dimensional setting, where data are available in three spatial dimensions and one temporal dimension. Thus, it is necessary to not only consider a spatial model but also a temporal model. A spatial wavelet model will be used, while the temporal model will be a linear model of varying complexity depending on the type of neuroimaging modality considered.

Theory

Estimation in image space

Let $\beta_I(x, y, z)$, where x, y, z represent voxel indices, be the spatial distribution of a parameter of interest. This parameter may result from a dynamic acquisition $Y(x, y, z, t)$ in either fMRI or PET. This acquisition consists of M discrete dynamic frames $Y(x, y, z, t_1),\ldots,Y(z, y, z, t_M)$. Typical parameters of interest include task-related changes in signal for fMRI, and blood flow, metabolism, and receptor concentrations for PET.

An assumption of these models is that there is a biological temporal process with associated spatial distribution of a parameter of interest that can be encapsulated in the mathematical model,

$$\beta_I(x,y,z) = \eta(Y(x,y,z,t)) \tag{1}$$

where $\eta()$ can be linear or nonlinear. In this paper, only the case where $\eta()$ is linear will be considered, for reasons explained below. The linear case covers many commonly used models in fMRI and PET. As $\eta()$ is linear, a design matrix X can be associated with the model, where X has dimension $M \times p$, the number of time frames by the number of parameters in the model.

For convenience, the spatial dependence in the notation will be dropped, but it should be noted that all the following equations refer to a voxel by voxel analysis for each x, y, z. Let

$$Y(t) = X\left[\beta_1..\beta_p\right]' + \epsilon(t) \tag{2}$$

where $'$ represents the transpose operator. Here $\eta()$ is assumed to be a simple linear model (see below for a removal of this assumption). $\beta_1,\ldots,\beta_p$ represent all the parameters in the model, and β_I is a linear (including any individual parameter $\beta_1,\ldots,\beta_p$) combination of these parameters. The least squares estimate of the parameters $\beta = [\beta_1,\ldots,\beta_p]'$ is then given by

$$\hat{\beta} = (X'X)^{-1}X'Y(t) \tag{3}$$

where here $\hat{}$ represents estimate.

Let

$$r(t) = Y(t) - X\hat{\beta} \tag{4}$$

be the residuals to the fitted model. Then

$$\hat{\sigma}^2\left(\hat{\beta}\right) = (X'X)^{-1}\frac{\sum_{i=1}^{M} r(t_i)^2}{df} \tag{5}$$

where df are the degrees of freedom associated with the model.

This can of course be extended to general least squares estimation, with the matrices $(X'X)^{-1}$ replaced by the suitable weighted least squares counterpart, for example,

$$\hat{\sigma}^2\left(\hat{\beta}\right) = (X'X)^{-1}X'\Sigma X(X'X)^{-1}\frac{\sum_{i=1}^{M} r(t_i)^2}{df} \tag{6}$$

where Σ is the weights for the assumed error model. This is a method of estimation that can be used for models with an autocorrelated error component.

Estimation problem in wavelet space

Due to the linear nature of the wavelet transform, the estimation of $\beta_I(x, y, z)$ in image space has an isomorphic representation in wavelet space, namely, the problem of estimating $\beta_I^W(w_x,w_y,w_z)$. Let W represent the discrete wavelet transform (DWT) in either two or three dimensions (the number of dimensions being dependent on the imaging type as will be discussed later, see fMRI spatiotemporal analysis section), and let w_x,w_y,w_z be the indices of the transformed volume. It should be noted here that W represents a linear transform from the sampled data to the wavelet coefficients. DWT can be seen as a combination of High (H) and Low (L) pass filters applied to the data. There are many good texts and papers on wavelet transform and further details can be found in, among others, Chui (1992), Percival and Walden (2000), Mallat (1999), and Daubechies (1992). An important part of the method is the splitting of the image into resolution levels, which will have importance in the analysis later. $\beta_I^W(w_x,w_y,w_z)$ can be found in exactly the same way as $\beta_I(x,y,z)$ except that the linear temporal model is applied to the temporal sequence

$$Y^W\left(w_x, w_y, w_z, t_i\right) = WY(x,y,z,t_i)\ i : i = 1,\ldots,M. \tag{7}$$

It must be remembered that W is a spatial transform only, with no effect on the temporal dimension. However, for equivalence

between image space and wavelet space when no thresholding is applied, the temporal model must commute with the spatial model. Thus, linear models are assumed for the temporal model. Associated parameters can also be found, namely,

$$\hat{\beta}_I^W(w_x, w_y, w_z), \hat{\sigma}^2\left(\hat{\beta}_I^W(w_x, w_y, w_z)\right) \text{ and } r^W(w_x, w_y, w_z, t).$$

One of the major advantages of using the wavelet transform is the ability to perform shrinking of the parameters to gain a better signal to noise ratio. This is due to the multiresolution nature of the transform in that the signal tends to be concentrated in a few coefficients while the noise is spread through all coefficients. Indeed, the wavelet coefficients in the highest (or coarsest) level are not shrunk as they are assumed to be entirely signal. This property can alternatively be considered as an approximation (highest level) to the signal and the details (lower levels) that make up the actual signal, some of the details being noise components.

Again, for ease, spatial indexing is assumed. Before the transformation of $\hat{\beta}^W$ back into image space, through the inverse wavelet transform, a thresholding method $\lambda()$ is applied. This results in an estimate in image space $\hat{\beta}_{Iw}$, with superior MSE properties to the original estimate in image space $\hat{\beta}_I$. It has been demonstrated (Turkheimer et al., 2003) that linear filters can provide better properties for the reconstructed images than non-linear thresholders, and these will be used here. These linear threshold techniques are calculated using the estimate $\hat{\sigma}^2(\hat{\beta}_I^W)$. Alternative threshold estimators can also be used and the method outlined will not be affected. The technique chosen here is merely an example.

The threshold used is determined from the James–Stein estimator and is given by

$$\lambda\left(\hat{\beta}_I^W\right) = \left(1 - \frac{n-2}{\tau}\right)_+ \hat{\beta}_I^W \tag{8}$$

where $\tau = \sum_{\text{quadrant}} (\hat{\beta}_I^W / \hat{\sigma}(\hat{\beta}_I^W))^2$ and $x_+ = \max(x, 0)$. The quadrant is the index of all the wavelet coefficients in a particular level and quadrant.

Thus, the estimator back in image space is given by

$$\hat{\beta}_{Iw} = W^{-1}\left(\lambda\left(\hat{\beta}_I^W\right)\right) \tag{9}$$

and if $\lambda() = \text{Id}()$, the identity, then the analysis in the image domain and in the wavelet domain after transformation will be identical for a spatially invariant temporal linear model.

Variance relationship between image space and wavelet space

A key problem in the use of wavelet transform has been the inability to estimate the variance of the reconstructed parameter images after wavelet transformation. While it has been shown through simulation that the MSE properties are superior, it has not been possible to give variance values on an individual basis. The method proposed here calculates this variance.

A previous study of this problem (Aston et al., 2003) yielded a technique that was based on the assumed decorrelation properties of the wavelet transform. This technique transformed the variance parameters calculated in wavelet space back into image space. However, while these properties may well be true in large samples, it will be seen below that these do not hold well enough for the technique to be readily used in the small temporal samples available in common imaging studies. In addition, the technique was time consuming to calculate.

The method proposed here is both fast and efficient, and also does not require any assumptions about the decorrelating properties of wavelet transform. It also has the appeal of being conceptually simple.

As was shown in Eq. (9), the parameter of interest is inverted from image space after thresholding. It is not possible to apply this inversion directly to standard errors as they are a second-order estimator (matrix) as opposed to the first order (vector) parameter estimate. However, there does exist a first-order estimate that can be used to calculate the standard errors, namely the residuals. The residuals can be thresholded using the parameter thresholder and then inverted back into image space. These can then be used to calculate the variance of the image space parameters, that is,

$$r_w(w_x, w_y, w_z, t) = W^{-1}\left(\lambda_{\beta_I}\left(r^W(w_x, w_y, w_z, t)\right)\right) \tag{10}$$

$$\hat{\sigma}^2(\hat{\beta}_{Iw}(x,y,z)) = (X'X)^{-1}\frac{\sum_{i=1}^{M} r_w(x,y,z,t_i)^2}{df} \tag{11}$$

where $\lambda_{\beta_I}()$ is the thresholding operator calculated and used for the image of the parameter of interest. Eq. (11) can be compared with Eq. (5) and extended in the same manner as with Eq. (6).

The procedure

The procedure for analysis of the image data is therefore as follows

Step 1. consists of calculation of $Y^W(w_x, w_y, w_z, t)$ as in Eq. (7).
Step 2. consists of application of $\eta()$ using the design matrix X to each wavelet vector in time to produce a parametric image $\beta_I^W(w_x, w_y, w_z)$.
Step 3. is to calculate the residuals $r^W(w_x, w_y, w_z, t)$ as in Eq. (4).
Step 4. is to calculate the variance using Eq. (5) from the residuals $r^W(w_x, w_y, w_z, t)$ and X, the model matrix. This is required for the thresholding estimation, specifically the estimate of τ in Eq. (8).
Step 5. applies $\lambda_{\beta_I}()$ to the parameter image $\beta_I^W(w_x, w_y, w_z)$, and to the residuals $r^W(w_x, w_y, w_z, t)$.
Step 6. applies the inverse wavelet filter W^{-1} to both the thresholded parametric image and the thresholded residuals.
Step 7. calculates the variance of $\hat{\beta}_{I_w}(x,y,z)$ using $r_w(x,y,z,t)$ as in Eq. (11).

Data examples

Three examples from differing modalities are presented. The first is an fMRI activation study. Here we also look at whether the decorrelation assumption is adequate to transform variances directly, as well as validating the method without thresholding. The others involve PET in two different settings; a metabolism study and a neuroreceptor ligand study.

fMRI spatiotemporal analysis

There is growing interest in the application of spatiotemporal wavelet models in fMRI, as it is understood that there is often more

sensitive analysis available if local information is combined. Indeed, there has been investigation into the use of fMRI in the wavelet domain, and a preliminary study carried out (Desco et al., 2001). This study looked at how different wavelet bases performed at different noise and signal levels. However, these studies tend to focus on ideal conditions and further work is needed to evaluate methods using more complex analysis schemes with more realistic experimental noise conditions.

It should be noted here that there is an inherent problem with three-dimensional spatiotemporal modeling in fMRI. It has been shown in PET that there is a difference between two- and three-dimensional wavelet modeling (Turkheimer et al., 2000b). Slices in fMRI are not acquired simultaneously but rather in a sequence of the duration of the time frame (TR). This leads to problems when trying to infer a spatial model in the data, as data that are next to each other in the image may have been collected up to the TR (usually in the order of 2 s apart). If the model is focusing on effects that are very short in duration, this can cause considerable differences in the information sequence in two neighboring pixels across slices. Thus careful decisions must be made in choosing the models both in a spatial and temporal sense to take account of this problem. In the wavelet setting, it is convenient to only use a wavelet spatial model slice by slice (in effect a two-dimensional model) so the temporal information in each pixel and wavelet time sequence is not distorted. This is similar to the image domain approach in Worsley et al. (2002), where a modified linear model is applied at each slice to correct for slice timing.

A standard linear model along the lines of the general linear models used in SPM (Friston et al., 1995) was used for the temporal analysis of the data. The temporal data in fMRI are well known to be autocorrelated. However, depending on the method chosen to deal with these errors, the procedure described here will continue to be applicable. The methods that account for autocorrelation in SPM99 (Friston et al., 2000) and SPM2 (Friston et al., 2002) while being very different from one another both possess the feature that the same overall model is used at every voxel. Indeed, this is one of the important desires mentioned in Friston et al. (2000). This feature means that if this model is also used in wavelet temporal analysis, the procedure follows through as above (with the small modification for the weighting factor). However, if there is a voxel-by-voxel estimate of autocorrelation in the analysis (i.e., a different model is used at each voxel such as in the case of spatially variant prewhitening), then the results of using wavelet analysis and image analysis to estimate the temporal autocorrelation are no longer equivalent. Methods such as that suggested by Worsley et al. (2002) estimate the autocorrelation at each voxel and then prewhiten the data using the spatially smoothed but not globally pooled autocorrelations across the image. This will not give the same parameter values in wavelet space and image space even when the identity threshold is used, in contrast to the standard linear model or the SPM techniques mentioned above. As was noted before, the wavelet transform is a spatial transform and so any spatially invariant temporal linear model will commute with the wavelet transform. Here a temporal method similar to SPM99 was used to analyze the data.

The wavelet analysis in this fMRI study was carried out using orthonormal cubic-spline Battle–Lemarie wavelets, which have been shown to be suitable bases for functional images (Ruttimann et al., 1998; Unser et al., 1995). The length of the filters was chosen as the minimal that produced efficient signal representation. The actual lengths used were 16 coefficients with four resolution levels (depths). These were linearly thresholded as described previously (Turkheimer et al., 2003). The wavelet transform was calculated using the wavelet toolbox Uvi-wave 3.0 (Śanchez SG et al, Grupo de Teoria de al Señal, Universidad de Vigo, Spain).

Measured data

Measured data previously analyzed in the image domain (Chen et al., 2002; Worsley et al., 2002) were used to test the method. Wavelet analysis was carried out in two-dimensional slice by slice due to the slice timing effects as mentioned above. The fMRI data set came from a study with a pain stimulus. A hot stimulus and warm stimulus were applied to the skin of a subject, with interleaved periods of rest. The contrast of interest looked at the difference between the hot and the warm condition.

As can be seen in Fig. 1, the wavelet image of the effects shows the areas where there is a difference between the hot and cold stimulus. The identity threshold has been applied in these images, as the purpose of this comparison is to show that analysis in the image domain and wavelet domain can be transformed backwards and forwards. Three different methods of calculating the variance are shown. The figures have been displayed using a monochromatic color map as the primary purpose of this study is to show how signal and error estimates can be found. Fig. 1b contains the variance calculated directly in image space. This is the reference image to which comparison with Figs. 1c and d needs to be made. Fig. 1c contains the variance image if complete decorrelation between the wavelet coefficients is assumed. This approximate method is calculated by taking the matrix product of the diagonal variance matrix and the discrete wavelet transform matrix and its transpose (Aston et al., 2003). A detailed description of the method is omitted due to both its complexity and the unsatisfactory nature of its performance relative to the proposed method. As can be seen in Fig. 1c, the resulting image has both a loss of magnitude in the errors and also the loss of resolution in the image. This shows that while the wavelet transform is well known to decorrelate the data, it is not possible to assume that the coefficients have zero correlation between them, certainly if only a small temporal sample is available. This variance map does retain the same layout as the true image, but does not provide an adequate estimate of the variance. Fig. 1d, however, perfectly captures the variance image of Fig. 1b as they are analytically identical. Thus the residuals method can be used to generate variance images back in image space for parameters calculated in wavelet space.

Table 1 contains the shrinkage values for the fMRI study, which were calculated using the James–Stein procedure. Image maps in Fig. 2 were produced through shrinkage using these values and by ensuing transformation back into the image domain. As can be seen, the parameter estimates in Fig. 2c are not changed to a great extent from those in Fig. 2a, but the variance in Fig. 2d is reduced from that in Fig. 2b. This is mainly due to the fact that the wavelet coefficients in the highest level are not shrunk as they are assumed to contain only signal, whereas noise is distributed across the coefficients.

PET spatiotemporal analysis—metabolism

A dynamic [^{18}F]Fluorodeoxyglucose (FDG) study of glucose metabolism in normal brain was also considered. Arterial samples were available and were used as input function for the method.

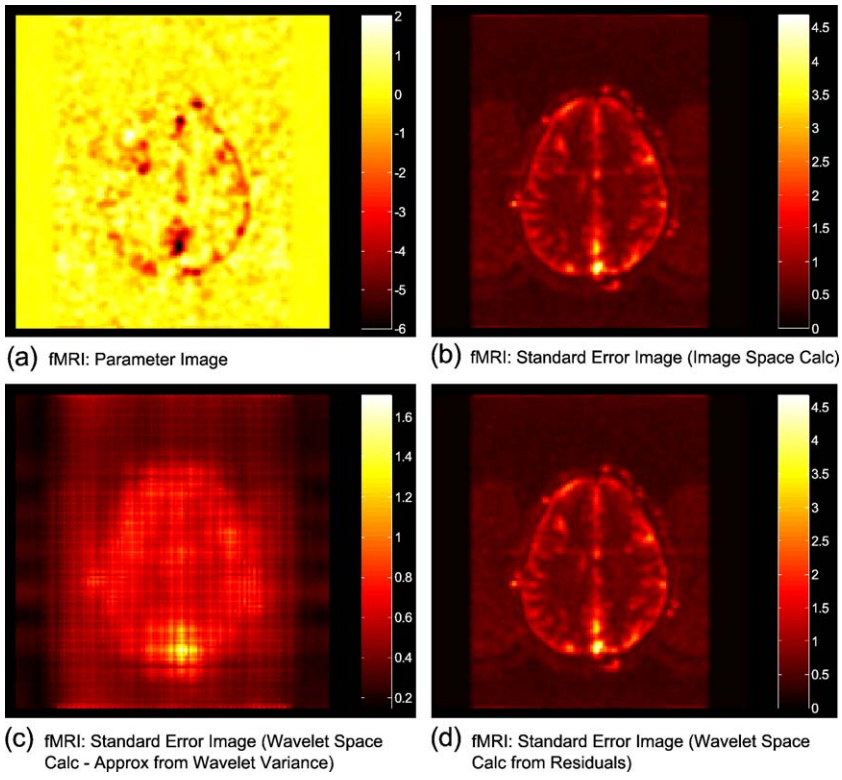


Fig. 1. fMRI activation study investigating the effects of warm and hot stimuli. These images are the differences between the warm and hot stimuli and the standard errors associated with the difference. Here no shrinking of the wavelet coefficients (i.e., the identity threshold) has been carried out to provide an exact equivalence between the methods in the image domain and the wavelet domain.

In the FDG model (Phelps et al., 1979; Sokoloff et al., 1977), transport of FDG across the blood–brain barrier is described by K_1 (ml/g/min), the reverse process by k_2 (1/min), and phosphorylation by k_3 (1/min). Total cerebral activity is then given by

$$C_T(t) = bC_p(t) + \frac{K_1k_2}{k_2+k_3}\int_0^t C_p(\tau)e^{-(k_2+k_3)(t-\tau)}d\tau + \frac{k_2k_3}{k_2+k_3}\int_0^t C_p(\tau)d\tau \quad (12)$$

where C_T and C_p are the tissue and input time activities, respectively, and b is the intravascular distribution volume of the tracer (either blood or plasma). (k_2k_3) / $(k_2 + k_3)$ is the FDG metabolic rate and is the parameter of interest.

After a certain time from tracer injection, the free tracer in the tissue reaches equilibrium with the plasma. Therefore, the model reduces to a simple one-compartment model and Eq. (12) reduces to

$$C_T(t) = bC_p(t) + \frac{K_1k_2}{(k_2+k_3)^2}C_p(t) + \frac{k_2k_3}{k_2+k_3}\int_0^t C_p(\tau)d\tau. \quad (13)$$

This allows estimation of the metabolic rate by linear least squares (Turkheimer et al., 2000a).

PET images do not suffer from the temporal autocorrelations that are present in the fMRI study and as such is not necessary to consider here. Wavelet analysis was again carried out using Battle–Lemarie wavelets with the same parameters (although a three-dimensional setup was used), and these were again linearly thresholded.

Table 1
Shrinkage values from fMRI study

x	H	L	H
y	L	H	H
k	Shrinkage values		
1	0.9442	0.9492	0.9411
2	0.9658	0.9697	0.9641
3	0.9835	0.9827	0.9812

x and y refer to the portions of the two-dimensional data structure involved and k is the level of the transform. H and L are high and low pass filtering, respectively.

Measured data

There are nonstationary noise conditions in FDG images, thus wavelet methods are needed over methods that assume stationarity. The data set is a full three-dimensional image volume and can be processed fully in three dimension due to the data being collected simultaneously. The shrinkage values for different portions of the wavelet image are given below in Table 2. These values indicate the proportion of the signal that is shrunk in the linear thresholding. Again, those wavelet coefficients in the highest level were assumed to be signal and consequently not shrunk.

The results of the measured FDG data are shown in Figs. 3 and 4 below. Again, the identity threshold wavelet analysis and the traditional analysis produced identical answers (not shown).

The figures below give the analysis when the linear shrinkage thresholds are applied in three dimension. The metabolic rate

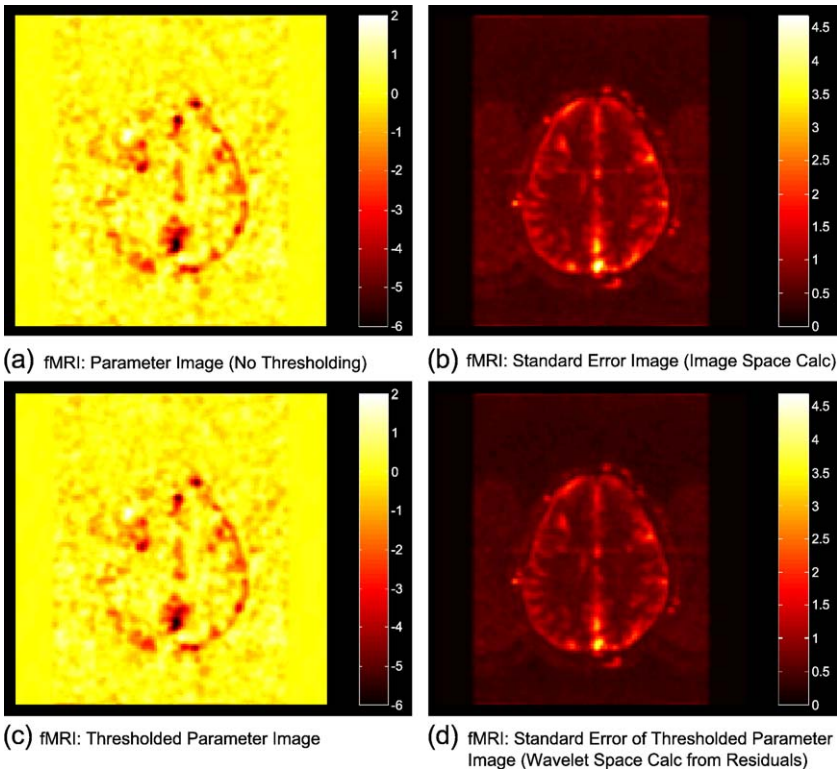


Fig. 2. fMRI activation study investigating the effects of warm and hot stimulus. These images are the differences between the warm and hot stimuli and the standard errors associated with the difference. These were calculated in wavelet space and a linear (James–Stein) threshold estimator used to linearly shrink the data. The values for the shrinkage are given in Table 1.

parameter image is not changed to a great extent when applying linear thresholding but the variance is again reduced as in the fMRI case. These demonstrate that the techniques described are independent of the dimensionality of the image and can thus be carried out in either two or three dimension.

PET spatiotemporal analysis—neuroreceptors

The second PET example considered a dynamic [^{11}C]raclopride study of D_2-receptor distribution in normal brain. Wavelets have been previously used to analyze neuroreceptor transmitter data (Cselenyi et al., 2002; Millet et al., 2000; Turkheimer et al., 2000a). Frames were acquired with an ECAT 953B PET camera (CTI/Siemens). Arterial samples were not available, and a reference region placed on the cerebellum was used as input function. Estimates of the distribution volume ratio (DVR) were obtained with the Logan plot (Logan et al., 1990, 1996). The original equation of this plot reads

$$\int_0^t C_i(\tau)\mathrm{d}\tau/C_i(t) = \mathrm{DVR}\int_0^t C_{\mathrm{ref}}(\tau)\mathrm{d}\tau/C_i(t) + c \qquad (14)$$

where $C_i(t)$ is the tissue time activity, $C_{\mathrm{ref}}(t)$ is the time activity in the reference region that is assumed to be devoid of D_2-receptors, and c is a constant. Although $C_{\mathrm{ref}}(t)$ can be assumed noise-free as being sampled on a large ROI, the same cannot be said for the term $C_i(t)$, which appears on both sides of the linear equation. However, Carson (1993) showed that the simple rearrangement of Eq. (14) as

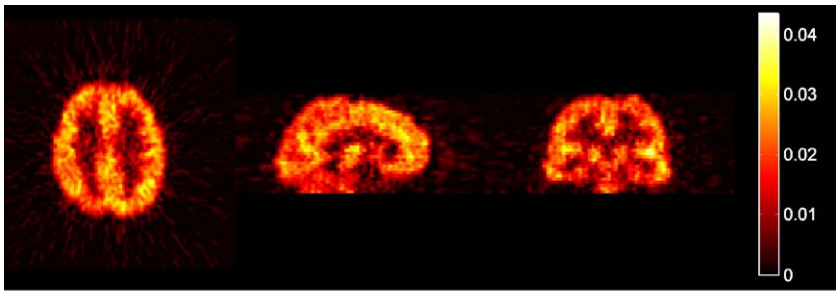
$$C_i(t) = -(\mathrm{DVR}/c)\int_0^t C_{\mathrm{ref}}(\tau)\mathrm{d}\tau + (1/c)\int_0^t C_i(\tau)\mathrm{d}\tau \qquad (15)$$

substantially reduces the bias of the estimated parameters even with noise levels typical of voxel time courses and allows the use

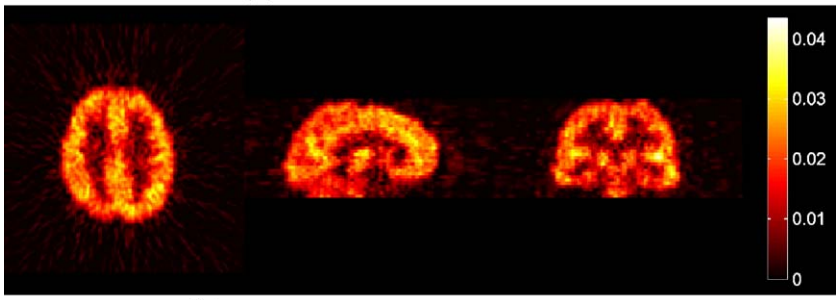
Table 2
Shrinkage values from PET FDG study

x	H	L	H	L	H	L	H
y	L	H	H	L	L	H	H
z	L	L	L	H	H	H	H
k	Shrinkage values						
1	0.8671	0.8564	0.8790	0.8916	0.8578	0.8480	0.8743
2	0.9125	0.9096	0.8973	0.9585	0.8897	0.8947	0.8829
3	0.9654	0.9687	0.9351	0.9940	0.9833	0.9829	0.8960
4	0.9961	0.9961	0.9688	0.9986	0.9967	0.9974	0.9538

x, y, and z refer to portions of the three-dimensional data structure involved and k is the level of the transform. Here H and L refer to high and low pass filtering, respectively.



(a) FDG Analysis Image of Metabolism Rate



(b) FDG Post Wavelet Analysis Image of Metabolism Rate

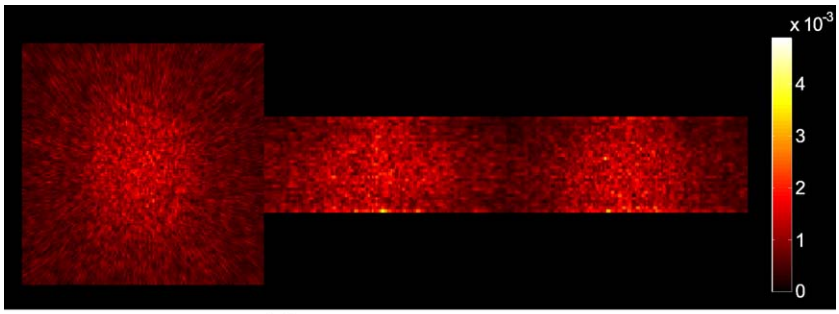
Fig. 3. Glucose metabolism rate as measured in a PET FDG scan. The image has been linearly thresholded in the wavelet domain and then returned to the image domain.

of the Cramer–Rao lower bound (Beck and Arnold, 1977) to estimate the standard deviation of the parameters.

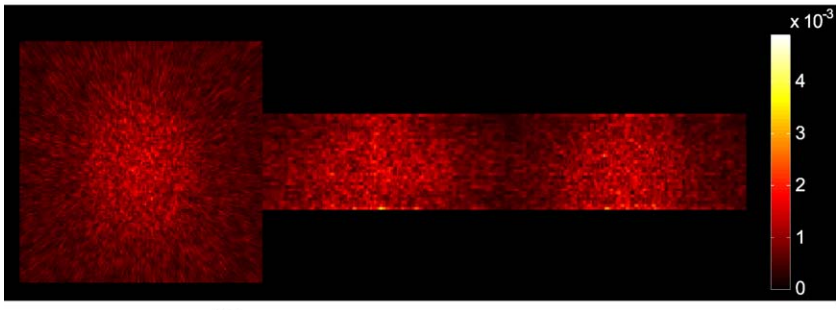
The frames for the time interval 5–60 min (nine frames) were used for the linear estimation of DVR after previous verification that the plot of Eq. (14) showed linear behavior in this time window. Again, Battle–Lemarie wavelets were used with the same parameters as before.

The model in Eq. (14) can be seen to be linear as is its rearrangement in Eq. (15). However, in the latter formulation, DVR is no longer a linear parameter of the model. The advantage of the formulation of Eq. (15) is that it provides more accurate measures of variance of the parameters due to the underlying nonlinear nature of the true model (Carson et al., 1993).

The methods developed in this paper are only of theoretical validity when applied to truly linear parameters in linear models. However, due to the approximately linear nature of the model in Eq. (15), it was deemed worth exploring the benefits of trying the method in this case. Further work would be required to determine



(a) FDG Analysis Standard Error Image



(b) FDG Post Wavelet Analysis Standard Error Image

Fig. 4. Glucose metabolism rate as measured in a PET FDG scan. The image has been linearly thresholded in the wavelet domain. This is an image of the associated errors.

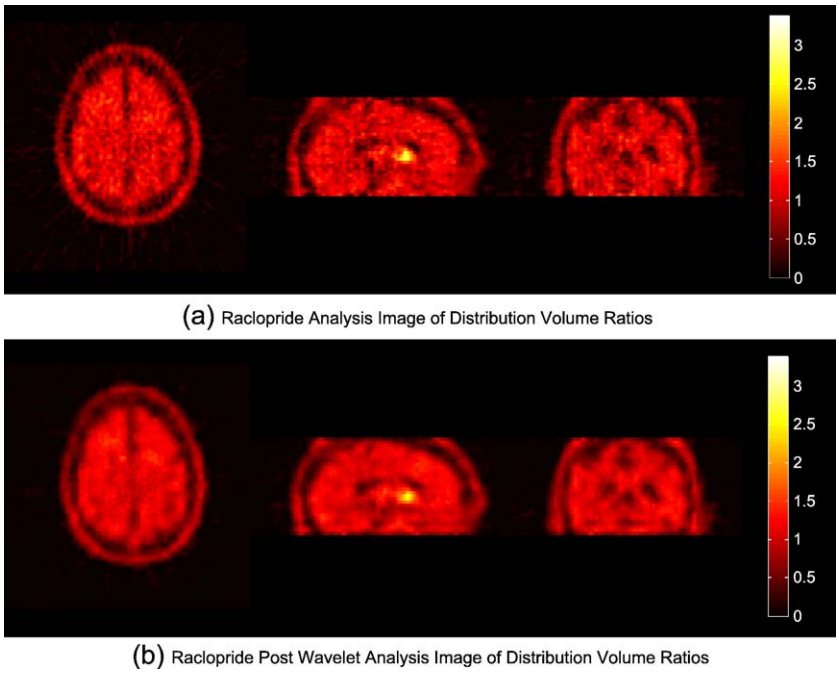


Fig. 5. [^{11}C]Raclopride PET image of DVR. The image has been linearly thresholded in the wavelet domain and then returned to the image domain.

how much of an approximation this is, probably through the use of a comprehensive simulation study.

Fig. 5 shows the results of parameter estimation using wavelet coefficients. As can be seen, the wavelet image (Fig. 5b) is smoother compared with the image space analysis (Fig. 5a). This is due to the smooth signal being present in the coarsest level, while noise is in the lower levels, which as can be seen from Table 3 are shrunk heavily. The image space analysis and wavelet space analysis with the identity threshold again produced identical parameter images (not shown) due to the means of the parameters being the same regardless of which linear model is used.

Fig. 6 shows the standard error images from the Raclopride study. Due to the linearization of the Logan plot, the variances between the image space calculation and the wavelet space calculation differ. However, as linearization is not spatially dependent, it can be argued that neither image space nor wavelet space estimation of the parameters is intrinsically more correct. Indeed, with better signal to noise properties in the wavelet coefficients for those containing signal, it is probable that the analysis will indeed be more accurate in wavelet space. This is borne out by looking at Figs. 6a and b. It is expected that the variance of DVR will be proportional to signal (Aston et al., 2000) and as such structure will be present in the DVR variance image. In Fig. 6a, the image space variance estimate, there is no structure present, whereas in Fig. 6b there are signs of structure. As mentioned above, further simulations and theoretical work will need to be carried out in order to validate this technique further.

Again, there is a reduction in variance in the wavelet smoothed image (Fig. 6c). The reduction is greater than in the FDG image, as the shrinkage values are much smaller for many components. There are greater amounts of structure present in the thresholded variance than either of the other images.

Conclusions

Wavelet-based methods are increasingly used to perform neuroimaging analysis. However, due to the need for comparison between methods, general acceptance of these methods has been restricted to special circumstances.

The methods described above yield an image of the errors on the estimates obtained using wavelet modeling. This allows error

Table 3
Shrinkage values from PET raclopride study

x	H	L	H	L	H	L	H
y	L	H	H	L	L	H	H
z	L	L	L	H	H	H	H
k	Shrinkage values						
1	0.3520	0.3675	0.2828	0.0353	0.4364	0.5420	0.5084
2	0.2075	0.3327	0.1370	0.7320	0.3212	0.0769	0.3960
3	0.8572	0.8842	0.4336	0.9451	0.8554	0.9017	0.2564
4	0.9263	0.9637	0.5420	0.9792	0.9818	0.9579	0.1474

x, y, and z refer to portions of the three-dimensional data structure involved and k is the level of the transform. H and L refer to high and low pass filtering, respectively.

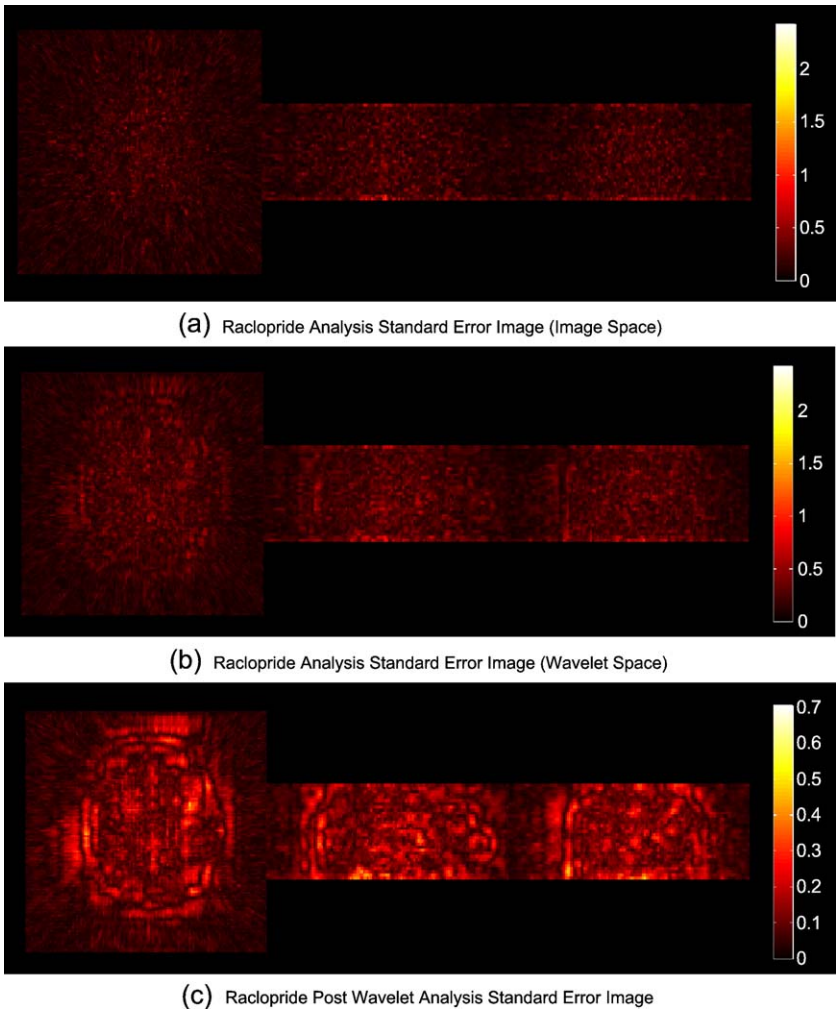


Fig. 6. [^{11}C]Raclopride PET image of DVR. The image has been linearly thresholded in the wavelet domain. This is an image of the associated errors along with those from the unthresholded studies.

estimates that have previously been calculated in image space to be compared with those in wavelet space, assuming a linear temporal model. These methods can be used with existing linear temporal modeling methods for neuroimaging while allowing a different form of noise suppression to be used to allow better MSE estimates of the parameters.

It is important to note that the methods here cannot be directly applied to hypothesis testing in the way BLU estimates can. In traditional SPM type methods, the parameters are subsequently t tested (using Student's t distribution) to evaluate experimental hypotheses. The distributional properties of the wavelet estimators are less clear cut, and as such direct hypothesis testing cannot be performed easily. However, often it can be of more interest to look at the signal itself as opposed to a binary yes/no decision (Turkheimer et al., 2004), and this can be facilitated by wavelet methods. Parametric maps regularized by wavelet methods and associated errors may play an important role in the newly proposed dissociation type inference for functional studies (Jernigan et al., 2003). In this framework, the statistics obtained from the temporal model are used for within-brain comparisons to assess significant differential association (i.e., dissociation) of one location compared to others. This approach can benefit from the regularization properties of wavelet filters that reduce the noise in the map but still allow inference when a reduced variance map can be produced. Future work will address this issue further.

Here the methods have been developed for linear temporal models. This is due to the need for linearity in the time domain to allow wavelet and image spaces to commute. However, it is possible that for some classes of nonlinear models, these techniques will provide good approximations even without true linearity being present. Indeed, a slight deviation from linearity in the third example suggests that this may provide a better estimation of variance than traditional methods, despite its approximate nature. It is also the case that the thresholding techniques used here are linear shrinkage operators. This technique can be used with nonlinear threshold techniques, and the linear techniques here are just an example. While the variance parameters calculated do not depend on the type of thresholding used, care must be taken in choosing appropriate thresholding to ensure that they do indeed lead to better MSE estimates of the parameters. Further investigation would also be of interest if random effect models were to be used in subsequent analysis, as the distribution of coefficients would no longer be from the standard linear model, and use of the variance estimates possibly more complex.

Hopefully the advantages of wavelet methods will continue to be recognized as more and more studies see the potential benefits of MSE estimates as opposed to traditional least squares estimates.

Acknowledgments

The authors would like to thank two anonymous referees for their very helpful comments. This work was partially funded by National Science Council (Taiwan) NSC-93-2118-M-001-031.

References

Ashburner, J., Friston, K., Holmes, A.P., Poline, J.B., 1999. Statistical Parametric Mapping, second ed. Wellcome Department of Cognitive Neurology, London.

Aston, J.A.D., Gunn, R.N., Worsley, K.J., Ma, Y., Evans, A.C., Dagher, A., 2000. A statistical method for the analysis of positron emission tomography neuroreceptor ligand data. NeuroImage 12, 245–256.

Aston, J.A.D., Turkheimer, F.E., Cunningham, V.J., Gunn, R.N., 2003. Wavelet variance components in image space for spatio-temporal neuroimaging data. Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing X, vol. 5207. SPIE, Bellingham, WA, pp. 849–857. San Diego CA, USA, Part II, August 3–8.

Beck, J.V., Arnold, K.J., 1977. Parameter Estimation in Engineering and Science. Wiley and Sons, New York.

Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., Brammer, M., 2003. Wavelets and statistical analysis of functional magnetic resonance images of the human brain. Stat. Methods Med. Res. 12, 375–399.

Carson, R.E., 1993. PET parameter estimation using linear integration methods: bias and variability considerations. In: Uemura, K., Lassen, N., Jones, T., Kanno, I. (Eds.), Quantitative of Brain Function. Excerpta Medica, pp. 499–507.

Carson, R.E., Yan, Y., Daube-Witherspoon, M.E., Freedman, N., Bacharach, S.L., Herscovitch, P., 1993. An approximation formula for the variance of PET region-of-interest values. IEEE Trans. Med. Imag. 12 (2), 240–250.

Chen, J.-I., Ha, B., Bushnell, M.C., Pike, B., Duncan, G.H., 2002. Differentiating noxiousand innocuous-related activation of human somatosensory cortices using temporal analysis of fmri. J. Neurophysiol. 88, 464–474.

Chui, C.K., 1992. An Introduction to Wavelets. Academic Press, New York.

Cselenyi, Z., Olsson, H., Farde, L., Gulyas, B., 2002. Wavelet-aided parametric mapping of cerebral dopamine D2 receptors using the high affinity PET radioligand [^{11}C]FLB 457. NeuroImage 17, 47–60.

Daubechies, I., 1992. Ten lectures on wavelets. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Desco, M., Hernandez, J.A., Santos, A., Brammer, M., 2001. Multiresolution analysis in fMRI: sensitivity and specificity in the detection of brain activation. Hum. Brain Mapp. 14 (1), 16–27.

Friston, K., Holmes, A.P., Worsley, K.J., Poline, J.B., Frith, C., Frackowiak, R., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Josephs, O., Zarahn, E., Holmes, A.P., Rouquette, S., Poline, J.-B., 2000. To smooth or not to smooth: bias and efficiency in fMRI time-series analysis. NeuroImage 12, 196–208.

Friston, K.J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: theory. NeuroImage 16, 465–483.

Jernigan, T.L., Gamst, A.C., Fennema-Notestine, C., Ostergaard, A.L., 2003. More "mapping" in brain mapping: statistical comparison of effects. Hum. Brain Mapp. 19, 90–95.

Logan, J., Fowler, J.S., Volkow, N.D., Wolf, A.P., Dewey, S.L., Schlyer, D.J., MacGregor, R.R., Hitzemann, R., Bendriem, B., Gatley, S.J., et al., 1990. Graphical analysis of reversible radioligand binding from time–activity measurements applied to [*n*-11c-methyl]-(−)-cocaine PET studies in human subjects. J. Cereb. Blood Flow Metab. 10 (5), 740–747.

Logan, J., Fowler, J.S., Volkow, N.D., Wang, G.J., Ding, Y.S., Alexoff, D.L., 1996. Distribution volume ratios without blood sampling from graphical analysis of pet data. J. Cereb. Blood Flow Metab. 16 (5), 834–840.

Mallat, S., 1999. A Wavelet Tour of Signal Processing. Academic Press.

Millet, P., Graf, C., Buck, A., Walder, B., Westera, G., Broggini, C., Arigoni, M., Slosman, D., Bouras, C., Ibanez, V., 2000. Similarity and robustness of PET and SPECT binding parameters for benzodiazepine receptors. J. Cereb. Blood Flow Metab. 20, 1587–1603.

Müller, K., Lohmann, G., Zysset, S., von Cramon, D.Y., 2003. Wavelet statistics of functional MRI data and the general linear model. J. Magn. Reson. Imaging 17, 20–30.

Percival, D.B., Walden, A.T., 2000. Wavelet Methods for Time Series Analysis. CUP, Cambridge, UK.

Phelps, M.E., Huang, S.C., Hoffman, E.J., Selin, C., Sokoloff, L., Kuhl, D.E., 1979. Tomographic measurement of local cerebral glucose metabolic rate in humans with [F-18]2-fluoro-2-deoxy-D-glucose: validation of method. Ann. Neurol. 6, 371–388.

Ruttimann, U.E., Unser, M., Rawlings, R.R., Rio, D., Ramsey, N.F., Mattay, V.S., Hommer, D.W., Frank, J.A., Weinberger, D.R., 1998. Statistical analysis of functional MRI data in the wavelet domain. IEEE Trans. Med. Imag. 17 (2), 142–154.

Sokoloff, L., Reivich, M., Kennedy, C., Des Rosiers, M.H., Patlak, C.S., Pettigrew, K.D., Sakurada, O., Shinohara, M., 1977. The ^{14}C-deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. J. Neurochem. 28 (5), 897–916.

Turkheimer, F.E., Brett, M., Visvikis, D., Cunningham, V.J., 1999. Multiresolution analysis of emission tomography images in the wavelet domain. J. Cereb. Blood Flow Metab. 19 (11), 1189–1208.

Turkheimer, F.E., Banati, R.B., Visvikis, D., Aston, J.A.D., Gunn, R.N., Cunningham, V.J., 2000. Modeling dynamic PET-SPECT studies in the wavelet domain. J. Cereb. Blood Flow Metab. 20 (5), 879–893.

Turkheimer, F.E., Brett, M., Aston, J.A.D., Leff, A.P., Sargent, P.A., Wise, R.J., Grasby, P.M., Cunningham, V.J., 2000. Statistical modeling of positron emission tomography images in wavelet space. J. Cereb. Blood Flow Metab. 20 (11), 1610–1618.

Turkheimer, F., Aston, J.A.D., Banati, R.B., Riddell, C., Cunningham, V.J., 2003. A linear wavelet filter for parametric imaging with dynamic PET. IEEE Trans. Med. Imag. 22, 289–301.

Turkheimer, F., Aston, J.A.D., Cunningham, V.J., 2004. On the logic of hypothesis testing in functional imaging. Eur. J. Nucl. Med. Mol. Imaging 31, 725–732.

Unser, M., Thevénaz, P., Lee, C., Ruttimann, U., 1995. Registration and statistical analysis of PET images using the wavelet transform. IEEE Eng. Med. Biol. Mag. 14 (5), 603–611 (September–October).

Van De Ville, D., Blu, T., Unser, M., 2003. Wavelets versus resels in the context of fMRI: establishing the link with SPM. Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing X, vol. 5207. SPIE, Bellingham, WA, pp. 417–425. San Diego CA, USA, Part I, August 3–8.

Van De Ville, D., Blu, T., Unser, M., 2004a. Integrated wavelet processing and spatial statistical testing of fMRI data. NeuroImage (in press).

Van De Ville, D., Blu, T., Unser, M., 2004b. Wavelet-based fMRI statistical analysis and spatial interpretation: a unifying approach. Proceedings of the Second 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'04). IEEE, pp. 1167–1170. Arlington VA, USA, April 15–18.

Worsley, K.J., Liao, C., Aston, J.A.D., Petre, V., Duncan, G., Evans, A.C., 2002. A general statistical analysis for fMRI data. NeuroImage 15 (1), 1–15.