Assessing the significance of focal activations using their spatial extent.

Friston KJ[1], Worsley KJ[2], Frackowiak RSJ[1], Mazziotta JC[3] & Evans, AC[2]

1 The MRC Cyclotron Unit, London, W12 OHS UK

2 The Montreal Neurological Institute, McGill University, Montreal, H3A 2B4 Canada

3 Reed Neurological Research Center, UCLA School of Medicine, CA 90024 USA

The MRC Cyclotron Unit

Hammersmith Hospital

DuCane Road

London W12 OHS

Tel: 44 81 740 3162

Fax: 44 81 743 3987

Key words: Statistical Parametric Mapping, Gaussian Fields, Activation, Significance, Thresholds, Excursion set, Functional Imaging.

Abstract

Current approaches to detecting significantly activated regions of cerebral tissue use statistical parametric maps, which are thresholded to render the probability of one or more activated regions *of one voxel, or larger,* suitably small (e.g. 0.05). We present an approximate analysis which gives the probability that one or more activated regions *of a specified volume, or larger,* could have occurred by chance.

These results mean that the detection of significant activations no longer depends on a fixed (and high) threshold but can be effected at any (lower) threshold, in terms of the spatial extent of the activated region. The substantial improvement in sensitivity which ensues is illustrated using a power analysis and a simulated phantom activation study.

Introduction

Functional images of the brain are, almost universally, compared using some form of statistical parametric mapping. Statistical parametric maps (SPMs) have voxel values that, under the null hypothesis, are distributed according to some known probability density function (Friston et al 1990). The most commonly employed statistics are the Students t (Friston et al 1990, Worsley et al 1992) and the correlation coefficient (e.g. Friston et al 1993).

Any continuous pdf (probability density function) can be transformed to the Gaussian distribution or z-statistic. If the degrees of freedom of the original distribution are reasonably high the resulting SPM approximates a Gaussian field or SPM(z). The analyzes presented in this paper are therefore restricted to the SPM{z}. It should be noted that some exact results for SPM{t}, SPM{F} and SPM{χ^2} have now been derived (Worsley 1994). The most common characterization of SPMs involves the identification of significant activation foci. This is achieved by thresholding. The problem of how to correct for the multiplicity of *non-independent* tests implicit in this approach has, in past years, been solved. The solution identifies a threshold such that for a SPM of given size, the probability of obtaining one or more *activation foci of at least one voxel*, by chance, is suitably small (for example 0.05). This approach uses the theory of level crossings in stochastic processes (Friston et al 1991), or the Euler characteristic (Worsley et al 1992), applied to Gaussian processes with a known (or measurable) auto-correlation.

The threshold identified, using current techniques, takes no account of the spatial extent of an activation. Activation foci are characterized not only by the threshold they reach but also by their spatial topography, for example their shape, spatial relationships to each other and their size. This paper deals only with size or volume. A spatially limited focus is usually considered less 'significant' than a very large one (see Poline & Mazoyer 1993 and Roland et al 1993 for empirical studies in this area). Clearly the detectability of significant foci would be enhanced if the volume of activated tissue was explicitly included when testing the null hypothesis that the activated region

could have occurred by chance. More precisely, for any region, one would like to ask: 'Over a SPM of given size, what is the probability of obtaining one or more *activation foci of the same size, or larger, than the one in question?* Another way of looking at this formulation is to compare current approaches, which provide a critical value for the maximum height of a peak, to approach proposed here, which asks "what is the critical value for the maximum size of a focus?". The purpose of this article is to present an approximate analysis of the probability theory that is needed to answer these questions. One important aspect of assessing significance in terms of spatial extent is that the analysis is freed from the arbitrary nature of fixed and high thresholds (in the sense that significance can be assessed at any threshold).

Below we provide a theory section which defines the problem addressed, the solutions obtained and the implications for detecting changes using a power analysis. We then describe a simple simulated phantom experiment which demonstrates the application, of the theory, to the detection of significant activations. More general applications of the results, to real data, will be deferred until later publications.

Theory

From a statistical perspective there are three things about a SPM that are of interest: (i) The number (N) of voxels above a threshold (the number of voxels in the excursion set which have values greater than a threshold u), (ii) the number (m) of activated regions (clusters or connected subsets of the excursion set) and (iii) the number (n) of voxels in each of these clusters. Each of these numbers has its own probability density function; $P(N = x)$, $P(m = x)$ and $P(n = x)$ [in this context $P(n = x)$ is strictly a conditional probability, given that the region exists = $P(n = x \mid m \geq 1)$]. These probability functions provide a fairly complete characterization of the SPM and allow one to address a number of hypotheses.

In what follows we apply results from the theory of Gaussian Fields to a *D*-dimensional lattice of continuous random variables (a voxelated SPM) by considering it as a good lattice representation of an underlying continuous Gaussian field. Although the parameters *N* and *n* are, in reality, discrete variable (numbers of voxels) they are treated here like continuous measures of volume. The volumes in question are the spatial volume occupied by voxels above a certain threshold (i.e. the number of voxels in a suprathreshold cluster).

The particular probability we are interested in is: *The probability of obtaining at least one activation with k voxels or more*. This is the same as the *probability that the largest region has k voxels or more* = $P(n_{max} \geq k)$, where n_{max} is the number of voxels in the biggest region. It should be noted that this probability is a more general case of that which is currently used, namely: The probability of getting at least one activation (with one voxel or more) = $P(n_{max} \geq 1)$. This is simply equal to the probability of getting at least one region $P(m \geq 1)$. $P(m \geq 1)$ is usually set to 0.05 by choosing an appropriate threshold or critical height (u). This threshold is chosen by estimating the expectation or mean of m ($E\{m\}$) and using the fact that

$$P(m \geq 1) \quad \leq \quad E\{m\} \qquad 1$$

tends to equality at high thresholds (Hasofer 1978). An approximate expression for $E\{m\}$ was presented in Friston et al (1991) using a somewhat heuristic argument. More exact results will be found in Hasofer (1978) which pertain to the number of maxima. Because at high threshold (u) the number of maxima (m_{maxima}) and m converge, these expressions should be equivalent (more generally $m \leq m_{maxima}$). Subsequently the Euler characteristic (χ) has been used as an estimate of $E\{m\}$ (Worsley at al 1992). Again the Euler characteristic and m converge at high u (more generally $\chi \leq m$). The Euler characteristic is more amenable to mathematical analysis than the number of maxima.

The problem addressed in this article is how to estimate $P(n_{max} \geq k)$ and therein obtain a critical threshold *for volume*. To do this one needs to know both the probability of obtaining an arbitrary number of regions $P(m = x)$ and the probability that these regions have less than k voxels $P(n < k)$. Unfortunately, in the context of Gaussian processes, there are no exact results for these probabilities. However, using a less formal analysis one can estimate the probability density functions using what is already known.

In the next three sections we introduce expressions for the expectations and probability density functions of m and n and combine these results to give the final expression for $P(n_{max} \geq k)$. In brief, we use a Poission form for $P(m = x)$ and a standard approximation for its expectation (Hasofer 1978; Adler 1981). The pdf for $P(n = x)$ uses an earlier observation (Nosko 1969a) that $n^{2/D}$ has an exponential distribution, in conjunction with the above result for $E\{m\}$.

Expectations of N, m and n

The three variables N, m and n have expectations which are related:

$$E\{N\} = E\{m\}.E\{n\} \qquad 2$$

The expectation of N is known exactly because of the Gaussian univariate assumptions and is simply the appropriate integral under the normal distribution or error function $\Phi(.)$ (see Friston et al 1990). As mentioned above we already know the approximate expectation of m. For any threshold u the expectations for a D-dimensional process of volume S are given by:

$$E\{N\} = S \cdot \Phi(-u) = S \int_u \cdot (2\pi)^{-1/2} e^{-x^2/2} dz \qquad 3$$

$$E\{m\} \approx S (2)^{-(D+1)/2} W^{-D} u^{D-1} e^{-u^2/2} \quad \text{(Hasofer 1978)} \qquad 4$$

$$E\{n\} = E\{N\} / E\{m\} \qquad 5$$

W is a measure of smoothness and is related to the full width at half maximum (FWHM) of the SPM's "resolution". Equivalently *W* is inversely related to the number of "resolution elements" or Resels (*R*) that fit into the total volume (*S*) of the *D*-dimensional SPM. ($R = S/FWHM_D$). More exactly, let SPM_i be the voxel value of the SPM as a function of the ith coordinate and similarly let ρ_i be the spatial autocorrelation of the SPM as a function of the *i*th coordinate, $i = 1....D$. Then:

$$W = \prod_{i=1}^{D} Var\{SPM_i'\}^{-1/(2D)} = \prod_{i=1}^{D} [-\rho_i''(0)]^{-1/(2D)} \qquad 6$$

where ' denotes the derivative (see Friston et al 1991 and Worsley et al 1992). If the point spread function (convolved with any pre-processing filters) has a Gaussian shape with full width at half maximum $FWHM_1$ .... $FWHM_D$ in each of the *D* coordinate directions then it can be shown (Friston et al 1991 and Worsley et al 1992) that $W = FWHM/\sqrt{(4\log_e 2)}$ where:

$$FWHM = \prod_{i=1}^{D} FWHM_i^{1/D} \qquad 7$$

In practice *W* can be determined directly from the effective FWHM (if it is known) or estimated *post hoc* using the measured variance of the SPM's first partial derivatives according to eqn(6) (see Friston et al 1991 and Worsley et al 1992 for discussion and validation of this characterization of smoothness).

Approximate expressions for P(m = x) and P(n = x)

P(*m* = *x*) has been shown, in the limit of high thresholds, to have a Poisson distribution (Adler 1981 Theorem 6.9.3 page 161). A Poisson distribution is intuitively sensible, in the sense one can

regard the centers or maxima of activated regions as multidimensional point processes with 'no memory'. In other words when passing through the SPM maxima are encountered in much the same way as radioactive decay events occur in time:

$$P(m = x) \approx 1/x! \; E\{m\}^x \; e^{-E\{m\}} \qquad 8$$

Asymptotic results for the distribution and expectation of $n_{2/D}$ have been given by Nosko (1969a; 1969b; 1970) and are reported by Adler (1981), page 158. They show that for large thresholds $n_{2/D}$ has an exponential distribution with expectation:

$$E\{n_{2/D}\} \approx \frac{2\pi W^2}{u^2\Gamma(D/2+1)^{2/D}} \qquad 9$$

or equivalently:

$$E\{n\} \approx \frac{(2\pi)^{D/2} W^D}{u^D} \qquad 10$$

A simple way to derive these equations is provided at the end of this section, for the interested reader. Unfortunately eqn(10) is not robust at lower thresholds and substantially over-estimates E{n} when compared to estimates based on eqn(5), or indeed empirical simulations (see below). Figure 1 illustrates this point by plotting the percentage over-estimation against threshold (u). This is important because we want to apply the approximations at relatively low thresholds. To obtain a better approximation for P($n = x$) we use the following device: Assume a form for P($n = x$) which is asymptotically correct and determine the parameters of the distribution by reference to its known moments. The form of P($n = x$) we assume is:

$$P(n = x) \approx \frac{2\beta}{D} x^{2/D-1} e^{-\beta x^{2/D}}$$

or: $P(n \geq x) \approx e^{-\beta x^{2/D}}$ 11

By a simple change of variables it is easy to show that $n^{2/D}$ is exponentially distributed with an expectation $E\{n^{2/D}\} = 1/\beta$. Similarly the expectation of n, $E\{n\} = \Gamma(D/2 + 1).\beta^{-D/2}$. β is determined according to the expression for $E\{n\}$ in eqn(5):

$$\beta = [\Gamma(D/2 + 1).E\{m\}/E\{N\}]^{2/D} \quad 12$$

In summary we have approximate expressions for both $P(m = x)$ and $P(n = x)$. These approximations allow us to estimate $P(n_{max} \geq k)$.

In fact the difference between Nosko's result and the conjecture adopted here is relatively simple (but not necessarily small - see Figure 1), and is resolved by using a standard approximation for the error function $\Phi(.)$ that works for (*and only for*) high thresholds:

$$\Phi(-u) \approx \frac{e^{-u^2/2}}{u\sqrt{2\pi}} \quad 13$$

Substituting this into eqn(3), the expression for $E\{N\}$ and expanding eqn(5) $E\{n\} = E\{N\}/E\{m\}$ gives the limiting value of $E\{n\}$ which is given by eqn(10). Recalling that $E\{n^{2/D}\} = 1/\beta$ [eqn(11)], the same substitution can be used to verify eqn(9) using the expression for β eqn(12).

Estimating $P(n_{max} \geq k)$

To calculate $P(n_{max} \geq k)$ we simply compute one minus the probability that all m regions have less than k voxels, times the probability of getting m regions. These probabilities are summed over all possible values of m:

$$P(n_{max} \geq k) = \sum_{i=1}^{\infty} P(m = i).[1 - P(n < k)^i] = 1 - e^{-E\{m\}.P(n \geq k)}$$
$$= 1 - \exp(-E\{m\}.e^{-\beta k^{2/D}}) \qquad 14$$

In the limiting case of small $E\{m\}$, or high threshold, $P(n_{max} \geq 1) \approx E\{m\}$ [note $P(n \geq 1) = 1$]. This special case is precisely the one adopted in current approaches (Friston et al 1991 and Worsley et al 1992).

Eqn(14) gives an estimate of the probability of finding at least one region with k or more voxels in an SPM. For any fixed value of $P(n_{max} \geq k)$ there are a whole family of activation foci with an equally improbable chance of occurrence. The interaction between threshold (u) and volume (k) is demonstrated in Figure 2 which plots the iso-probability contours of $P(n_{max} \geq k)$ for a 2-dimensional SPM with 128^2 pixels (2 dimensional voxels) and FWHM = 9.42. It can be seen that a region of one or more pixels at a threshold of ~3.9 is as "improbable" (p = 0.05) as a focus with 60 or more pixels at a threshold of ~3. This is very important from the perspective of detecting regions of cerebral activation: Although a spatially extensive region of activation may not necessarily reach a high threshold (e.g. 3.9), it may be significant (at p = 0.05) if assessed at a lower threshold (e.g. 3). This is the application of the above results we pursue in this article.

Choice of threshold - a power analysis

The question we address here is how to choose the threshold (u). The optimum threshold should maximize sensitivity. In this section we derive an approximate expression for the sensitivity to a 'random' signal. We then demonstrate how the sensitivity, or power, depends on an interplay between the shape of the signal and the threshold used. Roughly speaking our results will show

that broader signals are best detected by low thresholds and that sharp focal signals are best detected by high thresholds.

We now give details. Suppose the 'signal' resembles Gaussian kernels or foci, of random height, distributed continuously throughout the SPM. The shape of the signal is characterized by the width (f) of these foci expressed in units of W. This signal can be modeled by a continuous ensemble of kernels with randomly distributed heights, or equivalently by convolving an uncorrelated random process with a 'kernel' of the same height. The resulting signal will be a Gaussian process of smoothness $f.W$. Although this form of signal was chosen for theoretical convenience, it is not an unreasonable model of distributed physiological signals in the brain. Following convolution with the point spread function (and any pre-processing filters) the resolution of the signal will be $W.\sqrt{(1 + f^2)}$. Let the convolved signal have a standard deviation σ, where σ corresponds to the amplitude of the measured signal. Note that σ characterizes the amount of *measured* signal not the underlying physiological signal (which would be subject to partial volume effects if f was small).

Power is the probability of getting at least one true positive, while using criteria that protect against false positives. The power at a particular threshold can be estimated with the probability of detecting at least one 'signal' using a critical region size (k_α) such that $P(n_{max} \geq k) = \alpha$, where α is suitable small, say 0.05. k_α is found by inverting eqn(14):

$$k_\alpha \approx [\, \log(-E\{m\}/\log(1 - \alpha))/\beta \,]^{D/2} \qquad 15$$

A small table of k_α is provided as a reference for anyone wishing to duplicate our calculations (Table 1). This size criterion k_α, at threshold u, is now applied to the process representing noise plus signal. This process will have zero mean and variance $1 + \sigma^2$, which means the original threshold u is effectively:

$$u^* = u/\sqrt{1+\sigma^2} \qquad 16$$

The process with signal will itself be a smooth Gaussian field with an effective resolution (W^*),

$$W^* = W \cdot \sqrt{\frac{1+\sigma^2}{1+\sigma^2/(1+f^2)}} \qquad 17$$

The validity of eqn(17) is easily established by deriving the auto-correlation function of the signal plus noise, in terms of W, and using eqn(6). The expectations of N and n for signal plus noise ($E\{N\}^*$ and $E\{m\}^*$) are given by substituting W^* and u^* in eqn(3) and eqn(4). β^* is similarly derived according to eqn(12). The probability of getting at least one region bigger than k_α in the new process is given by eqn(14). Because this region is unlikely ($p \leq \alpha$) to be due to the noise component, one can interpret the result as the probability of getting at least one true positive. Therefore, on substituting the above:

$$\text{power} = 1 - \exp[-E\{m\}^* . \exp(-\beta^* k_\alpha^{2/D})] \qquad 18$$

The results of this sort of analysis are presented graphically in Figure 3 which shows how power depends on threshold and region size for a three dimensional search volume of 65536 voxels with FWHM = 6.12, $\sigma = 0.7$ and a false positive rate $\alpha = 0.05$. It can be seen that for $f <$ 0.7 (the underlying signals are narrower than 70% the resolution of the SPM) then *power increases with threshold*, so that the most powerful test is just that based on the maximum value at a voxel, as advocated by Friston et al (1991) and Worsley et al (1992). If $f > 0.7$ (the activated region is broader than 70% of the resolution) then power increases as the *threshold is lowered.* It should be born in mind that the results presented here are only good approximations for large thresholds, so

it seems prudent to keep the thresholds high enough for the p values to remain valid. Our simulations (see below) suggest that the approximations hold reasonably well for thresholds as low as 2.4.

Empirical verification of the pdfs

This section presents an empirical validation of the above results. This is necessary because the expressions introduced depend on a number of approximations and should be regarded as (mathematically) rough estimates, especially at lower thresholds.

We present 1-, 2- and 3-dimensional simulations to assess the adequacy of eqn(8) and eqn(11) to describe $P(m = x)$ and $P(n = x)$ and to validate eqn(14), the expression for $P(n_{max} \geq k)$. The 1-dimensional simulations used processes with 4096 voxels ($S = 4096$), a smoothness (FWHM) of 9.4 and a threshold (u) of 2.6. The 2-dimensional simulations used processes with 256 x 256 pixels, FWHM = 9.2 and $u = 2.58$. The parameters for the 32 x 32 x 64 voxel 3-dimensional simulated processes were FWHM = 5.7 and $u = 2.8$. For each of the three simulations, 10^4 realizations were created using uncorrelated Gaussian processes, which were convolved with (1, 2 or 3 dimensional) Gaussian kernels. The FWHM given above are based on *post hoc* empirical estimates of W using eqn(6). For each realization, the number of supra-threshold regions (m), the numbers of voxels in each region (n) and the size of the largest region (n_{max}) were recorded. Using this data we compared the empirical and theoretical estimates of $P(m = x)$, $P(n = x)$ and $P(n_{max} \geq k)$.

The agreement between the empirical (dotted lines) and theoretical (solid lines) distributions is evident (Figure 4 - 1 dimensional, Figure 5 - 2 dimensional and Figure 6 - 3 dimensional). In all simulations the agreement is particularly good for large clusters of voxels. Note that the values of $P(n_{max} \geq k)$ are too conservative in the one dimensional case. However at higher dimensions they cannot be distinguished from the empirically determined values.

The small discrepancies between the empirical and theoretical distributions may derive from a number of sources: (i) Many of our approximations are only asymptotically true, in the limit of high thresholds. (ii) The effect of using discretized processes (voxels). (iii) The fact our simulated process were not 'very large' and had edges.

Simulated Phantom experiments

To provide a concrete and clear illustration of how these equations can be applied to detecting activation foci we performed a simulated phantom activation study. We compared two images from a phantom, 108 pixels in diameter, with 6 circular wells. In the *baseline* condition the wells contained the same activity as in the main body of the phantom. This background activity was 100 counts per pixel. The counts per pixel in a typical human study are about 6 - 12, therefore the simulated images can be thought of as coming from about 10 subjects. In the *activation* phantom the wells contained activity which was 10% higher than the background. The wells increased in size with diameters of $W.(2 + j)/\sqrt{2}$, where j ran from 1 to 6. The smallest well was therefore about the size of the FWHM and subject to partial volume effects.

Matrix manipulations and computations were performed in Matlab (MathWorks Inc. Sherborn MA, USA). The 128x128 voxel baseline and activated images were constructed assuming Poisson counting statistics (variance equal to mean activity) and a Gaussian point spread function with σ = 3. The simulated resolution corresponded to a FWHM of 7.06 pixels. Figure 7 (top) shows the two images. The baseline and activated images were subtracted and scaled to unit variance. Figure 7 (bottom) shows the subtraction image and the underlying signal (signal from the wells following convolution). The significance of the 6 activation foci were assessed using two methods:

(i) A threshold was applied to the normalized difference image according to current approaches, which rendered $P(m \geq 1) = P(n_{max} \geq 1) \approx 0.05$. In this instance the threshold was 4.108. This analysis only detected the three largest activation foci. The results of this conventional and high thresholding are shown in Figure 8 (left).

(ii) The normalized difference image was thresholded at a much lower and arbitrary level (2.8). The significance of each focus was assessed with the $P(n_{max} \geq k)$ calculated using eqn(14). The

results of this analysis are shown in Figure 8 (right). Although detected, the smallest focus is not significant [$P(n_{max} \geq k) = 0.75$]. Indeed something this size or bigger would occur at least once on about 75% of occasions by chance. The second smallest focus had a $P(n_{max} \geq k)$ of 0.17. The remaining 4 foci were found to be highly significant.

These results demonstrate a substantial increase in sensitivity to actual change when the activation topography is taken into account. This approach frees the analysis of significant focal change from the arbitrary nature of high thresholds.

The configuration of activation foci we chose to report here is rather arbitrary. We obtained equivalent results with simulated activation foci of constant shape but varying intensity and using combined differences in size and intensity.

Discussion

We have used approximate expressions for the probability density functions of (i) the number of regions above threshold and (ii) the number of voxels in each region. This characterization of Gaussian processes is rough but fairly complete and allows a significant advance in the detection of activation foci in SPMs. Previous approaches have taken an activation focus, at some threshold, and estimated the probability that one or more such foci, of any size, could have occurred by chance. The threshold is usually set such that this probability is 0.05. We are now in a position to estimate the significance of an activation focus in terms of the probability that one or more foci, of the *same or greater size*, could have occurred by chance [by using eqn(14)]. This is important because information about the spatial extent or volume of the activation is explicitly included and the analysis of significant focal change is freed from the arbitrary nature of high thresholds. The improved sensitivity which results allows the threshold to be lowered to much more realistic levels with no increase in the experiment-wise probability of a false positive.

Thresholds and signal width

A power analysis based on the results presented here suggests that narrow focal activations are most powerfully detected by high thresholds (as currently implemented) whereas broader more diffuse activations are best detected by low thresholds. Alternatively for a fixed threshold *power increases with resolution*. The fundamental importance of this for functional MRI studies is obvious, however it also raises the issues about the optimum resolution for PET data, where current practice is to use low resolutions.

The proposal to use lower thresholds is vindicated by the power analysis, but is predicated on the assumption that real signals are broader than the resolution. There is empirical evidence to suggest that this is the case: The autocorrelation of physiological changes measured with PET is substantially larger (~12mm) than the autocorrelation of noise (~8mm) (Friston et al 1992).

Future work by Siegmund and Worsley will attempt to resolve this 'dependency on assumptions', by searching over tuning parameters (like smoothing) as well as voxels.

It should be noted that the apparent volume of activation in an SPM is not the real volume of activated cerebral tissue. Furthermore one must also be aware that small but intense physiological activations will be subject to partial volume effects.

We re-iterate that the topography of activation foci has been analyzed only with respect to size. There are other features (shape, symmetrical distribution etc.) which may render a particular focus, or set of foci, sufficiently improbable that they can be accepted as real. The scope of analyses which could be brought to bear on statistical parametric maps is clearly extensive. The results presented have, of course, many other potential applications, which will be pursued in subsequent publications. The expressions presented are approximations (usually exact in the limit of high thresholds). It is possible that more exact expressions may become available in the next year or two, or at least other approximate answers should appear to substantiate or supersede the present ones. In the interim the results presented here appear to perform well enough to justify practical application.

We have deliberately not limited the analysis or simulations to three dimensional processes because the results are applicable in any dimension. Although we envisage such an approach being applied to three dimensional SPMs there are other important applications. For example in the assessment of peri-stimulus-joint-histograms, obtained from multi-unit electrode recording data, the problem is to assess the significance of conjoint unit activity at some temporal distance from stimulus onset. This problem can be formulated in terms of a smooth 1-dimensional Gaussian process and can be addressed using the equations presented above. At the other extreme, functional or dynamic magnetic resonance imaging (MRI) studies of the hemodynamic response to

a repeating stimulus can be assessed using SPMs which are also a function of time from stimulus onset (e.g. an SPM of the cross-correlation function between the functional MRI signal and some time-dependent sensorimotor or cognitive parameter). In this instance an SPM is computed at a series of temporal offsets from the start of stimulation. Because the repeat time of scans is typically less than the hemodynamic time constants there is an effective *temporal point spread function* which introduces smoothness in the time domain. A time series of SPMs constitutes a 4-dimensional SPM which could be modeled as a 4-dimensional Gaussian process according to the expressions presented here.

Acknowledgments

Part of this work was undertaken while KJF was funded by the Wellcome Trust.

References

Adler RJ & Hasofer AM (1981) The geometry of random fields. Wiley, New York.

Belyaev Yu K (1972) Point processes and first passage problems. Proceedings of the sixth Berkeley Symposium on Mathematical Statistics and Probability 2:1 - 17 University of California at Berkeley

Friston KJ Frith CD Liddle PF Dolan RJ Lammertsma AA & Frackowiak RSJ (1990) The relationship between global and local changes in PET scans J. Cereb. Blood Flow Metab. 10:458-466

Friston, K.J Frith, C.D Liddle, P.F and Frackowiak, R.S.J. (1991) Comparing functional (PET) images: The assessment of significant change. J. Cereb. Blood Flow Metab. 11:690-699.

Friston KJ Jezzard P Frackowiak RSJ & Turner R (1993) Characterizing focal and distributed physiological changes with MRI and PET. In Functional MRI of the Brain, Society of Magnetic Resonance in Medicine, Berkely CA p207-216

Friston KJ, Frith CD, Passingham RE, Dolan RJ, Liddle PF & Frackowiak RSJ (1992) Entropy and cortical activity: Information theory and PET findings. Cerebral Cortex 2:259-267

Hasofer AM (1978) Upcrossings of random fields. Suppl. Adv Appl. Prob. 10:14-21

Nosko VP (1969a) Local structure of Gaussian random fields in the vicinity of high level shines. Soviet Mathematics Doklady, 10:1481-1484

Nosko VP (1969b) The characteristics of excursions of Gaussian homogeneous random fields above a high level. Proceedings of the USSR-Japan Symposium on Probability (Harbarovsk, 1969) Novosibirak, 216-222

Nosko VP (1970) On shines of Gaussian random fields (in Russian) Vestnik Moscov. Univ. Ser. I Mat. Meh., 1970:18-22

Poline JB & Mazoyer BM (1993) Analysis of individual positron emission tomography activation maps by detection of high signal-to-noise-ratio pixel clusters. J. Cereb. Blood. Flow Metab. 13:425-437

Roland PE, Levin B, Kawashima R and Ackerman S (1993) Three dimensional analysis of clustered voxels in ^{15}O-Butanol brain activation images. Human Brain Mapping 1:3-19

Worsley KJ Evans AC Marrett S & Neelin P (1992) A three-dimensional statistical analysis for rCBF activation studies in human brain. J. Cereb. Blood. Flow Metab. 12:900-918

Worsley KJ (1994) Local Maxima and the expected Euler characteristic of excursion sets of ^{2}, F and t fields. Advances Appl. Prob. in press

Table 1

Title: A Table of critical sizes k

A small table of approximate critical values k_α of k, the number of voxels in a suprathreshold cluster, chosen such that $P(n_{max} \geq k) \approx \alpha$. where n_{max} is the number of voxels in the largest cluster. These values are specific to the tabulated threshold (u), resolution (FWHM), search volume (S) and dimensionality (D). They are provided as a reference for those who wish to reproduce our calculations.

Legends for figures

Figure 1

Graphical illustration of how the Nosko result for the $E\{n\}$ - the expectation of n (spatial extent of a region in voxels) fails at low thresholds. The over-estimation is relative to $E\{n\}$ as defined in eqn(5). The relationship presented here does not depend on dimensionality, search volume or smoothness.

Figure 2

Iso-probability contours calculated according to eqn(14) which give the relationship between the threshold (u) and the number of voxels (x) an activation focus should contain to maintain a certain level of "probability". More strictly $P(n_{max} \geq k)$ (the chance probability of obtaining one or more regions of at least k voxels) has been calculated over a range of voxels and thresholds and contoured at 4 levels (0.1, 0.05, 0.01, 0.001). In this example S (volume) = 128^2, FWHM (smoothness) = 9.42, D (dimensionality) = 2.

Figure 3

2 dimensional plot of the power as a function of threshold (u) and the width of a signal (f) expressed in units of the SPM's smoothness (W). The key think to note here is the marked dissociation between the optimum thresholds (thresholds which maximize power) at high and low signal widths (f).

Figure 4

Empirically (dotted lines) and theoretically (solid lines) determined probability density functions for (i) the number of regions (m) per realization - top left, (ii) the number of voxels per region (n) - top right, (iii) and $P(n_{max} \geq k)$ the probability that the largest region, per realization, has k or more voxels - bottom left.

In this 1-dimensional simulation FWHM (smoothness) = 9.4, u (threshold) = 2.8 and S (volume) = 4096. The empirical estimates are from 10000 realizations. The most marked failure in these 1-dimensional simulations was an over-estimation of the probability of getting a large region by chance. This failure is in the conservative direction.

Figure 5

As for Figure 2 but using 2-dimensional processes. FWHM (smoothness) = 9.2, u (threshold) = 2.58 and S (volume) = 256 x 256. The empirical estimates are from 10000 realizations. The agreement between the empirical and theoretical values for $P(n_{max} \geq k)$ is remarkably good, particularly at large regions sizes (empirical - dotted lines, theoretical - solid lines).

Figure 6

As for Figure 2 but using 3-dimensional processes. FWHM (smoothness) = 5.7, u (threshold) = 2.8 and S (volume) = 32 x 32 x 64. The empirical estimates are from 10000 realizations. As for the 2-dimensional case the agreement is remarkable (empirical - dotted lines, theoretical - solid lines).

Figure 7

Simulated 128^2 phantom data. Top left - baseline phantom image 108 pixels in diameter with uniform activity of about 100 counts per pixel and a FWHM of 7.06 pixels. Top right - activated image where 6 wells, of increasing size, have been filled with activity 10% higher than in the main body of the phantom. Bottom left - difference image obtained by subtracting the activated image from the baseline image. Bottom right - the underlying signal following convolution with the point spread function.

Figure 8

Results of thresholding to detect significant activations. Left - conventional threshold [setting $P(n_{max} \geq 1)$ to 0.05] of 4.108. Right - a lower and arbitrary threshold (2.8) has been applied to the data and the significance of each activation focus has been assessed using $P(n_{max} \geq k)$. This new approach to detecting significant activations is more sensitive because it is not threshold dependent and takes account of the size of each activated region.