Scaling up Directed Graph Models for Resting-State fMRI with Stepwise Regression

Ruth Harbord[1], Lilia Carolina Carneiro da Costa[2], Jim Smith[3], Janine Bijsterbosch[4], Sonia Bishop[4] and Thomas E. Nichols[3]

EPSRC Engineering and Physical Sciences Research Council

[1] MOAC Doctoral Training Centre, University of Warwick, Coventry, United Kingdom, [2] Universidade Federal da Bahia, Salvador/BA, Brazil
[3] Dept. of Statistics, University of Warwick, Coventry, United Kingdom, [4] FMRIB Centre, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom

Abstract

The **Multiregression Dynamic Model (MDM)** is a graphical model for estimating dynamic, directed connectivity. The model evidence has a closed form and factors by node, allowing a fast, parallelised model search. Using a resting-state fMRI data set, we show that applying stepwise methods to approximate the model search can dramatically reduce computation time with a negligible effect on the estimated connectivity structures. This will allow the extension of the methodology to much larger numbers of nodes.

The Multiregression Dynamic Model

The model equations are

$$\begin{aligned} \textit{Observation Equation} \quad & Y_t(r) = \mathbf{F}_t^T(r)\boldsymbol{\theta}_t(r) + v_t(r) \quad v_t(r) \sim \mathcal{N}(0, V_t(r)) \\ \textit{System Equation} \quad & \boldsymbol{\theta}_t(r) = \boldsymbol{\theta}_{t-1}(r) + \boldsymbol{w}_t(r) \quad \boldsymbol{w}_t(r) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{W}_t(r)) \\ \textit{Initial Information} \quad & (\boldsymbol{\theta}_0|y_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0) \end{aligned}$$

where $Y_t(r)$ is the time series for node r at time t. The matrix $\mathbf{F}_t(r)^T$ contains a 1 for an intercept and the values of the parent nodes at time t. The coefficients $\boldsymbol{\theta}_t(r)$ represent connectivity strength and may vary over time. This variability is controlled by a scalar *discount factor* δ where if $\delta = 0.5$, $\boldsymbol{\theta}_t(r)$ evolves as a random walk and if $\delta = 1$ the model is static (Costa et al., 2015).

Data

We applied exhaustive and stepwise MDM-DGM searches to 32 subjects with 15-minute resting-state fMRI scans (TR=1140 ms, 2 x 2 x 2 mm). ROIs were defined either functionally or based on the Harvard-Oxford atlas: ventromedial prefrontal cortex (VMPFC), orbitofrontal cortex (OFC, L,R), dorsolateral prefrontal cortex (DLPFC, L,R), Amygdala (L,R), Anterior Insula (AntIns, L,R), Posterior Insula (PostIns, L,R), anterior mid-cingulate cortex (aMCC), primary and secondary somatosensory cortex (SI and SII, mean over left and right hemispheres) and periaqueductal gray (PAG) (Bijsterbosch et al., 2015).

The MDM-DGM Model Search

In a **Directed Graph Model** search **(MDM-DGM)**, the optimal model m (the optimal set of parents $Pa(r)$) is found for each node r individually by maximising the Log Predictive Likelihood

$$\text{LPL}(m) = \sum_{r=1}^{n} \sum_{t=1}^{T} \log p(y_t(r)|\mathbf{y}^{t-1}, Pa(r))$$

As the number of nodes n increases, the total number of models (2^{n-1}) grows exponentially, making an exhaustive model search infeasible for $n > 20$.

Two models, m_1 and m_2, may be compared using the log Bayes Factor

$$\text{logBF} = LPL(m_1) - LPL(m_2)$$

where logBF $> \pm 1$ indicates there is evidence to prefer one model over another and logBF $> \pm 2$ indicates that this evidence is *strong* (Harrison and West, 1997).

Stepwise Methods

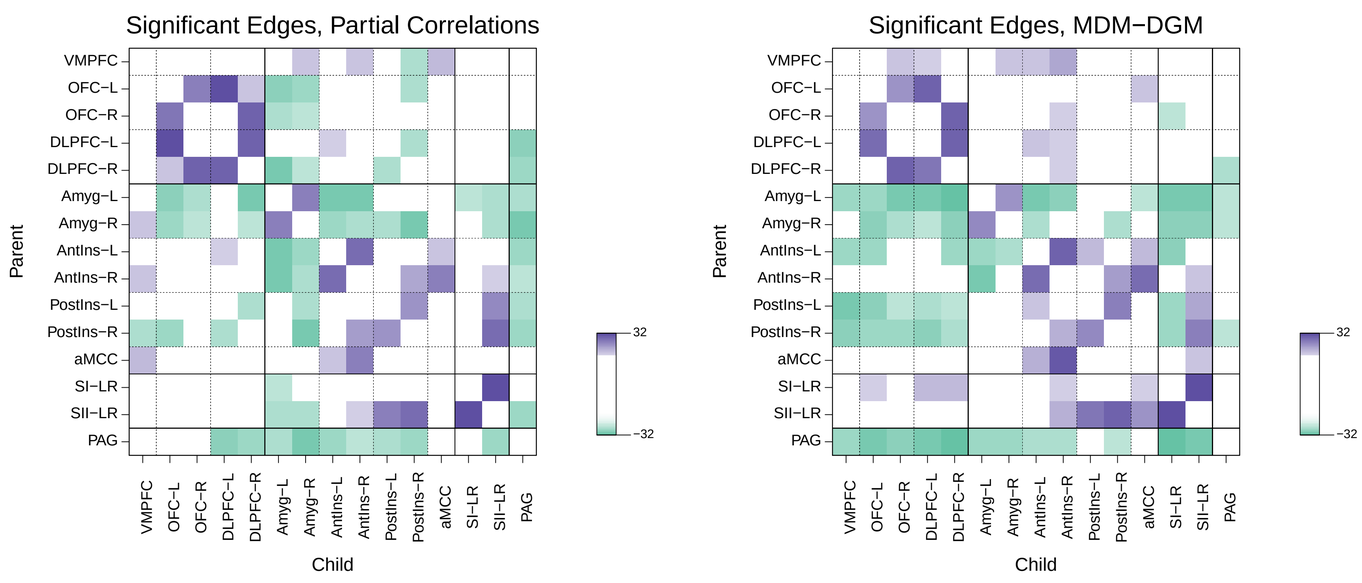
Forward selection starts from the no-parent (intercept only) model. The first step scores all the one parent models, selecting for inclusion the parent which most increases the LPL. Parents are added one at a time until there is no increase in the LPL.

Backward elimination begins with the model that includes all the possible parents, maximising the LPL as each parent is removed until this removal fails to improve the LPL.

An improved performance may be achieved by combining the results of the forward and backward approaches: when they disagree the model with the higher score is selected.

Code was run on Dell PowerEdge R410 servers with 12 2.80 GHz processors (X5660) and 64GB RAM, using R version 3.2.4 with C++ implementation.

The MDM-DGM Search Identifies Consistent, Directed Networks



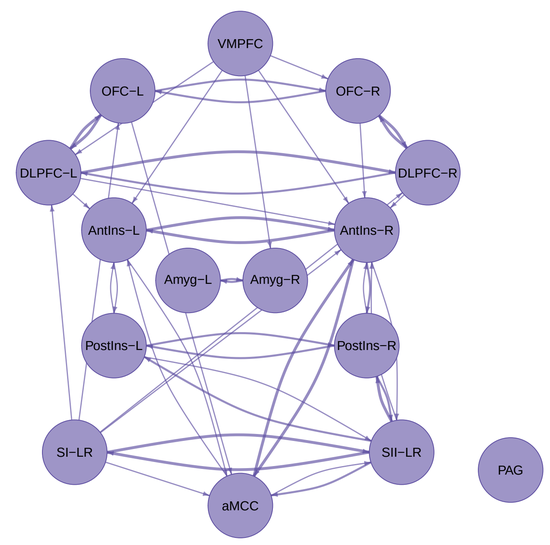
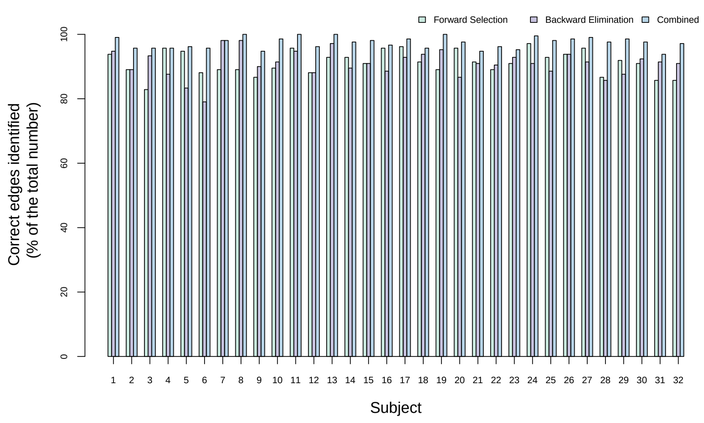


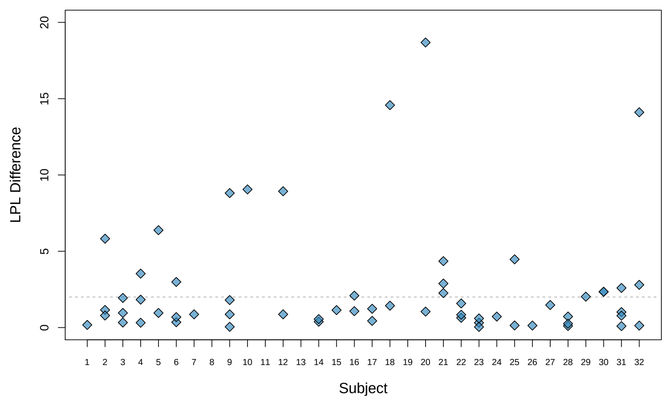
Figure 1: Estimation of individual network structure. MDM-DGM networks are consistent with networks based on partial correlations but can also estimate directionality.

(a) An MDM-DGM was fit for each subject s and the total number of edges n_s calculated. Equivalent adjacency matrices based on partial correlations were then created by selecting the n_s edges with the highest absolute partial correlation (excluding the diagonal). One-sided Binomial tests were performed to identify edges occurring more (purple) or less (green) frequently than would be expected in a homogenous network (defined as the average proportion of subjects with an edge over the 210 possible edges ($p = 0.37$ (presence), $p = 0.63$ (absence)). Edges surviving a 5 % false discovery rate threshold are shown.

(b) An alternative representation of the MDM-DGM network (of present edges) shown in **(a)**, emphasising the strongest edges, using the igraph package (Csardi and Nepusz, 2006).

Stepwise Methods Provide an Accurate Approximation to the MDM-DGM Search





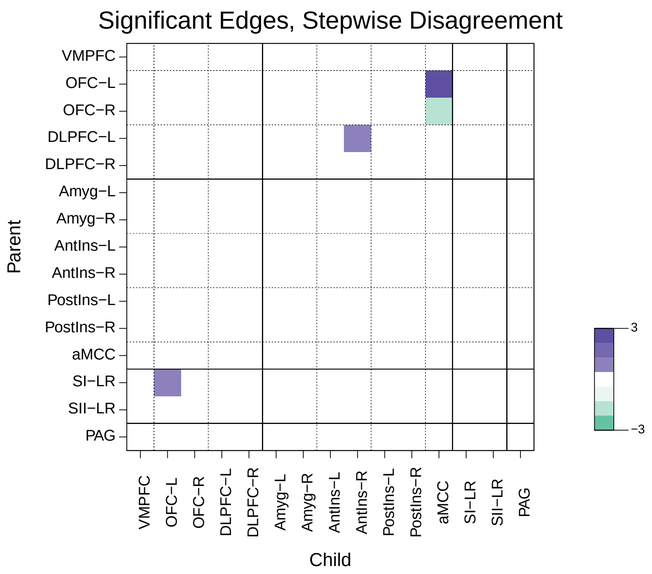


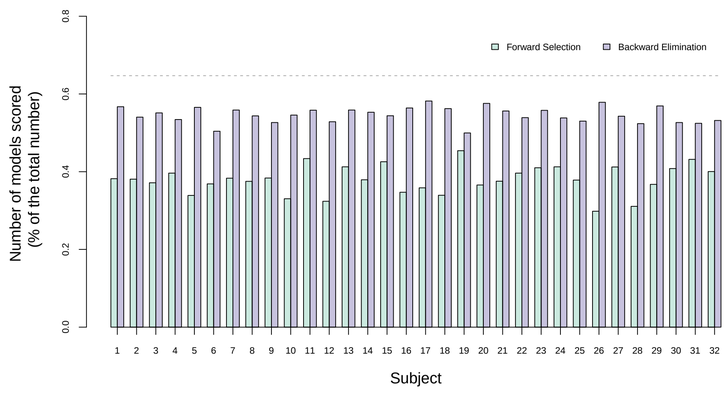
Figure 2: Comparison of an exhaustive MDM-DGM search over 2.46×10^5 candidate models to a stepwise search over a maximum of 1590 models.

(a) When used in isolation, forward selection correctly identifies an average of 91 % of edges (min 83 %, max 97 %, sd 4 %). Backward elimination performs comparably, with mean correct edge identification 91 % (min 79 %, max 98 %, sd 4 %). In combination, performance is improved to an average of 97 % (min 94 %, max 100 %, sd 2 %). For 4 subjects, the exhaustive and stepwise networks are identical.

(b) Using the log Bayes Factor to compare the Log Predictive Likelihoods when the stepwise search fails to find the true maximum. For 48 % of the models incorrectly specified by the stepwise search, logBF < 1, indicating there is insufficient evidence to distinguish between the models selected by the exhaustive and stepwise searches. For 67 % , logBF < 2, indicating there is no strong evidence for a difference.

(c) Identical analysis to that performed in **Figure 1** but using a stepwise search. Edges that were significant or non-significant in both approaches have been removed. The stepwise search failed to identify the edges SI-LR -> OFC-L, DLPFC-L -> AntIns-R (both -1 subject) and OFC-L -> aMCC (-3 subjects) as significant but identified OFC-R -> aMCC ($+2$ subjects). Note that over all edges the two networks differ by a maximum of 3 subjects.

Stepwise Methods Significantly Reduce Computation Time



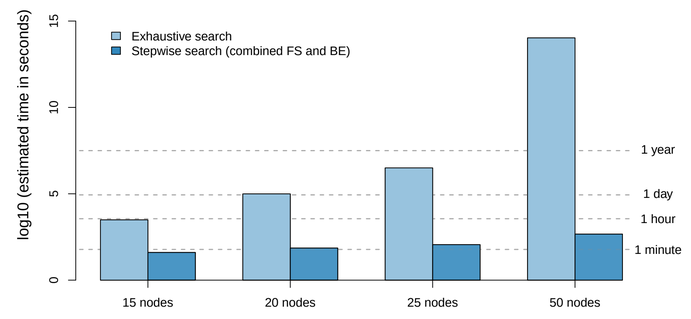


Figure 3: Stepwise methods score a fraction of the total number of candidate models, making models with large numbers of nodes computationally tractable.

(a) The number of models scored as a percentage of the total (2.46×10^5). This is on average 0.38 % (sd 0.04 %) and 0.55 % (sd 0.02 %) for forward selection and backward selection respectively. For 15 nodes, the maximum percentage of models that can be scored using these methods is 0.65 %.

(b) Estimated computation times (per subject, per node) for increasing numbers of nodes (assuming a run time of 0.2 seconds per model and a time series with 790 time points), plotted on a $\log_{10}$ scale. Stepwise estimates assume both forward selection and backward elimination are run and the maximum number of models is scored. The computation time of a 15 node network may be reduced from hours to less than a minute while a 50 node network is now feasible in approximately 8 minutes (assuming parallelisation by node).

Conclusion

The MDM-DGM search estimates directed, physiologically-interpretable networks, consistent with those obtained by analysing partial correlations. We have shown that stepwise methods can dramatically reduce computation time for a small trade-off in accuracy. Future work will assess the stability of these methods on larger (50-100) numbers of nodes.

Acknowledgements

The authors gratefully acknowledge Simon Schwab for his contribution to the MDM-DGM code.

Ruth Harbord is funded by EPSRC grant number EP/F500378/1

References

Bijsterbosch et al., (2015) Functional Connectivity under Anticipation of Shock: Correlates of Trait Anxious Effect versus Induced Anxiety *Journal of Cognitive Neuroscience* 27(9):1840-1853

Costa et al. (2015) Searching Multiregression Dynamic Models of Resting-State fMRI Networks using Integer Programming. *Bayesian Analysis* 10(2):441-478.

Csardi and Nepusz (2006) The igraph Software Package for Complex Network Research, *InterJournal, Complex Systems* 1695(5):1-9 http://igraph.org

Harrison and West (1997) Bayesian Forecasting and Dynamic Models. Springer. pp. 328

R packages for C++ implementation

Eddelbuettel and Francois (2011) Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software* 40(8):1-18

Eddelbuettel and Sanderson (2014) RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis* 71:1054-1063.