Less Than the Sum of its Parts: Complex Models of Cognition Struggle to Capture Regional Activity within Otherwise Well-Fitting Model Structures

Sönke Steffen (s.steffen@rug.nl) **Dr. Catherine Sibert (c.l.sibert@rug.nl)**
Artificial Intelligence, University of Groningen
Groningen, The Netherlands

Abstract

Dynamic Causal Modeling is a widely-used method for examining brain connectivity. Most commonly, it is applied to brain regions showing strong responses to experimental tasks, comparing different network configurations based on the temporal dynamics of the neural signals. It can further be applied to models employing a theory-driven selection of brain regions, showing a weaker experimental effect. However, it is unclear if these effects provide sufficient temporal information for Dynamic Causal Modeling to reliably identify the best-fitting model. This study investigated the regional predictive fit in a theory-driven model which has been found to consistently outperform alternatives using Dynamic Causal Modeling. Results revealed issues with the fit of some regions and subjects, raising concerns regarding the reliability of model comparisons using Dynamic Causal Modeling with regions selected based on theory instead of a strong experimental effect.

Keywords: fMRI; Dynamic Causal Modeling; Brain Connectivity; Common Model of Cognition; Model Comparisons

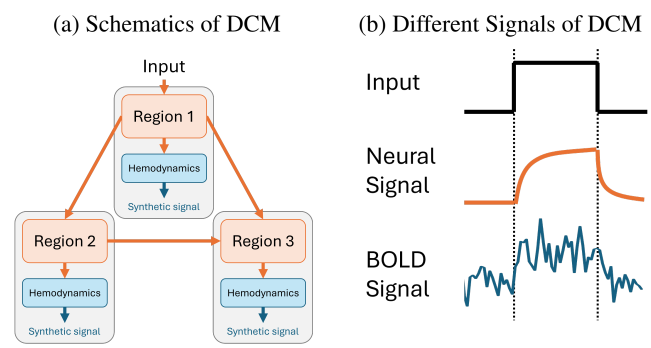
Introduction

Cognition in the human brain is dynamic, arising from the flow of information through a network of pathways between specialized regions that carry out various elements of cognitive processes, like the processing of visual stimuli in the visual cortex or the planning and execution of motor movements in the motor cortex (Pessoa, 2014; Parks & Madden, 2013; Bajaj, Butler, Drake, & Dhamala, 2015). These processes do not take place in isolation, and the outcome of any particular aspect of perception, planning, or information retrieval can be modulated by the activity in other brain regions, both those nearby and ones that are connected more distantly. While understanding the mechanisms of each individual region will be vital to understanding the brain as a whole, it is equally important to consider the connectivity between regions when forming models of cognition. Brain connectivity can help to describe how behavior results from cognition, provide insight into different stages of processing, explain how plasticity in the brain and differences between humans affect cognitive outcomes, or even categorize and help understand neurological and psychiatric disorders (de Schotten & Forkel, 2022; Kelly & Castellanos, 2014; Parks & Madden, 2013; Seghier, Zeidman, Neufeld, Leff, & Price, 2010).

Dynamic Causal Modeling

Dynamic Causal Modeling (DCM) is a method used to investigate effective connectivity in brain data, such as functional

Figure 1: Dynamic Causal Modeling (DCM)



Notes. **Left**: The neural signal traverses from the input throughout the whole network, as indicated in orange. For each region, the neural signal is further transformed into a biologically plausible synthetic BOLD-signal via the hemodynamic model. **Right**: Representation of the different signals created and used in Dynamic Causal Modeling.

magnetic resonance imaging (fMRI) (K. Friston, Harrison, & Penny, 2003; K. E. Stephan et al., 2008; K. E. Stephan, Penny, Daunizeau, Moran, & Friston, 2009). It is usually based on a conventional experimental setup where subjects are instructed to respond to different categories of stimuli while their brain activity is recorded. The General Linear Model (GLM) is used to identify brain regions where the activity differs significantly between experimental conditions, with those regions being interpreted as relevant for the cognitive processing of the task. While such a General Linear Model is useful to identify relevant regions, it does not indicate their individual roles, relationships, or what types of networks and pathways between the identified regions exist.

Dynamic Causal Modeling is based on the idea that changes in neural activity driven by experimental conditions traverse through networks within the brain over time, with individual regions representing different stages of processing. The timing of changes in the observed signal of one region in response to experimental conditions should be affected by changes in other regions of the network.

Traditionally, multiple plausible network configurations based on the regions identified in the GLM are created and compared (K. Stephan et al., 2010). For each, a biophysically grounded artificial network is generated, and a signal is traversed through it over time according to the experimental conditions as inputs representing the different neural states (schematically depicted in Figure 1). The weights associated with each connection within the network are adjusted in an effort to generate a predicted time-series of neural activity for each region, which should optimally match the observed time-series. The different network configurations lead to slight differences in timing of the activity relating to the conditions, reflecting the timing dynamics of the network. The generated model should therefore capture the timing of activity in different regions better when the network configuration resembles the actual underlying network best. Factoring in complexity, in the model comparison, the model configuration that best fits the observed data is considered to be the winner, and supports that model's structure as representative of actual structure in the brain.

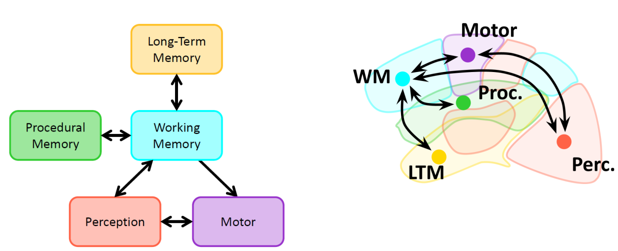
The aforementioned approach assumes that the selection of regions included in the model is based on a strong experimental effect. While this is a common and appropriate method, alternate approaches to modelling cognition take a top-down approach, and to be implemented in DCM, require the inclusion of regions which are driven by theoretical considerations. In that instance, a strong experimental effect in the data which DCM can fit different models to might not always be present for all regions.

The Common Model of Cognition

One example of a theory-driven model that has been investigated using Dynamic Causal Modeling is the Common Model of Cognition (CMC). The Common Model of Cognition (CMC) was proposed as a sparse, high-level, consensus model that aims to capture the common elements across nearly five decades of cognitive model research (Laird, Lebiere, & Rosenbloom, 2017). The model identifies five functional modules (Perceptual, Motor, Working Memory, Procedural Memory, and Long Term Memory), localizes each in a particular brain region, as shown in Figure 2 and posits a network of connections that represent the flow of information between them. As such, the CMC is intended to serve as a common-ground framework for cognition across a wide variety of modalities and domains.

Figure 2: The Common Model of Cognition (CMC)

(a) The CMC Schematically (b) The CMCs Mapping



Previous research has validated the CMC on various tasks using Dynamic Causal Modeling (Stocco et al., 2021; Steine-Hanson, Koh, & Stocco, 2018). The pathways posited by the CMC and its resulting network were compared to models using plausible alternate networks between the same regions. While the selection of the five regions is theory-driven, and does not always equate to the regions showing the strongest peaks in the experimentally-based contrasts, the General Linear Model was still used to confirm that there were detectable differences between conditions in all regions which DCM could fit to.

Traditionally, it is assumed that the observed signal in the regions is strong enough for DCM to fit a representative predictive signal, and the fit is not directly tested. For DCM, it is not necessary to have a perfectly fitting prediction for the model to be valid, but if no effect of the observed signal is captured at all, the required aspects of timing to determine the best-fitting model cannot be captured either. While the General Linear Model being applied to the regions of the CMC does provide some detectable difference, it remains unclear whether the effect in the observed data is sufficient in strength to fit a meaningful prediction to it.

This research examines the fit of the individual regions when applying DCM to the Common Model of Cognition, and explores subject- and region-specific differences, providing deeper insight into what it means for a complex model to fit complex data, highlighting important challenges when interpreting high-level results, and balancing the contribution of theoretical frameworks and the limitations of a technical method.

Methods

Data

The following analyses were conducted on task-fMRI data from the Human Connectome Project (HCP) (Van Essen et al., 2013), focusing on a subset of 100 subjects participating in the Social Cognition task (Wheatley, Milleville, & Martin, 2007). The analysis was conducted in line with previous research on the CMC (Stocco et al., 2021).

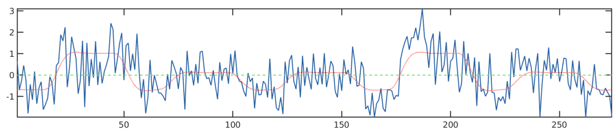
Measuring Regional Fit

The BMS results capture differences in **overall** model fit, but this metric is derived by combining the **regional** fits of all regions included in the model network. These fits are not usually examined individually, but because the CMC selects regions using their theorized importance to the model and not their known signal differences, comparing the predicted signal for each region to the extracted observed signal allows for a closer examination of this top-down approach. Namely, exploring the underlying patterns of regional model fit address questions like: Are there regions that consistently show a poor fit across subjects? Do some subjects show an overall better or worse fit? And lastly, are there recurring patterns of well-fitting regions in subjects which are more prevalent than others, and might therefore impact and skew the model fitting and comparison procedure?

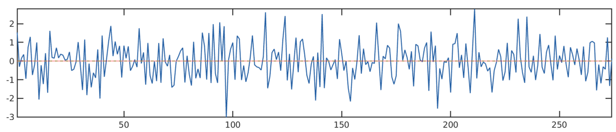
Figure 3 shows the difference between an informative and

Figure 3: Time-Series Prediction vs. Observed vs. Average

(a) Example of an Informative Prediction



(b) Example of an Uninformative Prediction



Notes. Example time-series of ROIs. The blue, solid line represents the observed data. The red, dotted line represents the predicted signal according to the DCM. The green, dashed line represents the mean of the observed signal.

uninformative prediction within a region over time. Figure 3a shows a prediction, shown by the red line, which generally follows the dynamics of the recorded signal, shown in blue, and suggests that the model that produced the prediction is doing a good job of capturing the related brain activity. By contrast, the predicted signal in Figure 3b is almost completely flat, and does not differ significantly from the mean activity, and seems unlikely to have been produced by a model that represents the underlying mechanisms which would help discern the temporal dynamics of the region.

These kinds of quality judgments of a predicted signal can be made easily by eye, but to systematically explore predictions across subjects and tasks, a measure needs to be defined that quantifies how well the predicted signal of each region in the DCM of an individual subject fits the observed BOLD-signal in their fMRI data. A simple measure of this can be calculated as:

$$r(t) = abs(o(t) - p(t))$$

With $r(t)$ representing the residuals of the observed signal o and the predicted signal p at time t. We can recover the average deviation r_m of the predicted signal from the observed signal throughout the session by averaging the residuals of all time points T.

$$r_m = \frac{\sum_{t=1}^{r} abs(o(t) - p(t))}{T}$$

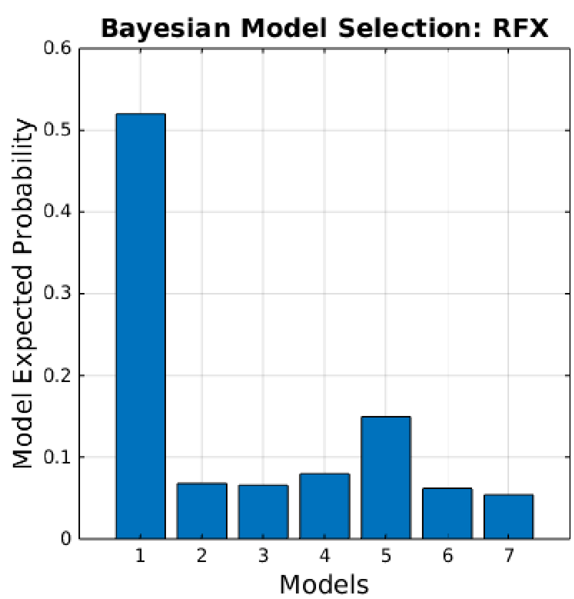
The goal of DCM is not to fit the prediction as accurately as possible, with Bayesian priors constraining model complexity. However, a complete lack of effect from conditions in a region means that the data does not provide any temporal information for DCM to determine the region's temporal dynamics and therefore effective connectivity.

A good point of comparison whether DCM was capable of predicting the signal better than "random" is to check whether its prediction fits the observed data better than just predicting the mean of the observed signal.

$$r_{mean}(t) = abs(o(t) - mean(o))$$

By applying a one-sided t test between the residuals over time we can check if the prediction by DCM is any better than predicting the mean of the observed data, and therefore, if any experimental effect has been captured at all. It does not check how strong this effect is, or how accurately it fits the observed data overall, as this should not necessarily be optimized by DCM due to its Bayesian constraints. Applying this metric to all subjects on all regions, we can deduce which regions have been able to provide any fit to an experimental effect for the different participants. This does not indicate whether any temporal dynamics have been correctly captured, but instead if it was possible to capture them from the respective region's experimental effect at all.

Figure 4: BMS Expected Probabilities CMC vs. Alternates

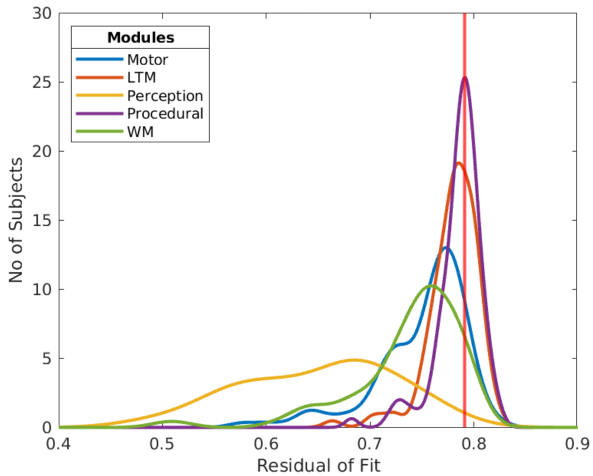


Notes. Models reported: **(1)** CMC; **(2)** Hierarchical 1; **(3)** Hierarchical 2; **(4)** Hierarchical 3; **(5)** Hub Prefrontal Cortex; **(6)** Hub Procedural; **(7)** Hub Temporal; see Stocco et al. (2021) for full descriptions of the alternate architectures.

Model Comparison Replication

To ensure comparability with previous DCM research, after the initial model fitting, the model comparison conducted by Stocco was repeated on the data used in this study. Bayesian Model Selection was applied to the fit of the Common Model of Cognition as compared to the previously used alternate network structures. As expected, the Common Model of Cognition outperformed alternate structures in regards to overall model fit, as visible in Figure 4. The CMC models created in this process were used as a basis for the regional analyses performed in the following.

Figure 5: Residual Distributions per Region in the CMC



Results

The fit of the predicted BOLD signal to the observed signal between regions was assessed by examining the distribution of the residuals across subjects, as shown in Figure 5. For comparison, predicting the mean of the signal for each region would yield residuals of about .791, as indicated by the red vertical line. The fit varies strongly between regions: Perception shows a clearly superior fit to the mean prediction at .65, whereas Procedural Memory performs poorly, with the residuals not fitting much better than predicting the mean, with a mean residual of .786. Long-Term Memory exhibits a similarly bad fit with mean residuals of .779, though its skewed distribution suggests that DCM predicts the data well for some of the subjects. Both Working Memory and Motor show an intermediary fit (mean residuals of .737 and .75 respectively), fitting noticeably worse then perception, but much better than Procedural and Long-Term Memory.

In addition to the differences in mean residuals across regions, there is a clearly visible variance between subjects in how well the respective predictions fit.

Table 1: Significant Subjects per Model per Region

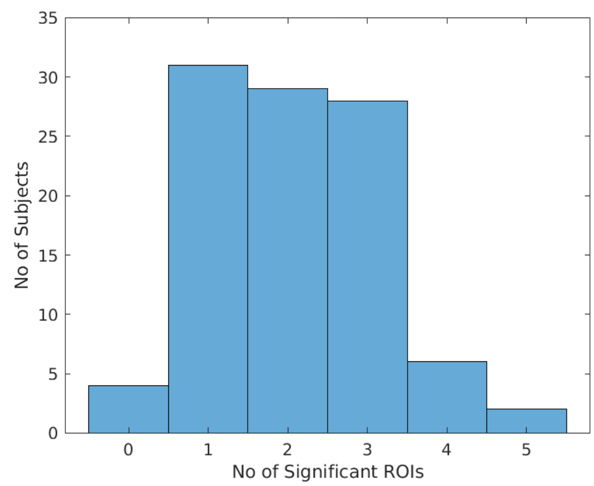
Model	Motor	LTM	Perc.	Proc.	WM	Avg.
CMC (Dir)	45	5	94	9	54	41.4
H1	39	7	97	6	49	39.6
H2	42	5	97	5	48	39.4
H3	43	10	97	4	58	42.4
Hub PFC	42	6	95	7	53	40.6
Hub Proc.	40	6	96	9	56	41.4
Hub Temp.	43	10	97	6	55	42.2
Average	42	7	96.1	6.6	53.3	41

Notes. The number of subject (from a set of 100) whose predicted signal was significantly better than the mean prediction in each region of the CMC, compared to the six alternate structures from Stocco et al. (2021).

As noted, the goal of DCM is not to maximize the fit of individual regions, but to maximize fit within the constraints imposed by the Bayesian Priors. However, for the experimental condition acting as an input to have an effect on model fit in a specific region, there must be at least some kind of detectable signal present. To determine whether the DCM predictions outperformed the mean predictions, t-tests were conducted comparing the two time-series for each region in each participant's model.

The results, summarized in Table 1, show that across models, for the vast majority of subjects, DCM predicts Perception significantly better than the mean. For Working Memory and Motor, this is true for about half of the subjects. In contrast, only a small minority of subjects achieve a significant fit for both Long-Term Memory, as well as Procedural Memory.

Figure 6: Distribution of Significant Regions per Subject in the CMC

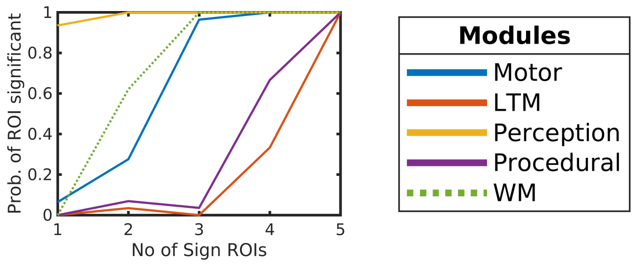


The substantial deviation in how well different regions fit raises the question of how the fit of regions is distributed across subjects. Therefore, the number of regions of interests with significant fits for each subject was examined.

As shown in Figure 6, only 2 subjects have DCM models with significantly-better-than-random fits on all five regions. An additional 6 subjects show significant fits for 4 out of the 5 regions. This means that for 92% of the subjects, their best fitting DCM model consists of three or less regions achieving significant fits (28 subjects with 3 significant regions, 29 with 2, and 31 with only 1 significant region). For 4 subjects, none of the regions achieved significant fits.

The unequal distribution of significant regions has implications for the combinations of significant regions more and less prevalent in subjects. Figure 7 shows the probability of each region being significant for subjects with a certain number of significant regions. If subjects have three or fewer significant regions, as is the case for the vast majority of subjects, these are most commonly a combination of the same three regions: Perception, Motor, and Working Memory. Only rarely will Long-Term Memory or Procedural Memory be significant in

Figure 7: Order Effects of Regions in the CMC



these cases. These regions typically only become significant for the minority of subjects with four or five significant regions, in which case the three aforementioned regions are significant as well.

Discussion

Traditionally, task-based Dynamic Causal Modeling is applied to networks among regions which show a strong experimental effect in the task. However, regions which do not exert a strong effect can still be involved in the cognitive processing of the task, and can be relevant or even necessary for a complete model of brain networks. If the observable difference between conditions, and therefore states, in the data is too small or nonexistent, DCM cannot capture the temporal dynamics of the respective region and cannot extract its relation to other regions. In its traditional setup, the quality of temporal dynamics in the General Linear Model is commonly assumed to be sufficient to be captured, and therefore the regional fit of the predicted signal is not usually investigated. Here, using a theory-driven setup with only some effect in the General Linear Model present, we demonstrate that the regional fit in the DCMs of the Common Model of Cognition, is often not capturing the observed signal well, even if the overall model fit strongly outperforms all other network configurations. Indeed, it was found that there were significant issues with the regional fit, specific to some regions and subjects.

Potential Causes for Fit Inconsistency Between Regions

The individual regions in the DCM showed a wide range of fit, with some, such as Perception, fitting much better than others, such as Procedural Memory. There are multiple possible explanations for this, some of which relate to the quality of the data. If, for example, the location selected from which to extract the time-series does not correctly reflect the cognitive processing attributed to it, one might end up with uninformative or misleading observed data. This is a particular risk when regions are derived from theory, instead of from previously identified strong signals. Further, some regions are difficult to record with modalities such as fMRI. Procedural Memory, the worst performing region, is mapped to the Basal Ganglia, which often poses problems for various brain imaging techniques due to its small size, varying location between individual brains, and its deep, subcortical location in the brain (de Hollander, Keuken, van der Zwaag, Forstmann, & Trampel, 2017; Wall, 2023). Further, they are susceptible to vascular artifacts affecting the recorded BOLD signal, which makes the detection of task-related neural changes more difficult. This means that even when the hypothesized region is relevant for task processing, and when there is generally processing to be recorded in the area, it might sill be difficult to extract a meaningful signal from a modality such as fMRI, making the temporal dynamics difficult to capture for DCM. Task-related changes in Perception on the other hand were comparably easy to detect in modeling the CMC using fMRI, as perceptual processing spreads over a large area, within which the task-related signal tends to be rather strong.

Another plausible cause relates not to issues with the data quality, but the modality with which the data is recorded. A region's activity might simply not experimentally differ, even if it is involved in the task. If the region is active regardless of condition, without any difference between experimental conditions showing in the recorded data, the temporal dynamics cannot be captured by simply using the different conditions as onsets for neural change.

The Role of Regions Within Larger-Scale Models

When using a tool like DCM, it is tempting to take results like these and determine that poorly fit regions should be removed from the overall model framework. However, this should not be enough to ignore the theoretical reasons that a region or component should be considered. Theoretical models may use these regions to explain processing steps expected during the task, and including the connections between these regions and those that are fit well with methods like DCM may still help to understand the network as a whole. Even more problematically, removing a region relevant for the cognitive processing of the task might lead to a misrepresentation of the network, causing DCM to incorrectly estimate the timing of the cognitive stages. While these issues are particularly relevant for theory-driven models, they also have the potential to impact results for traditional contrast-derived models as well, as a region relevant for processing might simply not show in the contrasts and not be included in the network analysis.

But while understanding the reasons behind a region's lack of fit in DCM might be helpful in selecting networks for comparisons or interpreting results, the issue of inconsistencies in regional fit will likely persist. Reliably detecting task-related changes in areas like the Basal Ganglia is challenging with most available imaging techniques, and even those that seem more promising, like High-Field MRIs or specific pulse sequences (de Hollander et al., 2017), are not feasible for collecting the scope of data required for this kind of general, whole brain network analysis (Wall, 2023). Simply ignoring the Basal Ganglia as a factor, however, is equally implausible as a path forward, given the central role it plays in many cognitive modeling paradigms. Future investigations in this direction, then, must balance the limitations of their methodologies with the demands of their theoretical frameworks. In the case of DCM, this includes better understanding the effects a non-fitting or badly fitting region has on the model fit and reliability of model comparisons.

The Impact of Region Fit Inconsistencies

A primary result of these analyses was the finding that the majority of participants had only 3 or fewer regions which were significantly better than predicting the mean. Differences in quality between participants can be caused by differences in either the general quality of the scans, or by the aforementioned issues in the localization and noise of specific regions. If DCM is only able to fit appropriately to the temporal dynamics of a subset of regions, two issues might occur. First, there is a risk that DCM might effectively end up "ignoring" the other regions, with the winning model resembling the best connectivity of only the factored in subset. Second, while DCM is intended to compensate for model complexity, validations of this are based on data which indeed has significant differences in regions between experimental conditions. There is a risk that the lack of an appropriate fit in some regions might skew the fit, with a more complex model having an advantage over a less complex one.

It is not trivial to investigate whether and how badly fitted regions affect the retrieval of the correct model, as there is no access to the ground truth underlying the brain's processing. It is possible that a badly fitting region throws off the model retrieval and comparison. However, it is also possible that DCM has some resiliency if a limited number of regions are not fitting well. For example, if a region between two regions cannot be captured well, but the other two can, the timing effects of the badly captured region on the others can still be sufficient for the generatively created DCM to correctly retrieve the best model. While this cannot be directly observed, some indicators should be investigated.

One possible approach could be to investigate differences of model fit between groups of subjects with well-fitting and badly-fitting regions. The amount of well-fitting subjects in this research was not sufficient to retrieve meaningful results, but with more balanced datasets or models, this might be achievable. Another possible approach could be to take a model with well-fitting regions and replace one or multiple of them with badly-fitting regions, investigating the effect this has on the model comparisons. Finally, the resiliency of DCM could be tested using simulated data, where the ground truth is known and any deviations in the model can be directly examined.

In addition to investigating the role of regional fit, exploring the impact of complexity could also be useful for contextualizing these kinds of analyses. The slightly higher number of connections of the CMC compared to its theory-driven alternatives might have given it an advantage in the model comparison, due to a higher level of complexity. While the current methods of comparison are supposed to be able to account for this degree of difference, this assumption has not considered how poorly fitting regions affect the effective complexity of a model. If complexity is the determining factor of the results, either due to the amount of overall connections, or the connections from the input-areas, such as perception or working memory, any randomly connected model with the same complexity should outperform the other models as much as the CMC does. This could be investigated by fitting randomly connected models with the same level of complexity that the CMC has and see if, and by how much, they outperform the competing models relative to the advantage the CMC has.

Finally, if a region suffers from a lack of capturable difference between conditions, but is otherwise a valid representation of the processing in the respective area over time, DCM might simply need a different, non-experimental signal it can fit to. This is already a standard with modeling the connectivity of resting-state fMRI data (rsfMRI) (K. J. Friston, Kahan, Biswal, & Razi, 2014), and has even been applied to investigating the CMC using rsfMRI (Sibert, Hake, & Stocco, 2022). During resting-state fMRI scans, participants lie in a scanner with open eyes without any task, so there is no experimental input that could be used. To model the spontaneous, resting-state related brain activity, inputs of different frequencies are used for DCM, which represent the fluctuations found in brain activity. This allows DCM to effectively fit its generative model to resting state data in a plausible way, and still capture the underlying temporal dynamics of the networks. Such inputs could be used for modeling an experimental task, in addition to the task-based inputs, which are currently the only drivers of activity considered by the model. This way, regions showing no discernible patterns in activity related to experimental inputs can still be fit to frequency-based fluctuations in the recorded brain activity. This cannot compensate for regions with a general lack of signal, or too much noise to discern any signal on a frequency basis, but these challenges do not differ from those already existent with the methods applied in resting state fMRI. This is also true for the risk of overfitting to patterns not related to brain activity, but rather noise, which would apply to all models fitted to the data.

Conclusion

A close investigation of the components of an otherwise well fitting model of brain connectivity using Dynamic Causal Modeling reveals important inconsistencies in regional fit. While we do not consider this to invalidate previous results, it provides important context both for current results and future investigations, particularly in cases where DCM is applied to theory driven, rather than data driven networks. Several paths forward are proposed for better understanding the effects of poor regional fits of brain data, but we also consider these results worth discussing in a broader sense. The investigation here is specific to theory driven network models being used to fit brain signaling data, but there are many projects across the wide field of cognitive science that are using similarly complex methods to explore similarly complex data. Examining the component elements of complex methods not only provides additional metrics through which to measure success, they also provide vital context through which to interpret the overall results.

References

Bajaj, S., Butler, A. J., Drake, D., & Dhamala, M. (2015). Functional organization and restoration of the brain motor-execution network after stroke and rehabilitation. *Frontiers in Human Neuroscience*, *9*. doi: 10.3389/fnhum.2015.00173

de Hollander, G., Keuken, M. C., van der Zwaag, W., Forstmann, B. U., & Trampel, R. (2017). Comparing functional mri protocols for small, iron-rich basal ganglia nuclei such as the subthalamic nucleus at 7 t and 3 t. *Human Brain Mapping*, *38*(6), 3226-3248. doi: 10.1002/hbm.23586

de Schotten, M. T., & Forkel, S. J. (2022). The emergent properties of the connected brain. *Science*, *378*(6619), 505-510. doi: 10.1126/science.abq2591

Friston, K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273-1302. doi: 10.1016/S1053-8119(03)00202-7

Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). A dcm for resting state fmri. *NeuroImage*, *94*, 396-407. doi: j.neuroimage.2013.12.009

Kelly, C., & Castellanos, F. X. (2014, March). Strengthening Connections: Functional Connectivity and Brain Plasticity. *Neuropsychology Review*, *24*(1), 63–76. doi: 10.1007/s11065-014-9252-y

Laird, J., Lebiere, C., & Rosenbloom, P. (2017, 12). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine*, *38*, 13. doi: 10.1609/aimag.v38i4.2744

Parks, E. L., & Madden, D. J. (2013). Brain connectivity and visual attention. *Brain Connectivity*, *3*(4), 317-338. (PMID: 23597177) doi: 10.1089/brain.2012.0139

Pessoa, L. (2014). Understanding brain networks and brain organization. *Physics of Life Reviews*, *11*(3), 400-435. doi: 10.1016/j.plrev.2014.03.005

Seghier, M. L., Zeidman, P., Neufeld, N. H., Leff, A. P., & Price, C. (2010). Identifying abnormal connectivity in patients using dynamic causal modelling of fmri responses. *Frontiers in Systems Neuroscience*, *4*. doi: 10.3389/fnsys.2010.00142

Sibert, C., Hake, H. S., & Stocco, A. (2022). The structured mind at rest: Low-frequency oscillations reflect interactive dynamics between spontaneous brain activity and a common architecture for task control. *Frontiers in Neuroscience*, *16*. doi: 10.3389/fnins.2022.832503

Steine-Hanson, Z., Koh, N., & Stocco, A. (2018). Refining the common model of cognition through large neuroscience data. *Procedia Computer Science*, *145*, 813-820. doi: 10.1016/j.procs.2018.11.026

Stephan, K., Penny, W., Moran, R., den Ouden, H., Daunizeau, J., & Friston, K. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage*, *49*(4), 3099-3109. doi: 10.1016/j.neuroimage.2009.11.015

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E., Breakspear, M., & Friston, K. J. (2008). Nonlinear dynamic causal models for fmri. *NeuroImage*, *42*(2), 649-662. doi: https://doi.org/10.1016/j.neuroimage.2008.04.262

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *NeuroImage*, *46*(4), 1004-1017. doi: j.neuroimage.2009.03.025

Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., & Rosenbloom, P. (2021). Analysis of the human connectome data supports the notion of a "common model of cognition" for human and human-like intelligence across domains. *NeuroImage*, *235*, 118035. doi: 10.1016/j.neuroimage.2021.118035

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The wu-minn human connectome project: An overview. *NeuroImage*, *80*, 62-79. (Mapping the Connectome) doi: 10.1016/j.neuroimage.2013.05.041

Wall, M. B. (2023, dec 22). Multiband acquisition sequences for fMRI: Proceed with caution. *Aperture Neuro*, *3*. doi: 10.52294/001c.91292

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science*, *18*(6), 469-474. (PMID: 17576256) doi: 10.1111/j.1467-9280.2007.01923.x