# Bayesian method of testing for differences in the responses to two stimuli

**Hasson R, Braeutigam S, and Swithenby SJ**
*The Open University, Milton Keynes, United Kingdom*

## Introduction

Bayesian approaches to the biomagnetic inverse problem have been suggested explicitly by several authors [1,2,3] and are implicit in other work [e.g. 4,5]. We have reported previously a Bayesian approach to the comparison of experimental data with source models [6]. Briefly, this method provides a map of the statistically significant differences between measured data and an *a-priori* source model. The first use was to test the hypothesis that focal sources underlie both simulated and real experimental data.

In this paper, the basic result is stated and then used in a novel way to test whether the source configurations associated with the responses to two stimuli are significantly different. The data used is from an experiment in which the subject is presented with a random sequence of images of human faces, motorbikes, animals and random dot patterns [7]. Analysis of data in the signal domain identified significant localised differences between the *signals* from human faces and the rest. In our Bayesian analysis, the model source pattern is determined empirically from the experimental data from one stimulus group (e.g. human faces) and the null hypothesis is that this model underlies the data from a second group (e.g. motorbikes). What emerges are the temporal and spatial distributions of sources where this hypothesis is invalid, i.e. the significant differences in source space. This approach is robust. It is of widespread applicability and provides an unbiased means of classifying responses.

## Methods

The following is a short precis of the Bayesian method proposed in [6]. The detector outputs at a given time are collected into a data vector $\tilde{m} \in \mathbb{R}^N$. The method makes no use of time continuity and treats each such snapshot independently. If $\vec{j}(\vec{r}) \in J$ is a current density in a suitable current space then the measurement process can be represented by a functional $z : J \to \mathbb{R}^N$. A subscript notation will be used to identify the sensor, i.e. $z_i$ is the ideal reading from the $i$th sensor. So the basic equation is:

$$\tilde{m}_i = z_i(\vec{j}) + e_i$$

where the $e_i$ are the measurement errors. These are assumed to be normally distributed with zero mean and a covariance matrix $D$.

The measurement geometry enters the algorithm via the Green's functions $\vec{L}_i$ which are defined by

$$z_i(\vec{j}) = \int_Q \vec{L}_i(\vec{r}) \cdot \vec{j}(\vec{r})\, \mathrm{d}\vec{r}$$

The approach is based on Bayes's theorem:

$$\mathcal{P}(\vec{j} \in A | m \in B) = \frac{\mathcal{P}(m \in B | \vec{j} \in A)\mathcal{P}(\vec{j} \in A)}{\mathcal{P}(m \in B)}$$

where $A$ is a set of currents and $B$ is a set of measurements. This equation reads, the *a posteriori* probability of a current set $A$ after the measurement $B$ is proportional to the probability of producing the measurement $B$ given that the current is in the set $A$ times the *a priori* probability of the current set $A$. ($\mathcal{P}(m \in B)$ is a constant for any measurement set $B$).

In this paper the probability distributions (both the prior probability and the errors) are assumed to be gaussian and so it is permissible to work with probability density functions. A further simplification is achieved by shrinking the measurement set $B$ to a single point $\{\tilde{m}\}$. This ignores the finiteness of the precision of the measurements. Equation 1 then becomes:

$$\rho_{\tilde{m}}(\vec{j}) \propto \rho(\vec{j}) \ \times \ \epsilon(\tilde{m} - z(\vec{j}))$$

where $\rho$ is the *a priori* distribution, $\rho_{\tilde{m}}$ is the *a posteriori* distribution and $\epsilon$ is the error distribution. It should be noted that, throughout the paper, probability density functions will only be determined up to a constant.

The constant of proportionality is found by requiring that the probability is normalised to one. The error distribution, $\epsilon$, is assumed to be Gaussian, i.e.

$$\epsilon(e) \propto \exp\left\{-\frac{1}{2}e^{\mathrm{T}}D^{-1}e\right\}$$

How is the prior distribution $\rho(\ )$ defined? Firstly, an inner product on $J$ is defined by:

$$\left\langle \vec{j}_1\ ,\ \vec{j}_2 \right\rangle = \int_Q \frac{\vec{j}_1(\vec{r})\cdot\vec{j}_2(\vec{r})}{\omega(\vec{r})}\ \mathrm{d}\vec{r}$$

where $\omega(\vec{r})$ is a weighting distribution defined on the source space $Q$. This provides a method of inputing prior information of the location of sources (e.g. gained from MRI images) into the algorithm.

Secondly, an arbitrary prior current $\vec{j}^p$ is introduced as a parameter of the method. The *a priori* probability distribution on $J$ is defined using $\vec{j}^p$ and the inner product:

$$\rho(\vec{j}) \propto \exp\left\{-\frac{1}{2\beta^2}\left\langle \vec{j}-\vec{j}^p\ ,\ \vec{j}-\vec{j}^p \right\rangle\right\}$$

where $\beta$ is the assumed *a priori* standard deviation and is a parameter of the method.

In order to produce images of the *a posteriori* current density, it is necessary to find the distribution of a single statistic $\lambda$ that can be computed. The statistic is defined through a 'test current' $\vec{t} = \chi_{V_k}(\vec{r})\hat{e}_\alpha$ where $\chi_{V_k}(\vec{r})$ is the characteristic function of a voxel in the brain and $\hat{e}_\alpha$ is a unit vector. An image is then produced by scanning $\vec{t}$ over the source space voxels.

With these definitions and some algebraic manipulation [6] gives the following formula for the *a posteriori* distribution of $\lambda$.

$$\text{Change in expectation of } \lambda = u^{\mathrm{T}}(\zeta D + P)^{-1}\left(\tilde{m} - z(\vec{j}^p)\right)$$

$$\text{variance of } \lambda = u^{\mathrm{T}}\left(PD^{-1}P + \zeta P\right)^{-1}u + \frac{1}{\zeta}\left(\langle\vec{t}\ ,\ \vec{t}\rangle - u^{\mathrm{T}}P^{-1}u\right)$$

where $u$ is a vector with components: $u_i = \left\langle \vec{t}\ ,\ \omega(\vec{r})\vec{L}_i(\vec{r})\right\rangle$ and $\zeta = 1/\beta^2$ plays the role of a regularization parameter.

From these equations it is possible to calculate a spatio-temporal probability distribution for differences between the *a-priori* and *a-posteriori* currents. There is a simple method of achieving this. The data set $\tilde{m}(1)$ associated with one stimulus class (e.g. faces) may be used as an implicit prior source current distribution by replacing $z(\vec{j}^p)$ with $\tilde{m}(1)$ in the above equations and then computing the change in expectation associated with the data $\tilde{m}(2)$ from a second stimulus class (e.g. motor bikes). The symmetry of the comparison is preserved in this formulation. The noise in the first measurement, which is neglected by using data as an implicit prior source current, may be allowed for by scaling the noise used in the subsequent calculation of the probability distribution of the null hypothesis.

## Results

The data we use to illustrate the method comes from a face recognition experiment [7]. MEG measurements were made using a whole head magnetometer (Neuromag-122$^{TM}$ [8]) covering all cortical regions (Fig. 1). The subjects viewed briefly presented and standardised grayscale images of faces, motor-bikes, dot patterns and coffee mugs while fixating at a point on the centre of the screen. After artefact rejection, the responses were digitised using a sampling rate of 397 Hz, high and low-pass filtered at 0.03 Hz and 130 Hz respectively and averaged separately for each image class.

In the initial analysis [7], the strength of the evoked activity within a specific region and latency span was parameterised by calculating the signal power, integrated over a group of channels and a range of latencies, and normalised to the variance of the relevant channel group. Such calculations indicated that brain activity in the right occipito-temporal region following face presentation is significantly different ($p = 0.05$) from activity following presentation of each of the non-face images during the latency range between 110 and 170 ms. No consistent statistically significant differences between face and non-face responses were seen at other latencies or in other regions although it was difficult to carry out an exhaustive search because of the large number of candidate possibilities.

For simplicity in the Bayesian analysis, we have used a homogeneous sphere conductor model and a source space that consists of a part spherical shell (Fig. 2) of radius 8 cm. This shell covers the majority of the posterior regions of the neocortex. The significance of differences between the face and motor bike responses were computed as a function of latency and position using the method outlined above. The regularization parameter $\zeta$ was determined by using the L-curve method [9] and an optimal value was determined to be 1.4 times the trace of the Gram-Schmidt matrix (which is always a useful rough guide). The standard deviation of the change in $\lambda$ was evaluated from the pre-stimulus interval.

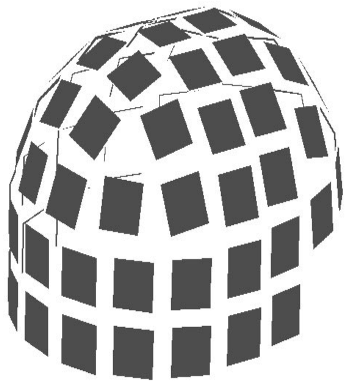

Fig. 1 The arrangement of the 122 transverse gradiometer detectors on 61 sites.

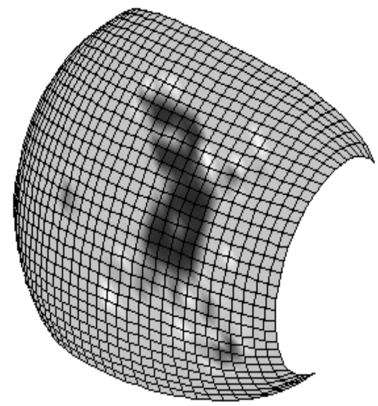

Fig. 2 The source space is a part spherical shell (radius 8 cm, solid angle $2\,\mathrm{rad} \times 2\,\mathrm{rad}$) covering posterior regions of the brain and including candidate visual processing areas. Superimposed on the space is the result for latency 138 ms where there is the most significant difference p=0.997.

Fig. 3 shows in one figure the results of the analysis of the motor bike/face comparison for all latencies across the entire source space. The null hypothesis is that the same generators underly both sets of data. The square regions show the probability that the null hypothesis is false mapped out along the surface of the source space. Time progresses from left to right and from top to bottom and is measured with respect to the stimulus presentation at 0 ms.

## Discussion

In Fig. 3, the results of the Bayesian analysis shown are consistent with the existence of a face specific system in right occipito temporal cortex with maximum specificity at 138 ms, confirming the earlier analysis. The new approach has some clear advantages. It provides an automatic means of scanning all the source space (rather than the signal space) and all latencies and it provides quantitative measures of differences in the activity. In this illustration it suggests that the specificity is predominantly in the right hemisphere but extends to the midline and involves a considerable latency span. Similar results have been obtained for comparisons of faces and other stimulus classes, and for other subjects.

The method has general applicability to any investigation involving stimulus or task dependence of responses. It is robust to model and noise assumptions.

## References

[1] Phillips JW, Leahy RM and Mosher JC: MEG-based imaging of focal neuronal current sources IEEE Trans. Med. Imaging, 16(3): 338-348 1997

[2] Baillet S and Garnero L: A Bayesian approach to introducing anatomo- functional priors in the EEG/MEG inverse problem. IEEE Trans. Biomed. Eng. 44(5): 374-385 1997

[3] Schmidt DM, George JS and Wood CC: Bayesian inference applied to the electromagnetic inverse problem. Technical Report LA-UR-97-4813, Los Alamos National Laboratory. 1997

[4] Khosla D and Singh M: A maximum-entropy method for MEG source imaging. IEEE Trans. Nucl. Sci. 44(3) Part2: 1368-1374 1997

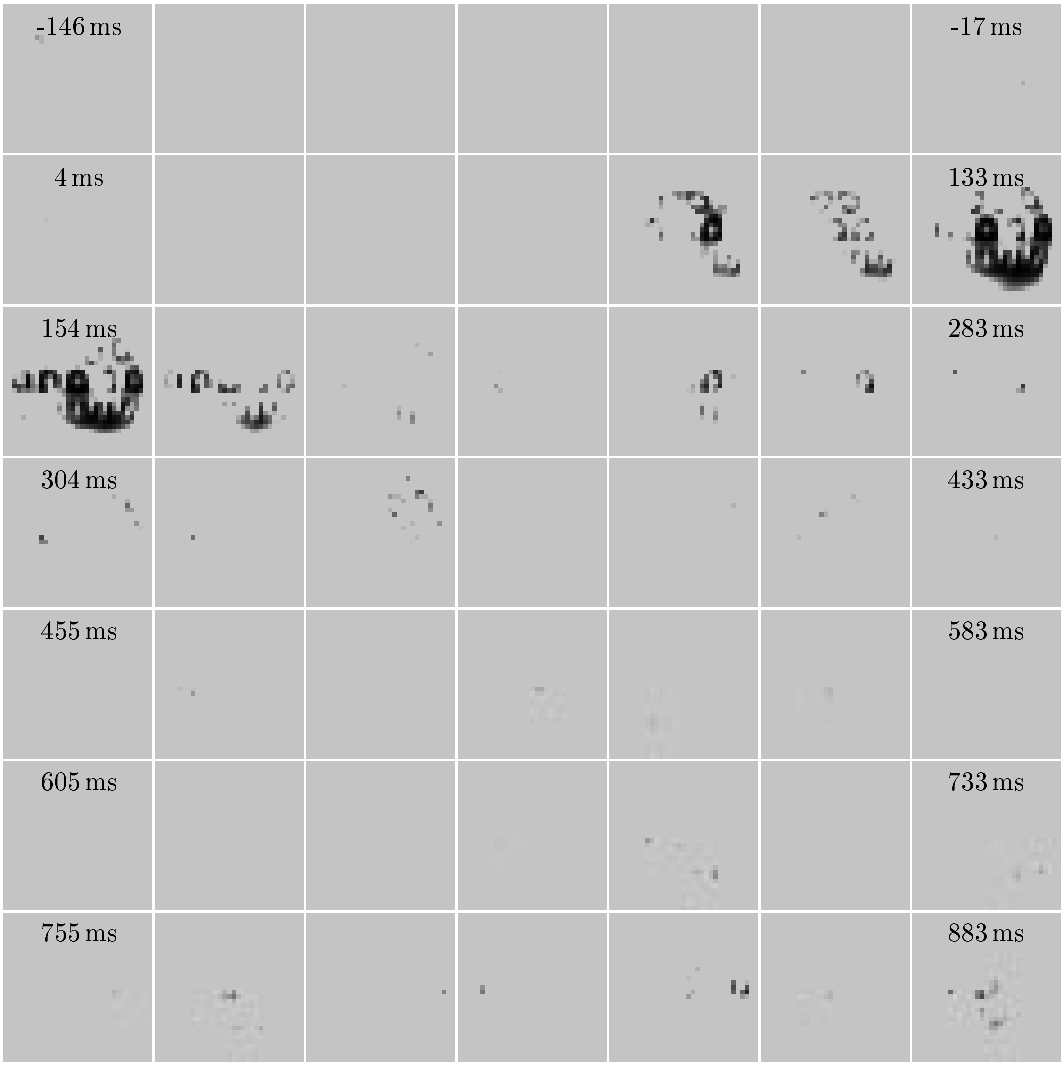


Fig. 3 An illustration for one subject. The probability that the null hypothesis is incorrect (face versus motor bike) for the source currents underlying the responses to face and motor bike stimuli as a function of latency (shown every 21.4 ms). Each map is a projection of the 2 rad×2 rad spherical shell source space. In order to highlight areas of robust difference, probabilities below 0.9 are mapped to mid grey the more significant probabilities are shown darker with p=1.0 shown as black.

[5] Clarke CJS and Janday BS: The solution of the biomagnetic inverse problem by maximum statistical entropy. Inverse Problems 5(4): 483-500 1989

[6] Hasson R, Swithenby SJ and Braeutigam S: A Bayesian test for the acceptability of a given source current distribution. Conference Digest for ISBET'97, 8th World Conf. of the Int. Soc. for Brain Electromag. Tomography: Abstract P2.19 1997

[7] Swithenby SJ, Bailey AJ, Braeutigam S, Josephs OE, Jousmaki V and Tesche CD: Neural processing of human faces: a magnetoencephalographic MEG study. Exp Brain Res. 118: 501-510 1998

[8] Ahonen AI, Hämäläinen MS, Kajola MJ, et al: 122-channel SQUID instrumentation for investigating the magnetic signals from the human brain. Physica Scripta 49: 198-205 1993

[9] Hansen PC, Regularization tools. Numerical Algorithms 6: 1–35 1994