NeuroImage 257 (2022) 119321

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Reliability and subject specificity of personalized whole-brain dynamical models

Justin W.M. Domhof[a,b], Simon B. Eickhoff[a,b], Oleksandr V. Popovych[a,b,*]

[a] Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Centre Jülich, Jülich, Germany
[b] Institute for Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

ARTICLE INFO

Keywords:
Brain connectome
Reliability
Resting-state brain dynamics
Subject specificity
Whole-brain model

ABSTRACT

Dynamical whole-brain models were developed to link structural (SC) and functional connectivity (FC) together into one framework. Nowadays, they are used to investigate the dynamical regimes of the brain and how these relate to behavioral, clinical and demographic traits. However, there is no comprehensive investigation on how reliable and subject specific the modeling results are given the variability of the empirical FC. In this study, we show that the parameters of these models can be fitted with a ”poor” to ”good” reliability depending on the exact implementation of the modeling paradigm. We find, as a general rule of thumb, that enhanced model personalization leads to increasingly reliable model parameters. In addition, we observe no clear effect of the model complexity evaluated by separately sampling results for linear, phase oscillator and neural mass network models. In fact, the most complex neural mass model often yields modeling results with ”poor” reliability comparable to the simple linear model, but demonstrates an enhanced subject specificity of the model similarity maps. Subsequently, we show that the FC simulated by these models can outperform the empirical FC in terms of both reliability and subject specificity. For the structure-function relationship, simulated FC of individual subjects may be identified from the correlations with the empirical SC with an accuracy up to 70%, but not vice versa for non-linear models. We sample all our findings for 8 distinct brain parcellations and 6 modeling conditions and show that the parcellation-induced effect is much more pronounced for the modeling results than for the empirical data. In sum, this study provides an exploratory account on the reliability and subject specificity of dynamical whole-brain models and may be relevant for their further development and application. In particular, our findings suggest that the application of the dynamical whole-brain modeling should be tightly connected with an estimate of the reliability of the results.

1. Introduction

The neuroscientific literature generally distinguishes between three types of macroscopic connectivity in the human brain: the structural (SC), functional (FC) and effective connectivity (Deco et al., 2014a; Robinson, 2012). Here, the SC assumes an anatomical viewpoint and reflects how different parts of the brain are connected via axonal projections bundled into white matter fibers (Maier-Hein et al., 2017; Sotiropoulos and Zalesky, 2019; Yeh et al., 2021). FC, on the other hand, uses synchronized co-activations as proxies for stable functional connections (Bolt et al., 2017; Deco et al., 2013; van den Heuvel and Hulshoff Pol, 2010). Finally, effective connectivity considers the causality or the directionality of the information flow between various parts of the brain (Friston, 2011; Gilson et al., 2016; Robinson et al., 2014; Valdes-Sosa et al., 2011). Studies have shown that SC and FC exhibit a complex relationship, which is demonstrated by the relatively low correlations between them, and many approaches have been proposed to infer the FC from the SC or vice versa (Honey et al., 2009; Larson-Prior et al., 2013; Saggio et al., 2016; Suárez et al., 2020; Woolrich and Stephan, 2013).

Dynamical whole-brain models are one of the main methodologies used to link SC and FC together into one comprehensive framework (Breakspear, 2017; Deco et al., 2011; Popovych et al., 2019; Sanz-Leon et al., 2015; Suárez et al., 2020). These models explain an additional amount of variance beyond a direct correlation between SC and FC, and have been used to study the dynamical properties of the resting-state human brain (Deco et al., 2017; Ghosh et al., 2008; Honey et al., 2009). Moreover, the models can be employed to study the mechanisms underlying neurobiological phenomena and neural disorders at a personalized level and suggest an approach for hypothesis testing in silico (Deco et al., 2019; Deco and Kringelbach, 2014; Hahn et al., 2019; Jirsa et al., 2017; Ritter et al., 2013; Zimmermann et al., 2018b).

* Corresponding author.
E-mail address: o.popovych@fz-juelich.de (O.V. Popovych).

https://doi.org/10.1016/j.neuroimage.2022.119321.
Received 8 February 2022; Received in revised form 6 May 2022; Accepted 12 May 2022
Available online 14 May 2022.

Studies have assessed how varying preprocessings of magnetic resonance imaging (MRI) data influence the results of dynamical whole-brain models. They showed that the models are sensitive to variations in the pipelines reconstructing the SC and the FC from diffusion-weighted MRI (dwMRI) and resting-state functional MRI (fMRI) images, respectively (Aquino et al., 2022; Jung et al., 2021), and how the model fitting may depend on the properties of the empirical data used for model derivation and validation (Domhof et al., 2021b; Popovych et al., 2021). Nevertheless, the methodological aspects of the (test-retest) reliability and the subject specificity of the modeling results and their relation to the empirical data have not been extensively investigated so far. In contrast, the reliability of the FC derived from fMRI data, which is used for model validation, has been scrutinized in many studies over a period longer than a decade (Birn et al., 2013; Noble et al., 2019; 2017; Pannunzi et al., 2017; Shehzad et al., 2009; Van Dijk et al., 2010). Its subject specificity reflected by, for example, fingerprinting analysis has received much attention as well (Amico et al., 2018; Finn et al., 2015; Peña Gómez et al., 2018; Li et al., 2021; Sarar et al., 2021; Waller et al., 2017). Also the subject specificity of the empirical structure-function relationship has been considered in the literature (Messé, 2020; Zimmermann et al., 2018a). Hence, comprehensive assessments of the reliability and the subject specificity of the modeling results and their relation with the empirical data are due.

This study therefore critically assesses the reliability and subject specificity of the results of the model validation and their relations with the empirical connectomes across a wide variety of conditions for model construction such as model definition and wielded parcellation. In short, it demonstrates that the results of the model fitting may be more reliable and subject specific than the empirical data. However, our results also show that this finding highly depends on the modeling conditions. In fact, for some of the tested circumstances we found a reliability and subject specificity that are substantially lower for the modeling results than for the empirical data. Moreover, we explicitly show that the models can integrate various types of subject-specific information extracted from empirical data into their output. This makes our study relevant for application, especially, given the current focus on the involvement of dynamical whole-brain models in clinical investigations, for example, in the framework of precision medicine.

2. Materials and methods

In the current study, we assessed the reliability and the subject specificity of the fits of the dynamical whole-brain models to the empirical FC. We first constructed such models on the basis of the *empirical SC* derived from dwMRI data (Fig. 1). Subsequently, we independently fitted them to different realizations of the *empirical FC* (the FC derived from resting-state fMRI data) of individual subjects by optimizing the (global) model parameters through a grid search paradigm (Fig. 1). By doing so, we obtained the optimal model parameters that were used by the models to generate the associated *simulated FCs* that provided the best fits of the model to these separate realizations of empirical FC.

We subsequently calculated the intraclass correlation (ICC) of the individual optimal model parameters as characterizations of their reliability. Additionally, we calculated the same quantity for the individual (undirected) edges of the empirical and fitted simulated FCs, and inspected the distribution of these ICCs across connections to examine the reliabilities of those connectomes. We also computed the *single-modal connectome correlations*, where different realizations of the empirical and the fitted simulated FCs were separately compared with each other for the same subject (within-subject) or different subjects (between-subject) to determine how variable the connectivity patterns are for the same and different subjects (Fig. 1, blue arrows). Furthermore, we determined the within- and between-subject, cross-modal connectome correlations to study how the different types of connectivity related to one another. Here we named the correlations calculated between the empirical SC and both types of FC (empirical and simulated) the *structure-function correlations* (Fig. 1, red arrows), and those computed between the empirical and the simulated FC the *model-fit correlations* (Fig. 1, brown arrows). Finally, we considered all values of similarity (Pearson correlation) between the empirical and simulated FC established at the model validation by the parameter grid search, which are referred to as a similarity map. We examined how the similarity maps relate to one another within and between subjects.

We repeated our calculations using 8 distinct parcellations for the reconstruction of the empirical SC and FC from the MRI data to determine whether a change of brain atlas could critically alter the conclusions. In addition, we repeated our computations for 6 distinct dynamical whole-brain model implementations to investigate whether varying model personalization and model complexity may yield qualitatively different results.

Below we describe the wielded procedures in detail. The code used for the simulation of the brain network dynamics, the analysis and the visualization can be found here: https://jugit.fz-juelich.de/inm7/public/specificity-modeling.

2.1. Empirical connectomes

In this work, we used the empirical connectomes that we have already published elsewhere (Domhof et al., 2021a). This repository contains the empirical SC and FC matrices of 200 healthy, unrelated subjects (96 males, 104 females, aged 28.5 ± 3.5 years) from the Human Connectome Project (HCP) S1200 release dataset (Van Essen et al., 2013; 2012). The local ethics committee of the HCP approved the study, and the informed consent of all subjects was collected. The connectomes were reconstructed for 19 different parcellation schemes, where the original parcellation images were first modified to increase the comparability of results across brain atlases. In particular, the modifications ensured the images only included cortical parcels and were sampled to the MNI152 non-linear template space (Grabner et al., 2006).

We used the empirical connectomes of 8 representative parcellations out of the available 19 brain parcellations in order to put more emphasis on varying the parcellation method rather than the granularity. Table 1 displays the final selection of parcellations. Below we provide a brief explanation on the derivation of the empirical SC and FC from the dwMRI and fMRI data, respectively. For a detailed description of the connectome data, we refer to the data descriptor included in the repository (Domhof et al., 2021a) and to the associated paper (Domhof et al., 2021b).

2.1.1. Empirical structural connectivity

The reconstruction of the empirical SC matrices from dwMRI data was carried out by a workflow developed in-house (Jung et al., 2021). The pipeline can be regarded as a wrapper around functions included in the software packages of ANTs (Tustison et al., 2010), FreeSurfer (Dale et al., 1999), FSL (Jenkinson et al., 2012) and MRtrix3 (Tournier et al., 2019), and is publicly available (https://github.com/inm7/vbc_dwmri). The result of the reconstruction consisted of the empirical SC matrix with the number of streamlines between all pairs of brain regions and the empirical path length (PL) matrix, which included the average lengths of those streamlines. For the details of the reconstruction process, we refer to the above repository hosting the workflow, to the data descriptor of the data repository (Domhof et al., 2021a) and to the associated paper (Domhof et al., 2021b).

In addition to the subjects' own (personalized) empirical SC and PL matrices, we also derived their grand-averages per parcellation, which is a common practice in modeling studies (Aquino et al., 2022; Cabral et al., 2011; Deco et al., 2018b; Donnelly-Kehoe et al., 2019; Iravani et al., 2021; Messé et al., 2014; 2015). However, by a straightforward averaging, the unconnected brain regions may bias the grand-averaged path lengths to lower values. Instead, we considered each edge of the empirical SC and PL matrices separately, and determined the medians

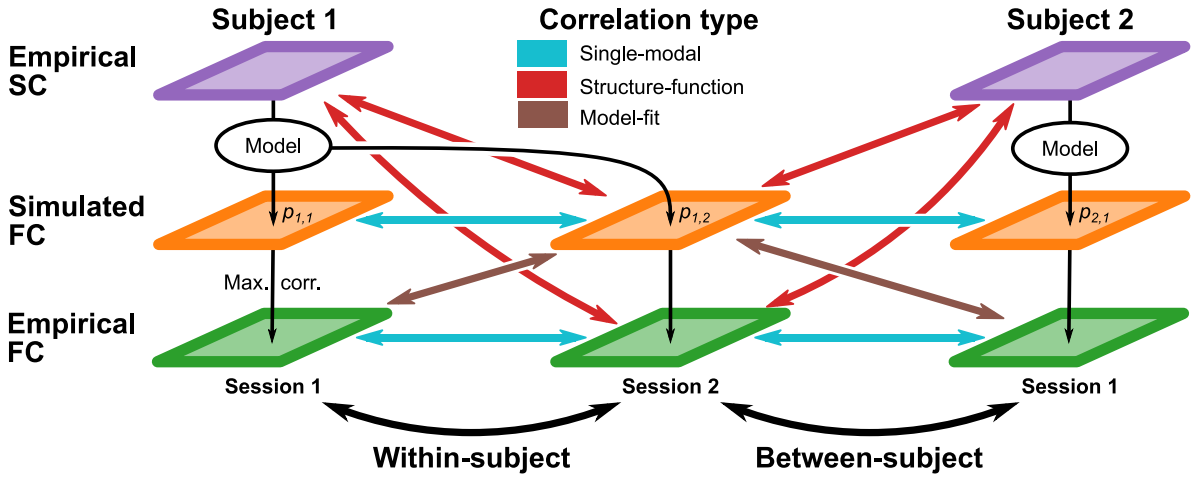

Fig. 1. Schematic illustration of the methodology used in this study. The empirical structural connectivity (SC) and the empirical functional connectivity (FC) were calculated from dwMRI and resting-state fMRI data, respectively. Dynamical whole-brain models were used to sample the simulated FC matrices that replicated each individual empirical FC as close as possible for every fMRI session by using the optimal model parameter configuration $p_{subject,session}$. This particular configuration was obtained by validating the model (fitting simulated FC to empirical FC) using a grid search in the parameter space. Subsequently, the upper triangles of the empirical SC, empirical FC and the corresponding fitted simulated FC matrices were correlated between different resting-state fMRI sessions or subjects to determine their similarities. Here, a distinction was made between three types of correlations. (1) The correlations evaluated between the same type of FC were named single-modal correlations (blue arrows). (2) Cross-modal structure-function correlations (red arrows) were calculated between the empirical SC and the empirical or simulated FC. (3) Cross-modal correlations between the empirical and simulated FC were termed model-fit correlations (brown arrows). All sessions participate in within- and between-subject comparisons, but arrows in the figure are only fully shown for session 2 of subject 1 (center column of simulated and empirical FC). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the connected edges across subjects. This variation from a straightforward averaging does not yield qualitatively different empirical SC matrices, but yields a more accurate estimation of the grand-averaged physical distance the signals have to travel; see supplementary Fig. S1 for illustration.

2.1.2. Empirical functional connectivity

The empirical FC matrices were calculated from the ICA-FIX preprocessed resting-state fMRI data as included in the HCP dataset (Griffanti et al., 2014). First, the mean intensity of the resting-state blood-oxygen-level-dependent (BOLD) signal was calculated across all voxels of a given brain region included in the considered parcellation, which resulted in one BOLD signal time series for each parcel. The resulting BOLD signals were recently published in a separate dataset as well (Domhof et al., 2022). Subsequently, the time series were linearly detrended and z-scored. Eventually, the empirical FC was derived from the time series by calculating the Pearson correlation coefficients across the time series for all pairs of brain regions.

For all considered subjects, the HCP dataset provided 4 resting-state fMRI sessions (left-to-right and right-to-left phase encoding directions scanned on 2 days) comprising 1200 volumes each (TR = 720 ms). We thus calculated 4 different realizations of the empirical FC per subject. These separate instances of the empirical FC for every individual subject enabled us to estimate the reliability of the empirical FC and hence that of the corresponding fitted simulated FC and the respective fitted model parameters as well.

2.2. Simulated functional connectivity

After the acquisition of the empirical connectomes, the simulated FC matrices were generated by dynamical whole-brain models. In these models, the brain was considered to be a network of nodes corresponding to the brain regions included in a particular parcellation. The mean-field activities of the brain regions were subsequently described by models for local dynamics that interact with one another according to the connectivity profile prescribed by the empirical SC. Here, the empirical SC and PL matrices were used to determine the strengths of the network connections and their associated time delays of signal propagation, respectively.

We performed our simulations for 6 different dynamical whole-brain model implementations to study how the distinct facets of *model personalization* and *model complexity* affect the results. The influence of model personalization was studied by considering multiple versions of the Kuramoto model of coupled phase oscillators (Kuramoto, 1984). In particular, the model could be constructed either on the basis of the grand-averaged or the personalized empirical SC, and could be simulated using either group-averaged or subject-specific region-specific oscillation frequencies; see below. Taken together, we considered the Kuramoto model

(1) using averaged empirical SCs and averaged frequencies,
(2) using personalized empirical SCs and averaged frequencies,
(3) using averaged empirical SCs and personalized frequencies and
(4) using personalized empirical SCs and personalized frequencies.

The first and the last modeling conditions define the least and the most personalized models considered, respectively.

The influence of model complexity was studied using three different models with similar personalizations (personalized SC). As the least complex model, we employed (5) a fully linear model. In addition, we used the results of the Kuramoto model that was simulated using the group-averaged frequency profiles (case (2) above) as a moderately complex model. Furthermore, we used (6) a Wilson-Cowan neural mass model (Wilson and Cowan, 1972), which has the most complex model description and implementation among all models wielded in this study. As mentioned above, these models were all constructed on the basis of the personalized empirical SC and PL matrices, but there were no other personalized data included in them.

The non-linear models have two free parameters: the global coupling G and the global delay τ; see below. We simulated the models, which yielded the activity time series of all N brain regions (network nodes), for broad ranges of these parameters sampled from a dense grid in the (G, τ)-parameter space. Subsequently, we derived the simulated FC from the sampled time series via the same procedure that we wielded to construct the empirical FC from the empirical BOLD signals. Con-

versely, the considered linear model had an analytical solution in which the global coupling is the only (relevant) free parameter (Saggio et al., 2016). Hence, we determined the simulated FC of the linear model via that solution for a broad range of global coupling values.

The correspondence between the empirical and the simulated FC matrices was then quantified by comparing both matrices through the Pearson correlation coefficient. Hence, we determined the similarity between the empirical and simulated FC as a function of the model parameters

$$\psi(G,\tau) = \mathrm{corr}[\mathbf{FC}_{emp.}, \mathbf{FC}_{sim.}(G,\tau)]. \tag{1}$$

Here, $\mathbf{FC}_{emp.}$ and $\mathbf{FC}_{sim.}(G,\tau)$ are vectors containing the upper triangular elements of the empirical FC and the simulated FC, respectively. In the case of the linear model, the delay parameter τ was dropped from Eq. (1). The function $\psi(G,\tau)$ is henceforth also referred to as the *similarity mapping* mentioned above.

For every individual realization of the empirical FC, we selected the parameter setting and associated simulated FC that provided the best fit of the model with that particular empirical FC (Fig. 1). In other words, for the four realizations of the empirical FC of every subject (four resting-state fMRI sessions per HCP subject), we acquired the four simulated FCs and the accompanying optimal model parameter settings that resulted in the highest value of the similarity $\psi(G,\tau)$. The actual maximum value of Eq. (1) is subsequently referred to as the *goodness-of-fit*. The selected optimal model parameter configurations and the corresponding fitted simulated FC matrices were subjected to further analyses together with the empirical connectomes. Below we describe the models used in this study in more detail and provide an explanation on their implementation and simulation.

2.2.1. Linear model

As a linear model, we used the well-known Ornstein-Uhlenbeck model approximating the diffusion of noise over the anatomical structure (Galán, 2008). Saggio et al. (2016) demonstrated that this model has an analytical solution, as it gives rise to the covariance matrix $\mathbf{K}$ via the equation

$$\mathbf{K} = -\frac{\sigma_L^2}{2}\left(-\mathbf{I} + G\cdot\overline{\mathbf{SC}}\right)^{-1}. \tag{2}$$

Here, σ_L is the intensity of the noise, G is the global coupling parameter and $\mathbf{I}$ is the identity matrix. Additionally, $\overline{\mathbf{SC}}$ is the personalized empirical SC matrix normalized by the maximum of its eigenvalues. This normalization makes the global coupling parameter range more comparable across subjects and parcellations as it ensures that $G = 1$ coincides with the critical coupling; for $G \geq 1$ the solution loses its stability (Saggio et al., 2016). The derived covariance matrix was converted to a (functional connectivity) correlation matrix by using the definition of the Pearson correlation coefficient $\rho_{X,Y} = \mathrm{cov}_{X,Y}/\left(\sigma_X\sigma_Y\right)$. Evidently, the noise parameter σ_L then becomes irrelevant, and hence the global coupling parameter G remains the only free parameter. We determined the simulated FC for a broad range of values for this parameter to maximize the fit between the simulated and the empirical data as given by Eq. (1) (without delay parameter); see below for details on this variation.

2.2.2. Kuramoto phase oscillator model

The Kuramoto model approximated the phase dynamics of the mean-field activity of brain region $i \in \{1,2,\ldots,N\}$ (N being the number of brain regions in a particular brain atlas), where the corresponding phase $\varphi_i(t)$ was governed by the differential equation

$$\dot{\varphi}_i(t) = 2\pi f_i + \sum_{j=1}^{N} C_{ij}\sin(\varphi_j(t-\tau_{ij}) - \varphi_i(t)) + \sigma_p\nu_i(t). \tag{3}$$

Here, f_i is a region-specific natural frequency, and $\nu_i(t)$ is zero-mean Gaussian white noise with an intensity $\sigma_p = 0.17$. In addition, the individual coupling strengths and delays were characterized by C_{ij} and τ_{ij}, respectively. They were determined from the personalized or grand-averaged empirical SC and PL matrices:

$$C_{ij} = \begin{cases} 0 & \text{if } i = j \\ G\cdot\frac{\mathrm{SC}_{ij}}{N\langle\mathbf{SC}\rangle} & \text{otherwise} \end{cases} \quad \text{and} \quad \tau_{ij} = \begin{cases} 0 & \text{if } i = j \\ \tau\cdot\frac{\mathrm{PL}_{ij}}{\langle\mathbf{PL}\rangle} & \text{otherwise} \end{cases}, \tag{4}$$

where $\langle\cdot\rangle$ returns the mean over all the elements in the matrix, and G and τ are the free parameters of the global coupling and delay scaling the individual coupling strengths and delays, respectively. The normalizations of the empirical SC and PL matrices by their mean values ensured that the coupling and delay parameter values were within similar ranges across subjects and parcellations. Other studies used a similar approach (Deco et al., 2019; 2017).

In our study, $\varphi_i(t)$ directly modeled the ultra-slow phase dynamics of the BOLD signals, which is similar to the paradigm described by Ponce-Alvarez et al. (2015) but different from Messé et al. (2014). The signal $\cos(\varphi_i(t))$ then was considered as a proxy for the simulated BOLD signals, and hence used to construct the simulated FC. We determined the simulated FC for a broad range of the parameters G and τ, which are sampled on a dense grid in the parameter space, to maximize the fit between the simulated and the empirical data; see below for details on this variation.

The oscillation frequencies f_i were determined via spectral density estimations calculated from the empirical BOLD time series. To better estimate the frequency spectra of a given subject, we first concatenated the four z-scored BOLD signals of the individual fMRI data acquisitions. The concatenated signals were analyzed using Welchs method (`welch` function in the SciPy module; Virtanen et al. (2020)), where we used a Hamming window function of 1024 time points and 95% overlap between segments (972 time points). We then used the peak frequencies within the [0.01, 0.10] Hz frequency range; see supplementary Fig. S2 for the distributions of the frequencies across all regions and subjects that were obtained by following this procedure for each individual parcellation. We added Gaussian white noise with zero mean and 0.002 Hz standard deviation to make the peak frequencies more heterogeneous and to avoid duplicate frequencies due to discretization of the frequency values. Following this approach, a vector of frequencies was obtained for each subject separately reflecting the peak BOLD frequencies of the N individual brain regions. Two considered versions of the phase oscillator model used these subject-specific frequencies. We also repeated our calculations while using the same group-averaged, region-specific frequencies for all subjects. These frequencies were calculated as the median frequencies of the brain regions across subjects and correspond to two other considered versions of the phase oscillator model.

2.2.3. Neural mass model

We used a neural mass model similar to the one used by Deco et al. (2009), which was an adaptation of the model described by Wilson and Cowan (1972). The activity of brain region i was modeled by pooling the activities of the excitatory and inhibitory neurons in that region together into the variables $E_i(t)$ and $I_i(t)$, respectively. The temporal dynamics of these activities were governed by the equations

$$\mu_E\dot{E}_i(t) = -E_i(t) + \kappa S\left(\sum_{j=1}^{N} C_{ij}E_j(t-\tau_{ij}) - c_{EI}I_i(t) + I_b\right) + \sigma_n\nu_i(t) \quad \text{and} \tag{5}$$

$$\mu_I\dot{I}_i(t) = -I_i(t) + \kappa S\left(c_{IE}E_i(t)\right) + \sigma_n\nu_i(t), \tag{6}$$

where $\mu_E = \mu_I = 20$ ms represented the decay time constants of the excitatory and inhibitory activity, respectively. The same independent Gaussian white noise with a mean of zero and an intensity of $\sigma_n = 0.002$ was received by both neuronal populations. $c_{EI} = 1.5$ and $c_{IE} = 0.6$ scaled the inhibition of the excitatory neurons by the inhibitory population and the excitation of the inhibitory neurons by the excitatory pool, respectively. Parameter $\kappa = (1+\exp(\lambda\gamma))/\exp(\lambda\gamma)$ scaled the sigmoid function

$$S(x) = \frac{1}{1+\exp(-\lambda(x-\gamma))} - \frac{1}{1+\exp(\lambda\gamma)} \tag{7}$$

so that $\kappa S(x) = 1$ as $x \to \infty$. Here, $\lambda = 20.0$ and $\gamma = 0.3$ were the parameters determining the width and the position of the inflexion point of $S(x)$, respectively. Finally, $I_b = 0.10$ was a constant external input arriving at the excitatory population, and C_{ij} and τ_{ij} were the individual coupling strengths and delays, respectively. Different from the Kuramoto model (Eq. 4), they were only derived from the personalized (hence not grand-averaged) empirical SC and PL matrices via

$$C_{ij} = \begin{cases} c_{EE} & \text{if } i = j \\ G \cdot \frac{\mathrm{SC}_{ij}}{N\langle \mathbf{SC} \rangle} & \text{otherwise} \end{cases} \quad \text{and} \quad \tau_{ij} = \begin{cases} 0 & \text{if } i = j \\ \tau \cdot \frac{\mathrm{PL}_{ij}}{\langle \mathbf{PL} \rangle} & \text{otherwise} \end{cases}. \tag{8}$$

In this equation, the parameter $c_{EE} = 1.0$ regulated the self-excitation of the excitatory neurons, and G and τ are the global coupling and delay parameters, respectively. These were considered as the free parameters of the model and required optimization; see below.

The model exhibited limit-cycle oscillatory behavior in the alpha frequency band when the brain regions were coupled in a network by a sufficiently large coupling parameter $G > 0$, and remained at a low activity state when the network was disconnected ($G = 0$). The modeled oscillations had alpha-band frequencies on purpose: Alpha oscillations have been associated with BOLD responses (Mayhew et al., 2013), and they dominate in human resting-state EEG (Fraga González et al., 2018; Spitoni et al., 2013).

The activities of the two neuron populations were sampled by simulating the model. However, as the fluctuations in the modeled neuronal activity took place on a much shorter time scale (~10 Hz) than the BOLD dynamics (<0.1 Hz), the simulated time series cannot be compared directly with the empirical BOLD signals. Instead, a Balloon-Windkessel model (Friston et al., 2003) was employed to convert the activities of the excitatory population to BOLD-like responses which were then used to construct the simulated FC matrix.

2.2.4. Model implementation and simulation

The models were implemented using the Python (Python Software Foundation, https://www.python.org) and C++ (Standard C++ Foundation, https://isocpp.org) programming languages, where we also made use of the SciPy (Virtanen et al., 2020) and NumPy (van der Walt et al., 2011) modules for Python. The extensive computations required to evaluate the model simulations and their subsequent analyses were performed on the JURECA high-performance computing cluster (Jülich Supercomputing Centre, 2018). The temporal integrations of the phase oscillator, neural mass and Balloon-Windkessel models were implemented according to Heun's method.

The linear model only required optimization of the global coupling parameter. This parameter was varied using the collection of global coupling values described by

$$G \in \{0.0005, 0.0010, 0.0015, \ldots, 1.0000\}. \tag{9}$$

Because the model had an analytical solution, the correlation matrix could directly be calculated from the empirical SC matrix using Eq. (2), and no computationally intensive model simulations were needed for this model.

We maximized the correspondence between the empirical and simulated FC for both the phase oscillator and the neural mass models by evaluating a dense grid search of 64 × 48 different parameter values for the global coupling G and delay τ, respectively. The phase oscillator model was simulated for the collection of global parameter values described by

$$G \in \{0.000, 0.015, 0.030, \ldots, 0.945\} \quad \text{and} \tag{10}$$

$$\tau \in \{0\ \mathrm{s}, 1\ \mathrm{s}, 2\ \mathrm{s}, \ldots, 47\ \mathrm{s}\}. \tag{11}$$

We simulated 70 minutes of phase dynamics in steps of 60 ms, and the first 10 minutes were disregarded as transient. When considering the neural mass model, the dense grid corresponded to all combinations between the collections of global coupling and delay values described by

$$G \in \{0.000, 0.018, 0.036, \ldots, 1.134\} \quad \text{and} \tag{12}$$

$$\tau \in \{0.0\ \mathrm{ms}, 1.5\ \mathrm{ms}, 3.0\ \mathrm{ms}, \ldots, 70.5\ \mathrm{ms}\}. \tag{13}$$

Also the configuration of the temporal integration was different for this model. For every parameter setting, 510 s of network activity were simulated in steps of 2 ms, and we omitted the first 150 s. These diverging simulation conditions were adapted to the alpha-frequency and ultra-slow time scales of the neural mass and the phase oscillator model, respectively.

The simulations above were performed individually for each combination of the 200 subjects, the 8 parcellations listed in Table 1 and the 6 model implementations; see above. These simulation conditions accumulated to over 15M model simulations used for the model validation (fitting) against empirical data on a dense parameter grid. Out of these simulations several optimal parameter settings of the closest correspondence between the simulated and empirical data were selected for further analysis of reliability and subject specificity: 4 (fMRI sessions) × 8 (parcellations) × 6 (models) × 200 (subjects) = 38,400 parameter points and the respective simulated FCs generated by the models for these parameters.

2.3. Reliability and subject specificity

As mentioned above, for every parcellation we acquired the empirical SC and FC of $S = 200$ individuals, where $M = 4$ different realizations of the empirical FC were available for each subject. Furthermore, after the simulations of a given model we additionally had to our disposal the $200 \times 4 = 800$ optimal model parameter configurations and the associated simulated FC matrices that provided the best replications of the individual empirical FC matrices. We subsequently performed additional analyses to evaluate the reliability and the subject specificity of the empirical data and the modeling results. We performed the analyses independently for each combination of the 8 parcellations listed in Table 1 and the 6 model implementations described in "Simulated functional connectivity" to estimate their influence on the results.

2.3.1. Intraclass correlation

We first used the intraclass correlation (ICC) to characterize the reliability of the model parameters of the global coupling and delay as well as the connectomes. In the latter case, the ICCs were calculated for the weights (correlation coefficients) of every $N(N-1)/2$ undirected edges of the functional connectomes (empirical and simulated). The calculated ICC reflects the between-subject variance of these quantities relative to the total variance (between- and within-subject), and was given by the following expression (Chen et al., 2018; Liljequist et al., 2019; Noble et al., 2019; Shrout and Fleiss, 1979):

$$ICC = \frac{\sigma^2_{subject}}{\sigma^2_{subject} + \sigma^2_{\epsilon}}. \tag{14}$$

Here, $\sigma^2_{subject}$ is the variance of the considered quantity (parameter or connectome edge weight) that is related to the variance among the subjects, and σ^2_{ϵ} is the residual variance induced by the different fMRI acquisitions; see "Empirical functional connectivity". Such an implementation of the ICC has been recommended for the case when no convincing argument can be made that the residual noise contains additional consistent effects (Chen et al., 2018; Noble et al., 2019). We wielded the equations proposed by Liljequist et al. (2019) in order to calculate the ICC directly from the data.

2.3.2. Connectome correlations

We also examined the single- and cross-modal connectome correlations within and between subjects. Here, we first vectorized the off-diagonal upper triangles of the individual connectivity matrices corresponding to all subjects and realizations (according to the different fMRI

Table 1
Names and abbreviations of the brain parcellations used in this study together with the number of cortical parcels and associated publications. The top and bottom blocks correspond to parcellations derived using data reflecting structural and functional brain organization, respectively.

Name (abbreviation)	Parcels	References
Desikan-Killiany (DK)	70	(Desikan et al., 2006)
von Economo-Koskinas (EK)	86	(von Economo and Koskinas, 1925; Scholtens et al., 2018)
AAL (version 2) (AAL)	92	(Rolls et al., 2015; Tzourio-Mazoyer et al., 2002)
Harvard-Oxford (HO)	96	(Desikan et al., 2006; Frazier et al., 2005; Goldstein et al., 2007; Makris et al., 2006)
Shen 2013 (Shen)	79	(Shen et al., 2013)
Schaefer (Sch.)	100	(Schaefer et al., 2018)
MIST (MIST)	103	(Urchs et al., 2019)
Craddock (CD)	108	(Craddock et al., 2012)

Table 2
Number of distinct values comprising each type of connectome correlation for S subjects that each have M distinct empirical FC realizations.

	Between-subject	Within-subject
Single-modal	$S \cdot (S-1) \cdot M^2/2$	$S \cdot M \cdot (M-1)/2$
Structure-function	$S \cdot (S-1) \cdot M$	$S \cdot M$
Model-fit	$S \cdot (S-1) \cdot M^2$	$S \cdot M \cdot (M-1)$

sessions) of the FC. Subsequently, we calculated the Pearson correlation coefficients between the resulting vectors, where we distinguished between three types of correlations (Fig. 1). The first type is the *single-modal correlations* comprising the correlations between FCs of the same modality, i.e. empirical FC vs. empirical FC or simulated FC vs. simulated FC (Fig. 1, blue arrows). The second type is the *structure-function correlations*, where the cross-modal correlations of the empirical SC with the empirical or the simulated FC were calculated (Fig. 1, red arrows). When a model was constructed on the basis of a grand-averaged SC, the structure-function correlations nevertheless involved the correlations between the empirical or simulated FC and the personalized SC matrix of the subject to compare with the personalized simulations. The third type is the *model-fit correlations* consisting of the correlations between the empirical and the simulated FC (Fig. 1, brown arrows). The calculated correlations quantified the extent to which the connectomes of the same or different modalities had similar patterns for the same or different subjects.

On top of these three different types of correlations, we distinguished between within- and between-subject correlations. Here, the *between-subject correlations* included all correlations calculated between two different subjects (Fig. 1). In addition, the *within-subject correlations* included the correlations computed between the connectomes of the same subject (Fig. 1). However, the correlations calculated for the same subject and the same FC realization (fMRI session) equal one in the case of the single-modal correlations, and they correspond to the goodness-of-fit values for the model-fit correlations which means that they are maximized and may thus bias the results; see the section "Simulated functional connectivity" above. They were therefore omitted from the analyses. Table 2 clarifies how many distinct values each type of correlation comprised. The within- and between-subject correlations were used to subsequently characterize the reliability and the subject specificity of the (cross-modal) connectome correlations.

Within-subject correlations

The models were fitted to the empirical data by maximizing the similarity between the connectivity patterns of the empirical and simulated FC (Eq. 1). We therefore investigated the reliability of the empirical and simulated FCs, that is, the reproducibility of the connectivity patterns for the same subject. The approach based on the calculation of the ICC (Eq. 14) quantified the reliability of each individual FC edge in isolation, but did not reflect whether the entire patterns of the functional connections were congruent. Such a reliability of the connectome patterns was characterized in this study by the within-subject single-modal connectome correlations (Fig. 1). For the empirical FC, these correlation coefficients reflected how similar the connectivity patterns were to one another when the fMRI data used for their construction were sampled for the same subject but on different days or with different phase encodings. Analogously, for the simulated FC, these correlations characterized the replicability of the simulated connectome under (potential) variations of the empirical FC. By comparing the replicability of the empirical and simulated FCs, we may evaluate whether the considered simulation condition (model, parcellation, etc.) led to an intra-subject variability of the simulated FC that is either enhanced or reduced relative to that of the empirical FC.

Specificity index

As mentioned above, the within-subject, single-modal correlations characterized whether model fits are realized through converging connectivity patterns of simulated FC. However, these patterns may be more similar in general, that is, also across different subjects. We therefore calculated the specificity index $Specificity$, where the mean between-subject correlation $\overline{Corr}_{between}$ was subtracted from the mean within-subject correlation $\overline{Corr}_{within}$

$$Specificity = \overline{Corr}_{within} - \overline{Corr}_{between}, \tag{15}$$

which is similar to the approach of Amico et al. (2018) and Zimmermann et al. (2018a). The specificity index reflects whether connectomes are indeed reproduced better (more similar to each other) within than between subjects and can be used to quantify the subject specificity. In practice, it fluctuates around zero when the considered type of correlation is not subject specific, and is (significantly) larger than zero when it is subject specific.

To assess the variations in this specificity index, we bootstrapped both mean correlations 50,000 times. Here, one bootstrap involved the resampling of the vectors containing all within- and between-subject correlations with replacement and the subsequent calculation of the means from the resampled vectors. The specificity index was then calculated for each bootstrap so that its 95% confidence interval could be constructed. If the lower bound of this interval was larger than zero, the within-subject correlations were significantly larger than the between-subject ones, and the considered relation was considered significantly subject specific. We performed this analysis separately for the single-modal, structure-function and model-fit correlations.

Connectome fingerprinting

We also adapted the fingerprinting analysis from Finn et al. (2015) to provide an additional measure for the subject specificity (or subject identifiability). The rationale behind this analysis is that a connectome is subject specific if a single subject can be identified from the full cohort on the basis of the connectome (cross-modal) correlations. For one particular connectivity matrix, we first evaluated either the single-modal, structure-function or model-fit correlations. Subsequently, we determined whether the maximum of these correlations involved a within-subject or a between-subject correlation, which implied a correct and false identification of the subject, respectively. By repeating this procedure for all connectivity matrices of that modality, we could determine the portion of correct identifications or *fingerprinting accuracy*. In addi-

tion to the fingerprinting accuracy, we calculated the *fingerprinting confidence*. Here, we first determined which subject provided the next highest correlation coefficient for each identification attempt. Subsequently, we subtracted these correlation coefficients from the maxima. Finally, we calculated the fingerprinting confidence by averaging these differences across all identification attempts. The fingerprinting confidence thus characterizes how dissimilar the next closest connectomes are to the identified connectivity matrices. In other words, larger fingerprinting confidences indicate facilitation of (correct) subject identification.

When using the structure-function correlations, a subject could be identified by the strongest correspondence between a given (empirical or simulated) FC and all empirical SC with known subjects. However, the analysis could also be performed using the opposite directionality, i.e. comparing one empirical SC with the empirical or simulated FC matrices of all subjects. Analogously, model-fit correlations were used to identify subjects by correlating one empirical FC with all simulated FC or by correlating one simulated FC with all empirical FC.

2.3.3. Inter- and intra-individual correspondences of the similarity maps

We also investigated how the similarity maps (Eq. 1) calculated between the empirical and simulated FC during the parameter grid search may relate to one another within and between subjects. In other words, we investigated how strongly these mappings change across subjects and across different empirical FC realizations of the same subject. For this analysis, we simply calculated the within- and between-subject correlations of these maps across all tested parameter settings, and inspected their distributions. Furthermore, we calculated the specificity indices and fingerprinting accuracies corresponding to these correlations analogous to the single-modal connectome correlations; see above.

3. Results

In this study, we used the empirical SC and FC matrices of 200 healthy subjects that were constructed on the basis of the 8 parcellations listed in Table 1. The empirical SCs were then used to construct dynamical whole-brain models that were based on the 6 distinct model implementations described in ”Simulated functional connectivity” (Materials and Methods). We optimized the free model parameters so that the similarity between the simulated FC and the empirical FC (Eq. 1) was maximized. Examples of this similarity as a function of the model parameters are shown in supplementary Figs. S3-S10, which are examples of similarity maps. These similarity maps provide information that may help the interpretation of our other findings; see below. For instance, the maps can have multiple regions of high similarity within the parameter space, in particular, when the neural mass model is considered. The latter indicates that the global optimum may be unstable, which could considerably impact the reliability of the fit of the model to the empirical FC.

This maximization procedure was performed individually for each combination of subject, empirical FC (4 realizations per subject), model implementation and parcellation. The corresponding goodness-of-fit values are shown in supplementary Fig. S11 for every combination of parcellation and model individually. In addition, the distributions of the optimal model parameter settings are shown in supplementary Figs. S12-S19. Subsequently, we investigated the reliability and the subject specificity of the empirical data and the modeling results by performing the analyses described in ”Reliability and subject specificity” in Materials and Methods.

3.1. Reliability of model parameters

We first investigated the reliability of the optimal model parameters by examining the distributions of their absolute differences between different subjects (inter-subject) and between different empirical FC realizations of the same subject (intra-subject). These distributions often appeared to be shifted closer to zero when the differences were calculated within subjects than between subjects (Fig. 2A-B; Fig. 3A-B). This might be an initial indication that the parameter variability between subjects is larger than the one within subjects. We further quantified this observation by computing the ICCs (Eq. 14) reflecting the variance between subjects relative to the total variance of the fitted model parameters. The results showed that the reliability of the coupling and delay parameters could range from ”poor” to ”good” depending on the atlas and model implementation (Fig. 2C-D; Fig. 3C-D).

We draw a specific attention to the positive influence of model personalization on the reliability of the fitted model parameters: Simulating the phase oscillator model using subject-specific frequency profiles yielded higher reliability than using group-averaged frequency profiles practically irrespective of whether the group-averaged or personalized SC was used (Fig. 2C-D, green vs. red and orange). However, when considering the phase oscillator model simulated using group-averaged frequencies, the model parameters were also fitted with higher reliability when the personalized instead of the group-averaged SCs were used for model construction (Fig. 2C-D, dark vs. light green). Hence, model personalization appears to promote the reliability of the model fit to the empirical data.

More complex models seemed to yield a reliability of the model parameters that was less variant across parcellations, and higher model complexity was not immediately more reliable at the same level of personalization (Fig. 3C-D). In addition, the linear model fitted the coupling parameter with higher reliability than the non-linear models in most cases (Fig. 3C-D). We however verified whether this could be explained by the absence of the signal latency in the network of the linear model. Hence, we considered the non-linear models with zero global delay $\tau = 0$ in Eq. (4) and Eq. (8). Subsequently, we determined the optimal coupling parameter values under this constraint and calculated their ICCs. The results of this investigation confirmed that model complexity did not exert an influence on the reliability of the coupling parameter in isolation that was consistent across parcellations (supplementary Fig. S20). Hence, the model complexity per se does not seem to systematically influence the reliability of the fitted model parameters.

We checked whether our results critically depended on the choice of the intraclass correlation for the characterization of the reliability of the optimal model parameters. For this investigation, we calculated the (non-parametric) test-retest Spearman correlation coefficient of the optimal model parameters. The results showed a strong covariation across parcellations and model implementations (supplementary Fig. S21; see figure caption for specifics), which indicated that our results did not qualitatively depend on the intraclass correlation as the reliability measure.

Taken together, these findings demonstrate that whole-brain dynamical models can be fitted to the empirical FC with a ”poor” to ”good” reliability depending on the implementation of the modeling paradigm. Furthermore, we explicitly demonstrated the positive influence of the model personalization on the reliability of the fitted model parameters. Moreover, higher model complexity reduces the parcellation-induced variations in the reliability of the optimal model parameters, but it cannot credibly be associated with systematic tendencies (enhancement or reduction) of the parameters’ reliability.

3.2. Reliability of functional connectivity edges

We also examined the reliability of the empirical and the simulated FC. First, we calculated the ICCs of all empirical and simulated FC edges (individual functional connections between brain regions) and inspected their distributions. The ICCs of the empirical functional connections remained approximately at the same (”fair”) level across parcellations (Fig. 2E, gray). In contrast, the edge reliability of the simulated functional connectomes varied considerably across parcellations, and ranged from ”poor” to ”good” (Fig. 2E; Fig. 3E). These findings indicate that the reliability of the empirical FC is rather stable across

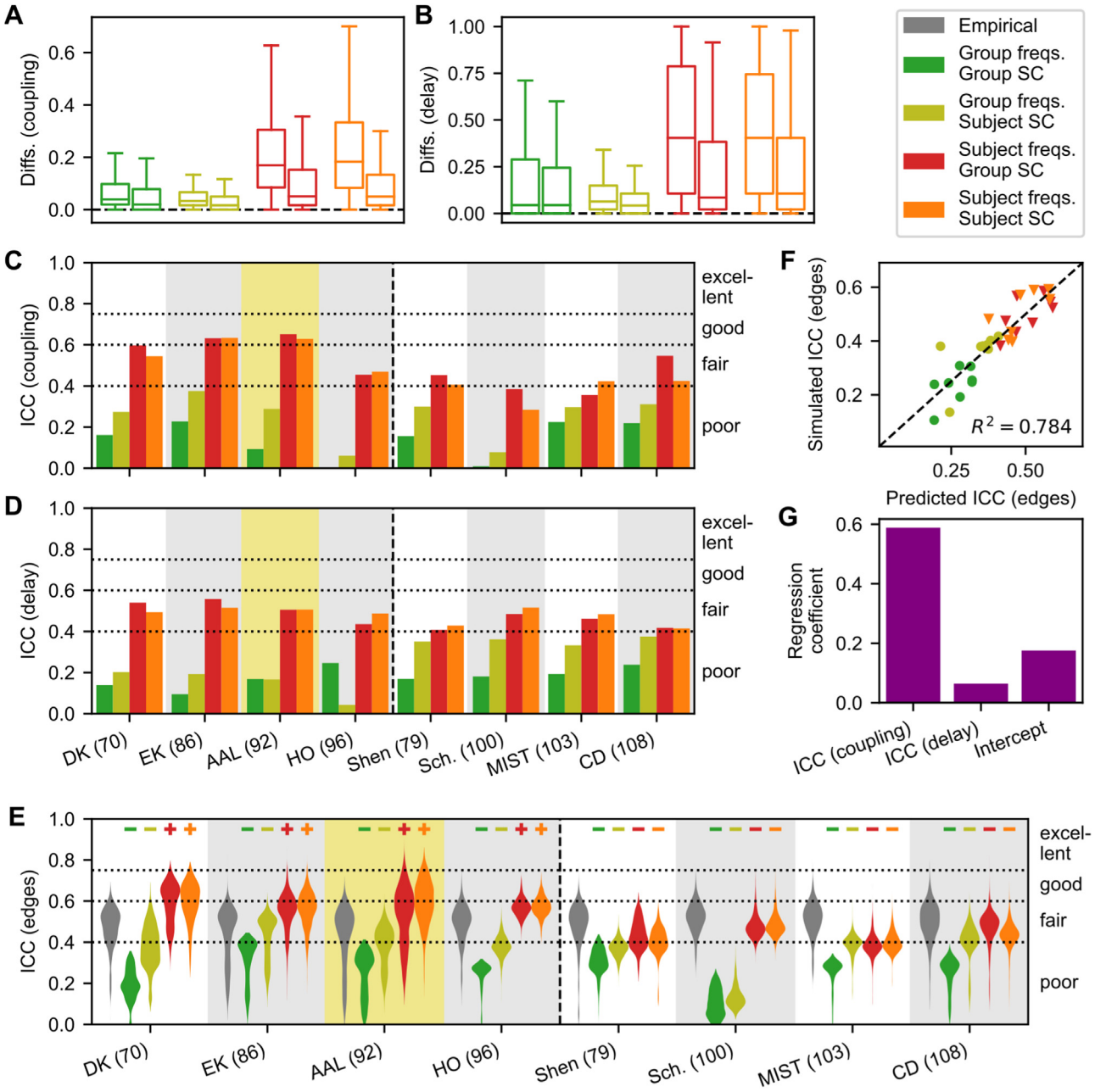

Fig. 2. Reliability of modeling results for varying personalization of the phase oscillator model. **(A-B)** Absolute differences (diffs.) of (A) the optimal coupling parameters and (B) the optimal delay parameters for the AAL atlas, which is also highlighted in yellow in panels C to E. The extent of the model personalization as given by the combinations of the subject-specific or group-averaged natural frequencies (freqs.) and SC is reflected by color as indicated in the legend. Left and right boxes of the same color in the plots correspond to inter- and intra-individual differences per model implementation, respectively. The differences were normalized using the maximum across all (inter- and intra-subject) parameter differences per model. **(C-D)** Intraclass correlations (ICCs; Eq. 14) of (C) the coupling parameters and (D) the delay parameters for all the atlases considered in this study (Table 1). The labels "poor", "fair", "good" and "excellent" correspond to those proposed by Cicchetti and Sparrow (1981). The vertical dashed black lines separate the brain atlases constructed on the basis of structural data (left blocks) from those based on functional data (right blocks). **(E)** Distributions of the ICCs of individual functional connections, edges of the empirical (gray) and simulated functional connectome for all the atlases considered in this study. Plus and minus signs at the top of the plot signify significantly increased and decreased ICC distributions for the respective simulated FC with respect to the one for the empirical FC, respectively ($p < 0.05$, two-sided Wilcoxon paired signed-rank test, Bonferroni corrected). **(F)** Scatter plot of the intraclass correlations (ICCs) calculated from and averaged across simulated FC edges (simulated ICC, vertical axes) and their predicted values obtained from a linear regression with the ICCs of the model parameters (predicted ICC, horizontal axes). The plotted symbols represent parcellations and models as indicated in the legend. The dashed black line represents $x = y$ for comparison. **(G)** Regression coefficients corresponding to the results shown in panel F. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

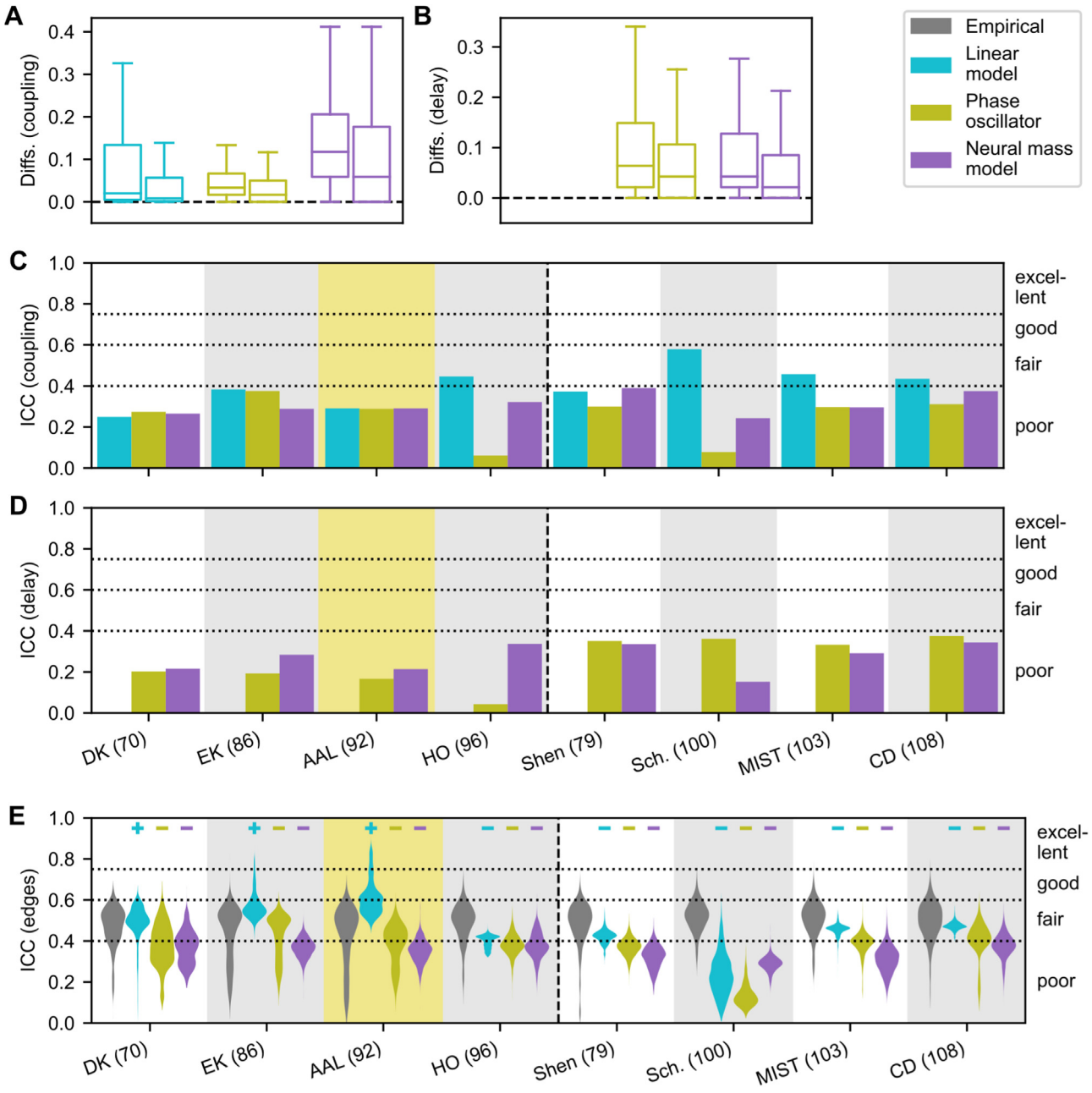

Fig. 3. Reliability of modeling results for varying model complexity with similar model personalization. The model implementations (2), (5) and (6) are considered; see "Simulated functional connectivity" in Materials and Methods. Here (5) the linear model (blue) corresponds to a low complexity, (2) the phase oscillator model (green) to a moderate complexity and (6) the neural mass model (purple) to a high complexity. **(A-B)** Absolute differences (diffs.) of (A) the optimal coupling parameters and (B) the optimal delay parameters for the AAL atlas, which is also highlighted in yellow in panels C to E. Left and right boxes of the same color in the plots correspond to inter- and intra-individual differences per model implementation, respectively. The differences were normalized using the maximum across all (inter- and intra-subject) parameter differences per model. The results of the delay parameter are not shown for the linear model as this model did not include this parameter (Eq. 2). **(C-D)** Intraclass correlations (ICCs; Eq. 14) of (C) the coupling parameters and (D) the delay parameters for all the atlases considered in this study (Table 1). The labels "poor", "fair", "good" and "excellent" correspond to those proposed by Cicchetti and Sparrow (1981). The vertical dashed black lines separate the brain atlases constructed on the basis of structural data (left blocks) from those based on functional data (right blocks). **(E)** Distributions of the ICCs of the empirical (gray) and simulated functional connectome edges for all the atlases considered in this study. Plus and minus signs at the top of the plot signify significantly increased and decreased ICC distributions for the respective simulated FC with respect to the one for the empirical FC, respectively ($p < 0.05$, two-sided Wilcoxon paired signed-rank test, Bonferroni corrected). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parcellations, while that of the simulated FC is more sensitive to the utilized brain parcellations.

Additionally, we found that the conclusions derived for the reliability of the fitted model parameters (Fig. 2C-D; Fig. 3C-D) can also be confirmed for the FC edges. Indeed, model personalization often led to an increase in the reliability of the connectome edges (Fig. 2E). The simulated FCs of the phase oscillator model using subject-specific regional frequencies clearly exceeded the empirical FC in terms of edge reliability for all considered structurally-derived atlases irrespective of the personalization of the empirical SC, and reached the "good" level (Fig. 2E,

red and orange). We regressed the ICCs of the optimal model parameters with the mean ICCs of all simulated FC edges and found that this linear regression could explain 78% of the variance for the phase oscillator model (Fig. 2F). The regression coefficients demonstrated a high contribution of the global coupling to the reliability of the simulated FC edges as compared to the optimal delay parameter (Fig. 2G). Moreover, the positive intercept of the regression also indicated that the reliability of the simulated FC edges was enhanced by the personalized phase oscillator model as compared to that of the optimal model parameters (Fig. 2G).

In addition, as for the reliability of the optimal model parameters, we again observed that enhanced model complexity (e.g., for the neural mass model) led to a reliability of the simulated FC edges that varied less across parcellations (Fig. 3E). On the other hand, the reliability of the simulated FC generated by the linear model varied considerably and significantly exceeded that of the empirical FC for the Desikan-Killiany, von Economo-Koskinas and AAL atlases (Fig. 3E). We also compared the reliability (ICC values) of the optimal model parameters and the simulated FC edges of the models in Fig. 3 by linear regression and found no consistent, strong dependencies across models (not shown).

In sum, the variation of the reliability of the simulated FC edges with respect to the extent of model personalization and the brain parcellation well agrees with that of the optimal model parameters (Fig. 2F). Evidently, the observed relationship between these two types of model reliability implies that they exhibit similar variations across model personalization, where an enhancement of the latter led to an improvement of the reliability of the simulated FC, possibly outperforming that of the empirical FC. Furthermore, higher model complexity has a positive effect on the consistency of the reliability of the simulated FC across parcellations, but it may not contribute to an enhancement of the reliability of individual FC edges, and a simple linear model sometimes performed better (Fig. 3E).

We also checked whether the different ICCs (Fig. 2C-E; Fig. 3C-E) could be related to the goodness-of-fits of the model to the empirical data. With regard to the reliability of the optimal model parameters, the tested regressions varied considerably across models in terms of the relationship (positive vs. negative) as well as variance explained (supplementary Fig. S22A-G). Hence, the goodness-of-fit is not a good predictor for the reliability of the model parameters when considering a particular modeling condition (parcellation and model implementation) at random. Additionally, the edge-wise reliability of the simulated FC exhibited a positive correlation with the quality of the model fit for all considered models, though also with varying fractions of explained variance (supplementary Fig. S22H-K).

3.3. Reliability and subject specificity of functional connectivity patterns

Several modeling conditions yielded simulated FCs with edges' reliability being lower than for the empirical FC (Fig. 2E; Fig. 3E). We therefore investigated whether the whole connectivity patterns of the simulated FCs were nevertheless similar given that they were fitted to different empirical FCs of the same subject. For this purpose, we evaluated the within-subject, single-modal connectome correlations (Fig. 1, blue arrows). A considerable number of the modeling conditions and subjects yielded simulated FC matrices that had strongly diverging connectivity motifs, which is reflected by low intra-subject correlations between simulated FCs compared to the empirical FCs (Fig. 4A-B). In particular, increased model complexity led to more dissimilar simulated FCs for most parcellations, especially, for the functionally-derived parcellations, where strong bimodalities were elicited in the within-subject, single-modal correlation distributions (Fig. 4B). Enhancing the model personalization could reduce or smooth these bimodalities (Fig. 4A). Hence, the fit of the model to the empirical data could on average enhance the within-subject variability of the empirical FC depending on the particular combination of model implementation and parcellation (Fig. 4A-B, minus signs on top of the plots).

We also checked whether the within-subject, single-modal correlations could be related to the goodness-of-fit of the model to the empirical data. Here, we found strong relationships between these two quantities for the non-linear models (supplementary Fig. S22M-O), but not for the linear model (supplementary Fig. S22L). This indicates that the FC patterns simulated by non-linear models can exhibit higher within-subject similarity when they are fitted better to the empirical FC, but also that such a relationship is not evident for the linear model.

For most modeling conditions, we nonetheless observed that the simulated FC matrices had connectivity patterns that were significantly more similar to one another than those of the empirical FC (Fig. 4A-B, plus signs on top of the plots). We investigated whether these enhancements of the within-subject, single-modal correlations were realized by a general increase in the similarity of the connectivity patterns, that is, both within and between subjects. We therefore calculated the (single-modal) specificity indices (Eq. 15) and fingerprinting accuracies to determine the gain of the within- relative to the between-subject, single-modal correlations. We observed that enhanced model personalization induced a clear increase in the specificity index and the fingerprinting accuracy, where both these specificity measures could exceed those of the empirical FC (Fig. 4C-E). Fingerprinting confidences were also increased for stronger model personalization and apparently exceeded those of the empirical FC (supplementary Fig. S23A), which resulted in a precise and robust subject identification. On the other hand, the least personalized model with the averaged frequencies and SC exhibited an extremely low subject specificity, fingerprinting accuracy and confidence (Fig. 4C-D and supplementary Fig. S23A, dark green) at a relatively high reliability as given by the intra-subject correlation of simulated connectomes (Fig. 4A, dark green).

Conversely, varying the model complexity did not result in differences of the specificity indices that were consistent across parcellations (Fig. 4C, blue, light green and purple). The same observation held for the fingerprinting accuracies of the non-linear models, but not for those of the linear model, which were enhanced relative to the non-linear models of the same personalization and could exceed those of the empirical FC in some cases (Fig. 4E). Interestingly, the fingerprinting confidences of the linear model were systematically much lower than those for the neuronal mass model and in many cases also lower than for the empirical FC (supplementary Fig. S24A). Therefore, subject identification by a simple linear model is less erroneous but also less robust than by more complex non-linear models. Hence, model personalization (but not model complexity) had a positive effect on both single-modal subject specificity measures (specificity index and fingerprinting accuracy) that was consistent across parcellations.

In sum, most of the model implementations yielded within-subject, single-modal correlations of the simulated FC that were significantly enhanced relative to the empirical FC. However, these significant enhancements actually reflected a general increase in both the within- and between-subject single-modal correlations such that the specificity index remained comparable with that of the empirical data. This is in particular true for the linear and non-linear models with a low and moderate extent of personalization (Fig. 4). Only an enhanced model personalization can lead to much improvement of both the subject specificity and the subject identifiability of the simulated FC as a modeling result (Fig. 4).

3.4. Subject specificity of cross-modal connectome correlations

So far we observed that dynamical whole-brain models produce simulated FCs with a particular subject specificity. We subsequently investigated the extent to which these subject-specific connectivity patterns agree with those of the empirical SC and FC by determining the specificity indices and fingerprinting accuracies corresponding to the structure-function and model-fit correlations (Fig. 1, red and brown arrows). We observed that the empirical structure-function relationship was only significantly subject specific for the functionally-

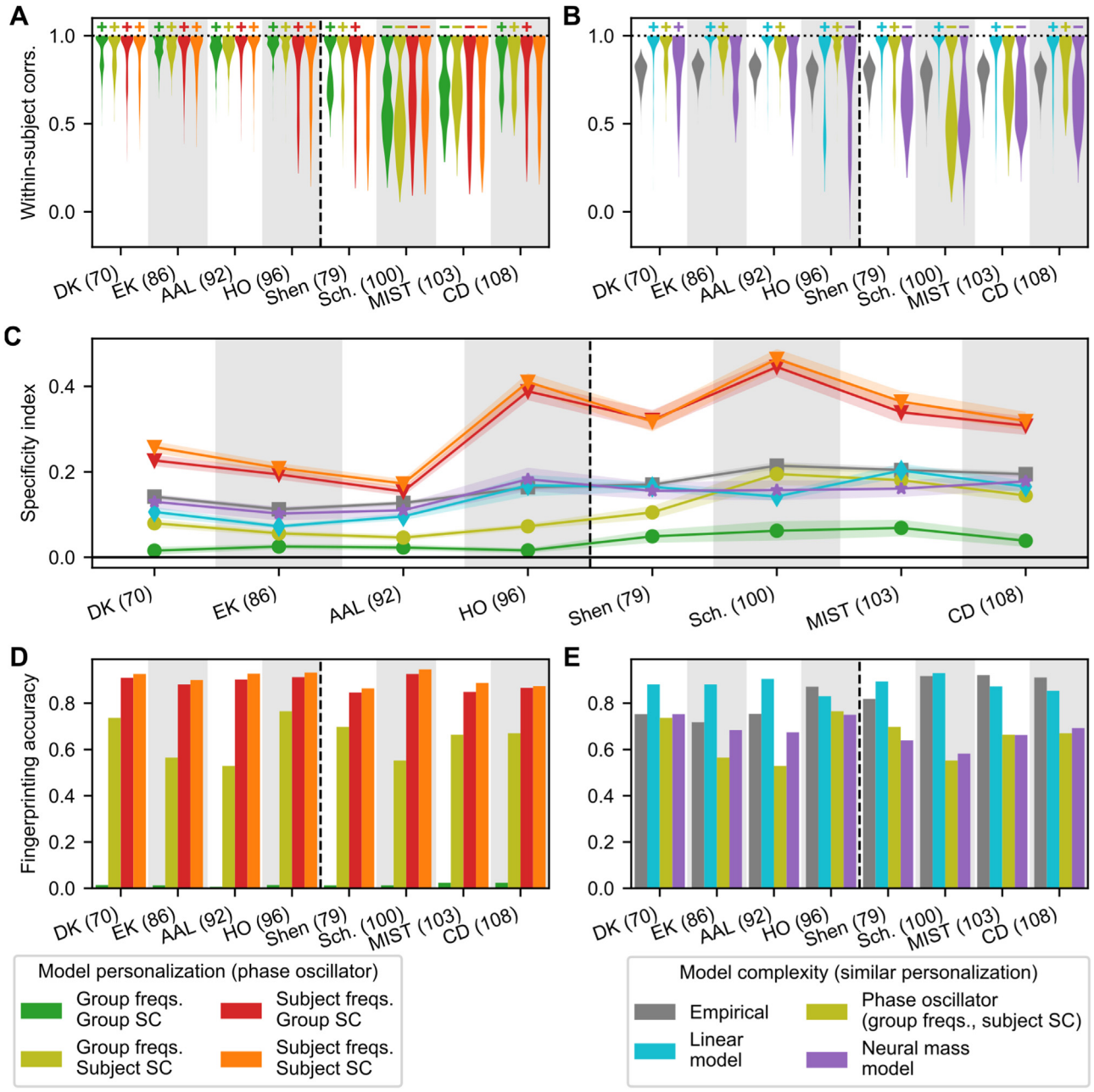

Fig. 4. Impact of the brain atlas, model personalization and model complexity on the reliability and subject specificity of the connectivity patterns of the empirical (gray) and simulated FC. **(A-B)** Distributions of the within-subject, single-modal correlations (corrs.) as a reliability measure of the empirical and simulated FC patterns for the various parcellations considered in this study (Table 1) and for varying levels of (A) model personalization and (B) model complexity. The extent of model personalization as given by the combinations of the subject-specific or group-averaged natural frequencies (freqs.) and SC is indicated in the legend shown in the lower left corner of the plot. Analogously, the level of model complexity as reflected by the linear (least complex), phase oscillator (moderately complex) and neural mass (most complex) models with similar personalization levels is indicated in the legend shown in the lower right corner. The vertical dashed black lines separate the brain atlases constructed on the basis of structural data (left blocks) from those based on functional data (right blocks). Plus and minus signs at the top of the plots indicate significantly increased and decreased within-subject correlation distributions for the respective simulated FC with respect to the one for the empirical FC (gray; panel B), respectively ($p < 0.05$, two-sided Wilcoxon paired signed-rank test, Bonferroni corrected). **(C)** Specificity indices (Eq. 15) calculated from the single-modal correlations of the empirical FC and the simulated FC. The symbols and shaded areas mark the medians and the (Bonferroni corrected) 95% confidence intervals across the 50,000 bootstrapped specificity index estimations, respectively. **(D-E)** Fingerprinting accuracy when identifying individual subjects by comparing one of their empirical (simulated) FCs against all other empirical (simulated) FCs for varying levels of (D) model personalization and (E) model complexity.

derived atlases, although being very small (Fig. 5D, gray). Model personalization through the use of the personalized SC yielded simulated FCs that had structure-function specificity indices significantly higher than zero for all functionally-derived parcellations as well as for some of the structurally-derived atlases (Fig. 5A, light green and orange). Conversely, deriving the models from a grand-average SC resulted in structure-function specificity indices indistinguishable from zero (Fig. 5A, dark green and red). This finding indicated that models constructed on the basis of a personalized SC can embed subject-specific aspects from these structural connectomes into the simulated FC that in fact was fitted to the empirical FC. The fingerprinting accuracies further supported this claim (Fig. 5B-C). We also observed that the magnitude

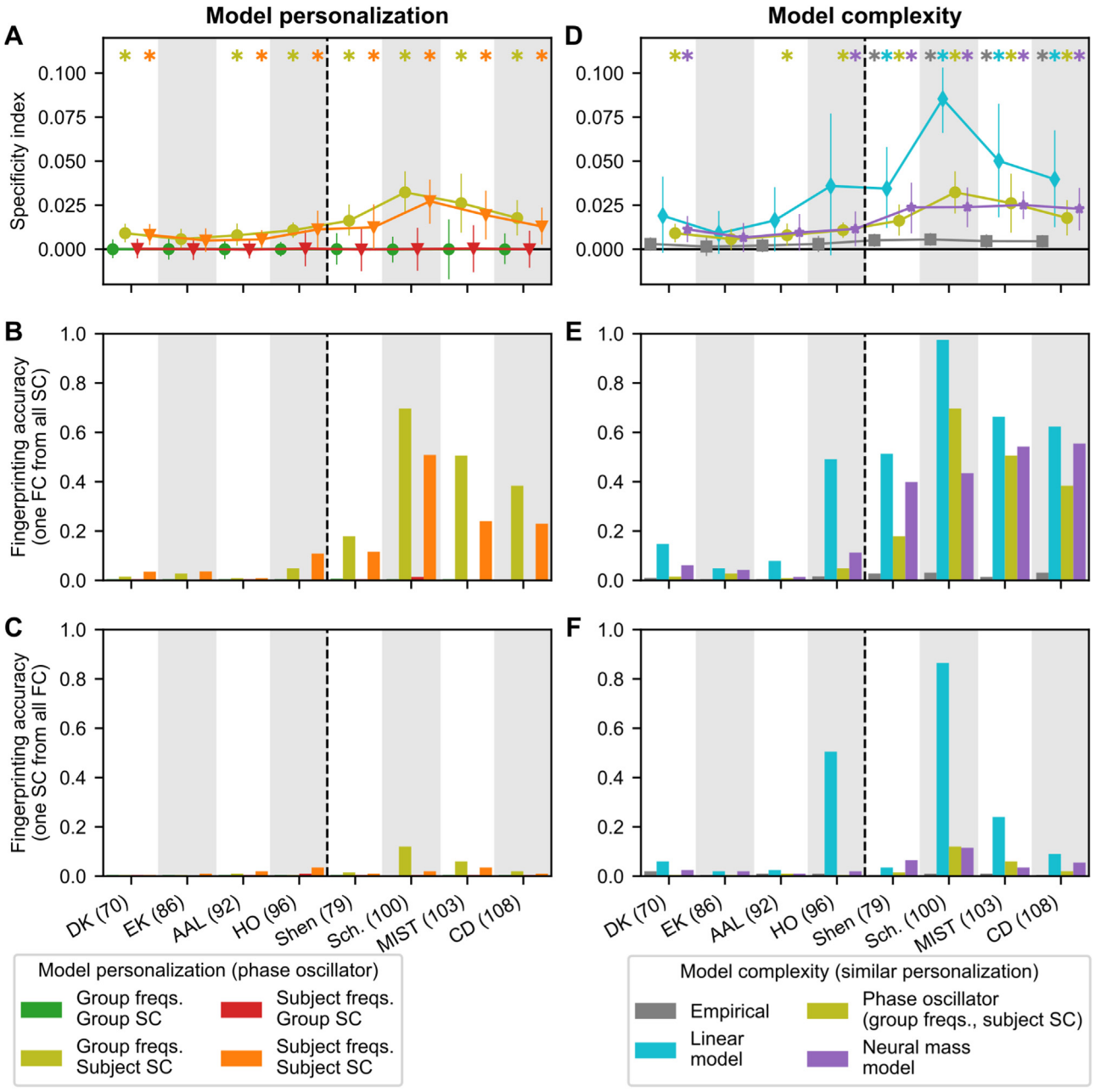

Fig. 5. Impact of the brain atlas, model personalization and model complexity on the subject specificity of the structure-function relationships being the correlations between the empirical SC and the empirical and simulated FC. **(A)** Specificity indices (Eq. 15) of the cross-modal correlations of the empirical SC with the simulated FC for the parcellations considered in this study; see Table 1. The extent of model personalization as given by the combinations of the subject-specific or group-averaged natural frequencies (freqs.) and SC is indicated in the legend (bottom left). The vertical dashed black line separates the brain atlases constructed on the basis of structural data (left block) from those based on functional data (right block). The symbols and error bars mark the medians and the (Bonferroni corrected) 95% confidence intervals across the 50,000 bootstrapped specificity index estimations, respectively. Asterisks indicate whether the lower bounds of these confidence intervals are higher than zero. **(B-C)** Fingerprinting accuracies determined by (B) identification of one simulated FC from all empirical SC based on the largest correlation between them and (C) by identification of one empirical SC from all simulated FC of the same modality for the parcellations considered in this study. **(D-F)** Same as panels A to C, but for varying levels of model complexity as reflected by the linear (least complex), phase oscillator (moderately complex) and neural mass (most complex) models with similar personalization levels and indicated in the legend shown in the lower right corner. The results for empirical FCs are also shown (gray).

of the structure-function specificity indices were much lower than for the single-modal case (Fig. 4C vs. Fig. 5A,D).

Model personalization through the use of subject-specific frequency profiles induced a negative effect on the subject specificity of the structure-function relationship (Fig. 5A-C, orange vs. light green), which is very different from the single-modal FC correlations (Fig. 4). Model complexity did not seem to exert a clear effect on the structure-function subject specificity when considering the non-linear models (Fig. 5D-F). In addition, the fingerprinting accuracies hint towards a particular directionality of the structure-function fingerprinting concept when considering the non-linear models: The identification of the simulated FC from the empirical SC (Fig. 5B,E) resulted in much higher accuracies than the inverted case (Fig. 5C,F). Simultaneously, the fingerprinting confidences were mostly lower for the former than for the latter case (supplemen-

tary Fig. S23B-C and Fig. S24B-C). Interestingly, the subject specificity with which the simulated FCs incorporate the empirical SC patterns is larger for the functionally- than for the structurally-derived parcellations (Fig. 5).

The latter observation regarding the impact of parcellations was also true for the specificity indices and fingerprinting accuracies of the other considered models, in particular, the linear model that exhibited enhanced specificity relative to those of the non-linear models (Fig. 5D-F, blue). However, there appeared to be a less pronounced directionality with respect to the identification of individual subjects for the linear model. In this modeling case, the simulated FC can be identified from empirical SC, and also empirical SC can be identified from the simulated FC with high accuracy (Fig. 5E-F, blue). Such a simple model thus established very strong connections between empirical SC and simulated FC such that the connectome identification in both directions becomes equally possible.

Subsequently, we performed the same analyses for the model-fit correlations (Fig. 1, brown arrows). Even though the specificity indices of these correlations were significantly larger than zero for all tested modeling conditions (supplementary Fig. S25A), the values of the specificity indices and the fingerprinting accuracies determined from the model-fit correlations were relatively low (supplementary Fig. S25B-C). Thus, the models were not fitted so subject specific to the empirical data that individual subjects can be identified from their model-fit correlations with great accuracy. Model personalization but not model complexity could have a positive influence on the subject specificity of the model-fit correlations, although this effect was little consistent across both measures of subject specificity (supplementary Fig. S25A-C).

3.5. Subject specificity of similarity maps

Finally, we investigated how the model personalization may lead to the enhanced reliabilities of the model parameters. We hypothesized that model personalization has an effect on the similarity mappings (Eq. 1) that characterize the agreement between the empirical and simulated FC patterns as a function of the model parameters. Hence, we evaluated how well these similarity maps corresponded to one another between subjects and between distinct empirical FC realizations of the same subject by calculating their within- and between-subject correlations across parameter settings. The results showed that model personalization did not alter the within-subject correlations of the similarity maps consistently across parcellations, and correlation-based reliability of the similarity maps was very high for most of the parcellations (Fig. 6A). On the other hand, the influence of enhanced personalization on the specificity index and fingerprinting accuracy was positive for all atlases (Fig. 6B-C). Combined, these findings suggest that model personalization strongly enhanced subject-specific properties of the similarity maps that became less comparable across (but not within) subjects.

Given our previous findings, one might suspect that model complexity would then exert no consistent effect on the subject specificities of the similarity mappings (Eq. 1). However, we actually found that the distributions of the within-subject similarity map correlations could diverge considerably between model complexities depending on the parcellation (Fig. 7A). Moreover, the specificity indices of the similarity maps of the most complex model (the neural mass model) exceeded those of the less complex ones for all parcellations except for the Harvard-Oxford and Schaefer atlases (Fig. 7B), which are also characterized by a higher variability of the intra-subject correlations (reliability) of similarity maps (Fig. 7A). The fingerprinting accuracies of the similarity mappings were also increased for incrementing levels of model complexities for all atlases (Fig. 7C). These findings indicate that higher model complexity could lead to an enhanced subject specificity with respect to the similarity maps (Fig. 7B-C).

Analogous to the ICC of the model parameters (Fig. 3; supplementary Fig. S20), we verified whether these enhancements for more complex models could be explained by the absence of the signal latency in the network of the linear model. Again, we considered the non-linear models with zero global delay $\tau = 0$ in Eq. (4) and Eq. (8). Subsequently, we determined the similarity maps under this constraint and calculated the specificity indices and fingerprinting accuracies from the correlations between these one-parameter (global coupling) similarity maps. The results of this analysis showed that the ordering of the specificity index and fingerprinting accuracy for varying model complexity was preserved (supplementary Fig. S26). Hence, enhanced model complexity indeed yields similarity maps that are more subject specific.

We also checked whether the different types of specificity indices (single-modal, structure-function, model-fit, similarity maps) could be related to the goodness-of-fits of the model to the empirical data. Even though the tested relationships varied considerably in terms of variance explained, almost all of them were negative (supplementary Fig. S27). This indicates that a higher goodness-of-fit is more likely to reflect a less subject-specific model fit.

4. Discussion

In this study, we showed that dynamical whole-brain models may be fitted to the empirical data with a reliability ranging from ”poor” to ”good” depending on the exact implementation of the dynamical whole-brain modeling paradigm and brain parcellation utilized (Fig. 2; Fig. 3). Subsequently, we showed that the fits of the models might be established through diverging or converging simulated FC patterns, where the variability of the empirical data (FC) used for the model validation can either be enhanced or suppressed by the fitting process. We also demonstrated that simulated FC represented by individual edges or the entire connectivity patterns can be more reliable and subject-specific than the empirical FC (Fig. 2; Fig. 3; Fig. 4). We additionally demonstrated that the simulated FC may exhibit correlations with the empirical SC and empirical FC that exhibit significant subject specificity (Fig. 5; supplementary Fig. S25).

We observed that model personalization positively influences the reliability and subject specificity of the modeling results (Fig. 2; Fig. 4; Fig. 5; Fig. 6; supplementary Fig. S25). Furthermore, model complexity often did not affect the reliability or the subject specificity consistently across parcellations and measures when the fitted model parameters and simulated FCs were considered. A simple linear model can in some cases have enhanced reliability and subject specificity relative to more complex, non-linear models. Nevertheless, the similarity mappings were more subject specific for more complex models consistently for almost all considered parcellations (Fig. 7). We sampled all our results for 8 distinct, state-of-the-art brain atlases and demonstrated the pronounced parcellation-induced variation in the modeling results relative to the purely empirical results. Here, we discuss these findings in the broader scientific context and emphasize their relevance.

4.1. Reliability of modeling results

Even though the (test-retest) reliability has been actively investigated for the empirical FC (Birn et al., 2013; Noble et al., 2019; 2017; Pannunzi et al., 2017; Shehzad et al., 2009; Van Dijk et al., 2010), the literature lacks a comprehensive assessment of it for dynamical whole-brain models. One study nevertheless demonstrated their ”excellent” reliability for multiple realizations of the empirical SC (but not FC) for the same subject (Cicchetti and Sparrow, 1981; Muldoon et al., 2016), while another considered the within-subject correspondences of the fitted model parameters for only one subject, one parcellation and one type of model (Donnelly-Kehoe et al., 2019). Our study provided a comprehensive investigation of the reliability of the modeling results for a cohort of 200 subjects by considering the ICCs of several realizations of the optimal model parameters and simulated FCs fitted to the corresponding different realizations of the empirical FCs for the same subject.

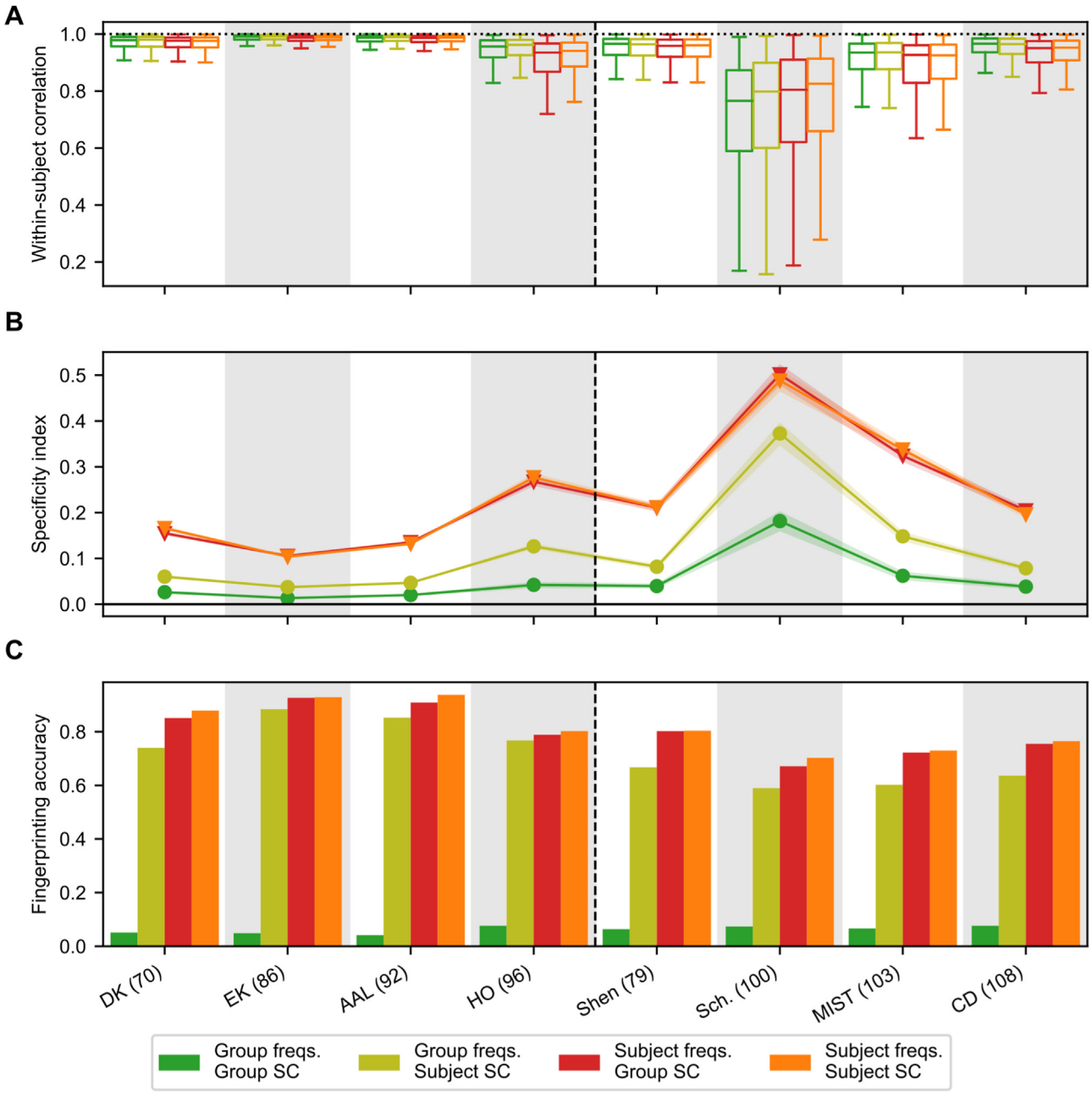

Fig. 6. Influence of parcellation and personalization of the phase oscillator model on the correspondences of the similarity mappings (Eq. 1). **(A)** Distributions of the within-subject correlations of the similarity maps calculated across parameter settings for the various parcellations considered in this study (Table 1). The extent of the model personalization as given by the combinations of the subject-specific or group-averaged natural frequencies (freqs.) and SC is indicated in the legend. The vertical dashed black line separates the brain atlases constructed on the basis of structural data (left block) from those based on functional data (right block). **(B)** Specificity indices (Eq. 15) calculated from the similarity mapping correlations. The symbols and shaded areas mark the medians and the (Bonferroni corrected) 95% confidence intervals across the 50,000 bootstrapped specificity index estimations, respectively. **(C)** Fingerprinting accuracy when identifying individual subjects by comparing the similarity mappings of one particular empirical FC with one another.

The obtained results demonstrated that the reliability of the simulated FC can be larger (and also smaller) than that of the empirical FC depending on the parcellation and exact model implementation (Fig. 2; Fig. 3). Here, our findings of the ”fair” reliability of the empirical FC agreed with the literature (Noble et al., 2019; 2017). As a next step, future studies may investigate how the simultaneous variation of the empirical SC and FC impacts the reliability of dynamical whole-brain modeling results. Our study and the study by Muldoon et al. (2016) may be used as a starting point for such an investigation, where our study in particular could be exploited for the selection of the modeling conditions to consider.

The results of this study, however, primarily suggest that the use of dynamical whole-brain models should be tightly connected with an estimate of the reliability of their results in order to enhance the interpretability of the observations. Despite the reported enhanced reliability of the modeling outcomes, our findings clearly indicate that the ICCs of the modeling results depend highly on the exact implementation of the dynamical whole-brain modeling paradigm. In fact, the reliability of the

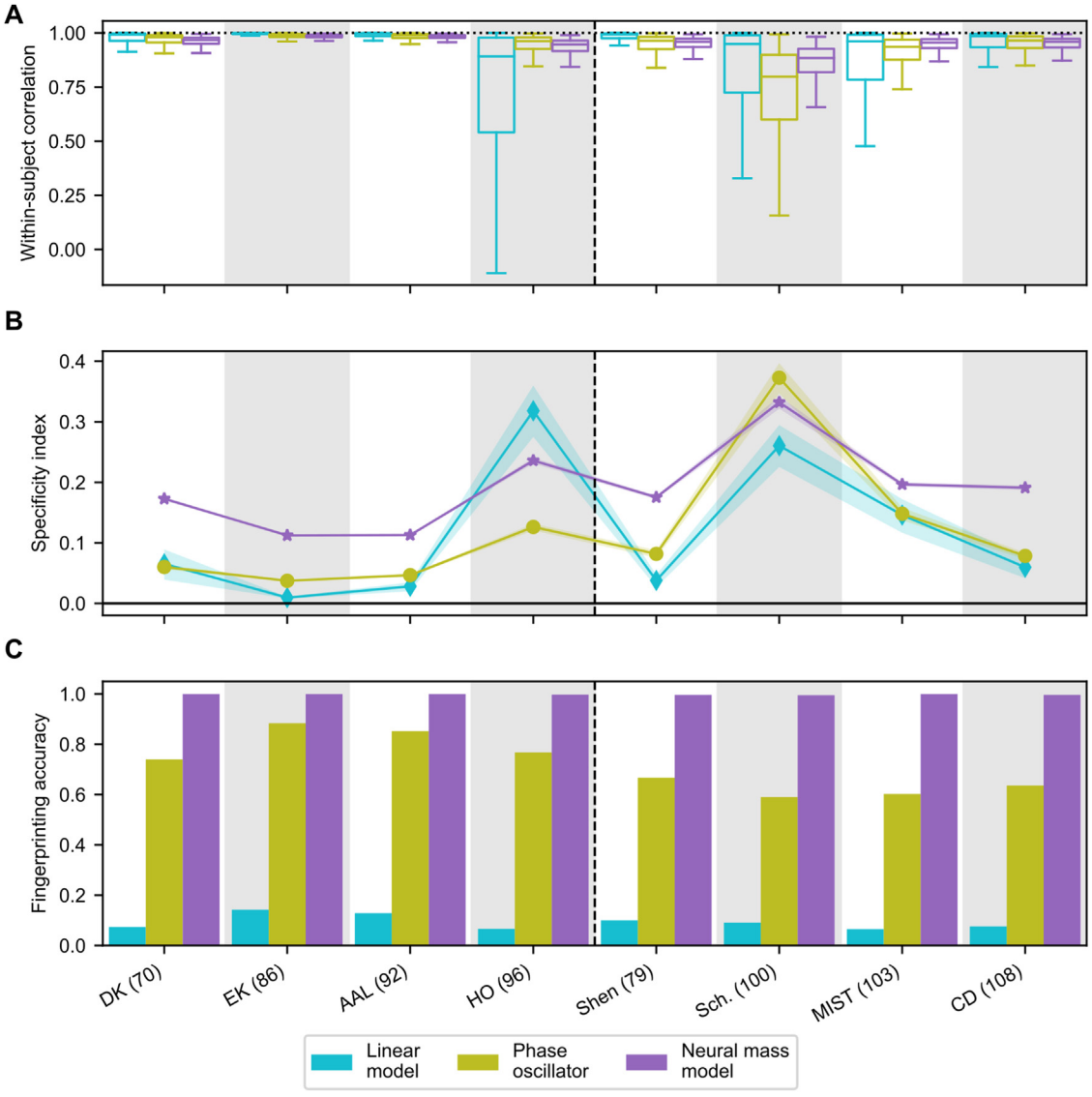

Fig. 7. Influence of parcellation and model complexity with similar model personalization on the correspondences of the similarity mappings (Eq. 1). **(A)** Distributions of the within-subject correlations of the similarity maps calculated across parameter settings for the various parcellations considered in this study (Table 1). The linear model (blue) corresponds to a low complexity, the phase oscillator model (green) to a moderate complexity, and neural mass model (purple) to a high complexity. The vertical dashed black line separates the brain atlases constructed on the basis of structural data (left block) from those based on functional data (right block). **(B)** Specificity indices (Eq. 15) calculated from the similarity mapping correlations. The symbols and shaded areas mark the medians and the (Bonferroni corrected) 95% confidence intervals across the 50,000 bootstrapped specificity index estimations, respectively. **(C)** Fingerprinting accuracy when identifying individual subjects by comparing the similarity mapping of one particular empirical FC with one another. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulated FC edges was lower than that of the empirical FC edges when considering many of the tested conditions (Fig. 2; Fig. 3). Moreover, the model parameters often exhibited ”poor” reliability, which may also sometimes be the case for simulated FC, indicating they exhibit substantial variance for distinct empirical FC realizations of the same subject. We frequently observed such an unreliability and unspecificity for little personalized models with e.g., the group-averaged SC that is widely used in the literature. In the absence of other personalized factors, e.g., subject-specific natural frequencies, such models are hardly reliable and specific. These results are of importance for the neuroscientific conclusions derived from the dynamical whole-brain modeling practices published in the literature, which we adapted and used in this study; see below. They therefore raise the question how reliable published dynamical whole-brain modeling studies actually are.

The literature on dynamical whole-brain modeling is highly heterogeneous with respect to both the reconstruction of the empirical SC and FC from empirical MRI data as well as the model implementations. We, for instance, only covered three of many possible model descriptions that have regularly been used in the whole-brain modeling literature, which also included the (Landau-Stuart) limit-cycle oscillator model (Deco et al., 2017; Ghosh et al., 2008), the (reduced) Wong-Wang model (Deco et al., 2014b; Wong and Wang, 2006) and other (more complex) biophysically-oriented models (Abeysuriya et al., 2018; Bick et al., 2020; Deco and Jirsa, 2012; Hansen et al., 2015; Honey et al., 2007; Naskar et al., 2021). Hence, the methodological procedures may vary considerably between dynamical whole-brain modeling studies, and for most of these variations it is still unclear whether they produce reliable modeling results. This notion further strengthens our recommendation

that dynamical whole-brain modeling studies should explicitly estimate the reliability of the reported results. A consistent reporting of the reliability of results may also help identify best practices in dynamical whole-brain modeling.

4.2. Subject specificity of simulated functional connectivity patterns

Various studies validated the dynamical whole-brain models on the basis of a variety of statistics (Cabral et al., 2011; Deco et al., 2019; 2017; 2013; Hansen et al., 2015; Naskar et al., 2021). Nevertheless, the correlation between the empirical and simulated FC still seems to be the current state-of-the-art in whole-brain modeling (Abeysuriya et al., 2018; Aquino et al., 2022; Naskar et al., 2021; Saggio et al., 2016), and so we used this particular measure for model validation as well. However, by computing the within-subject, single-modal correlations, we demonstrated that this model fitting procedure can yield strongly diverging simulated FC patterns depending on the model implementation and parcellation (Fig. 4). Moreover, even when simulated FCs had similar connectivity motifs across different empirical FC realizations of the same subject, this could still reflect an unspecific increase in both the within- and between-subject single-modal correlations (Fig. 4).

On a positive note, the reliability and subject specificity of the model parameters and simulated FC can essentially be improved by enhancing the model personalization. Furthermore, the correspondences between the simulated FC and the empirical SC were subject specific, that is, their specificity indices were statistically distinguishable from zero, only if the personalized empirical SC was used for model construction (Fig. 5). Hence, this result demonstrated that some of these subject-specific SC patterns are embedded in the simulated FC after the model simulations. We also found that the model-fit correlations can be significantly subject-specific (supplementary Fig. S25). The dynamical whole-brain models thus seem to have the ability to integrate connectivity patterns from both the (personalized) empirical SC and FC, which may (in part) explain how they replicate resting-state brain activity at a personalized level (Bansal et al., 2018; Deco et al., 2017; Jirsa et al., 2017; Ritter et al., 2013; Sanz-Leon et al., 2015), and how they yield good subject classification results (Iravani et al., 2021; Zimmermann et al., 2018b).

Nevertheless, our results merely showed that model construction on the basis of the personalized empirical SC can introduce subject-specific subtleties in the simulated FC; they do not explicitly reveal to which (clinical) purposes this may be beneficial other than subject identification (Fig. 4; Fig. 5). Furthermore, the specificity indices of the structure-function and model-fit correlations involving a simulated FC had comparable and small scales, especially when comparing them to the much higher single-modal specificity indices (Fig. 4). This indicates that the models do not straightforwardly map the empirical SC to the simulated FC with high specificity. We therefore propose that the simulated FC assimilating a diversity of personalized information should be regarded as a separate connectome modality together with the empirical SC and empirical FC.

For the single-modal and structure-function correlations, we could apply the subject specificity analyses also to purely empirical data. Here, the specificity indices of the empirical structure-function relationship (Fig. 5) roughly agree with the study by Zimmermann et al. (2018a). In addition, we identified individual subjects based on the structure-function correlations by identifying one FC (empirical or simulated) from all empirical SC and by identifying one empirical SC from all FC. For the empirical FC, we found the computed fingerprinting accuracies resembling the results reported by Messé (2020). Also the identification of one FC from all SC manifested much higher success rates than vice versa when considering the non-linear models (Fig. 5). The latter result agrees with the problematic inference of the empirical SC from the empirical FC reported by Honey et al. (2009). The linear model was particularly different from the non-linear models with respect to the structure-function correlations. In particular, it exhibited large values of the structure-function specificity index, and could have enhanced fingerprinting accuracy irrespective of whether one simulated FC was identified from all empirical SC or the other way around (Fig. 5). Such a rigid connection between structure and function, which may impair the flexibility of a variety of functions emerging from the same structure, was observed neither in the brain (Deco et al., 2011; Hansen et al., 2015; Honey et al., 2009; Ponce-Alvarez et al., 2015) nor in the non-linear models considered in this study.

Also the fingerprinting accuracies for the single-modal correlations of the empirical FC (Fig. 4) are in agreement with the literature (Finn et al., 2015; Li et al., 2021). However, the latter were rather variable across parcellations with a difference of up to 20% (Fig. 4). Even though the atlas granularity is known to influence the fingerprinting accuracy (Peña Gómez et al., 2018; Li et al., 2021), we minimized this effect by selecting parcellations that contained roughly the same number of parcels. Our study therefore demonstrates the considerable effect of the parcellation technique in isolation on the fingerprinting analysis, which has not been assessed previously. With respect to the fingerprinting analysis, we also acknowledge that the limited number of subjects used in our study may lead to some positive bias in the fingerprinting accuracy (Li et al., 2021; Waller et al., 2017). Nevertheless, as we performed the same fingerprinting analysis for different modalities, this bias (if any) should be included in all results and hence does not render the comparison invalid.

4.3. Model implementations

Enhanced model personalization influenced the within-subject correlations of the similarity maps mildly at best, while it increased the specificity indices of these mappings (Fig. 6). Given Eq. (15), this implies a decrease in the correspondence of the similarity maps across subjects. Qualitatively, the latter finding agrees with the similarity maps shown in supplementary Figs. S3-S10 as well. The observed decrease in the inter-subject correspondence of the similarity mappings can also induce additional variation in the location of the maxima of these similarity mappings (optimal model parameters) between subjects. This is explicitly demonstrated by the enhancements of the between-subject variance in the optimal model parameters for enhanced model personalization (Fig. 2). Despite the relatively untouched within-subject correlations of the similarity maps, increased model personalization also somewhat enhanced the within-subject variance of the optimal model parameters, but not as much as the between-subject variance (Fig. 2). Given Eq. (14), this then leads to the higher ICC for enhanced model personalization. Taken together, we discovered that the higher reliability of the model parameters for more personalized dynamical whole-brain models is induced by a decrease in the comparability of the similarity maps between subjects. Future studies should confirm whether enhanced model personalization indeed improves the differentiability of modeling results across subjects in, for example, classification studies.

We sampled our results for 6 different model implementations that were based on two non-linear models and one linear model, which were all adapted from the literature (Deco et al., 2009; Galán, 2008; Ponce-Alvarez et al., 2015; Saggio et al., 2016). Here, we note that not all parameters of the considered models can straightforwardly be interpreted and associated with brain dynamics. We therefore consider them in the first approximation as model properties that may influence the quality of the model validation, reliability and specificity. Moreover, the literature also inspired the use of the grand-averaged and personalized empirical SCs for model construction and the wielding of the subject-specific frequency profiles in the simulations of the phase oscillator model. Iravani et al. (2021) and Zimmermann et al. (2018b), for instance, constructed their models on the basis of group-averaged and personalized empirical SCs, respectively, before using the modeling results for subject classifications. In addition, several recent studies embedded additional region-specific and potentially subject-specific data (among others, regional frequency profiles) in the dynamical whole-brain mod-

eling workflow (Deco et al., 2018b; 2019; Demirtaş et al., 2019; Domhof et al., 2021b; Donnelly-Kehoe et al., 2019; Jung et al., 2021; Kringelbach et al., 2020; Popovych et al., 2021). The investigation presented in this study may be extended by considering, for example, limit-cycle models (Deco et al., 2017), where the additional model parameters controlling the oscillator amplitudes can be varied for further model personalization. A recent study also demonstrated that subject- and region-specific data can be incorporated into a neural mass model (Demirtaş et al., 2019), indicating that model personalization is in active investigation nowadays.

The two non-linear models (the Kuramoto and Wilson-Cowan models) were selected for three particular reasons. First, both models have frequently been used in previous investigations involving dynamical whole-brain models (Abeysuriya et al., 2018; Daffertshofer and van Wijk, 2011; Deco et al., 2009; Hellyer et al., 2016; Jung et al., 2021; Messé et al., 2014; Muldoon et al., 2016; Ponce-Alvarez et al., 2015; Popovych et al., 2021). Second, their dynamical behaviors under different parameter conditions can be understood and controlled well, because they are sufficiently reduced in terms of complexity and provided with good documentation (Kuramoto, 1984; Wilson and Cowan, 1972). Third, their underlying concepts and implementations in whole-brain modeling studies diverge considerably, which makes it more probable that differences are found between models. In particular, published studies used the Kuramoto model (just like a network of Landau-Stuart limit-cycle oscillators) to model the BOLD signal dynamics directly from the empirical SC (Deco et al., 2018a; 2019; 2017; Domhof et al., 2021b; Jung et al., 2021; Ponce-Alvarez et al., 2015; Popovych et al., 2021), whereas the Wilson-Cowan model requires a haemodynamic conversion model since it specifically models interactions between neural masses. However, the modeling of BOLD signal variations directly from the empirical SC does not reflect the neural dynamics underlying the BOLD signal, and hence future studies could check whether a transformation of the empirical SC matrix might be more appropriate in this case.

Furthermore, there were two reasons for the selection of the linear model. First, its analytical solution ensured that we could estimate the reliability of its global coupling parameter in the absence of any specifications associated with model simulations (Saggio et al., 2016). We actually found this reliability to be at about the same level as those of the global couplings of the non-linear models if signal latency was disregarded (supplementary Fig. S20). Second, as mentioned in Materials and Methods, the model reflects the diffusion of noise across the empirical SC (Galán, 2008; Saggio et al., 2016). This process can be seen as a (linear) scaling of the direct dependencies included in the empirical SC to indirect dependencies, which are more compatible with the definition of the FC (Das et al., 2017; Liégeois et al., 2020). In most cases, we observed that model complexity did not exert a particular increasing or decreasing influence on the reliability or subject specificity of the results of the model fitting (Fig. 3; Fig. 4; supplementary Fig. S25). Also, there were no particular differences with respect to the goodness-of-fit (supplementary Fig. S11). In other words, with regard to (the reliability and specificity of) the wielded model fitting procedure, the Kuramoto and neural mass models in fact do not seem to outperform the linear model. On the contrary, the linear model sometimes demonstrated stronger reliability and subject-specificity than the non-linear models with the same level of personalization. However, the similarity mappings of more complex models appeared to exhibit much enhanced subject specificity as reflected by both the specificity index and the fingerprinting accuracy (Fig. 7). Also the results obtained for the structure-function relationship indicated that complex non-linear models can deliver more realistic results as discussed above. Therefore, the non-linear models appeared to have an increased potential in terms of modeling structure-function interactions and preserving and enhancing the reliability and subject specificity of empirical data as well as model personalization. However, the presented results may also indicate that the model validation procedure of fitting static empirical and simulated FCs is suboptimal in spite of being state-of-the-art as discussed above. Future studies may therefore scrutinize the influence of the model fitting procedure on the reliability and subject specificity of dynamical whole-brain modeling results and propose concrete procedures on how to improve this reliability. In particular, they could investigate whether multimodalities or degeneracy in the similarity maps, which can be observed in supplementary Figs. S3-S10 for some combinations of parcellation and model, can affect the reliability and specificity of the model and should therefore be considered more explicitly when selecting the optimal model parameters. Alternatively, they could examine whether a completely different model fitting strategy, such as fitting the models on the basis of the dynamics of the FC (Brovelli et al., 2017; Hansen et al., 2015; Heitmann and Breakspear, 2018; Hutchison et al., 2013; Kong et al., 2021; Preti et al., 2017), yields more reliable results.

4.4. Atlas variation

A vast number of methods for parcellating the brain have been proposed in the literature (Amunts and Zilles, 2015; Eickhoff et al., 2018a; 2018b). A lot of attention was devoted to the effect that the brain parcellation may have on the analyses of empirical data (Albers et al., 2021; Arslan et al., 2018; Messé, 2020; Wang et al., 2009; Zalesky et al., 2010) and recent modeling results (Domhof et al., 2021b; Jung et al., 2021; Popovych et al., 2021). In this study, we put more emphasis on varying the parcellation method rather than the granularity (number of parcels included in the atlas) when investigating the effect of the parcellation on the modeling results. Previous studies support this focus: Even though granularity is a determining factor when considering statistical analyses of empirical data (Messé, 2020; Wang et al., 2009; Zalesky et al., 2010), parcellation-induced variations in the modeling results could not be explained by only considering this property of the parcellations. Instead, variations in the model fitting quality were primarily related to graph-theoretical network properties extracted from the empirical connectomes (Domhof et al., 2021b) and to other data variables reflecting some statistical properties of the empirical data (Popovych et al., 2021).

In the variation of the parcellations, we balanced between parcellations derived from structural and functional data. Here, the included functionally-derived parcellations presumably optimize the regional homogeneity with respect to the voxel-wise FC (Craddock et al., 2012; Schaefer et al., 2018; Shen et al., 2013; Urchs et al., 2019). In contrast, the structurally-derived atlases have not been designed to do the same for the SC, but may for example follow the anatomical folding patterns of the cortex (Desikan et al., 2006; von Economo and Koskinas, 1925; Frazier et al., 2005; Goldstein et al., 2007; Makris et al., 2006; Rolls et al., 2015; Scholtens et al., 2018; Tzourio-Mazoyer et al., 2002). Our results portrayed distinctions between these structurally- and functionally-derived parcellations, especially when considering the structure-function correlations of the empirical SC with the simulated FC (Fig. 5). The reliability and the (single-modal) subject specificity of the simulated FC also demonstrated opposite tendencies for different parcellation groups, where the former is enhanced for the structurally-derived parcellations, while the latter is larger for the functionally-derived parcellations (Fig. 2; Fig. 3; Fig. 4). These distinctions between the structurally- and functionally-derived brain atlases demonstrate that the parcellation technique has a systematic impact on the modeling results, which can be organized according to particular parcellation principles. Even though relationships were found between the goodness-of-fits and the subject specificities across parcellations, the quality of these associations was rather variable across models (supplementary Fig. S27). Hence, the precision with which the reliability and subject specificity can be estimated from the goodness-of-fit, which is a proxy for the network properties of the empirical connectomes (Domhof et al., 2021b) as well as the empirical structure-function relationship and other statistical properties (Popovych et al., 2021), is highly model-dependent.

We also observed notable model-dependent, parcellation-induced differences in the within-subject correlations of the similarity maps (Fig. 6; Fig. 7); see, for instance, the elongated boxes shown in

Fig. 6 and Fig. 7 for the Schaefer atlas relative to the other parcellations. Since we found the reliability and specificity of the empirical FC to be relatively stable across parcellations (Fig. 2; Fig. 3; Fig. 4), we conclude that dynamical whole-brain models are sensitive to the choice of brain atlas; see also Domhof et al. (2021b) and Popovych et al. (2021). In particular, these results imply that, even though the model is constructed from the same empirical SC, variations of the empirical FC for the same subject may lead to considerably different similarity maps depending on the atlas (and model). We do not suggest that this is necessarily a negative facet of a particular brain atlas. When such parcellations are used for model construction, the fitted models may, for instance, characterize the distinct brain states or other information stored in the different resting-state empirical FC realizations of the same subject that are obscured by other parcellations (for a discussion, see Finn (2021); Finn and Rosenberg (2021)). The negative relationships between the goodness-of-fits and specificity indices support this notion (supplementary Fig. S27), because these findings indicate that a better fit of the model to the empirical data is, in fact, more likely to reflect a more generic (hence not subject-specific) fit. All things considered, our study clearly demonstrates that the proper selection of the brain parcellation appears to be even more important for research using dynamical whole-brain models than studies straightforwardly analyzing the empirical data.

In sum, we extensively assessed the (test-retest) reliability and the subject specificity of the modeling results and their relation to the empirical data. We showed that the model parameters may be fitted to the empirical data with a reliability ranging from ”poor” to ”good” depending on the implementation of the dynamical whole-brain modeling paradigm. In addition, we demonstrated that more personalized models yield increasingly reliable and subject-specific modeling results. For some modeling conditions, we even found that the modeling results were more reliable and subject specific than the results only involving empirical data. We additionally illustrated that the simulated FC may concurrently adopt subject-specific connectivity patterns from both the empirical SC and the empirical FC through the model fitting procedure, which could support considering simulated FC as a separate connectome modality. Finally, we sampled all our findings for 8 state-of-the-art parcellations and demonstrated the substantial impact that a change of parcellation can have on the modeling results, which by far exceeded the parcellation-induced deflections in the results of the empirical data. Taken together, our findings provide an exploratory account on relevant methodological aspects of dynamical whole-brain modeling results. They contribute to the mechanistic understanding of (the personalization of) these models and reveal best practices. Hence, the presented results can be relevant for application of the whole-brain dynamical models and their further development.

Data and Code Availability Statement

MRI data used in this study are publicly available from ConnectomeDB (https://db.humanconnectome.org) and from EBRAINS (https://doi.org/10.25493/81EV-ZVT; https://doi.org/10.25493/F9DP-WCQ). The code used for the simulation of the brain network dynamics, analysis and visualization can be found here: https://jugit.fz-juelich.de/inm7/public/specificity-modeling.

Ethics Statement

Acquisition of the data of human subjects acquired by the human connectome project (HCP; https://humanconnectome.org/ ; Van Essen et al. (2013)) was approved by the local ethics committee of the HCP WU-Minn. Written informed consent was given by all subjects.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Credit authorship contribution statement

Justin W.M. Domhof: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Simon B. Eickhoff:** Conceptualization, Funding acquisition, Project administration, Supervision, Validation, Writing – review & editing. **Oleksandr V. Popovych:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing, Visualization.

Acknowledgments

We thank M. Kollmann for his consultation on the manuscript. The authors gratefully acknowledge the computing time granted through JARA on the supercomputer JURECA at Forschungszentrum Jülich. This study was made possible through the Portfolio Theme Supercomputing and Modeling for the Human Brain of the Helmholtz Association (https://www.helmholtz.de/en), and through the European Union’s Horizon 2020 Research and Innovation Program (grant agreements 785907 (HBP SGA2), 945539 (HBP SGA3) and 826421 (VirtualBrainCloud)). Open access publication was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 491111487. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119321

References

Abeysuriya, R.G., Hadida, J., Sotiropoulos, S.N., Jbabdi, S., Becker, R., Hunt, B.A.E., Woolrich, M.W., 2018. A biophysical model of dynamic balancing of excitation and inhibition in fast oscillatory large-scale networks. PLoS Comput. Biol. 14 (2), e1006007. doi:10.1371/journal.pcbi.1006007.

Albers, K.J., Ambrosen, K.S., Liptrot, M.G., Dyrby, T.B., Schmidt, M.N., Mørup, M., 2021. Using connectomics for predictive assessment of brain parcellations. Neuroimage 238, 118170. doi:10.1016/j.neuroimage.2021.118170.

Amico, E., Goñi, J., 2018. The quest for identifiability in human functional connectomes. Sci Rep 8 (1), 8254. doi:10.1038/s41598-018-25089-1.

Amunts, K., Zilles, K., 2015. Architectonic mapping of the human brain beyond brodmann. Neuron 88 (6), 1086–1107. doi:10.1016/j.neuron.2015.12.001.

Aquino, K.M., Fulcher, B., Oldham, S., Parkes, L., Gollo, L., Deco, G., Fornito, A., 2022. On the intersection between data quality and dynamical modelling of large-scale fMRI signals. Neuroimage 256, 119051. doi:10.1016/j.neuroimage.2022.119051.

Arslan, S., Ktena, S.I., Makropoulos, A., Robinson, E.C., Rueckert, D., Parisot, S., 2018. Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. Neuroimage 170, 5–30. doi:10.1016/j.neuroimage.2017.04.014.

Bansal, K., Nakuci, J., Muldoon, S.F., 2018. Personalized brain network models for assessing structure-function relationships. Curr. Opin. Neurobiol. 52, 42–47. doi:10.1016/j.conb.2018.04.014.

Bick, C., Goodfellow, M., Laing, C.R., Martens, E.A., 2020. Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions: a review. The Journal of Mathematical Neuroscience 10 (1), 9. doi:10.1186/s13408-020-00086-9.

Birn, R.M., Molloy, E.K., Patriat, R., Parker, T., Meier, T.B., Kirk, G.R., Prabhakaran, V., 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. Neuroimage 83, 550–558. doi:10.1016/j.neuroimage.2013.05.099.

Bolt, T., Nomi, J.S., Rubinov, M., Uddin, L.Q., 2017. Correspondence between evoked and intrinsic functional brain network configurations. Hum Brain Mapp 38 (4), 1992–2007. doi:10.1002/hbm.23500.

Breakspear, M., 2017. Dynamic models of large-scale brain activity. Nat. Neurosci. 20 (3), 340–352. doi:10.1038/nn.4497.

Brovelli, A., Badier, J.-M., Bonini, F., Bartolomei, F., Coulon, O., Auzias, G., 2017. Dynamic reconfiguration of visuomotor-related functional connectivity networks. J. Neurosci. 37 (4), 839–853. doi:10.1523/JNEUROSCI.1672-16.2016.

Cabral, J., Hugues, E., Sporns, O., Deco, G., 2011. Role of local network oscillations in resting-state functional connectivity. Neuroimage 57 (1), 130–139. doi:10.1016/j.neuroimage.2011.04.010.

Chen, G., Taylor, P.A., Haller, S.P., Kircanski, K., Stoddard, J., Pine, D.S., Cox, R.W., 2018. Intraclass correlation: improved modeling approaches and applications for neuroimaging. Hum Brain Mapp 39 (3), 1187–1206. doi:10.1002/hbm.23909.

Cicchetti, D.V., Sparrow, S.A., 1981. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 86 (2), 127–137.

Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33 (8), 1914–1928. doi:10.1002/hbm.21333.

Daffertshofer, A., van Wijk, B.C.M., 2011. On the influence of amplitude on the connectivity between phases. Front Neuroinform 5, 6. doi:10.3389/fninf.2011.00006.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: i. segmentation and surface reconstruction. Neuroimage 9 (2), 179–194. doi:10.1006/nimg.1998.0395.

Das, A., Sampson, A.L., Lainscsek, C., Muller, L., Lin, W., Doyle, J.C., Sejnowski, T.J., 2017. Interpretation of the precision matrix and its application in estimating sparse brain connectivity during sleep spindles from human electrocorticography recordings. Neural Comput 29 (3), 603–642. doi:10.1162/NECO_a_00936.

Deco, G., Cabral, J., Saenger, V.M., Boly, M., Tagliazucchi, E., Laufs, H., Kringelbach, M.L., 2018. Perturbation of whole-brain dynamics in silico reveals mechanistic differences between brain states. Neuroimage 169, 46–56. doi:10.1016/j.neuroimage.2017.12.009.

Deco, G., Cruzat, J., Cabral, J., Knudsen, G.M., Carhart-Harris, R.L., Whybrow, P.C., Kringelbach, M.L., 2018. Whole-brain multimodal neuroimaging model using serotonin receptor maps explains non-linear functional effects of LSD. Current Biology 28 (19), 3065–3074. doi:10.1016/j.cub.2018.07.083. E6

Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N.K., Kringelbach, M.L., 2019. Awakening: predicting external stimulation to force transitions between different brain states. Proceedings of the National Academy of Sciences 116 (36), 18088–18097. doi:10.1073/pnas.1905534116.

Deco, G., Jirsa, V., McIntosh, A.R., Sporns, O., Kötter, R., 2009. Key role of coupling, delay, and noise in resting brain fluctuations. Proceedings of the National Academy of Sciences 106 (25), 10302–10307. doi:10.1073/pnas.0901831106.

Deco, G., Jirsa, V.K., 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J. Neurosci. 32 (10), 3366–3375. doi:10.1523/JNEUROSCI.2523-11.2012.

Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat. Rev. Neurosci. 12 (1), 43–56. doi:10.1038/nrn2961.

Deco, G., Kringelbach, M., 2014. Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders. Neuron 84 (5), 892–905. doi:10.1016/j.neuron.2014.08.034.

Deco, G., Kringelbach, M.L., Jirsa, V.K., Ritter, P., 2017. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. Sci Rep 7 (1), 3095. doi:10.1038/s41598-017-03073-5.

Deco, G., McIntosh, A.R., Shen, K., Hutchison, R.M., Menon, R.S., Everling, S., Jirsa, V.K., 2014. Identification of optimal structural connectivity using functional connectivity and neural modeling. J. Neurosci. 34 (23), 7910–7916. doi:10.1523/JNEUROSCI.4423-13.2014.

Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G.L., Mantini, D., Corbetta, M., 2014. How local excitation-inhibition ratio impacts the whole brain dynamics. J. Neurosci. 34 (23), 7886–7898. doi:10.1523/JNEUROSCI.5068-13.2014.

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G.L., Hagmann, P., Corbetta, M., 2013. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. J. Neurosci. 33 (27), 11239–11252. doi:10.1523/JNEUROSCI.1091-13.2013.

Demirtaş, M., Burt, J.B., Helmer, M., Ji, J.L., Adkinson, B.D., Glasser, M.F., Murray, J.D., 2019. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. Neuron 101 (6), 1181–1194. doi:10.1016/j.neuron.2019.01.017. E13

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31 (3), 968–980. doi:10.1016/j.neuroimage.2006.01.021.

Domhof, J.W.M., Jung, K., Eickhoff, S.B., Popovych, O.V., 2021a. Parcellation-based structural and resting-state functional brain connectomes of a healthy cohort [Dataset]. EBRAINS doi:10.25493/81EV-ZVT.

Domhof, J.W.M., Jung, K., Eickhoff, S.B., Popovych, O.V., 2021b. Parcellation-induced variation of empirical and simulated brain connectomes at group and subject levels. Network Neurosci. 1–56. doi:10.1162/netn_a_00202.

Domhof, J.W.M., Jung, K., Eickhoff, S.B., Popovych, O.V., 2022. Parcellation-based resting-state blood-oxygen-level-dependent (BOLD) signals of a healthy cohort (v1.0) [Dataset]. EBRAINS doi:10.25493/F9DP-WCQ.

Donnelly-Kehoe, P., Saenger, V.M., Lisofsky, N., Kühn, S., Kringelbach, M.L., Schwarzbach, J., Deco, G., 2019. Reliable local dynamics in the brain across sessions are revealed by whole-brain modeling of resting state activity. Hum Brain Mapp 40 (10), 2967–2980. doi:10.1002/hbm.24572.

von Economo, C., Koskinas, G. N., 1925. Die Cytoarchitektonik der Hirnrinde des erwachsenen MenschenWien, Springer.

Eickhoff, S.B., Constable, R.T., Yeo, B.T.T., 2018. Topographic organization of the cerebral cortex and brain cartography. Neuroimage 170, 332–347. doi:10.1016/j.neuroimage.2017.02.018.

Eickhoff, S.B., Yeo, B.T.T., Genon, S., 2018. Imaging-based parcellations of the human brain. Nat. Rev. Neurosci. 19 (11), 672–686. doi:10.1038/s41583-018-0071-7.

Finn, E.S., 2021. Is it time to put rest to rest? Trends Cogn. Sci. (Regul. Ed.) 25 (12), 1021–1032. doi:10.1016/j.tics.2021.09.005.

Finn, E.S., Rosenberg, M.D., 2021. Beyond fingerprinting: choosing predictive connectomes over reliable connectomes. Neuroimage 239, 118254. doi:10.1016/j.neuroimage.2021.118254.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18 (11), 1664–1671. doi:10.1038/nn.4135.

Fraga González, G., Smit, D.J.A., van der Molen, M.J.W., Tijms, J., Stam, C.J., de Geus, E.J.C., van der Molen, M.W., 2018. EEG Resting state functional connectivity in adult dyslexics using phase lag index and graph analysis. Front Hum Neurosci 12, 341. doi:10.3389/fnhum.2018.00341.

Frazier, J.A., Chiu, S., Breeze, J.L., Makris, N., Lange, N., Kennedy, D.N., Biederman, J., 2005. Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. American Journal of Psychiatry 162 (7), 1256–1265. doi:10.1176/appi.ajp.162.7.1256.

Friston, K.J., 2011. Functional and effective connectivity: areview. Brain Connect 1 (1), 13–36. doi:10.1089/brain.2011.0008.

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

Galán, R.F., 2008. On how network architecture determines the dominant patterns of spontaneous neural activity. PLoS ONE 3 (5), e2148. doi:10.1371/journal.pone.0002148.

Ghosh, A., Rho, Y., McIntosh, A.R., Kötter, R., Jirsa, V.K., 2008. Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS Comput. Biol. 4 (10), e1000196. doi:10.1371/journal.pcbi.1000196.

Gilson, M., Moreno-Bote, R., Ponce-Alvarez, A., Ritter, P., Deco, G., 2016. Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. PLoS Comput. Biol. 12 (3), e1004762. doi:10.1371/journal.pcbi.1004762.

Goldstein, J.M., Seidman, L.J., Makris, N., Ahern, T., O'Brien, L.M., Caviness, V.S., Tsuang, M.T., 2007. Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. Biol. Psychiatry 61 (8), 935–945. doi:10.1016/j.biopsych.2006.06.027.

Grabner, G., Janke, A.L., Budge, M.M., Smith, D., Pruessner, J., Collins, D.L., 2006. Symmetric Atlasing and Model Based Segmentation: an Application to the Hippocampus in Older Adults. In: Larsen, R., Nielsen, M., Sporring, J. (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006 Medical Image Computing and Computer-Assisted Intervention - MICCAI, 2006. Springer, Berlin, Heidelberg, pp. 58–66. doi:10.1007/11866763_8.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Smith, S.M., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95, 232–247. doi:10.1016/j.neuroimage.2014.03.034.

Hahn, G., Skeide, M.A., Mantini, D., Ganzetti, M., Destexhe, A., Friederici, A.D., Deco, G., 2019. A new computational approach to estimate whole-brain effective connectivity from functional and structural MRI, applied to language development. Sci Rep 9 (1), 8479. doi:10.1038/s41598-019-44909-6.

Hansen, E.C.A., Battaglia, D., Spiegler, A., Deco, G., Jirsa, V.K., 2015. Functional connectivity dynamics: modeling the switching behavior of the resting state. Neuroimage 105, 525–535. doi:10.1016/j.neuroimage.2014.11.001.

Heitmann, S., Breakspear, M., 2018. Putting the "dynamic" back into dynamic functional connectivity. Network Neurosci. 02 (02), 150–174. doi:10.1162/netn_a_00041.

Hellyer, P.J., Jachs, B., Clopath, C., Leech, R., 2016. Local inhibitory plasticity tunes macroscopic brain dynamics and allows the emergence of functional brain networks. Neuroimage 124, 85–95. doi:10.1016/j.neuroimage.2015.08.069.

van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. 20 (8), 519–534. doi:10.1016/j.euroneuro.2010.03.008.

Honey, C.J., Kötter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proceedings of the National Academy of Sciences 104 (24), 10240–10245. doi:10.1073/pnas.0701519104.

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proceedings of the National Academy of Sciences 106 (6), 2035–2040. doi:10.1073/pnas.0811168106.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Chang, C., 2013. Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80, 360–378. doi:10.1016/j.neuroimage.2013.05.079.

Iravani, B., Arshamian, A., Fransson, P., Kaboodvand, N., 2021. Whole-brain modelling of resting state fMRI differentiates ADHD subtypes and facilitates stratified neuro-stimulation therapy. Neuroimage 231, 117844. doi:10.1016/j.neuroimage.2021.117844.

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62 (2), 782–790. doi:10.1016/j.neuroimage.2011.09.015.

Jirsa, V.K., Proix, T., Perdikis, D., Woodman, M.M., Wang, H., Gonzalez-Martinez, J., Bartolomei, F., 2017. The virtual epileptic patient: individualized whole-brain models of epilepsy spread. Neuroimage 145, 377–388. doi:10.1016/j.neuroimage.2016.04.049.

Jülich Supercomputing Centre, 2018. JURECA: Modular supercomputer at Jülich Supercomputing Centre. Journal of large-scale research facilities 4, A132. doi:10.17815/jlsrf-4-121-1.

Jung, K., Eickhoff, S.B., Popovych, O.V., 2021. Tractography density affects whole-brain structural architecture and resting-state dynamical modeling. Neuroimage 237, 118176. doi:10.1016/j.neuroimage.2021.118176.

Kong, X., Kong, R., Orban, C., Wang, P., Zhang, S., Anderson, K., Yeo, B.T.T., 2021. Sensory-motor cortices shape functional connectivity dynamics in the human brain. Nat Commun 12 (1), 6373. doi:10.1038/s41467-021-26704-y.

Kringelbach, M.L., Cruzat, J., Cabral, J., Knudsen, G.M., Carhart-Harris, R., Whybrow, P.C., Deco, G., 2020. Dynamic coupling of whole-brain neuronal and neurotransmitter systems. Proceedings of the National Academy of Sciences 117 (17), 9566–9576. doi:10.1073/pnas.1921475117.

Kuramoto, Y., 1984. Chemical oscillations, waves, and turbulence. Chemical oscillations, waves, and turbulence. Springer, Berlin.

Larson-Prior, L.J., Oostenveld, R., Della Penna, S., Michalareas, G., Prior, F., Babajani-Feremi, A., Snyder, A.Z., 2013. Adding dynamics to the human connectome project with MEG. Neuroimage 80, 190–201. doi:10.1016/j.neuroimage.2013.05.056.

Li, K., Wisner, K., Atluri, G., 2021. Feature selection framework for functional connectome fingerprinting. Hum Brain Mapp 42 (12), 3717–3732. doi:10.1002/hbm.25379.

Liégeois, R., Santos, A., Matta, V., Van De Ville, D., Sayed, A.H., 2020. Revisiting correlation-based functional connectivity and its relationship with structural connectivity. Network Neurosci. 4 (4), 1235–1251. doi:10.1162/netn_a_00166.

Liljequist, D., Elfving, B., Roaldsen, K.S., 2019. Intraclass correlation - a discussion and demonstration of basic features. PLoS ONE 14 (7), e0219854. doi:10.1371/journal.pone.0219854.

Maier-Hein, K.H., Neher, P.F., Houde, J.-.C., Côté, M.-.A., Garyfallidis, E., Zhong, J., Descoteaux, M., 2017. The challenge of mapping the human connectome based on diffusion tractography. Nat Commun 8 (1), 1349. doi:10.1038/s41467-017-01285-x.

Makris, N., Goldstein, J.M., Kennedy, D., Hodge, S.M., Caviness, V.S., Faraone, S.V., Seidman, L.J., 2006. Decreased volume of left and total anterior insular lobule in schizophrenia. Schizophr. Res. 83 (2), 155–171. doi:10.1016/j.schres.2005.11.020.

Mayhew, S.D., Ostwald, D., Porcaro, C., Bagshaw, A.P., 2013. Spontaneous EEG alpha oscillation interacts with positive and negative BOLD responses in the visual-auditory cortices and default-mode network. Neuroimage 76, 362–372. doi:10.1016/j.neuroimage.2013.02.070.

Messé, A., 2020. Parcellation influence on the connectivity-based structure-function relationship in the human brain. Hum Brain Mapp 41 (5), 1167–1180. doi:10.1002/hbm.24866.

Messé, A., Rudrauf, D., Benali, H., Marrelec, G., 2014. Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. PLoS Comput. Biol. 10 (3), e1003530. doi:10.1371/journal.pcbi.1003530.

Messé, A., Rudrauf, D., Giron, A., Marrelec, G., 2015. Predicting functional connectivity from structural connectivity via computational models using MRI: an extensive comparison study. Neuroimage 111, 65–75. doi:10.1016/j.neuroimage.2015.02.001.

Muldoon, S.F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S.T., Vettel, J.M., Bassett, D.S., 2016. Stimulation-based control of dynamic brain networks. PLoS Comput. Biol. 12 (9), e1005076. doi:10.1371/journal.pcbi.1005076.

Naskar, A., Vattikonda, A., Deco, G., Roy, D., Banerjee, A., 2021. Multiscale dynamic mean field (MDMF) model relates resting-state brain dynamics with local cortical excitatory-inhibitory neurotransmitter homeostasis. Network Neurosci. 5 (3), 757–782. doi:10.1162/netn_a_00197.

Noble, S., Scheinost, D., Constable, R.T., 2019. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. Neuroimage 203, 116157. doi:10.1016/j.neuroimage.2019.116157.

Noble, S., Spann, M.N., Tokoglu, F., Shen, X., Constable, R.T., Scheinost, D., 2017. Influences on the test-retest reliability of functional connectivity MRI and its relationship with behavioral utility. Cerebral Cortex 27 (11), 5415–5429. doi:10.1093/cercor/bhx230.

Pannunzi, M., Hindriks, R., Bettinardi, R.G., Wenger, E., Lisofsky, N., Martensson, J., Deco, G., 2017. Resting-state fMRI correlations: from link-wise unreliability to whole brain stability. Neuroimage 157, 250–262. doi:10.1016/j.neuroimage.2017.06.006.

Peña Gómez, C., Avena-Koenigsberger, A., Sepulcre, J., Sporns, O., 2018. Spatiotemporal network markers of individual variability in the human functional connectome. Cerebral Cortex 28 (8), 2922–2934. doi:10.1093/cercor/bhx170.

Ponce-Alvarez, A., Deco, G., Hagmann, P., Romani, G.L., Mantini, D., Corbetta, M., 2015. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. PLoS Comput. Biol. 11 (2), e1004100. doi:10.1371/journal.pcbi.1004100.

Popovych, O.V., Jung, K., Manos, T., Diaz-Pier, S., Hoffstaedter, F., Schreiber, J., Eickhoff, S.B., 2021. Inter-subject and inter-parcellation variability of resting-state whole-brain dynamical modeling. Neuroimage 236, 118201. doi:10.1016/j.neuroimage.2021.118201.

Popovych, O.V., Manos, T., Hoffstaedter, F., Eickhoff, S.B., 2019. What can computational models contribute to neuroimaging data analytics? Front Syst Neurosci 12, 68. doi:10.3389/fnsys.2018.00068.

Preti, M.G., Bolton, T.A., Van De Ville, D., 2017. The dynamic functional connectome: state-of-the-art and perspectives. Neuroimage 160, 41–54. doi:10.1016/j.neuroimage.2016.12.061.

Ritter, P., Schirner, M., McIntosh, A.R., Jirsa, V.K., 2013. The virtual brain integrates computational modeling and multimodal neuroimaging. Brain Connect 3 (2), 121–145. doi:10.1089/brain.2012.0120.

Robinson, P.A., 2012. Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory. Physical Review E 85 (1), 011912. doi:10.1103/PhysRevE.85.011912.

Robinson, P.A., Sarkar, S., Pandejee, G.M., Henderson, J.A., 2014. Determination of effective brain connectivity from functional connectivity with application to resting state connectivities. Physical Review E 90 (1), 012707. doi:10.1103/PhysRevE.90.012707.

Rolls, E.T., Joliot, M., Tzourio-Mazoyer, N., 2015. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. Neuroimage 122, 1–5. doi:10.1016/j.neuroimage.2015.07.075.

Saggio, M.L., Ritter, P., Jirsa, V.K., 2016. Analytical operations relate structural and functional connectivity in the brain. PLoS ONE 11 (8), e0157292. doi:10.1371/journal.pone.0157292.

Sanz-Leon, P., Knock, S.A., Spiegler, A., Jirsa, V.K., 2015. Mathematical framework for large-scale brain network modeling in the virtual brain. Neuroimage 111, 385–430. doi:10.1016/j.neuroimage.2015.01.002.

Sarar, G., Rao, B., Liu, T., 2021. Functional connectome fingerprinting using shallow feed-forward neural networks. Proceedings of the National Academy of Sciences 118 (15). doi:10.1073/pnas.2021852118. e2021852118

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-.N., Holmes, A.J., Yeo, B.T.T., 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cerebral Cortex 28 (9), 3095–3114. doi:10.1093/cercor/bhx179.

Scholtens, L.H., de Reus, M.A., de Lange, S.C., Schmidt, R., van den Heuvel, M.P., 2018. An MRI Von Economo - Koskinas atlas. Neuroimage 170, 249–256. doi:10.1016/j.neuroimage.2016.12.069.

Shehzad, Z., Kelly, A.M.C., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Milham, M.P., 2009. The resting brain: unconstrained yet reliable. Cerebral Cortex 19 (10), 2209–2229. doi:10.1093/cercor/bhn256.

Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415. doi:10.1016/j.neuroimage.2013.05.081.

Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86 (2), 420–428. doi:10.1037//0033-2909.86.2.420.

Sotiropoulos, S.N., Zalesky, A., 2019. Building connectomes using diffusion MRI: why, how and but. NMR Biomed 32 (4), e3752. doi:10.1002/nbm.3752.

Spitoni, G.F., Cimmino, R.L., Bozzacchi, C., Pizzamiglio, L., Di Russo, F., 2013. Modulation of spontaneous alpha brain rhythms using low-intensity transcranial direct-current stimulation. Front Hum Neurosci 7, 529. doi:10.3389/fnhum.2013.00529.

Suárez, L.E., Markello, R.D., Betzel, R.F., Misic, B., 2020. Linking structure and function in macroscale brain networks. Trends Cogn. Sci. (Regul. Ed.) 24 (4), 302–315. doi:10.1016/j.tics.2020.01.008.

Tournier, J.-.D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Connelly, A., 2019. MRTrix3: a fast, flexible and open software framework for medical image processing and visualisation. Neuroimage 202, 116137. doi:10.1016/j.neuroimage.2019.116137.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: Improved N3 bias correction. IEEE Trans Med Imaging 29 (6), 1310–1320. doi:10.1109/TMI.2010.2046908.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289. doi:10.1006/nimg.2001.0978.

Urchs, S., Armoza, J., Moreau, C., Benhajali, Y., St-Aubin, J., Orban, P., Bellec, P., 2019. MIST: A multi-resolution parcellation of functional brain networks [version 2; peer review: 4 approved]. MNI Open Research 1 (3). doi:10.12688/mniopenres.12767.2.

Valdes-Sosa, P.A., Roebroeck, A., Daunizeau, J., Friston, K., 2011. Effective connectivity: influence, causality and biophysical modeling. Neuroimage 58 (2), 339–361. doi:10.1016/j.neuroimage.2011.03.058.

Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J. Neurophysiol. 103 (1), 297–321. doi:10.1152/jn.00783.2009.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-minn human connectome project: an overview. Neuroimage 80, 62–79. doi:10.1016/j.neuroimage.2013.05.041.

Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E.J., Bucholz, R., Yacoub, E., 2012. The human connectome project: a data acquisition perspective. Neuroimage 62 (4), 2222–2231. doi:10.1016/j.neuroimage.2012.02.018.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., van Mulbregt, P., 2020. Scipy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17 (3), 261–272. doi:10.1038/s41592-019-0686-2.

Waller, L., Walter, H., Kruschwitz, J.D., Reuter, L., Müller, S., Erk, S., Veer, I.M., 2017. Evaluating the replicability, specificity, and generalizability of connectome fingerprints. Neuroimage 158, 371–377. doi:10.1016/j.neuroimage.2017.07.016.

van der Walt, S., Colbert, S.C., Varoquaux, G., 2011. The NumPy array: a structure for efficient numerical computation. Computing in Science Engineering 13 (2), 22–30. doi:10.1109/MCSE.2011.37.

Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., He, Y., 2009. Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. Hum Brain Mapp 30 (5), 1511–1523. doi:10.1002/hbm.20623.

Wilson, H.R., Cowan, J.D., 1972. Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. 12 (1), 1–24. doi:10.1016/S0006-3495(72)86068-5.

Wong, K.-.F., Wang, X..J., 2006. A recurrent network mechanism of time integration in perceptual decisions. Journal of Neuroscience 26 (4), 1314–1328. doi:10.1523/JNEUROSCI.3733-05.2006.

Woolrich, M.W., Stephan, K.E., 2013. Biophysical network models and the human connectome. Neuroimage 80, 330–338. doi:10.1016/j.neuroimage.2013.03.059.

Yeh, C.-.H., Jones, D.K., Liang, X., Descoteaux, M., Connelly, A., 2021. Mapping structural connectivity using diffusion MRI: challenges and opportunities. J. Magn. Reson. Imaging 53 (6), 1666–1682. doi:10.1002/jmri.27188.

Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yücel, M., Pantelis, C., Bullmore, E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? Neuroimage 50 (3), 970–983. doi:10.1016/j.neuroimage.2009.12.027.

Zimmermann, J., Griffiths, J., Schirner, M., Ritter, P., McIntosh, A.R., 2018. Subject specificity of the correlation between large-scale structural and functional connectivity. Network Neurosci. 3 (1), 90–106. doi:10.1162/netn_a_00055.

Zimmermann, J., Perry, A., Breakspear, M., Schirner, M., Sachdev, P., Wen, W., Solodkin, A., 2018. Differentiation of Alzheimer's disease based on local and global parameters in personalized Virtual Brain models. NeuroImage: Clinical 19, 240–251. doi:10.1016/j.nicl.2018.04.017.