Supplementary Data

Criteria for Evaluating the Optimal Number of Microstate Maps

Eleven criteria have been used to evaluate independently the quality of each clustering (see Supplementary Table S1 for an overview of the criteria). They were then merged together to make a single synthetic criterion, which has been used here. The idea was to increase the confidence in the right number of clusters, improving the previous way, which was to pick either the Cross-Validation criterion (Pascual-Marqui et al., 1995) or the Krzanowski-Lai Index (Krzanowski and Lai, 1988).

The chosen 11 criteria were [from review articles (Charrad et al., 2014; Milligan and Cooper, 1985; Murray et al., 2008; Pascual-Marqui et al., 1995)]:

- Cross-Validation (its second derivative),
- Cubic Clustering Criterion (its first derivative),
- Davies and Bouldin,
- Dunn,
- Frey and Van Groenewoud,
- Hartigan (its first derivative),
- Krzanowski-Lai Index,
- Marriott,
- Point-Biserial,
- Tau,
- Trace (W) (its second derivative).

Meta-Criterion

These 11 criteria were then merged into a single meta-criterion with the following steps:

- Each criterion is first singlehandedly ranked, by using the relative positions of the values sorted from the lowest to the highest.
- The ensuing merging formula aims at maximizing the average ranking, as well as favoring unanimity across all criteria, which is seen as important as a high average response. Maximum unanimity is defined here as the highest signal to noise ratio of all criteria. We therefore end up with the proposed merging formula (Eq. 5), when using robust estimators:

$$MetaCriterion = IQM \times SNR \qquad \text{(Eq. 1)}$$

with

$$IQM = Interquartile\ Mean\ (Criteria) \qquad \text{(Eq. 2)}$$

$$IQR = Interquartile\ Range\ (Criteria) \qquad \text{(Eq. 3)}$$

$$SNR = \frac{IQM}{IQR} \qquad \text{(Eq. 4)}$$

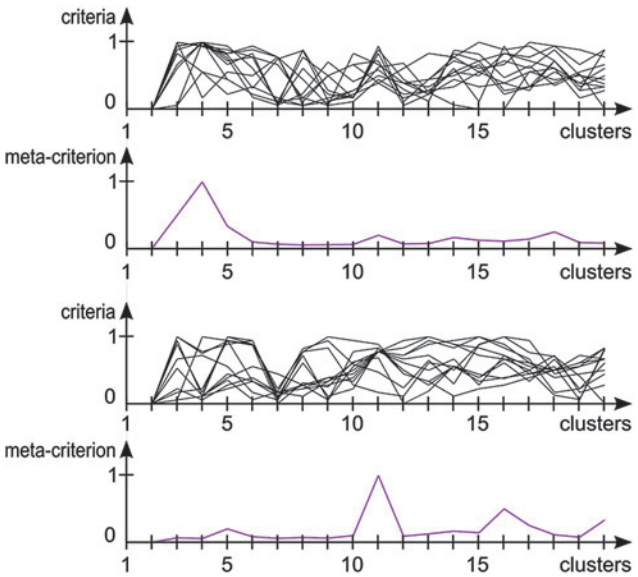
which finally gives:

$$MetaCriterion = \frac{IQM^2}{IQR} \qquad \text{(Eq. 5)}$$

The best number of clusters n according to the Meta-criterion is then:

$$\arg_n \max_{n \geq 4} [MetaCriterion(n)] \qquad \text{(Eq. 6)}$$

In Supplementary Figure S1 we show two examples of the meta-criterion applied on different data sets:



SUPPLEMENTARY FIG. S1. Two examples of the meta-criterion (magenta tracks) applied to the 11 criteria (black, superimposed tracks). The optimal number of clusters is at the peak of the meta-criterion, where the average of all criteria is high and the dispersion is low. The x axis is the number of clusters, and the y axis is the normalized criterion values, ranging from 0 to 1.

Criteria's Cost Functions

SUPPLEMENTARY TABLE S1. OPTIMAL CRITERIA OVERVIEW

Criterion	*Meaning*
Cross-Validation	A modified version of the predictive residual variance.
Cubic Clustering Criterion	A function assuming that clusters from a uniform distribution on a hyperbox are hypercubes.
Davies and Bouldin	A function of the sum ratio of within-cluster to between-cluster separation.
Dunn	An evaluation of how all clusters are well separated.
Frey and Van Groenewoud	A ratio of the mean of between-cluster distances to the mean of within-cluster distances.
Hartigan	A ratio of the Trace of within-cluster distances.
Krzanowski-Lai Index	A ratio of the relative difference of the within-cluster dispersion.
Marriott	A function proportional to the determinant of within-cluster distances.
Point-Biserial	A Point-Biserial correlation calculated between the distance matrix and a binary cluster index.
Tau	The difference between pairs of between- and within-cluster distances.
Trace (W)	The Trace of within-cluster distances.

Supplementary References

Charrad M, Ghazzali N, Boiteau V, Niknafs A. 2014. NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Softw 61:1–36.

Krzanowski WJ, Lai YT. 1988. A criterion for determining the number of groups in a data set using sum-of-squares clustering. Biometrics 44:23–34.

Milligan GW, Cooper MC. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–179.

Murray MM, Brunet D, Michel CM. 2008. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr 20:249–264.

Pascual-Marqui RD, Michel CM, Lehmann D. 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans Biomed Eng 42:658–665.