

## INFORMATION CRITERIA DETERMINE THE NUMBER OF ACTIVE SOURCES

T.R. Knösche, E.M. Berends, H.R.A. Jagers, and M.J. Peters

*Low Temperature Division, University of Twente, Postbus 217, 7500 AE Enschede, The Netherlands*  
*email: t.r.knosche@tn.utwente.nl*

### Abstract

With the neuroelectromagnetic inverse problem, the optimal choice of the number of sources is a difficult problem, especially in the presence of correlated noise. In this paper we present a number of information criteria that help to solve this problem. They are based on the probability density function of the measurements or their eigenvalues, make use of the Akaike or MDL (minimum description length) correction term and all employ some sort of noise information. By extensive simulations we investigated the conditions under which these criteria yield reliable estimations. We were able to quantify two major factors of influence: (1) the precision of the noise information and (2) the signal-to-noise ratio (SNR), here defined as the ratio of the smallest signal eigenvalues and the average of the noise eigenvalues. Furthermore, we found that the Akaike correction term tends to overestimate, due to its greater sensibility to the precision of the noise information.

### Introduction

The reconstruction of intracranial activity from EEG and MEG measurements, also referred to as the inverse problem, is generally non-unique. In order to obtain a solution anyway, it is necessary to make a model of the sources. Obviously, the number of free parameters of this model must not exceed the number of independent measurements. In order to achieve this, certain assumptions about the expected activity have to be made. Many inverse algorithms are based on *a priori* knowledge on the number of independent sources. For the well-known spatiotemporal dipole fit method [1] this means it must be known, how many current dipoles are simultaneously active.

If there was not any noise, the number of non-zero eigenvalues of the spatial covariance matrix of the measured data would equal the number of active sources (see example in figure 1.a). If we added Gaussian uncorrelated noise and were able to obtain the exact data covariance matrix, the spectrum would look like the one in figure 1.b. In this case, we are still able to determine the number of source components that stick out of the uniform noise floor. If, however, we have to be content with the sample covariance matrix and the noise is more realistic, meaning e.g. spatially

and temporally correlated, a picture like in figure 1.c occurs. If the noise is small, the largest eigenvalues that differ significantly from the rest, still represent the sources. However, in many cases the eigenvalues decrease smoothly, making the choice of the number of components very subjective.

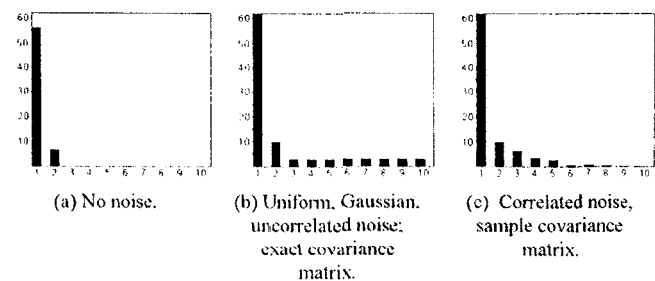


Figure 1: Example of eigenvalue spectra of noise measurement covariance matrices. First 10 eigenvalues shown, units arbitrary. There are 2 signal components.

It is obvious that a source model consisting of a larger number of components can explain more of the measurements. Therefore, if the residual variance was the only criterion for the choice of the number of sources in the presence of noise, it would always equal the rank of the data matrix. Many of the sources would then merely explain noise.

In this paper, we present information criteria that may help to make an more objective choice. Some of them have can be found in literature [2-4], others were and are being published elsewhere [5]. These criteria have been explored by means of simulations with respect to their reliability.

### Methods

A set of measurements may be described as superposition of a linear combination of the source signals and additive noise.

$$\underline{X} = A\underline{S} + \underline{N} \quad (1)$$

$\underline{X}$  represents the data matrix with one row for each of the  $m$  channels and one column for each of the  $n$  time samples.  $\underline{S}$  stands for the source strengths and has  $k$  rows and  $t$  columns.  $\underline{N}$  contains additive noise (Gaussian, uncorrelated in time) and  $A$  symbolises the transfer matrix between source

strengths and measured values, taking into account the source and sensor positions, the volume conduction of the head, etc. No assumptions are made about the spatial shape, position, and temporal behaviour of the sources.

The criteria explored in this paper generally consist of two terms: (1) a log likelihood function that reflects how well the observations  $\underline{x}$  can be explained by a certain statistical model, and (2) a correction term that somehow represents the complexity of the model.

$$Crit(\kappa) = -2 \log f_{\Theta_\kappa}(\underline{x}) + C(v_\kappa, n) \quad (2)$$

The first term is monotone decreasing with the number of sources  $\kappa$ . The maximum likelihood function  $\Theta_\kappa$  is the parameters vector of the statistical model that best explains the data with  $\kappa$  sources. The correction term depends on the number of parameters for the particular statistical model  $v_\kappa$  and the number of observations  $n$ . It is an increasing function. The minimum of the sum of both should be at the correct number of sources (see figure 2).

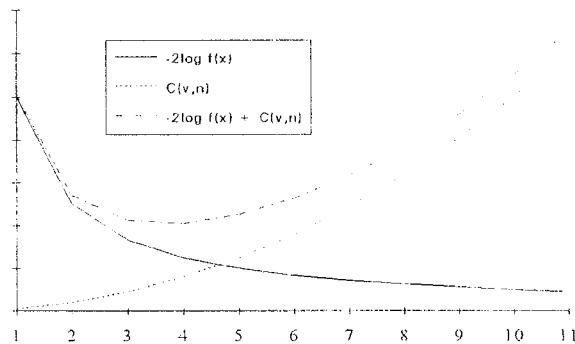


Figure 2: The combination of log likelihood function and correction term yields a criterion with a minimum at the correct number of sources (here 3).

The criteria differ in (1) the stochastic variables the probability density function is based upon, (2) the statistical model of the noise, and (3) the type of correction term.

Criteria based on the probability density function of	Spatial noise model		
	white noise		correlated noise
	noise level unknown	noise level known	
measurements	C1	C2	C3
all eigenvalues	C4a	C5a	C6b
noise eigenvalues	C7a	C8a	C9a/C9b
average noise eigenvalue	C10	C11	

Table 1: Overview of the information criteria. They can be used with Akaike (CxA) or MDL (CXR) correction term. Table 1 gives an overview. Each criterion can be used with the Akaike correction term [6], denoted CxA. Alternatively,

the minimum description length (MDL) [7] correction term can be used, denoted CXR.

$$C_{Akaike} = 2 v_\kappa \quad C_{MDL} = v_\kappa \log n \quad (3)$$

The lowercase letters  $a$  and  $b$  in the codes for the criteria 4, 5, and 6 stand for statistical assumptions on the sources ( $a$  - stochastic,  $b$  - deterministic). For the criteria 7, 8, and 9 the letter  $a$  indicates that the equality of the eigenvalues is tested against the alternative that they are not all equal to the noise level. The letter  $b$  means that the equality of the eigenvalues is tested against the alternative that any deviation from the noise level is positive. Although some of the criteria assume that the noise is spatially uncorrelated, they can also be used for correlated noise if the data covariance matrix is corrected in advance using the known noise covariance matrix [5].

## Results

The criteria have been tested on simulations of electrical brain activity. Current dipoles with fixed positions and time-varying moments have been placed in a volume conductor model consisting of three concentric spheres. The electric potential was computed at electrode positions distributed over the upper hemisphere and Gaussian noise was added.

In order to test the reliability of the criteria we varied noise level, accuracy of the noise information, number of observations ( $n$ ), number of channels ( $m$ ), noise type (correlated and uncorrelated), and dipole configuration. The signal-to-noise ratio (SNR) was defined as ratio of smallest signal eigenvalue to average noise eigenvalue. This value depends not only on the noise level but also on dipole configuration,  $m$ , and  $n$ . It turned out that the performance of the criteria was independently limited by two factors: the SNR and the accuracy of the noise information. For both, we found maximum values that must not be exceeded. Correlated and uncorrelated noise yielded the same results, indicating the correction procedure worked effectively. Figure 3 shows the minimum SNR for which the criteria estimated the correct value in 90 % of all cases (25 trials).

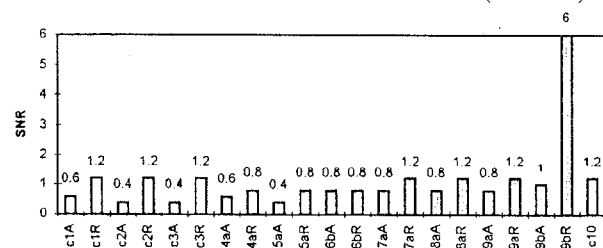


Figure 3: Minimum SNR for at least 90 % success (2 dipoles with 4 components, 27 electrodes, 550 time steps).

It can be seen that criteria using the MDL correction generally need a higher SNR to respond correctly.

In order to test the influence of  $n$  and  $m$  separately, we did not look at the SNR (since it depends on  $n$  and  $m$ ), but at the absolute variance of the noise. It was found that the maximum variance for which a criterion still works reliably, increases with both  $n$  and  $m$ . In figure 4 we plotted the maximum noise level that could be tolerated by the Akaike criteria against the number of time samples.

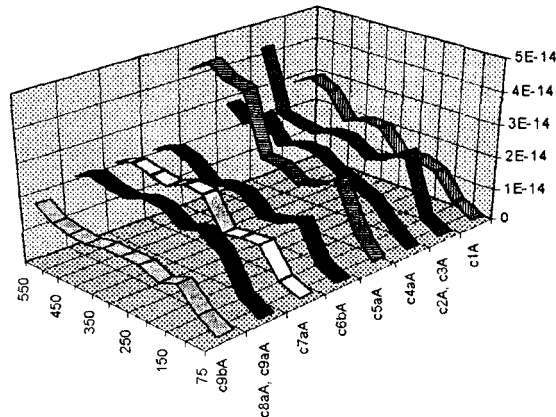


Figure 4: Maximum noise variance ( in  $V^2$ ) for which the Akaike criteria responded correctly (90 % of cases) against the number of time samples (4 components, 27 electrodes)

The other crucial factor for the performance of the criteria is the accuracy of the provided noise information. Since the noise covariances must be estimated from a limited number of samples, certain errors are inevitable. If the noise is Gaussian and the noise covariances is computed from certain signal-free parts of each trial (pre-stimulus intervals), then the relative error of the noise variance is defined as follows.

$$\varepsilon_r = \sigma_{\text{var}} / \sigma^2 \cdot 100\% = \frac{2}{\sqrt{nq-1}} \cdot 100\% \quad (4)$$

with  $n$  being the number of time samples in the pre-stimulus interval,  $q$  the number of trials,  $\sigma_{\text{var}}$  the standard deviation of the noise variance, and  $\sigma^2$  the noise variance itself.

We found that the maximum error that is tolerated (for 90 % of cases) by the criteria 1, 4, 7, and 10 (those criteria using noise information only for *a priori* correction, see table 1) is independent of both SNR and noise variance. The values are presented in table 2. For the other criteria, no general rule could be established so far.

criterion	c1A	c1R	c4aA	c4aR	c7aA	c7aR	c10
$\varepsilon_{\text{max}} / \%$	1	6	1	2	1	6	6

Table 2: Maximum error of the noise variance for which the criteria 1, 4, 7, and 10 estimated the right number of sources.

From table 2 we can see that the MDL criteria are more stable against inaccurate noise information.

## Discussion and Conclusions

We investigated the reliability of 12 different information criteria with respect to the numbers of time samples ( $n$ ) and channels ( $m$ ) in the data, the SNR (defined as ratio of the smallest signal and the average noise eigenvalue), and the accuracy of the provided noise information. We may summarise the following findings:

1. The maximum noise level that could be tolerated increased with growing  $m$  and  $n$ .
2. MDL based criteria generally need a higher SNR than Akaike based ones.
3. For all criteria that use noise information only for *a priori* correction of the data (1, 4, 7, 10) the maximum tolerable error in the noise variance was independent of the noise level and the SNR.
4. MDL criteria are more stable against inaccurate noise information.

In practice, the stability against inaccurate noise information seems the most crucial factor. From equation 4 and table 2 one can conclude that e.g. criterion c1R needs  $nq = 1112$  for a reliable estimate (e.g. 100 trials and 12 samples), while a typical Akaike criterion such as c1A needs  $nq = 40001$  (e.g. 500 trials and 80 samples). The latter case may prove problematic, especially if only a short pre-stimulus interval is available or if the noise information must be estimated from the average (just 1 trial then).

We also performed analyses of real measurements. The results, although not presented here due to limited space, suggest that the MDL versions of the criteria 2, 3, 6b, and 9b produce the most consistent and convincing results. These are criteria that use noise information explicitly, suggesting that the *a priori* compensation works less well here (possibly due to non-Gaussian distribution of the noise).

## References

- [1] M. Scherg and P. Berg, *Use of prior knowledge in brain electromagnetic source analysis*, Brain Topography, **4**(1991)
- [2] M. Wax and T. Kailath, *Detection of signals by information criteria*, IEEE Trans. Acoust. Speech Signal Processing, **33**(1985)
- [3] G.J.H. Uijen and A. van Oosterom, *The performance of information-theoretic criteria in detecting the number of signals in multilead ECG's*, Meth. Inform. Med., **31**(1992)
- [4] K.M. Wong, Q.T. Zhang, J.P. Reilly, and P.C. Yip, *On information theoretic criteria for determining the number of signals in high resolution array processing*, IEEE Trans. Acoust. Speech Signal Processing, **38**(1990)
- [5] H.R.A. Jagers, T. Knösche, M.J. Peters, *Estimation of number of independent sources: Information criteria can help to solve inverse problem*, submitted to IEEE Trans. Signal Processing, 1995
- [6] H. Akaike, *Extension of the maximum likelihood principle*, in B.N. Petrov and F. Csáki (eds.) *2nd International Symposium on Information Theory*, Akadémiai Kiadó, Budapest, 1973
- [7] J. Rissanen, *Modelling by shortest data principle*, Automatica, **14**(1978)