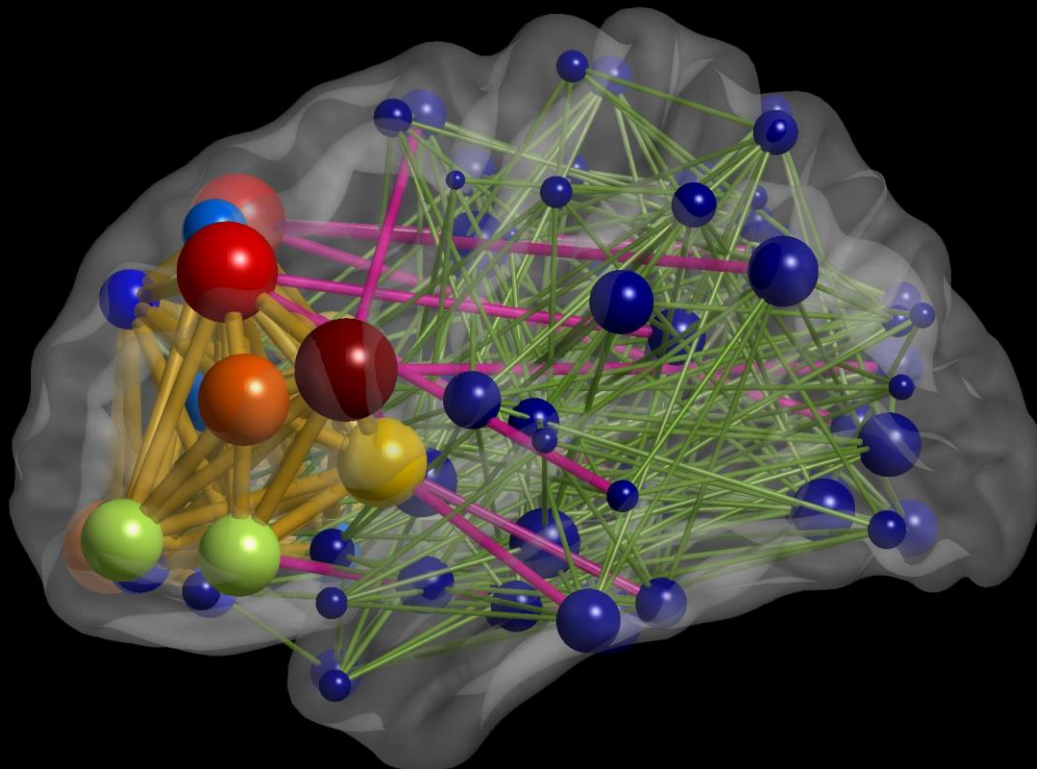


WFU_MMNET Manual



*Laboratory for Complex Brain Networks (LCBN)
Department of Radiology
Wake Forest School of Medicine
Winston-Salem, NC 27127, USA*

January 2022

Contents

1	Introduction	3
2	Installation.....	5
2.1	Matlab.....	5
2.2	Modeling software.....	5
2.2.1	Adding executable file path	5
2.3	Running WFU_MMNET	7
2.4	Modeling framework.....	8
2.5	Supported formats	8
3	Network model.....	9
3.1	NetMod Global – Network model for global network analysis.....	9
3.1.1	Imaging data files.....	10
3.1.2	Loading data files.....	10
3.1.3	Spatial distance	12
3.1.4	Adding spatial distance via atlas files.....	13
3.1.5	Adding spatial distance manually	13
3.1.6	Correlation	14
3.1.7	Network metrics	15
3.1.8	Output directory	16
3.1.9	Run.....	17
3.1.10	Clear.....	17
3.1.11	Running multiple network models	17
3.1.12	Adding spatial distance files manually.....	20
3.2	NetMod Regional – Network model for regional network analysis	21
3.3	NetMod Global Multitask – Network model for global multitask network analysis	21
3.4	NetMod Global Dynamics – Network model for dynamic global network analysis	22
4	Statistical Model.....	24
4.1	StatMod Global – Statistical model for global network analysis.....	24
4.1.1	Input directory	25
4.1.2	Output directory	25
4.1.3	Network metrics	26
4.1.4	Covariates	26
4.1.5	Interactions	29

4.1.6	Spatial Distance.....	29
4.1.7	Nodal propensities and Paired subjects	30
4.1.8	Modeling.....	30
4.1.9	Selecting modeling software	30
4.1.10	SAS (Glimmix)	31
4.1.11	R (lmerTest).....	32
4.1.12	Python (Statsmodels – GEE)	32
4.1.13	Data Size Reduction	32
4.1.14	Run.....	33
4.1.15	Clear.....	33
4.2	StatMod Regional – Statistical model for regional network analysis	33
4.2.1	Regional Covariates.....	34
4.3	StatMod Global Multitask – Statistical model for multitask network analysis.....	36
4.4	StatMod Global Dynamics – Statistical model for dynamic network analysis.....	36
5	Visualization – Regional Analysis.....	37
6	Acknowledgement.....	41
	References	42

1 Introduction

WFU_MMNET is a freely-available toolbox that provides a multivariate framework to study brain networks. This toolbox is specifically developed for testing hypotheses on topological changes of brain networks within a multivariate framework. It allows testing hypotheses on brain connectivity changes as well as simple and complex topological changes of brain networks. However, we caution the users that for more complex changes, the interpretation of results may not necessarily be straightforward. However, it still provides statistical inference about such complex changes of brain topology within a multivariate framework. This toolbox utilizes a mixed-effects regression framework (Simpson and Laurienti 2015) that allows assessing brain network differences between study populations as well as assessing the effects of covariates of interest such as age and disease phenotypes on brain connections in global (i.e., whole-brain) and local (i.e., subnetworks) brain networks. A variety of neuroimaging data such as fMRI, EEG, MEG, and DTI can be analyzed with this toolbox, which makes it useful for a wide range of studies examining the structure and function of brain networks. The new version of the toolbox includes additional GUIs to test hypotheses on local brain networks (i.e., subnetworks such as DMN) within the context of their whole brain networks (i.e., allowing for simultaneous analyses of local – global analyses) while accounting for the complex interactions between local brain networks with other brain regions (Bahrami, Laurienti et al. 2019), to test hypotheses on whole-brain networks across multiple tasks (Simpson, Bahrami et al. 2019), and to test hypotheses on whole-brain dynamic changes of (functional) brain networks through utilizing orthonormal polynomials for capturing the dynamic trends.

WFU_MMNET is developed in Matlab (The MathWorks Inc., Natick, MA, US) with a user-friendly GUI, under a 64-bit Linux platform. ***For whole-brain (i.e., global) static changes, the toolbox uses SAS, R, or Python (depending on software availability) to perform the statistical modeling. However, for the latter three analyses, including: regional network analysis, multitask network analysis, and dynamic network analysis, the user can only use SAS as R and Python will face computational issues for the latter three analyses and, thus, have not been***

implemented in the current version. Also, for the regional analysis, we have added a visualization tool to map the results into the brain space. The toolbox uses the BrainNet viewer toolbox (Xia, Wang et al. 2013) and modifies the node color and size to illustrate the significant changes. The users of this toolbox that use the visualization utility, should cite the BrainNet Viewer toolbox as well. This toolbox also uses the functions implemented in the brain connectivity toolbox (BCT) (Rubinov and Sporns 2010) to compute the topological network features. For the whole-brain static networks, this toolbox has been successfully tested under Linux operating systems with Matlab, SAS, R (and the required package – lmerTest), and Python (and the required modules – numpy, scipy, pandas, statsmodels), and for the latter three analyses, i.e., regional network analysis, multitask network analysis, and dynamic network analysis, the toolbox has been tested with Matlab and SAS. Any suggestions about the toolbox or the manual would be greatly appreciated.

Adding BCT and BrainNet Viewer. After downloading WFU_MMNET, download BCT (<https://drive.google.com/file/d/1DmMvRnferBfGe057O-sZwB5jL4j8w1Hu/view>) and BrainNet Viewer (https://www.nitrc.org/frs/shownotes.php?release_id=4161) and put the unzipped folders under 'BCT' and 'BrainNetViewer' folders in the 'Toolbox' folder of WFU_MMNET. You may skip adding BrainNet Viewer if you don't use regional network analysis or don't need the visualization tool.

Laboratory for Complex Brain Networks
Department of Radiology
Wake Forest School of Medicine
Winston-Salem, NC 27127, USA.

Contact Information:

Mohsen Bahrami: mohsenb6@vt.edu, mbahrami@wakehealth.edu

Copyright © 2022

2 Installation

2.1 Matlab

A version of R2021a or above is recommended for Matlab as the new version has been developed in R2021a version (If you have older versions above R2014b, you can use WFU_MMNET v1.0, but it only allows analyzing whole-brain static networks).

Add BCT path to Matlab search path. You can add the path through either the “addpath(genpath(...))” function or “Set Path” option in the Matlab HOME toolstrip:

- Type “addpath(genpath('/.../path to the BCT folder/..'))” in the Matlab command window.
- Click “Set Path” - > Click “Add with Subfolders...” button- > Select the BCT folder - > Click “OK” button- > Click “Save” button.

2.2 Modeling software

For the whole-brain static networks, WFU_MMNET uses SAS, R, or Python for performing the modeling part. Thus, the user needs to have one of SAS, R, or Python (preferably SAS or R) installed prior to using this toolbox. Python is usually installed on Linux by default. So, we have not provided the Windows option for Python. If you want to use Python, you must have it installed on Linux. Please install the ‘lmerTest’ package if using R, and install ‘numpy’, ‘scipy’, ‘pandas’, and ‘statsmodels’ if using Python. For regional network analysis, multitask network analysis, and dynamic network analysis, you can only use SAS.

2.2.1 Adding executable file path

Add the **path of the folder that includes the executable file of the modeling software** in the specified location to the “executable_file_path” text file. This text file is in the toolbox folder – i.e., ‘WFU_MMNET’ folder.

- Open the “executable_file_path” text file -> add the path of the folder including executable file in the specified location as shown in Fig 1.

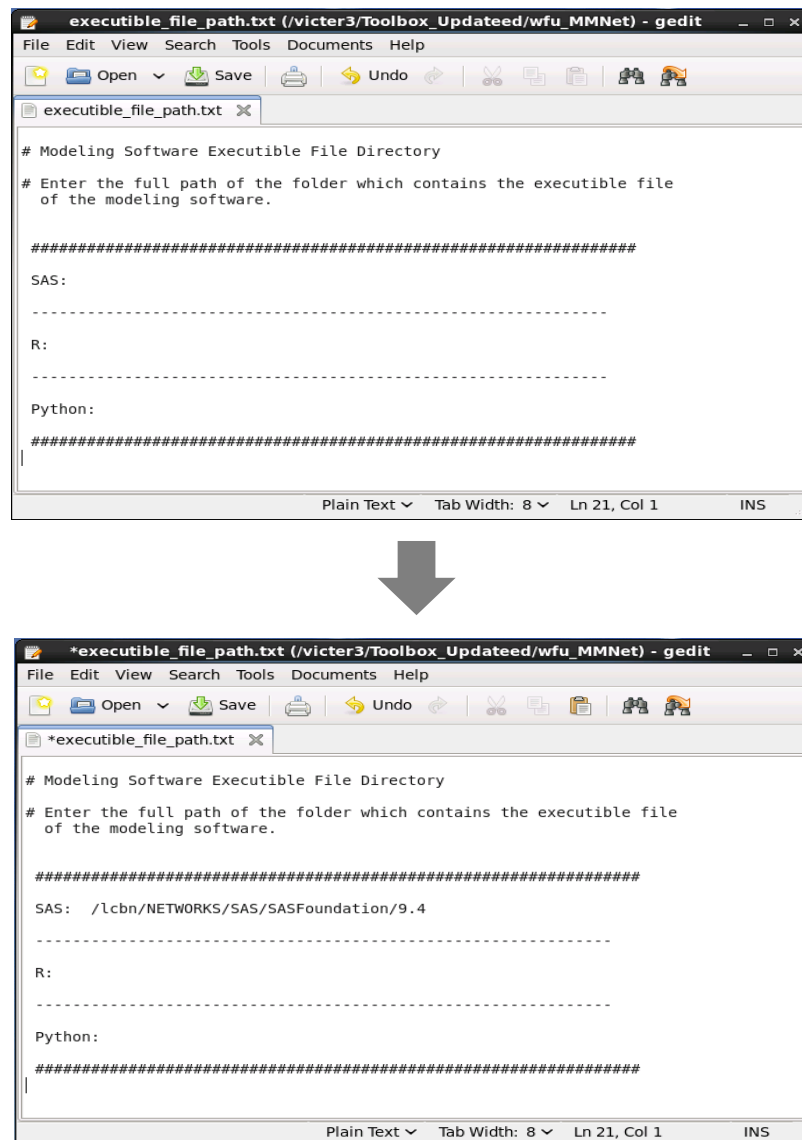


Fig 1 Adding path of the folder that includes SAS executable file. Add R and Python in the same way. The example shows the location of the SAS folder in the computer environment at the Laboratory for Complex Brain Networks.

You can add more than one modeling software (e.g., SAS and Python) if they are available, and select either one when modeling your datasets.

2.3 Running WFU_MMNET

Add WFU_MMNET path to Matlab search path:

- Click “Set Path” - > Click “Add with Subfolders...” button- > Select the “WFU_MMNET” folder - > Click “OK” button- > Click “Save” button.
- or
- Type “`addpath(genpath('/.../path to the WFU_MMNET folder/..'))`” in the Matlab command window.

Run ‘WFU_MMNET.m’ to start the toolbox - you can type ‘WFU_MMNET’ in the command window of Matlab.

You should see the window shown in Fig 2.

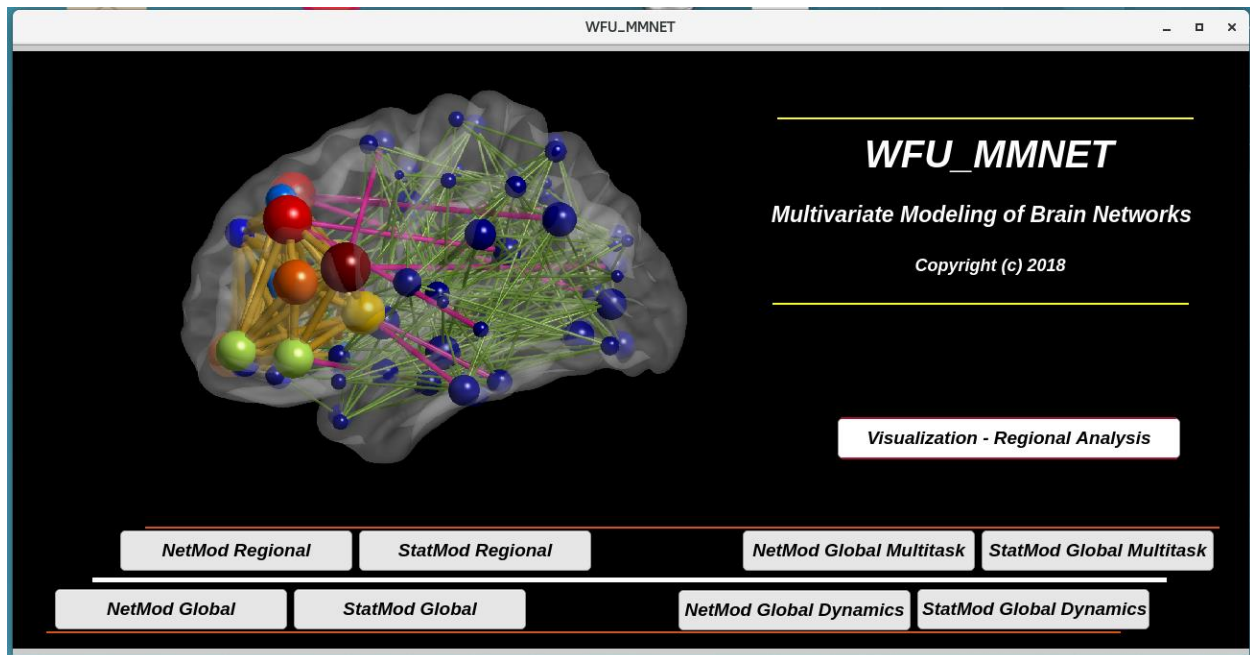


Fig 2. The main graphical user interface (GUI) of WFU_MMNET.

2.4 Modeling framework

For each analysis, modeling is done in two main steps using two different GUIs. We have divided it into two steps for an easier, flexible, and more efficient modeling. In the first step, using imaging data files or connectivity matrices, a modeling data frame is constructed through the network model GUI (by using one of NetMode Global, NetMod Regional, NetMod Global Multitask, or NetMod Global Dynamics GUIs). This step is independent of the second step and can be repeated for making different data frames of different datasets or making different data frames for different options of the same dataset. This step is described in section 3.

In the second step, the generated data frame is used by the statistical model GUI (by using one of StatMode Global, StatMod Regional, StatMod Global Multitask, or StatMod Global Dynamics GUIs) for constructing the appropriate modeling files and fitting statistical models. This step can also be done independently as long as modeling data frames are available. This step is described in section 4.

2.5 Supported formats

Several data files should be loaded as inputs when using this toolbox. Multiple formats including mat, nii, nii.gz, csv, txt, xls, xlsx, and xlsm are supported for any loaded data in this toolbox. Loaded 2D numeric files can have row names, column names, or neither. Data files can be loaded directly or via loading a list file that includes full directory paths and file names for the data files.

3 Network model

As mentioned above, for each analysis, modeling is done in two main steps. In the first step, the data frame will be generated. This file will be used in the next step for fitting the statistical models.

3.1 NetMod Global – Network model for global network analysis

Here, you generate the data frame through “NetMod Global” GUI.

Click the ‘NetMod Global’ button in ‘WFU_MMNET’ to start the ‘NetMod_Global’ GUI. You should see the window shown in Fig 3.

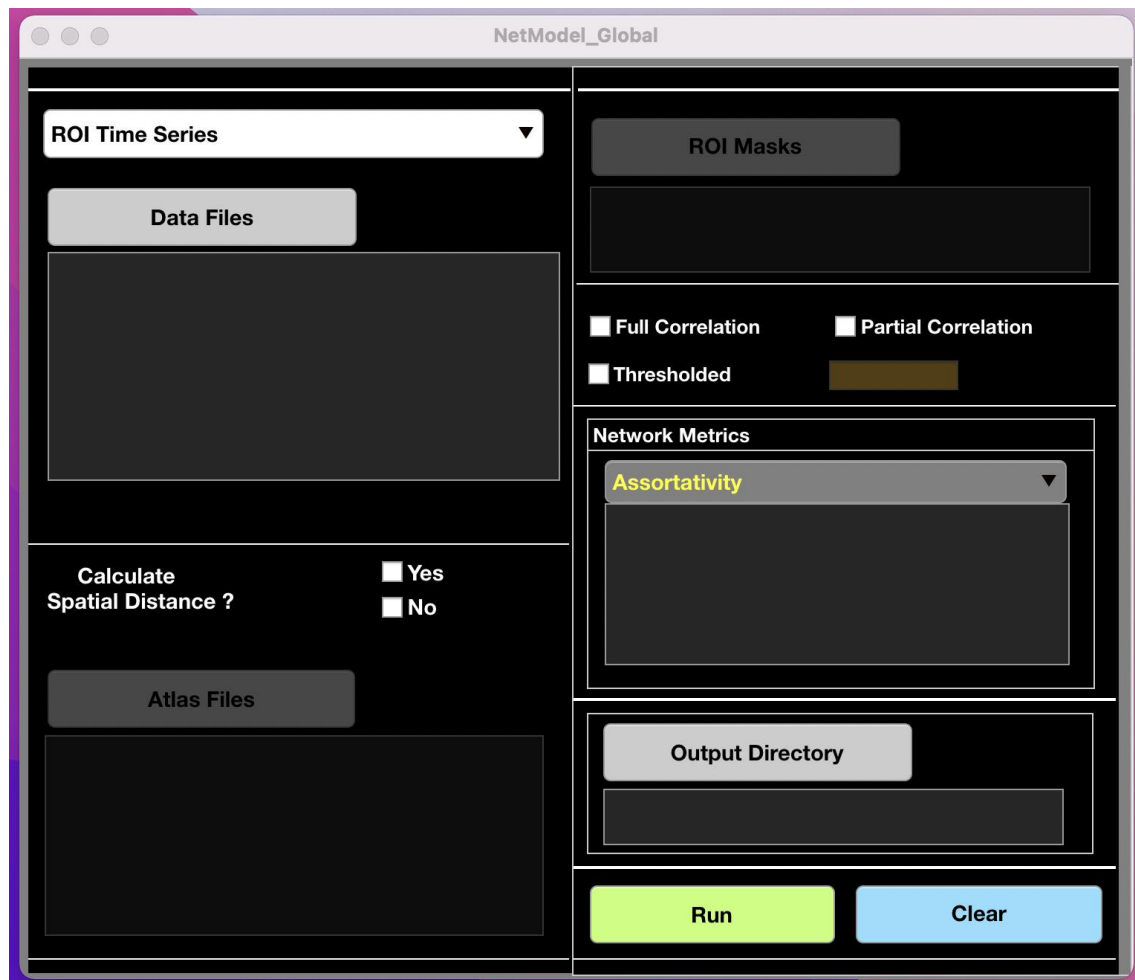


Fig 3. The network model GUI for global network analysis

3.1.1 Imaging data files

Imaging data files can be loaded as both matrices of ROI time series (with rows representing the ROIs and columns representing the time) and connection matrices. FMRI users can also use (preprocessed) 4d voxel-wise time series (and an atlas defining ROIs) instead of ROI time series or connection matrices.

Click the popup dialogue at the top left, and select the appropriate data type for your imaging data (Fig 4). Selecting “Voxel-wise Time Series (fMRI)” enables the “Atlas Files” button for loading the atlas file(s).

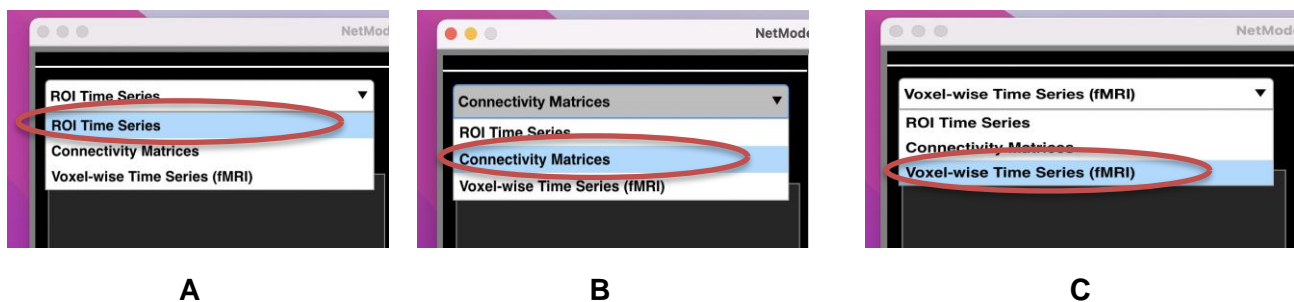


Fig 4. Imaging data type. **A.** ROI time series, **B.** Connection matrices, **C.** Voxel-wise time series (only for fMRI users). Selecting “Voxel-wise Time Series (fMRI)” enables the “Atlas Files” button. You should load atlas files prior to loading the data files if using 4d voxel-wise time series.

3.1.2 Loading data files

Click “Data Files” button to load the data files. Data files can be loaded as matrices of ROI time series (with rows representing the ROIs and columns representing time), connection matrices, or fMRI 4d voxel-wise time series. Load data files directly (Fig 5.A) or via loading a list file which includes data files directories (Fig. 5.B). A sample list file is shown in Fig 6.

Full directories of loaded files will be listed in the listbox, and the number of subjects and ROIs will be displayed (see Fig 5). If the length of loaded time series is different across subjects, a warning message showing the maximum difference will be displayed



Fig 5. Imaging data files. **A.** Loading files directly (files will be sorted). **B.** Loading files via loading the list file including data files' directories.

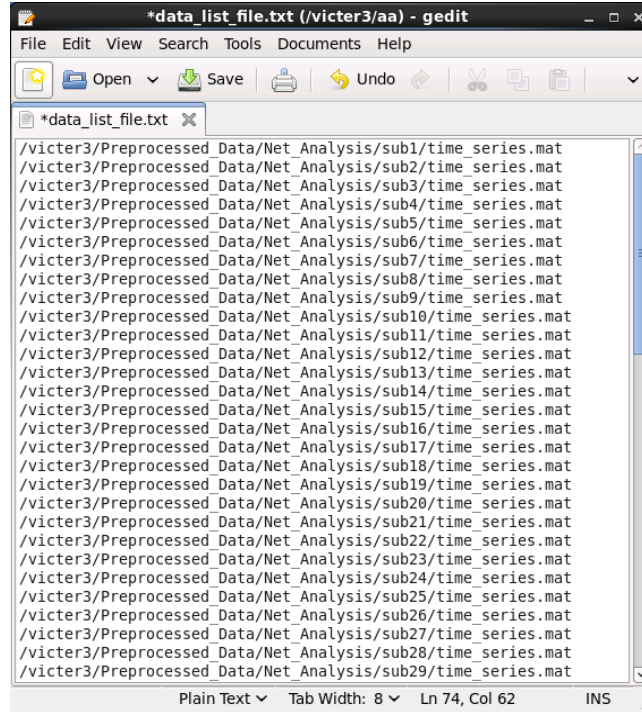


Fig 6. Sample list file including data files' directories. You can use list files with backslash (\) separated directories as well (i.e., Windows format).

in the command window of Matlab. If data files are selected directly (not via a list file), the toolbox sorts them and uses the sorted ones in the subsequent analyses. If analyzing paired subjects, the first and the second half of loaded data files must represent the paired subjects in correct order. For example, for a paired analysis with 50 subjects (e.g., pre-post treatment), data files 1 and 51, 2 and 52 , ..., and 50, 100, must each represent different networks of the same subject.

3.1.3 Spatial distance

As pointed out in (Simpson and Laurienti 2015), spatial distance (anatomical distance) and square of spatial distance between brain regions are important geometric features in predicting the connectivity between brain regions. Thus, we have made it possible to control for their likely confounding effects.

3.1.4 Adding spatial distance via atlas files

Spatial distance can be computed by loading the atlas file(s) defining the ROIs. The toolbox computes Euclidean distances between the centers of mass of the ROIs. To use this option, choose 'Yes' next to 'Calculate Spatial Distance?' (Fig 7); then, load your atlas file(s). You can either load a single atlas (i.e., a common atlas for all subjects) or multiple atlas files (i.e., a separate atlas for each subject) (Fig 8). The latter case is useful for situations where data are analyzed in the individual subjects' native space and different atlases are used for each study participant. If loading multiple atlas files, again, you can either load them directly or through a list file that includes their directories, as done for loading imaging data files. Loaded imaging files and atlas file(s) must have the same number of ROIs.

A computed spatial distance file for each subject will be saved as a symmetric numeric matrix.

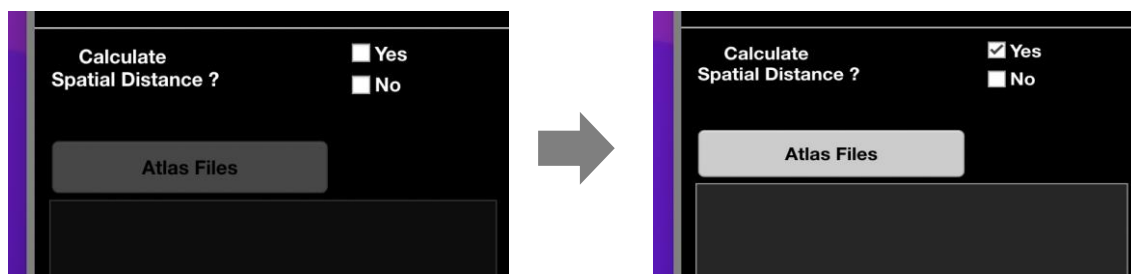


Fig 7. Select 'Yes' if you want to add the spatial distance via the toolbox. This will enable 'Atlas Files' button for loading atlas file(s).

3.1.5 Adding spatial distance manually

If you have already computed spatial distance or have access to it, choose 'no' next to 'Calculate Spatial Distance?' and skip loading atlas files.

Distance file should be a symmetric numeric matrix with element a_{ij} (and a_{ji}) representing the spatial distance between regions (nodes) i and j . You can add your distance files at the end (after running the model), as detailed in section 3.1.12.



Fig 8. Loading atlas files. You can either select a single atlas file or multiple atlas files. In the latter case, the number of loaded atlas files must be the same as the number of loaded imaging data files, and the number of ROIs must be the same across subjects.

3.1.6 Correlation

If you use connection matrices as your data files, you can skip this section. If you use (ROI or voxel-wise) time series, the connection matrices will be computed through either Pearson full or partial correlation. You can select either one (Fig 9.A). However, if you wish to use methods other than full or partial correlation, you should compute the connection matrices using your preferred method and then load the connection matrices instead of time series. Note that negative connections will be set to zero for the subsequent analyses. A density thresholding option is provided as well to remove weak connections (however there may be benefits to using fully, weighted matrices). If you do decide to use this option, enter the percentage of the strongest connections that you wish to preserve for each subject (Fig 9.B).



Fig 9. Correlation method. **A.** Select full or partial correlation by filling their corresponding checkboxes (negative correlations will be set to zero). **B.** Enter the percentage of the strongest connections that you wish to maintain. For example, here, 80 percent of correlations in each subject's connection matrix will be set to zero and the strongest 20 percent of positive correlation values will be maintained.

3.1.7 Network metrics

Click the popup dialogue and select all the network features that you would like to consider using in your models (Fig. 10). Selected network metrics here will be computed and available (but not required) for use in your statistical models.

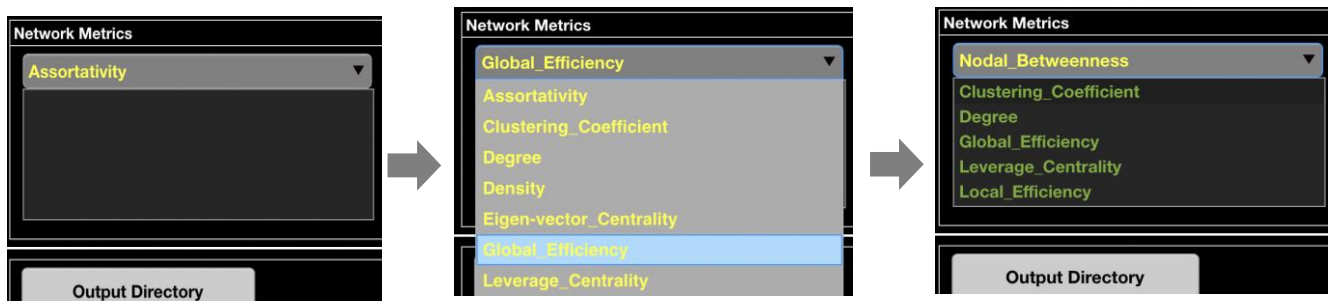


Fig 10. Topological network metrics. Click the popup dialogue and select network metrics that you might want use in your models. You can remove the selected metrics by clicking on them in the listbox.

Final network metrics will be selected in the “StatMod Global” GUI from those you select here. This approach is efficient since network features are computed once (here), while different sets of them can be used in your models later.

Selected features are displayed in the listbox. You can remove the selected metrics by clicking on them in the listbox. Network metrics are computed using BCT (Rubinov and Sporns 2010).

3.1.8 Output directory

Click ‘Output Directory’ and select an output directory (Fig 11). Several files, including *data frame*, *correlation matrices* (if starting with time series), *network metrics*, *spatial distance* (if you want to add the spatial distance though the toolbox), *extracted ROI time series* (if starting with 4d voxel-wise time series), and *network model summary*, will be stored in this folder. Generated files will be saved in appropriate auto-generated sub-folders in the selected output directory.

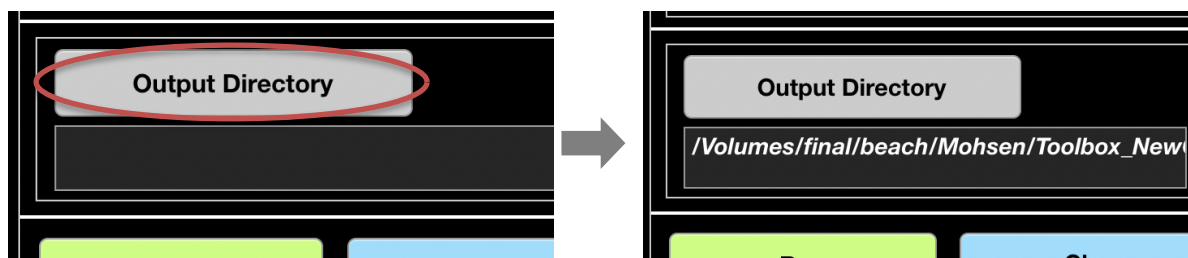


Fig 11. Output directory. Generated files will be saved in appropriate sub-folders under this directory.

If you use the same directory in different runs, older files will be overwritten by new generated files. So, if you want to keep the generated files of each run, you should use different directories to avoid this. A warning message will be displayed reminding you about this. (Note that the output directory of the ‘NetMod Global’ GUI will be the ‘Input

Directory' of the 'StatMod Global' GUI. However, you can run several network models and save them in different directories and start from either one in the statistical model.

3.1.9 Run

After loading all required data files and output directory, and selecting desired network metrics and correlation methods, Run the network model by clicking 'Run' (Fig 12).

Multiple sets of files will be generated during this step, including: extracted ROI time series (if starting with 4d voxel-wise time series), correlation matrices (if starting with time series), network metrics, spatial distance (if choosing 'Yes' for calculating spatial distance), data frame, network model summary, and a cell file which includes selected network features' names.

- Load('/... your selected output directory .../Network_Model_Summary.mat');

The loaded file will appear in the Matlab's Workspace at the top right. Double click on this file to see the network summary (Fig 14).

3.1.10 Clear

Click 'Clear' button to reset the network model.

3.1.11 Running multiple network models

As pointed out before, the 'NetMod Global' and 'StatMod Global' GUIs work independently of each other. So, if you wish to model different datasets or model the same dataset with different options (e.g., using full correlation and partial correlation), you can first run your multiple network models (and choose different output directories for each one); then, start with either one in the 'StatMod Global' GUI.

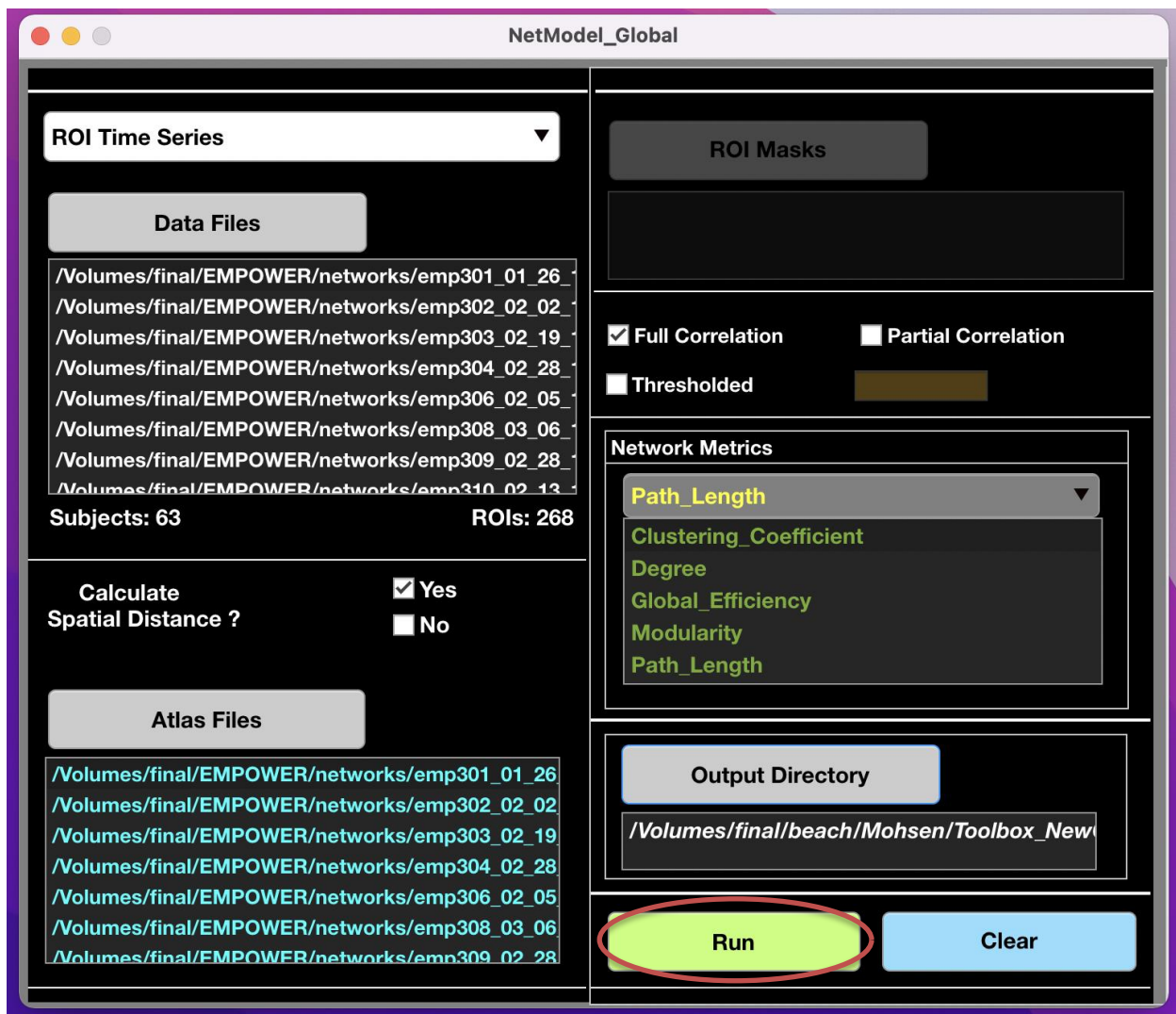


Fig 12. Running network model. Multiple sets of files will be generated and saved in the output directory.

All generated files will be saved in appropriate sub-folders in the output directory (Fig 13). If you use a previously used output directory, a question dialogue pops up asking you whether you want to continue to run or not, because, as mentioned earlier, the older files will be overwritten if you opt to continue.

'Network_Model_Summary.mat' is a matlab file containing a summary of loaded files and selected options of the network model. To see this file, you can type the following code in the command window of Matlab:

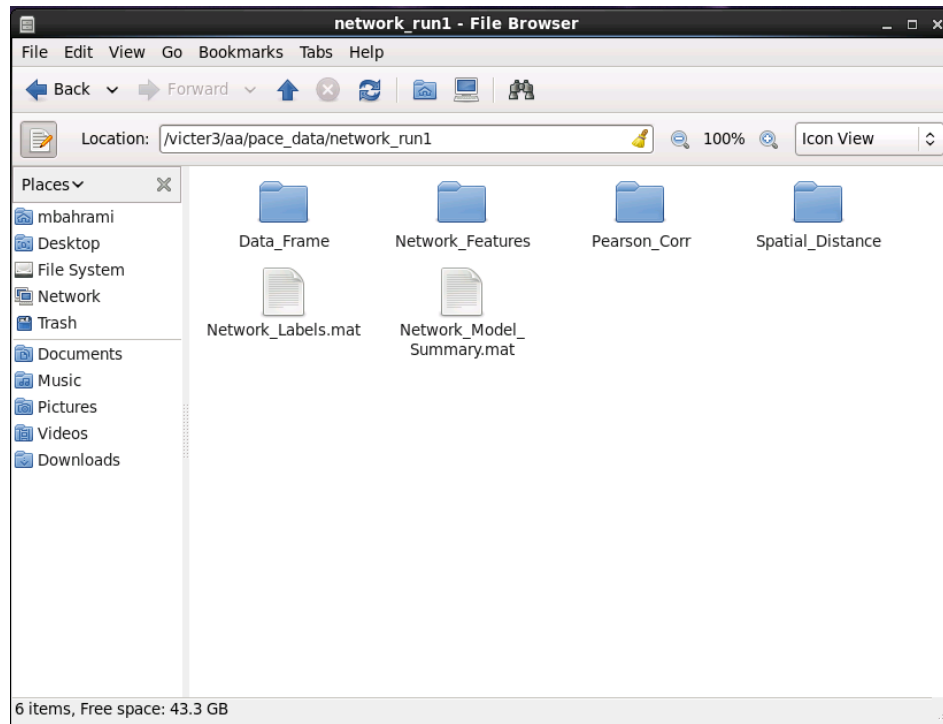


Fig 13. Generated files if running a network model with the options shown in Fig 12. 'Network_Labels.mat' is a cell file that includes the names of the selected network features. This file will be used by the 'StatiMod Global' GUI in the second step.

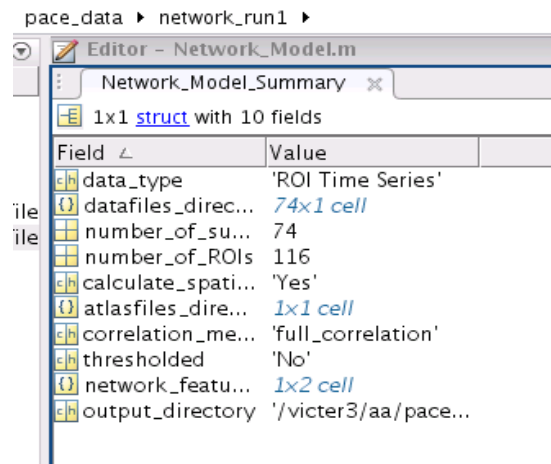
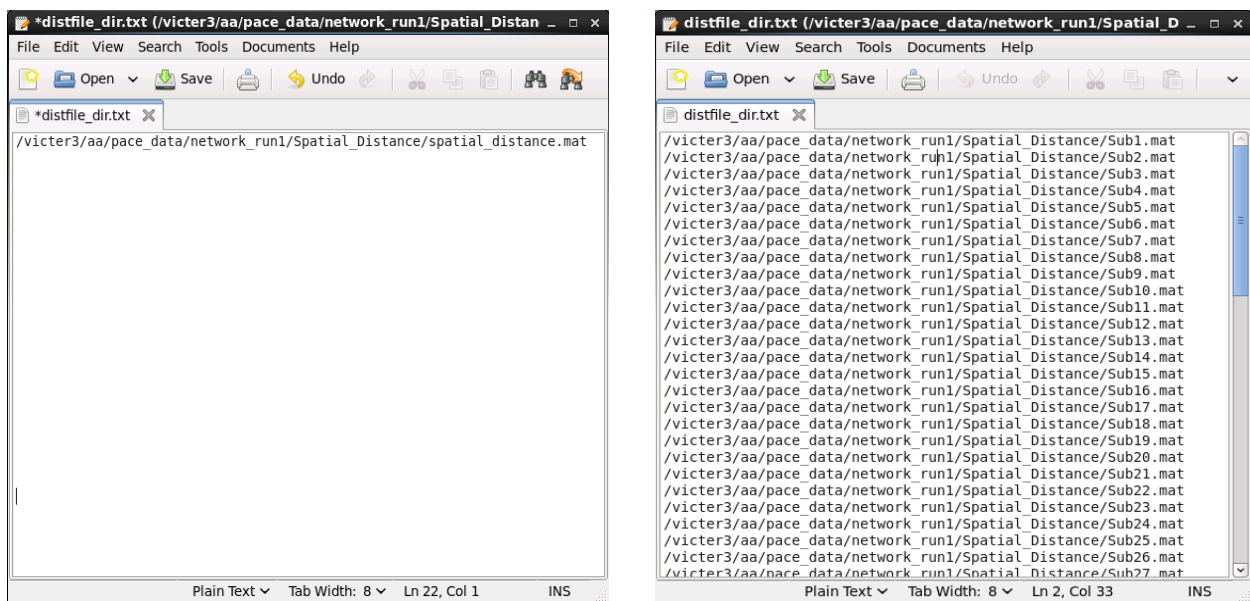


Fig 14. Network Model Summary. This file shows a summary of loaded files and selected options in the 'NetMod Global' GUI. This is useful for future reviews of the model set up.

3.1.12 Adding spatial distance files manually

If you have already computed spatial distance or have access to it, add your distance file(s) after you finish running the model as detailed below:

- Create a folder and name it 'Spatial_Distance' in the 'output directory' (the directory you have selected for your network model).
- Create a text file in this folder (i.e., 'Spatial_Distance' folder), and name it 'distfile_dir'; Then, add the directory of your spatial distance file(s) to this file (Fig 15).



A

B

Fig 15. Adding spatial distance manually. First, create a folder and name it 'Spatial_Distance' in the output directory. Next, create a text file and name it 'distfile_dir' in the 'Spatial_Distance' folder. Then, add the directory of your spatial distance file if it is a single file (A), or add the directories if you want to use a separate distance file for each subject (B).

Note that the distance file should be a symmetric numeric matrix with element a_{ij} (and a_{ji}) representing the spatial distance between regions (nodes) i and j .

3.2 NetMod Regional – Network model for regional network analysis

For regional network analysis, click ‘NetMod Regional’ button in ‘WFU_MMNET’ to start the ‘NetMod Regional’ GUI. This GUI is the same as the ‘NetMod Global’ GUI and should be filled in the exact same way. See section 3.1 for instructions about filling and running this GUI. Note that regional analyses of local brain networks is done in the statistical analysis section, i.e., when using ‘StatMod Regional’ GUI as will be detailed in section 4.

3.3 NetMod Global Multitask – Network model for global multitask network analysis

For multitask network analysis, click ‘NetMod Global Multitask’ button in ‘WFU_MMNET’ to start the ‘NetMod Global Multitask’ GUI. You should see the window shown in Fig 16.

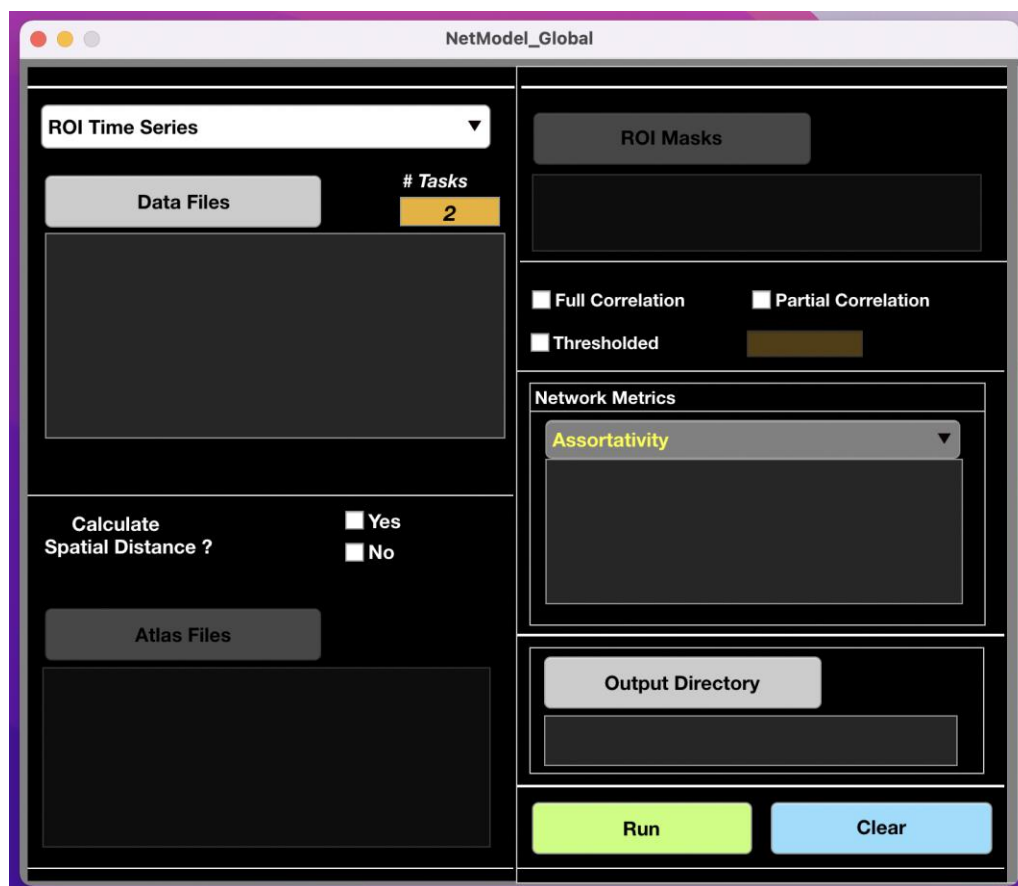


Fig 16. The network model GUI for global multitask network analysis

As shown in Fig 16, this GUI is also similar to the 'NetMod Global' GUI. However, the number of tasks should be specified here. For example, if analyzing resting and 2-back working memory networks, "# tasks" should be set to 2 (this is the default value), and for analyzing resting, 1-back, and 2-back working memory networks, "# tasks" should be set to 3.

Note that for this analysis, each subject must have at least 2 imaging data files. All imaging data from all subjects must be loaded simultaneously via a single list file containing directories of all data files or through directly selecting all data files. For more detail about loading data files see sections 3.1.1 and 3.1.2. All other sections of this GUI should be filled in the same way done for the 'NetMod Global' GUI as well.

3.4 NetMod Global Dynamics – Network model for dynamic global network analysis

For dynamic network analysis, click 'NetMod Global Dynamics' button in 'WFU_MMNET' to start the 'NetMod Global Dynamics' GUI. **Choose 'Connectivity Matrices' from the data type popup menu, and then enter the number of dynamic networks (per subject) by filling the "DynNets #" box before loading dynamic networks (Fig 17).**

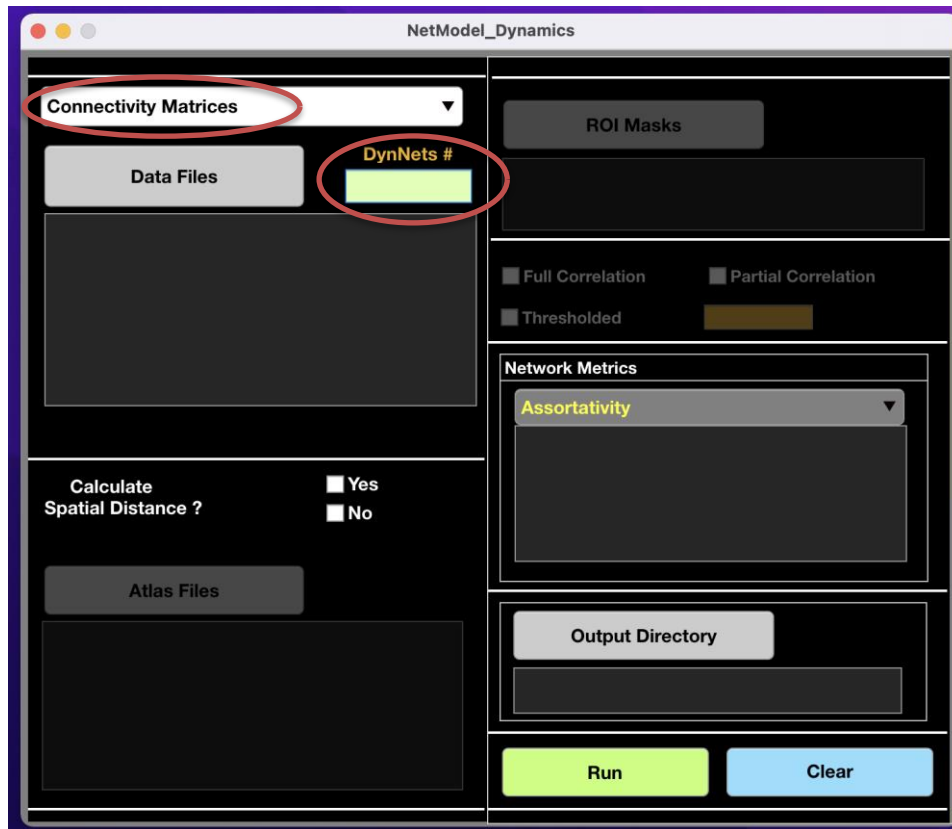


Fig 17. The network model GUI for dynamic global network analysis

For imaging data files, you must load all dynamic networks from all subjects through a single list file containing directories of their data files or directly selecting all data files. Load all required files and select desired options as instructed in section 3.1, and then click 'Run' to generate the data frame. Note that due size consideration, the data frame will be saved as a csv file which may take more time.

4 Statistical Model

As pointed out earlier, modeling is done in two main steps. The first step was described in section 3. In the second step, using the GUIs developed for statistical modes, and generated files from the first step, appropriate modeling files will be generated and mixed models will be fitted to produce the estimates and p-values.

4.1 StatMod Global – Statistical model for global network analysis

Click the ‘StatMod Global’ button in ‘WFU_MMNET’ to start the ‘StatMod Global’ GUI. You should see the windows shown in Fig 18. If you start this GUI from a directory that you ran a network model, the ‘Input Directory’ will be automatically filled (However, you can still change it).

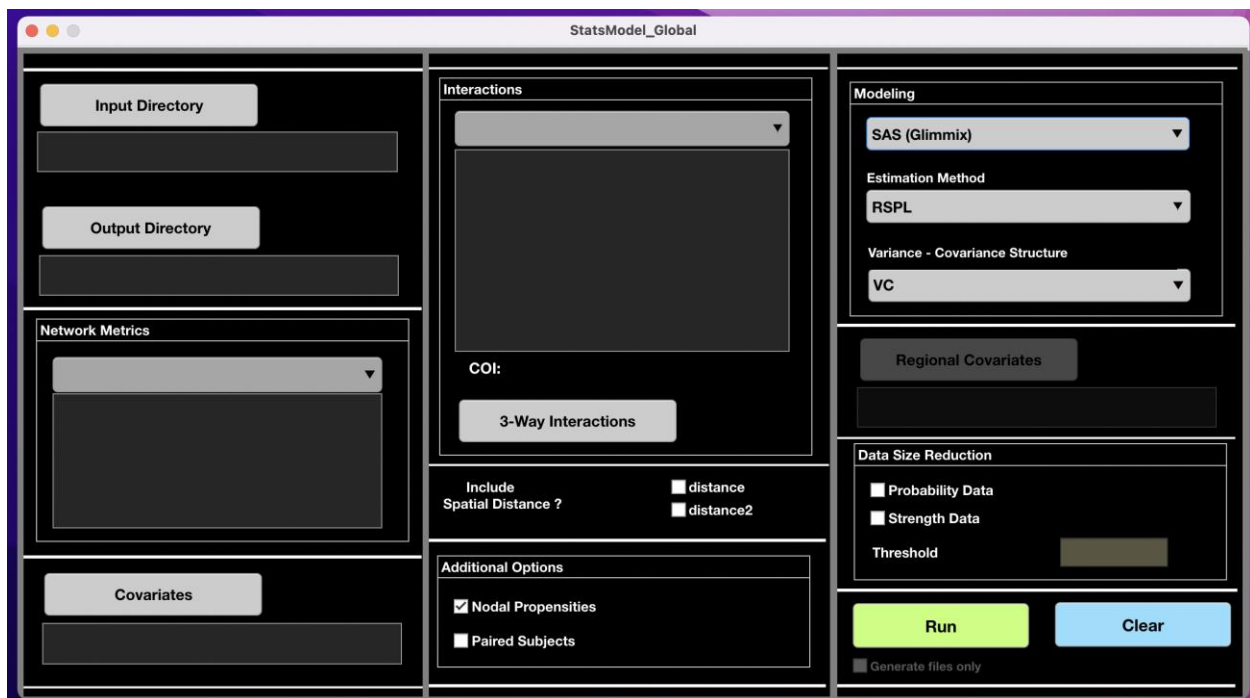


Fig 18 .The statistical model GUI for global network analysis. If you start from a directory that you ran a network model, the ‘Input Directory’ of this GUI will be set to the ‘Output Directory’ of the network model by default. However, you can change the ‘Input Directory’ if you have run several network models.

4.1.1 Input directory

Click 'Input Directory' button and select the input directory. If you start from a directory that you have run a network model, the input directory will be set to the output directory of the 'NetMod Global' GUI or any other network model ran in that directory by default. If you ran multiple network models before starting the 'StatMod Global' GUI, the input directory will be set to the output directory of the last network model. If the input directory was empty (i.e., if you started from a different directory), or if you ran several network models, you can simply select your desired input directory. Note that the input directory must be the output directory of a previously run network model.

After selecting the input directory, the selected network metrics in the network model will appear in the popup dialogue of the 'Network Metrics' panel. Also, checkboxes for 'distance' and 'distance2' would be filled if you added spatial distance in your network model, otherwise, they would be grayed out.

4.1.2 Output directory

Click the 'Output Directory' button and select an output directory. Generated files, including modeling datasets (used in modeling the probability and strength of brain connections), modeling files (modeling options and equations) used by the modeling software to fit the statistical models, modeling results, and a summary file of the statistical model will be stored in this directory. Again, use different output directories for different runs if you wish to keep the results of your runs, because the older files will be overwritten if you use the same output directory. A warning message will be displayed if using a previously used output directory. After selecting the output directory, the 'Network_Model_Summary', which is generated when running the network model, will be copied to the output directory. Although you can see the 'Network_Model_Summary' in the network model output as well, having this file in the statistical model output directory helps to know which network model was used in cases of running several network models. The toolbox doesn't allow using the same directory for the input and output.

4.1.3 Network metrics

Select the network metrics from the popup dialogue (Fig 19). Network metrics shown in the popup dialogue are those that you have selected in setting your network model (see section 3.1.7). Note that by changing the input directory, your network metric choices (i.e., metrics appeared on the popup dialogue) can differ based on the network metrics that you selected in your network models.

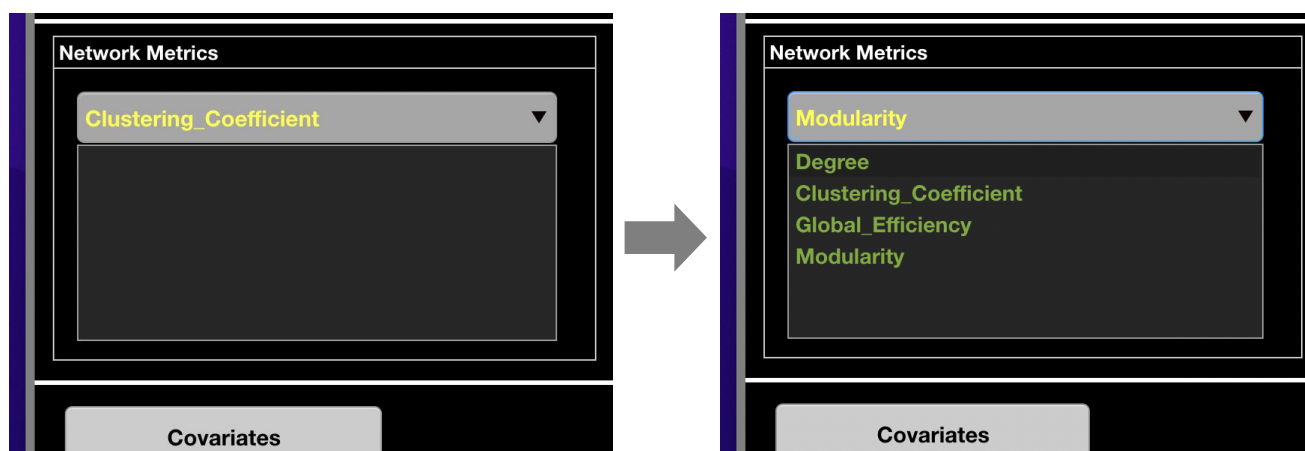


Fig 19. Topological network metrics. Click the popup dialogue and select the network metrics. You can delete the selected metrics by clicking on them in the list box.

You can remove the selected metrics by clicking on them in the list box.

4.1.4 Covariates

Covariates data is a 2d single matrix with rows representing the subjects and columns representing the exogenous variables (i.e., all covariates that **are not** network metrics) (Fig 20). This matrix **must** have row names and column names. The first Column must represent the row names (e.g., subject identifiers), and the second column must be assigned to the main covariate of interest (COI). The COI can be a binary, categorical or continuous variable. In a typical study, the COI is the variable defining study population groups. Continuous or categorical exogenous variables such as age, gender, education,

disease phenotypes, or risk factors can make up the other columns of this matrix. ***If paired subjects are analyzed, the covariates file must have one row for each network data. For example, if analyzing 50 subjects with each subject having 2 networks (e.g., pre-post treatment), the covariate file must have 100 rows. The covariate file must also have a column, named ‘Paired_COI’. This column should be a binary variable with zero for the first half of the covariate data (e.g., the first 50 rows from covariate file in the example above) and one for the second half of the covariate data (e.g., the last 50 rows from the covariate file in the example above). The ‘Paired_COI’ can be either the second column of the covariate file (as the main covariate file) or be a separate column if other desired covariate is of interest for a paired analysis. Figs. 21 and 22 show two examples for accepted covariate files for analyzing paired subjects.***

	<i>Study_Population</i>	<i>BP</i>	<i>Age</i>	<i>Smoking</i>	<i>...</i>
Subj1	0	95	67	No	--
Subj2	0	105	71	Yes	--
....
Subj (N-1)	1	140	69	Yes	--
Subj N	1	120	76	No	--

Fig 20. Covariates data format. Sample covariate data with COI as a binary variable defining two study populations.

	<i>Age</i>	<i>BP</i>	<i>Paired_COI</i>	<i>Smoking</i>	<i>...</i>
Subj1	67	95	0	No	--
...
SubjN	76	105	0	Yes	--
Subj1	67	85	1	No	--
...
SubjN	76	100	1	Yes	--

Fig 21. Covariates data format when paired subjects are analyzed. The main covariate of interest (COI) is Age here. This file has a column, named: “Paired_COI”.

	<i>Paired_COI</i>	<i>BP</i>	<i>Smoking</i>	<i>...</i>
Subj1	0	95	No	--
...
SubjN	0	105	Yes	--
Subj1	1	85	No	--
...
SubjN	1	100	Yes	--

Fig 22. Covariates data format when paired subjects are analyzed with the paired subject covariate as the main covariate of interest.

After appropriately making the covariate file, click 'Covariates' button and load the covariate file. Subjects in this matrix should be arranged in the same order as the loaded imaging files. The toolbox always changes the name of the second column to 'COI' and shows the actual variable name next to the 'COI' in the Interactions panel (Fig 23).

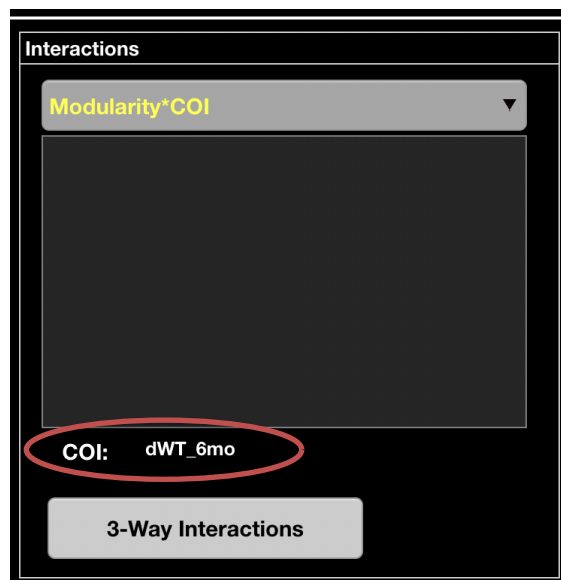


Fig 23. Covariates data file. Click "Covariates" button and load the covariates data matrix. The name of the second column, which must be assigned to the main covariate of interest, will always be changed to 'COI' for subsequent analyses. The actual COI name will be displayed next to COI as shown above.

4.1.5 Interactions

Interactions of the selected network metrics with COI, as well as other two-way interactions will be added to the popup dialogue in the “Interactions” panel after selecting the network metrics and loading the covariates data. Click the popup dialogue and choose your desired two-way interactions (Fig 24). We suggest using interactions of the network metrics with COI.

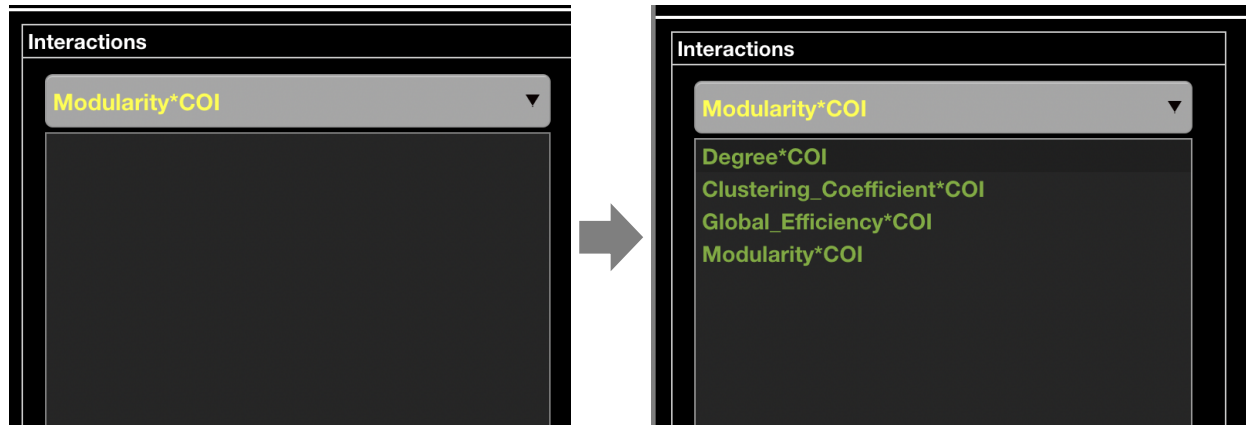


Fig 24. Selecting interaction variables.

You can also choose three-way interactions by clicking ‘3-way Interactions’ button. However, the use of three-way interactions should be carefully considered as interpreting them can get very difficult, and using them might also result in convergence issues in fitting statistical models.

4.1.6 Spatial Distance

If you skipped calculating spatial distance when setting up the network model, and didn’t add it manually either, the ‘distance’ and ‘distance2’ checkboxes will be grayed out by default. However, if spatial distance is available, the checkboxes will be filled by default. But, you can remove them if you would like to exclude spatial distance from your analysis.

4.1.7 Nodal propensities and Paired subjects

Nodal propensities are modeled as random effects by default. However, if you are modeling a dataset with a large number of ROIs (e.g., > 100), including the propensities may result in convergence problems, especially in modeling the probability of brain connections and when paired subjects are analyzed. If this occurs, we suggest removing them.

Fill the 'Paired Subjects' checkbox if you are using paired subjects in your analysis. The loaded covariates file must have a column, named 'Paired_COI' as detailed above (see section 4.1.4). *Also, if the Paired_COI is not the main covariate of interest (i.e., if it is not the second column of the loaded covariates file), additional two-way interactions with the Paired_COI will be added to the popup menu of the Interaction panel to allow using desired two-way interactions. Note that for any three-way interaction, all combinations of the underlying two-way interactions of the variables used in that three-way interaction must be included.*

4.1.8 Modeling

This toolbox uses SAS, R, or Python (depending on software availability) to perform the statistical modeling for this analysis. Thus, you should have one of them (preferably SAS or R) and required packages installed on your operating system before using this toolbox. Note that for 'StatMod Regional', 'StatMod Global Multitask', and 'StatMod Global Dynamics' you can only use SAS.

4.1.9 Selecting modeling software

Click the popup dialogue and select the modeling software. As described in section 2.2.1, you should add the path of the folder including its executable file to 'executable_file_path' as a part of the installation process (before starting the toolbox). The added modeling software will appear with a black color in the popup dialogue in the 'Modeling' panel. You can also add more than one modeling software. However, you only need one.

4.1.10 SAS (Glimmix)

SAS (Glimmix) uses general linear mixed models (GLMMs) to fit statistical models to datasets with correlated samples. We suggest using SAS due to its flexibility and fast computations. Click the popup dialogue and select 'SAS (Glimmix)'. The default is also set to 'SAS (Glimmix)'. Glimmix allows changing many modeling parameters which make it efficient and flexible for different situations. However, for simplicity and user-friendliness we have only provided two main options as the most important ones in our case: 1) Estimation method, and 2) variance-covariance structure for random effects. Additional modeling options may be provided in the future versions.

Click the 'Estimation Method' popup dialogue and select the estimation method. We suggest using the default estimation method (RSPL). RSPL is a pseudo-likelihood technique in which the estimation is based on a residual likelihood approach. Other available estimation methods are detailed in (Schabenberger 2005). Click the 'Variance – Covariance Structure' popup dialogue and select the variance-covariance structure for the random effects. Most applications use either variance-components (VC) or unstructured (UN) structures for the random effects. **VC is a simple diagonal matrix which models a different variance component for each random effect. We strongly suggest using this default option as using more complex structures can result in convergence issues.** In special cases, such as when nodal propensities are not included, you may use an unstructured matrix that is parameterized in terms of its Cholesky parameterization (i.e., using 'CHOL'). Other variance-covariance structures have been detailed in (Schabenberger 2005). Note that for paired subjects, the selected variance-covariance is ignored and an unstructured matrix that is parameterized in terms of its Cholesky parameterization is used to account for the covariance across paired networks as well. This is similar to the covariance structure used for analyzing multitask networks. See the referenced paper for the global multitask network analysis (Simpson, Bahrami et al. 2019).

4.1.11 R (lmerTest)

Click the popup dialogue and select 'R (lmerTest)'. If you use R, you need to install 'lmerTest' package (Kuznetsova, Brockhoff et al. 2015) before using this toolbox. lmerTest uses general linear mixed models as well. It fits a linear mixed model to the strength dataset via a restricted maximum likelihood approach, and fits a generalized linear mixed model to the probability dataset via a maximum likelihood approach. Two variance-covariance structures can be used for modeling the random effects: variance components (VC) and unstructured (UN). We strongly suggest using the 'VC' (it is the default option). In certain cases where nodal propensities are not included, you may use the 'UN' structure.

4.1.12 Python (Statsmodels – GEE)

Click the popup dialogue and select 'Python (Statsmodels - GEE)'. If using Python, you need to install several modules, including: "numpy", "scipy", "pandas", and "statsmodels" before using this toolbox (we suggest not using Python, especially for datasets with a small number of subjects or ROIs.) Unlike SAS (Glimmix) and R (lmerTest) which use GLMMs to fit the models, Python uses generalized estimating equations (GEE). Click on the 'Variance – Covariance Structure' popup dialogue and select the covariance structure. Here, only the correlation structure of a random error term is modeled. Three working correlation structures are provided: 1) Independence, 2) Exchangeable, 3) Autoregressive. We suggest using the Independence structure. This structure assumes equal correlation between observations.

4.1.13 Data Size Reduction

Although we do not suggest reducing the data size by default, modeling large datasets may result in convergence issues or increase the modeling time, especially when using R. Thus, we have provided a k-means clustering data size reduction approach in this toolbox to deal with this problem.

To use this approach, select the data (i.e., probability or strength) by filling its checkbox, and enter the percentage of the data that you wish to **remove**.

4.1.14 Run

After loading all required files and selecting desired options, click the 'Run' button to fit the statistical models. First, modeling datasets, equations, and options are generated and saved under appropriate folders in your selected output directory; then, using these files, the modeling software fits the statistical models, and saves the results in the output directory in the 'Results' folders.

Results for probability and strength of brain connections are saved under the 'Results' folder. The output directory's name will be appended to the names of the results.

4.1.15 Clear

Click 'Clear' button to reset your statistical model.

4.2 StatMod Regional – Statistical model for regional network analysis

For running this GUI, you must have SAS installed on your operating system. R and Python cannot be used at present due to computational considerations. Fill all required fields and load required data files as instructed for the 'StatMod Global' above. However, if analyzing paired subjects, you can only use 'Paired_COI' as your main covariate of interest in your covariates file, i.e., you can only use the format shown in Fig 22.

4.2.1 Regional Covariates

To analyze regional networks (i.e., subnetworks such as DMN), click 'Regional Covariates' button. You can then load your regional covariate file. The regional covariate file must be a table with binary columns separating the desired subnetworks (regional networks) from other brain regions. Regions within and outside the desired subnetwork should be labeled as one and zero, respectively. The column name of this file specifies the names of the desired subnetworks and no row name should be used. Fig 25 shows an example for the regional covariate file.

	A	B	C
1	DMN	ATN	
2	0	0	
3	0	0	
4	1	0	
5	0	0	
6	1	0	
7	1	0	
8	0	1	
9	0	1	
10	0	0	
11	1	0	
12	0	1	
13	1	0	
14	1	0	
15	0	1	
16	0	0	
17	0	0	
18	0	0	

Fig 25. Regional covariate file format. This file defines the DMN and ATN (Attention network) regions in the utilized atlas file. For more detail about generating this file from the ROI masks and atlas files, see the referenced paper for the regional covariate analysis.

After loading the regional covariate file, choose network metric(s) and just one subnetwork that you want to model (Fig 26). Note that for network metrics, you can choose one or multiple. A final question box will be asked to confirm the final selections. After confirming the final selections, you can click 'Run' to generate the required files, equations, and options, and fit the statistical models (Fig 27).

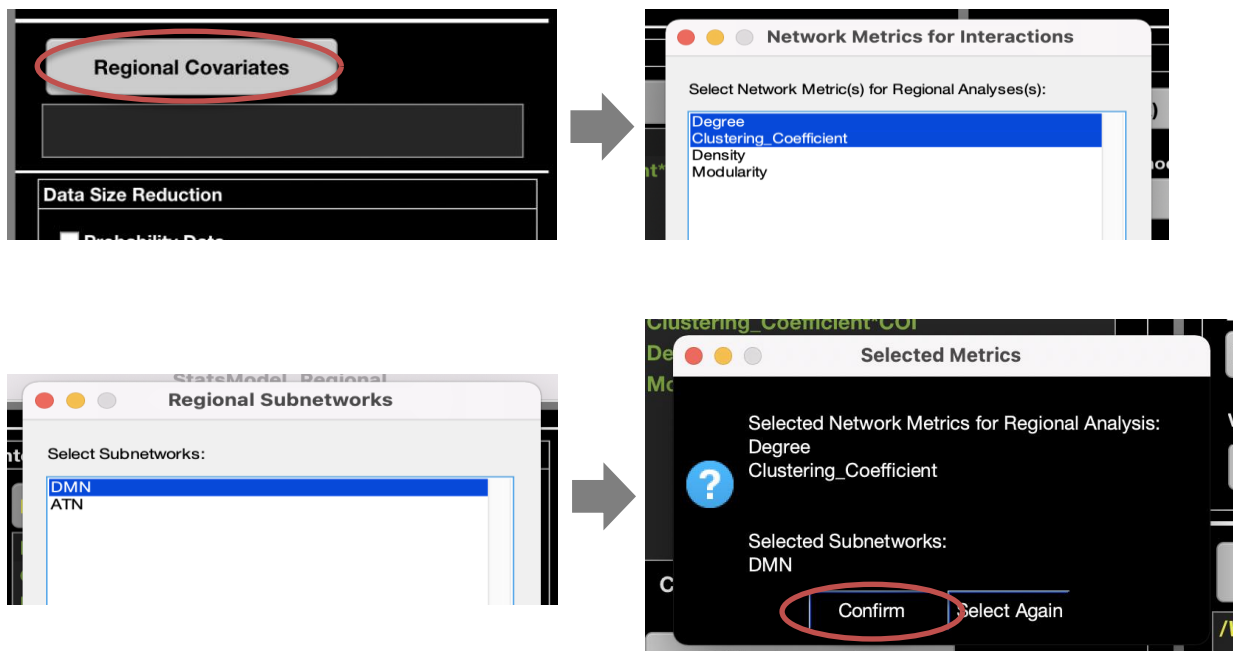


Fig 26. Loading and selecting regional covariate file(s) and network metric(s) for analyzing brain subnetworks.

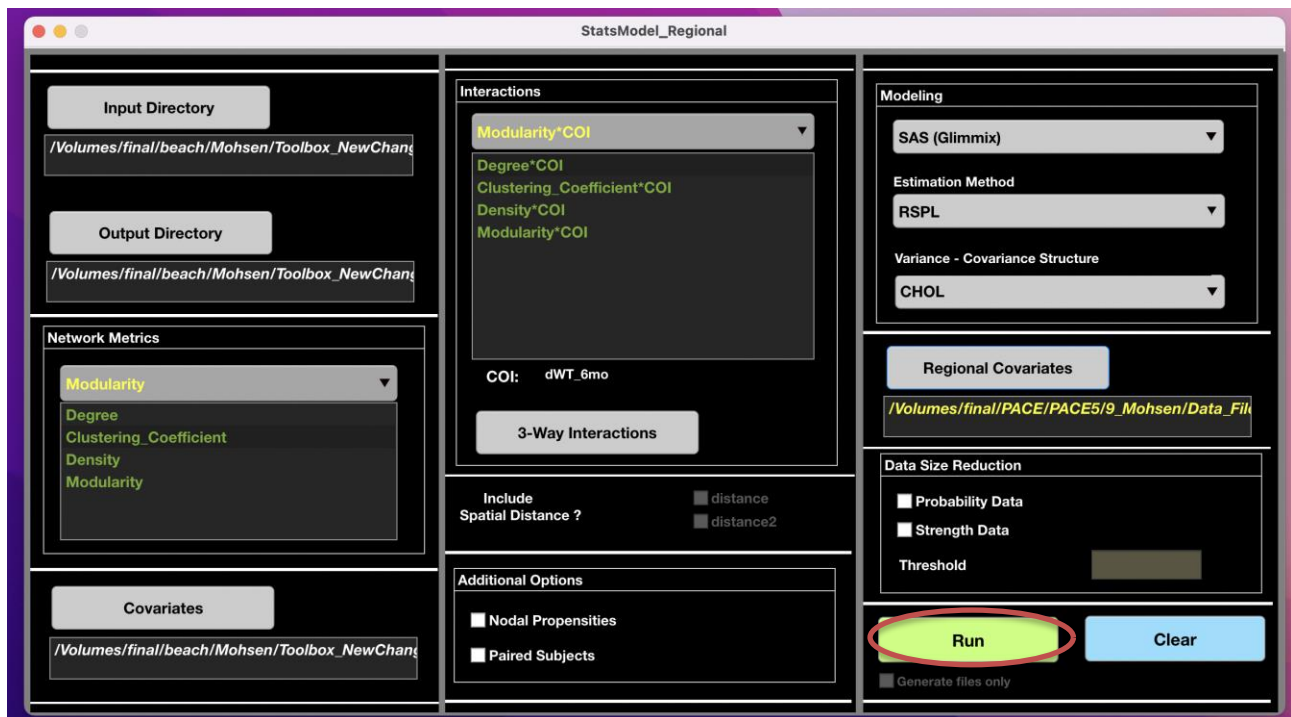


Fig 27. After loading all required files and selecting all required options, click 'Run'.

4.3 StatMod Global Multitask – Statistical model for multitask network analysis

Fill all required filed exactly as instructed for the 'StatMod Global' GUI (section 4.1). ***The covariates file must have one row for each network data. For example, if analyzing 50 subjects with each subject having 3 networks (e.g., resting, 1-back working memory task, and 2-back working memory task), the covariate file must have 150 rows.***

4.4 StatMod Global Dynamics – Statistical model for dynamic network analysis

Fill this GUI as instructed for the 'StatsMod Global' GUI. ***The covariates file must have only one row for each subject (each subject has multiple dynamic networks). Note that the number of dynamic networks (i.e., 'DynNets #' box) will be automatically filled if you start from the same directory that you ran the network***

model. If not, you must first enter the number of dynamic networks. For orthonormal polynomial degrees, if your dataset includes more than 12 dynamic networks per subject, we recommend leaving the default numbers (i.e., mindeg: 3 and maxdeg: 12). Otherwise, choose a degree range from 3 to one less than the number of your dynamic networks. You can then click 'Run' to run the models. Note that the orthonormal polynomials for all degrees that you choose will be added only once and then different degrees will be modeled.

5 Visualization – Regional Analysis

We have implemented visualization tools for the regional analysis. ***To use this tool, you need to first run the regional network ('NetMod RegionaI' GUI) and statistical ('StatMod Regional' GUI) models. You also need to make an excel file that includes the MNI coordinates of the atlas used in your analyses. Name this file 'Atlas_Coors' and add it to the WFU_MMNET folder. We have prpvoided an example Atlas_Coors in the WFU_MMNET folder (Coordinates of Shen Atlas (Shen, Tokoglu et al. 2013)). Replace this with your own atlas coordinates file.***

This tool allows representing the results as line (for binary COI) and surface (for continuous COI – however, the line plots are provided for this as well) plots, and projecting the desired results into the brain space using the BrainNet Viewer toolbox. The node size and color will be modified to illustrate the differences. We have included a version of BrainNet Viewer (Xia, Wang et al. 2013) in WFU_MMNET with modified functions for the node color and size. We will provide more detail in the following.

Click 'Visualization - Regional Anlalysis' in 'WFU_MMNET' to start the GUI, and then load the folder of the already ran statistical model (Fig 28). This folder is the exact same folder that you use as the 'Output Directory' in your 'StatMod RegionaI' GUI.

After loading the stastical model, the network metrics and regioanl covariates will be displayed in the corresponding popup menus. Change the toggle switch to choose either probability or strength models. After selecting the model, select the network metrics and regional subnetworks (Fig 29) and click 'Plot' to draw the 2D line plots (Fig 30).

For a continuous COI, click 'Display 3D' and then click 'Plot' to plot the surface plot. The COI values will be displayed in the corresponding panel. You can see the corresponding line plot of each COI value on the surface plot by filling its checkbox (Fig 31).

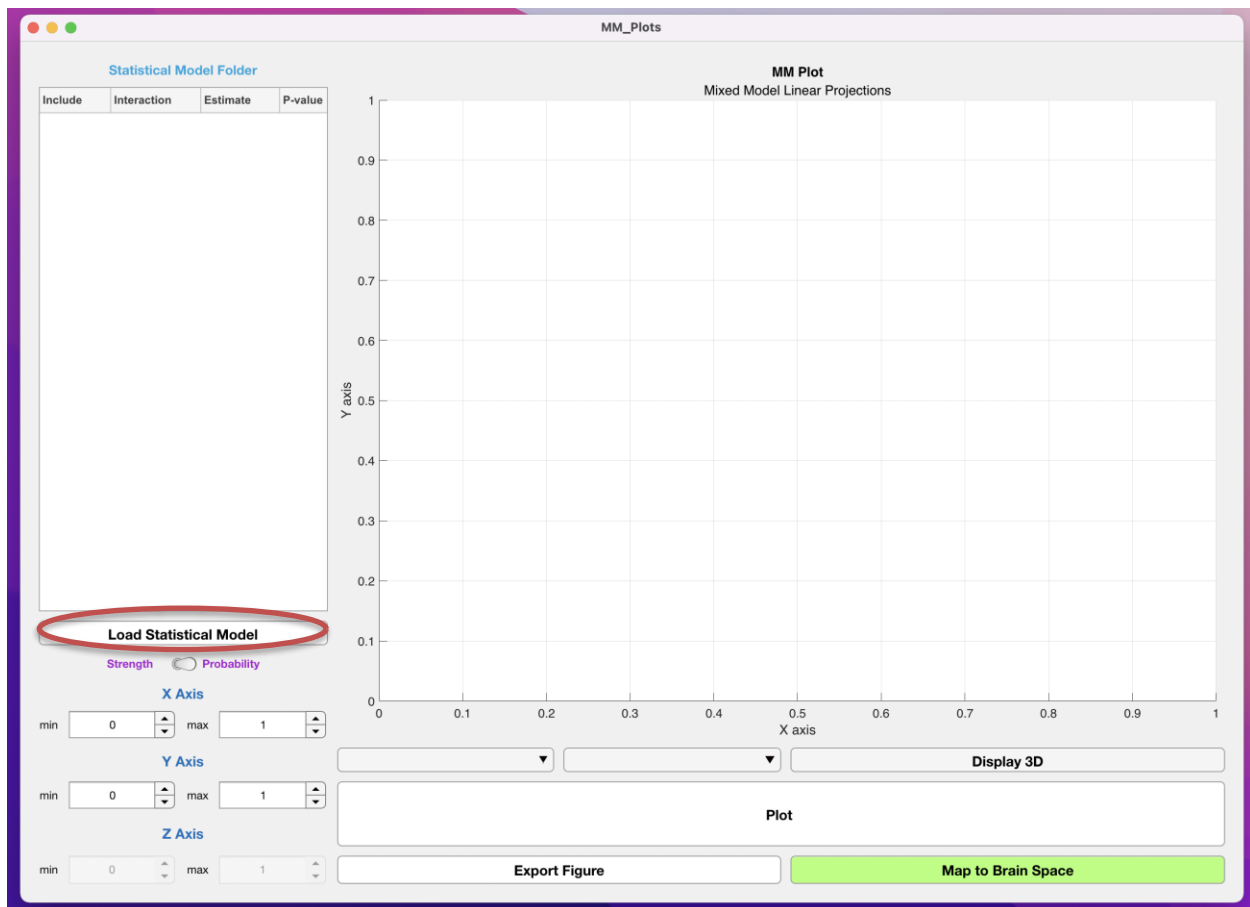


Fig 28. Click 'Load Statistical Model' and load the the directory (folder) that you used as the output directory for your stastical model. The network metrics and regioanl covariates (subnetwork names) will be displayed in the corresponding popup menus after successfully loading the stastical model.

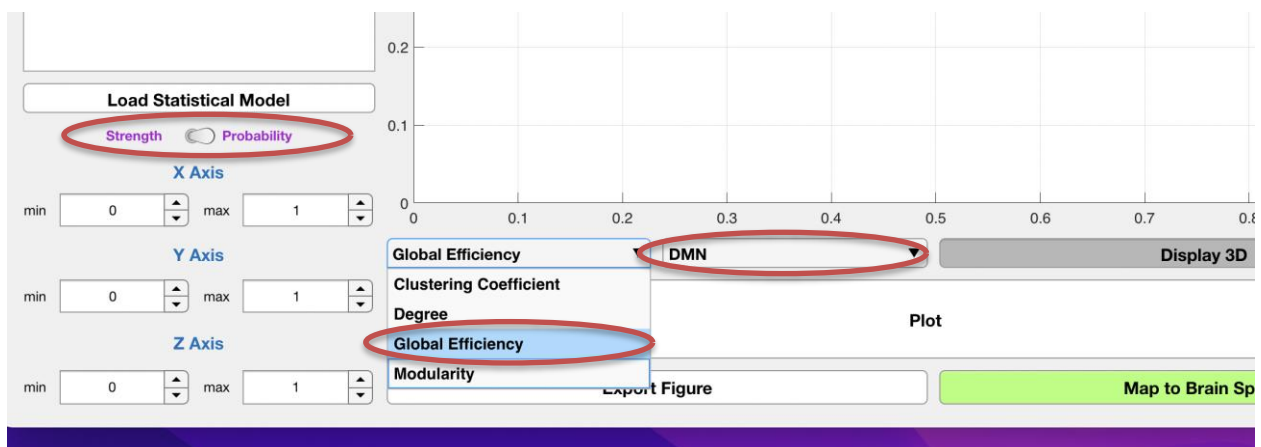


Fig 29. Select model (probability/strength), network metrics and regional covariate.

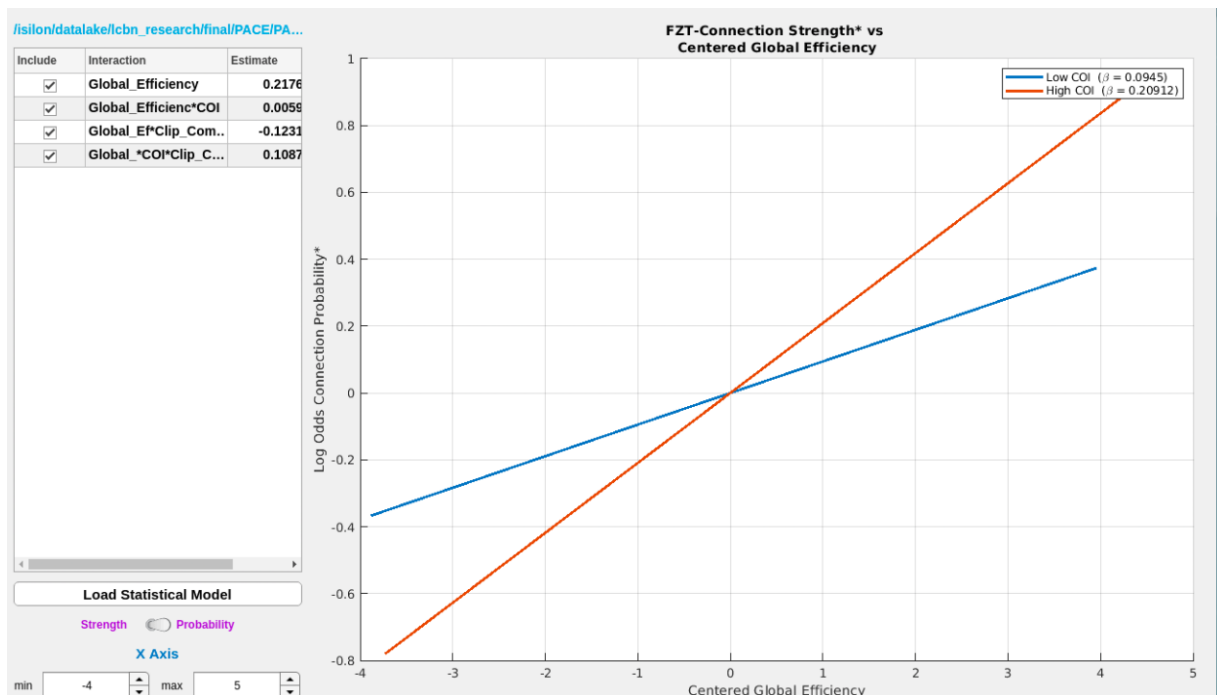


Fig 30. Line plots for representing results (the modifying effects of the COI on the global efficiency – connection probability relationship in the DMN here). COI can be both binary or continuous.

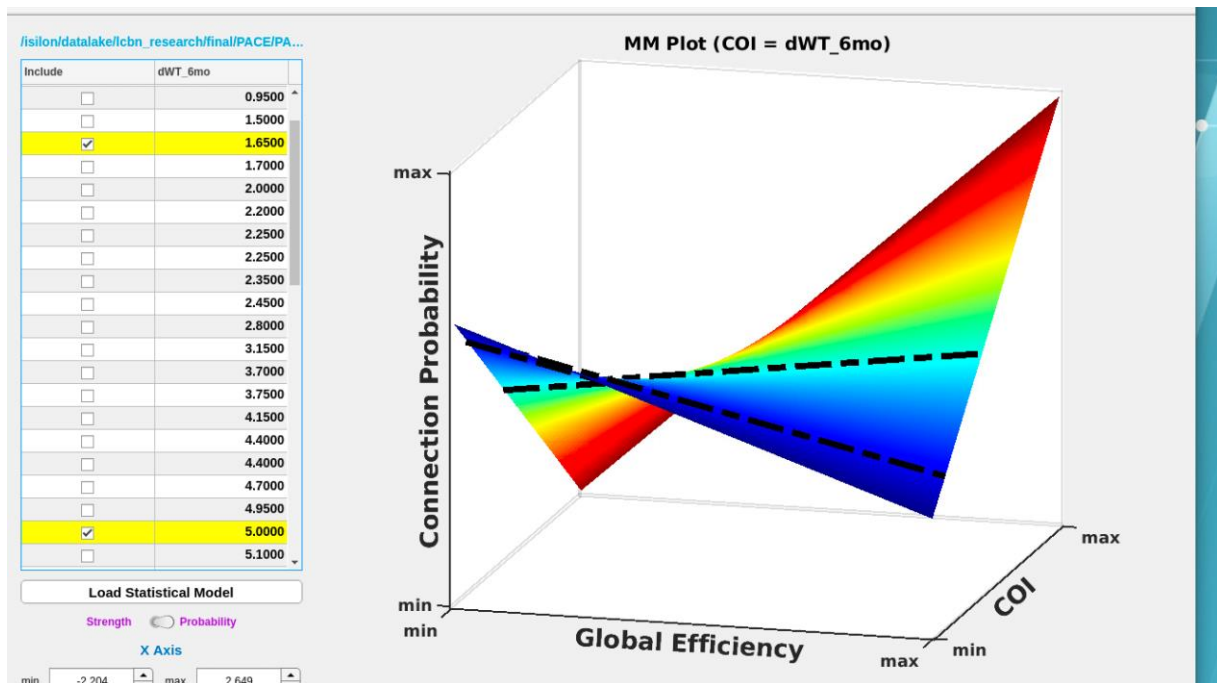


Fig 31. Surface plot and lines associated with the selected COI values on the surface plot.

To map the selected COIs into brain space, click 'Map to Brain Space'. The brain network mappings will be generated and saved in the 'Visualization' folder with the COI values, selected model (probability/strength), and selected network metrics, appended to the subfolder names. You need to choose exactly two COIs for each brain mapping run. If the COI is binary, again click 'Display 3D', select 0 and 1 for the COI, and click 'Map to Brain Space' without clicking the 'Plot' button.

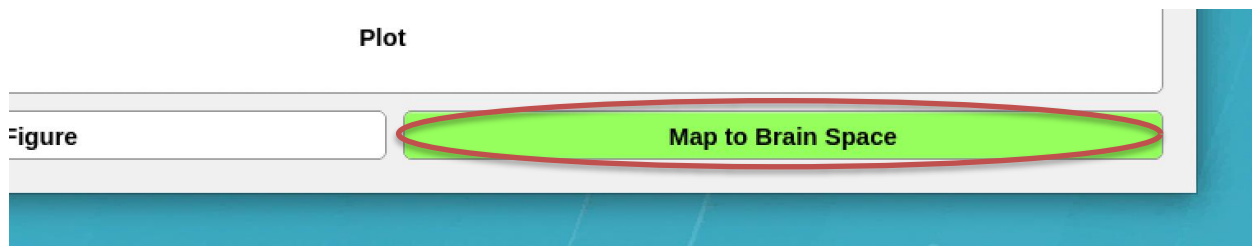


Fig 32. Click 'Map to Brain Space' to map the results for the selected COIs into the brain space.

Node color represents the slopes of selected relationships (i.e., selected network metrics – selected models (probability/strength)) in the selected subnetworks. Nodes with more red and blue colors represent the more positive and negative relationships (i.e., estimated beta values from the result tables), respectively. The node color is scaled across the study populations to better contrast the similarities and differences. The node size represents the actual network metric value for the networks averaged across the study population. **It is very important that you compare the node size, and particularly node color of the generated networks for visualization with the actual results to ensure that there are no discrepancies.**

6 Acknowledgement

A major part of the visualization GUI (MM_Plots) was developed by Joseph Pittman. We appreciate all the effort that he dedicated into advancing this GUI. We also thank the following colleagues for their instructive feedback and their kind help during the development of the WFU_MMNET toolbox:

Robert Lyday, Debra Hege, Fatemeh Mokhtari, Rhiannon Mayhugh, Haley Kirse, and Johnathan Burdette, Laboratory for Complex Brain Networks, Virginia Tech – Wake Forest School of Biomedical Engineering and Sciences, Winston-Salem, NC, USA.

This work was supported by NIBIB K25 EB012236, Wake Forest Clinical and Translational Science Institute (WF CTSI) NCATS UL1TR001420 (Simpson), and NIEHS (R01 ES008739).

References

- Fox, J., M. Friendly and S. Weisberg (2013). "Hypothesis Tests for Multivariate Linear Models Using the car Package." R Journal **5**(1): 39-52.
- Bahrami, M., P. J. Laurienti and S. L. Simpson (2019). "Analysis of brain subnetworks within the context of their whole-brain networks." Human Brain Mapping **40**(17): 5123-5141.
- Kuznetsova, A., P. B. Brockhoff and R. H. B. Christensen (2015). "Package 'lmerTest'." R package version **2**(0).
- Rubinov, M. and O. Sporns (2010). "Complex network measures of brain connectivity: Uses and interpretations." Neuroimage **52**(3): 1059-1069.
- Schabenberger, O. (2005). "Introducing the GLIMMIX procedure for generalized linear mixed models." SUGI 30 Proceedings: 196-130.
- Shen, X., F. Tokoglu, X. Papademetris and R. T. Constable (2013). "Groupwise whole-brain parcellation from resting-state fMRI data for network node identification." Neuroimage **82**: 403-415.
- Simpson, S. L., M. Bahrami and P. J. Laurienti (2019). "A mixed-modeling framework for analyzing multitask whole-brain network data." Network Neuroscience **3**(2): 307-324.
- Simpson, S. L. and P. J. Laurienti (2015). "A two-part mixed-effects modeling framework for analyzing whole-brain network data." Neuroimage **113**: 310-319.
- Xia, M. R., J. H. Wang and Y. He (2013). "BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics." Plos One **8**(7).