Assessing the effect of threshold on graph theory metrics

Joanne R. Hale[1], Stephen D. Mayhew[1], Izabela Przezdzik[1], Theodorous N. Arvanitis[2], Andrew P. Bagshaw[1]

[1]School of Psychology, University of Birmingham, Birmingham, United Kingdom
[2]School of Electronic, Electrical and Computer Engineering, University of Birmingham, Birmingham, United Kingdom

www.birmingham.ac.uk/mig

E-mail: j.hale.1@bham.ac.uk

Introduction

The application of graph theoretical methods in neuroscience [1] is increasingly popular, as it offers the potential to characterise and summarise the brain's complex network. However, derived graph theory metrics are sensitive to several underlying assumptions.

To date there has been little focus on the definition of the association matrix when examining functional networks [2]. In particular, matrix connections are generally constrained to those whose connectivity exceeds an arbitrary threshold.

Here we use a functional parcellation scheme to examine the effect of altering the correlation threshold (ct) (i.e. the connection density/ sparsity) on common graph theory metrics. We also consider the stability of graph metrics across different time-series lengths.

Methods

Graph Analysis

Data acquisition
Eight healthy volunteers (5 male, 32±6yrs) underwent a 15 minute resting-state fMRI scan (3x3x4mm voxels, TR=2s, 450 volumes, FA = 80°)

Pre-processing
Resting state fMRI data were pre-processed according to standard methodology [3]

Association matrix formation
Pearson correlation coefficient was computed between BOLD timecourses extracted from all pairwise combinations of network ROIs and averaged across subjects, forming the association matrix

Effect of time-series length
Association matrices were constructed using (1) whole 15 minute time-series; (2) time-series averaged over 2 minute epochs; (3) time-series averaged over 30 second epochs

Thresholding
The association matrices were each binarised by thresholding at a range of values from ct= 0.01-0.7

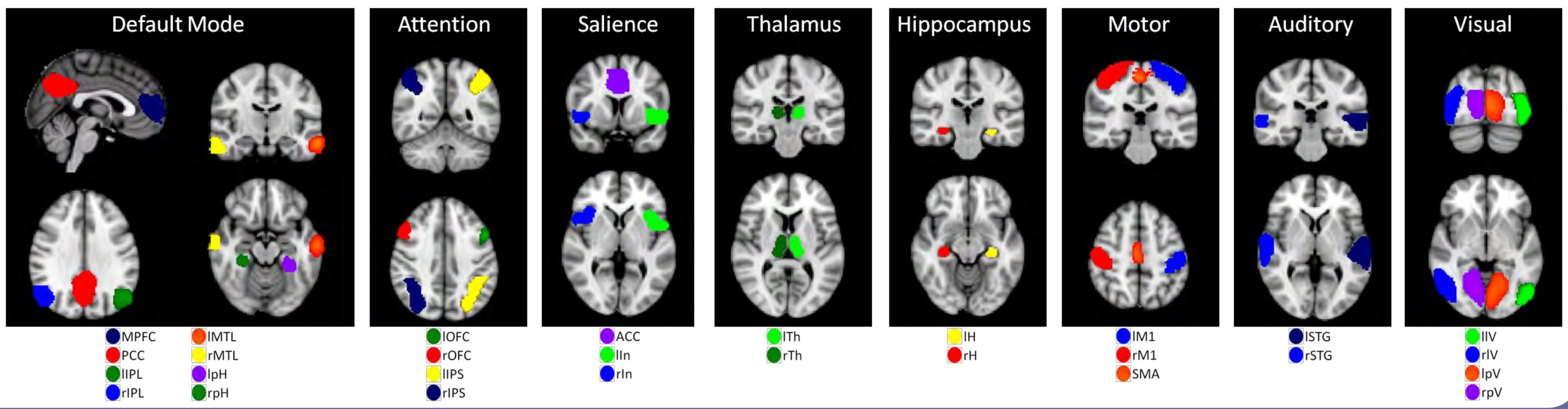
Graph metrics
The Brain Connectivity Toolbox [1] was used to compute graph theoretical metrics as a function of ct for each of the three time-series lengths

ROI extraction

RSN identification
Separate cohort of 55 healthy subjects (28 male, 25±4yrs), data from 6 minute resting fMRI scans (3x3x4mm voxels, TR=2s) were used to identify resting-state networks (RSNs) using MELODIC [4]

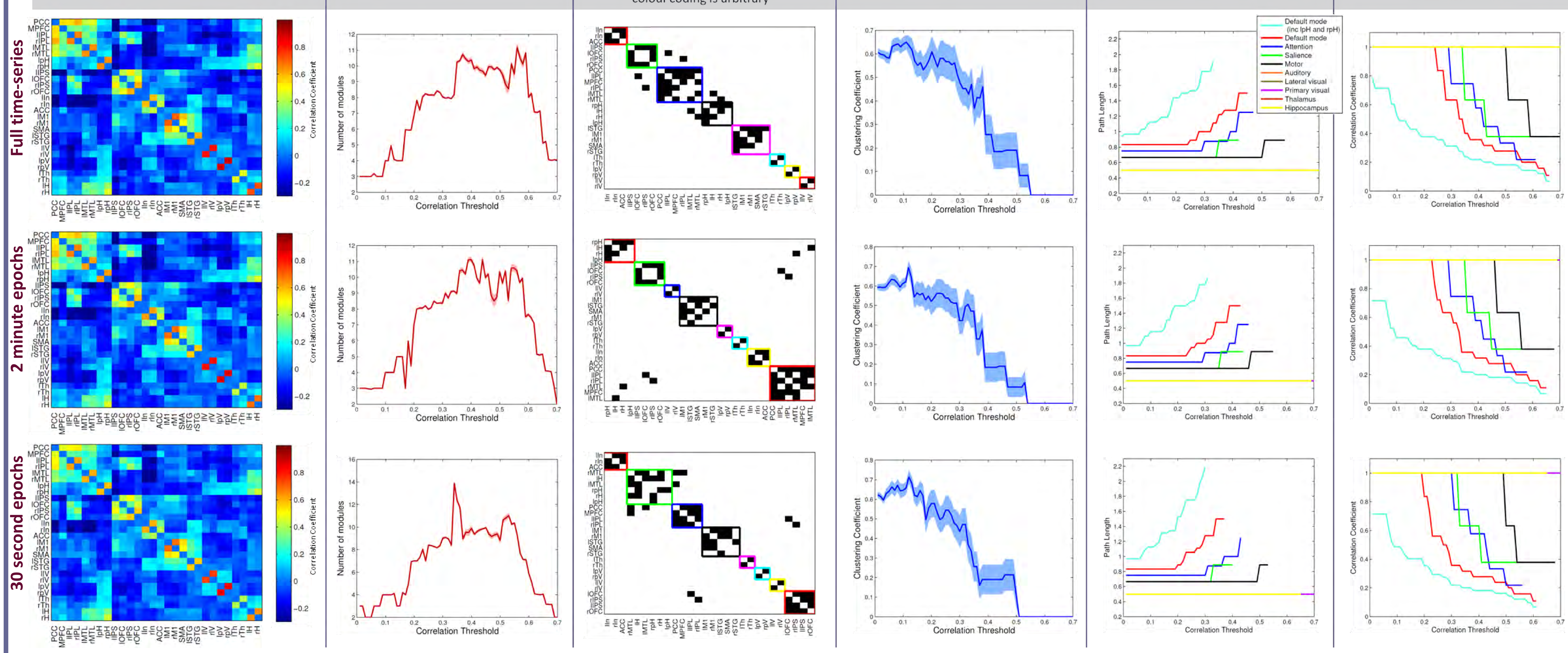
ROI definition
For each RSN, individual ROIs were defined from 3x3x3 voxel cubes centred on the maximum z-statistic voxel in each major node

Major resting-state network nodes identified with ICA and used as ROIs



Results

Correlation coefficient matrices	Variation in the number of modules with threshold. Shaded region represents standard error across initialisations of modularity code	Example of modular organisation Matrices are thresholded to correspond to a connection density of 0.1. Nodes are ordered by degree within modules and colour coding is arbitrary	Average clustering coefficient as a function of threshold Shaded region represents the standard error across nodes	Characteristic path length for each resting state network as a function of threshold	Pearson correlation coefficient between ideal network pattern (i.e. all nodes of each network connected) with the real networks from the thresholded matrix for each network



- Patterns of correlation coefficients are similar across the varying time-series lengths
- With increased threshold the number of modules increases, reaching the number of modules expected from the functional parcellation with ct≈0.2-0.3. As ct>0.3 the number of modules is more variable as matrix sparsity is increased.
- Modular organisation reveals 8 modules each representing individual RSNs (except motor and auditory which are combined in a single module, (full time-series and 2 minute epochs)).

- Average clustering coefficient and characteristic path length metrics show a fairly consistent value until ct= 0.3. Between ct=0.3 -0.4 clustering coefficient displays a sharp drop-off, whilst path length simultaneously increases to infinity (for RSNs with >3 nodes).
- Correlation between the thresholded matrix and the idealised version of each RSN was greatly reduced for the default mode, attention and salience networks for ct>0.35.

Discussion

- The choice of connectivity threshold greatly influences graph theoretical metrics that are intended to summarise network properties. This will have a major impact on subsequent interpretation of brain networks. For pre-processing and parcellation scheme applied, the optimal correlation threshold was found to occur between 0.2 and 0.3.
- Graph metrics displayed stability across time-series lengths, which is particularly noteworthy when considering the application of graph theory to study sleep as data are epoched before classifying different stages.
- These results have strong implications for the selection of appropriate connectivity thresholds for graph analysis of RSNs, and highlight the importance of the details of the construction of the adjacency matrix.
- An advantage of the functional parcellation scheme employed here is that it allows some internal validation of the graph theoretical quantities, e.g. a low ct that leads to a low number of estimated modules can be identified as inappropriate. This internal validity is particularly important when considering using graph theoretical metrics to summarise changes in functional connectivity between different stages of sleep, or between patient populations and control subjects.

References

[1] Rubinov, M., Sporns, O. (2010) Neuroimage; [2] Smith, S.M. (2012) Neuroimage; [3] Fox, M.D. , Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C. (2005), PNAS USA; [4] Beckmann, C.M., Smith, S.M. (2004) IEEE Trans Med Imaging

Acknowledgments: Funded by the UK Engineering and Physical Sciences Research Council (EPSRC Grant Number EP/J002909/1)