

**Efficient small-world and scale-free functional brain networks at rest using k-nearest  
neighbors thresholding**

*Caroline Garcia Forlim<sup>a</sup> \*, Siavash Haghiri<sup>b</sup> \*, Sandra Düzel<sup>c</sup>, Simone Kühn<sup>a,c</sup>*

<sup>a</sup> University Medical Center Hamburg - Eppendorf,

Clinic and Polyclinic for Psychiatry and Psychotherapy,

Martinistraße 52, 20246, Hamburg, Germany

<sup>b</sup> University of Tübingen

Department of Computer Science

Sand 14, 72076, Tübingen, Germany

<sup>c</sup> Max Planck Institute for Human Development,

Lise Meitner Group for Environmental Neuroscience

Lentzeallee 94, 14195 Berlin, Germany

\*indicates a shared first-authorship

## Highlights.

- A novel thresholding method for brain networks based on k-nearest neighbors (kNN)
- kNN applied on resting state fMRI from a big cohort of healthy subjects BASE-II
- kNN built networks present greater small world properties than density threshold
- kNN built networks present scale-free properties whereas density threshold did not

## Abstract

In recent years, there has been a massive effort to analyze the topological properties of brain networks. Yet, one of the challenging questions in the field is how to construct brain networks based on the connectivity values derived from neuroimaging methods. From a theoretical point of view, it is plausible that the brain would have evolved to minimize energetic costs of information processing, and therefore, maximizes efficiency as well as to redirect its function in an adaptive fashion, that is, resilience. A brain network with such features, when characterized using graph analysis, would present small-world and scale-free properties.

In this paper, we focused on how the brain network is constructed by introducing and testing an alternative method: k-nearest neighbor (kNN). In addition, we compared the kNN method with one of the most common methods in neuroscience: namely the density threshold. We performed our analyses on functional connectivity matrices derived from resting state fMRI of a big imaging cohort (N=434) of young and older healthy participants. The topology of networks was characterized by the graph measures degree, characteristic path length, clustering coefficient and small world. In addition, we verified whether kNN produces scale-free networks. We showed that networks built by kNN presented advantages over traditional

thresholding methods, namely greater values for small-world (linked to efficiency of networks) than those derived by means of density thresholds and moreover, it presented also scale-free properties (linked to the resilience of networks), where density threshold did not. A brain network with such properties would have advantages in terms of efficiency, rapid adaptive reconfiguration and resilience, features of brain networks that are relevant for plasticity and cognition as well as neurological diseases as stroke and dementia.

**Keywords:** brain network, resting-state fMRI, functional connectivity, k-nearest neighbor, small-world networks, scale-free networks

## 1. Introduction

Our brain can be thought as a complex network. This view is supported by histological, structural and functional studies, showing complex levels of interaction in the brain. Complex networks applied to neuroscience, hypothesize a network created from inter-connected brain areas, brain areas are called nodes and their connections are described as links (Fornito, Zalesky, and Bullmore 2016; Bullmore and Sporns 2009). Local interconnections can form clusters. Clusters play an important role as functional segregators, processing specialized information and keeping this information at a certain location (Rubinov and Sporns 2010; Bullmore and Sporns 2009). Nevertheless, these centers of specialized information processing are most likely not isolated from each other, thus, one expects intercommunication among the clusters, so that, the information can be exchanged in the brain. The connections involved in global intercommunication among different functionally segregated regions would play the role of so-called functional integrators. Moreover, it is of advantage that all processes described above are preserved in cases of external or internal perturbations.

An interesting scientific question is, what an optimal brain network would look like or phrased differently: What would be the most efficient way that the brain could work? The concept of small-world organization in networks was introduced by Watts and Strogatz (1998), when they investigated biological, technological and social networks (Watts and Strogatz 1998). This concept is very broad and appears in many other fields and disciplines where complex networks are used, e.g. the power grid networks, social networks, computer science, biology (Telesford et al. 2011; Amaral et al. 2000). It has been suggested that the optimal brain network is topologically reflected by having high clustering and short paths connecting the clusters among themselves (Bassett and Bullmore 2006). This points to the fact that it is equally important to have centers processing specialized information, as it is to have fast communication between them.

To keep these processes properly working in face of external or internal perturbation, the brain network should be resilient, that is maintaining its activity in the face of faults. Thus, besides the small-worldness characteristics, it is desirable that brain networks show scale-free properties. A network is scale-free if its degree distribution follows a power-law function. This property has been associated with network resilience (Cohen et al. 2000; Albert, Jeong, and Barabási 2000; Albert and Barabási 2002). There are multiple studies on brain networks which suggest the presence of scale-free networks (Eguíluz et al. 2005; van den Heuvel et al. 2008). A brain network with such characteristics would bring advantages as both segregate and integrated information processing and enables rapid adaptive reconfiguration, which is important for plasticity and cognition (Baronchelli et al. 2013). Besides, it is very plausible that the brain would evolve to minimize energetic costs of information processing, that is, maximize efficiency (Hilgetag and Goulas 2016; Bassett and Bullmore 2006).

Brain networks can be extracted using various neuroimaging measures: correlations of both functional (e.g. functional magnetic resonance imaging (fMRI), electroencephalography (EEG)) and structural data (e.g. diffusion tensor imaging (DTI), structural MRI). The focus of this article is on networks based on functional correlations from resting state fMRI (rsfMRI). In rsfMRI, subjects lie in the scanner at rest without performing a specific task so that the spontaneous activity of the brain can be measured. Brain regions are said to be functionally connected and therefore, exchanging information, if they present temporal correlation.

Typically, connectivity between areas or voxels is inferred using Pearson's correlation coefficient, resulting in a large connectivity matrix, sizing number of brain areas/voxels \* brain areas/voxels. The connectivity can serve as a distance function between brain regions, where high connectivity values are interpreted as short distance between pairs of nodes and vice-versa, low connectivity values are interpreted as long distances. Nevertheless, there is no consensus on how to treat these connectivity values, resulting in a major problem in brain network analysis: the network construction.

A common practice in network construction is to set a threshold to define where nodes are connected and where they are not. The most frequently used thresholds are called absolute and density threshold, where respectively, only connections that surpass a fixed given connectivity strength are kept or stronger connections matching a given density are kept. Once the full connectivity matrix is thresholded, the final network can be weighted or binary. Weighted networks are those where the connectivity strength is kept, whereas binary ones, the connectivity strength is set to 1. Binary networks are often used in neuroscience and therefore we focus our study in them. In addition the graph measurements, needed to compute the small-worldness, are also defined solely for the binary networks. Noteworthy is that, no thresholding method is unbiased, and there is no consensus about the optimal way to choose a certain

threshold. Usually an educated guess is recommended, based on the network features one aims to stress (van Wijk, Stam, and Daffertshofer 2010). The most frequently applied threshold in network neuroscience is the density threshold because it ensures that the networks are constructed with the same budget (the same wiring cost). However, networks emerging from this construction method do not always present the characteristic of being optimally efficient and resilient, namely the small-world and scale-free properties.

Here, we propose an alternative network construction method borrowed from machine learning, called k-nearest neighbor (kNN) graph construction (Luxburg 2007; Daitch, Kelner, and Spielman 2009). Numerous methods in machine learning are based on graphs. The fundamental issue of graph construction also emerges naturally: there is a set of items and a distance function between the pairs of items. The question is how to connect the items by edges and form a graph, in such a way that the resulting graph is informative for machine learning tasks such as clustering, classification, etc. This problem has been extensively explored in machine learning and various solutions were developed (Ng, Jordan, and Weiss 2002; Zhou et al. 2004; Daitch, Kelner, and Spielman 2009; Luxburg 2007). When using kNN to construct a network, the main difference from the density or absolute threshold is that the selection of links is based on local structure and similarities, that is, links that belong to the same neighborhood. In the kNN method, each node of the network is allowed to have exactly  $k$  connections to its nearest neighbours. In other words, for each node, the  $k$  strongest connections are kept. kNN had been previously used in neuroscience in algorithms for classification (Arbabshirani et al. 2013; Suárez Sánchez et al. 2014; Zhu et al. 2007), MRI segmentation of the brain (Gang et al. 2013), EEG-based assessment of neurophysiological changes (Kortelainen, Väyrynen, and Seppänen 2011) and in combination with minimum spanning tree graph construction (Alexander-Bloch et al. 2010).

In this paper, we formally introduce the kNN graph construction method for the construction of brain networks. Then, we show the advantages of this method. For that, we applied kNN to a big cohort of healthy participants (N=434) of the Berlin Aging Study II (BASE-II) subjected to resting state fMRI. We show that the kNN network construction leads to the optimally efficient network with small-world characteristics. Moreover, we show that these networks are scale-free, whereas the networks resulting from the traditional network construction do not show this property.

## **2. Methods**

### **2.1. Participants**

MR eligible participants were recruited within the Berlin Aging Study II (BASE-II, <https://www.base2.mpg.de/en>); for additional information about the cohort see (Bertram et al. 2014), resulting 438 healthy adults (20-80 years old). None of the participants was on medication that may have affected memory function or had a history of head injuries, medical (e.g., heart attack), neurological (e.g., epilepsy), or psychiatric disorders (e.g., depression). Four subjects were excluded due to excessive movement calculating using framewise displacement (FD) ( $FD > 0.5$  (Power et al. 2012)) during the MRI acquisition, resulting in a total of 434 healthy participants.

### **2.2. Data acquisition**

Brain images were collected on a Siemens Tim Trio 3T scanner (Erlangen, Germany) using a 12-channel head coil. A T1-weighted magnetization prepared gradient-echo sequence (MPRAGE) based on the ADNI protocol (TR=2500ms; TE=4.77ms; TI=1100ms, acquisition matrix=256×256×176; flip angle = 7°; 1x1x1mm<sup>3</sup> voxel size) was used to obtain structural

images. Whole brain functional resting state images were collected over a period of 10 minutes by using a T2\*-weighted EPI sequence sensitive to BOLD contrast (TR=2000ms, TE=30ms, image matrix=64×64, FOV=216mm, flip angle=80°, slice thickness=3.0mm, distance factor=20%, voxel size 3×3×3mm<sup>3</sup>, 36 axial slices). Participants were instructed to look at a fixation cross and relax during data acquisition.

### **2.3. Resting state fMRI preprocessing**

To ensure for steady-state longitudinal magnetization, the first 5 images were discarded. First, the acquired data was corrected for slice timing and then realigned, followed by coregistration between structural individual T1 images and functional images. Segmentation into gray matter, white matter, and cerebrospinal fluid were performed. Data was then spatially normalized to stereotactic space of the Montreal Neurological Institute (MNI) and spatially smoothed with a 6-mm full width at half maximum [FWHM]) to improve signal-to-noise ratio. Afterwards, motion and signals from white matter and cerebrospinal fluid were regressed. Data was filtered (0.01 – 1 Hz) to reduce physiological high-frequency respiratory and cardiac noise and low-frequency drift and, finally, detrended. All steps of data preprocessing were performed in MATLAB 2012b ([www.mathworks.com](http://www.mathworks.com)) using SPM12 (Wellcome Trust Centre for Neuroimaging, London, United Kingdom) except filtering that was applied using the REST toolbox (Song et al. 2011). Additionally, to control for the effect of motion on functional connectivity measures, the voxel-specific mean framewise displacement were calculated (FD; Power and colleagues (Power et al. 2012). Subjects with FD higher than the recommended threshold of 0.5 were excluded (N=4).

### **2.4. Functional connectivity matrix**



To build the resting state functional connectivity matrices, nodes and connectivity strength must be defined. Here the nodes were created based on the AAL atlas (Tzourio-Mazoyer et al. 2002), cerebellum was excluded, resulting in 90 regions of interest and therefore 90 nodes. The node-averaged time series were extracted for each subject using the REST toolbox (Song et al. 2011). As explained above, the BOLD signal was corrected for motion, white matter and cerebrospinal fluid, as nuisance signals of no interest. Connectivity strength was obtained using Pearson's correlation coefficient resulting in a 90 x 90 matrix. Connectivity values that were not statistically significant ( $p\text{-value} \geq 0.05$ ) were excluded.

## 2.5. Graph construction

The goal was to construct an informative binary graph from the connectivity matrix. Here we applied the kNN method to functional connectivity matrices constructed as described in 2.4. In addition, we compared graphs built with kNN to those built based on a density threshold, the most common method in neuroscience.

The density threshold method considers a particular threshold for each subject in such a way that only the highest connections are kept until the desired density is reached. In turn, the kNN method connects each node to its  $k$ -nearest neighbors. The  $k$ -nearest neighbors of a node are  $k$  nodes with the highest connectivity values to the node. This procedure can lead to non-symmetric graphs. When denoting the set of  $k$ -nearest nodes to node  $i$  by  $KNN(i)$ , if  $j \in KNN(i)$ , it does not necessarily mean that  $i \in KNN(j)$ . However, the connectivity matrix is symmetric, and it is desirable to have a symmetric graph as well. To solve this issue, we connected a pair of nodes  $i$  and  $j$ , if  $i \in KNN(j)$  or  $j \in KNN(i)$ . Note that the edges in the final graph have no weight and the final graphs in both density threshold and kNN methods are binary graphs.

In order to obtain a fair comparison, we built density threshold and kNN networks with the same number of edges, considering that the number of edges is the budget for graph construction. In other words, we assumed that the budget of the network to make connections (edges) was fixed and identical for both methods. The value of the threshold controls the number of connections in the density threshold graphs. The constant  $k$  plays the same role for the kNN graphs. By increasing  $k$ , we let each node have more connections and the total number of edges in the graph will increase. We chose the range of  $k$  and the density threshold, in a way that resulting graphs had the same number of edges.

Increasing  $k$  only by one drastically increases the number of edges in the graph, therefore we cannot fine-tune the number of edges in the kNN method. On the other hand, by slightly changing the threshold value of the density threshold graphs, we can obtain the desired number of edges. Therefore, we chose a fixed range for the parameter  $k$  of the kNN graphs. For each value of the parameter  $k$  and the resulting kNN graph we computed the total number of edges that the graph uses. Then, we tuned the threshold such that density threshold graphs also had the same number of edges. As the number of nodes was fixed, the average degree of graphs was the same across the two methods.

In order to thoroughly understand the behavior of graph constructions we checked a wide range of the parameter  $k \in \{6,7, \dots, 60\}$  equivalent to density thresholds from 0.14 to 0.80. In this way, we investigate a range from very sparse to almost fully connected graphs.

## **2.6. Graph Theory analysis**

Please note that both graph construction methods in section 2.5 produce binary (unweighted) graphs and the further analysis of small-world and scale-free property is defined only for binary graphs.

**Graph density:** Density of a graph with  $n$  nodes and  $E$  edges is defined as

$$\frac{2E}{n(n-1)}$$

**Degree:** Degree of a node is defined as the number of connected edges to the node.

**Clustering coefficient:** Clustering coefficient of a node was first introduced by Watts and Strogatz (Watts and Strogatz 1998). The clustering coefficient of the node  $i$  is defined as the number of triangles that it makes with its neighbors divided by the maximum possible triangles. If node  $i$  has  $k$  neighbors, then it can form at most  $k(k-1)/2$  triangles with its neighbors. Let  $\Delta(i)$  denotes the number of triangles that node  $i$  makes with its neighbors, then the following shows the formal definition of the clustering coefficient:

$$C(i) = \frac{2\Delta(i)}{k(k-1)}$$

The average clustering coefficient of a graph is defined as the arithmetic mean of the clustering coefficients of nodes in the graph.

**Characteristic path length:** If  $d(i, j)$  denotes the length of the shortest path between nodes  $i$  and  $j$ , then the characteristic path length of a graph is defined as the average of distance between all pairs of nodes, formally:

$$SP = \frac{2 \sum_{1 \leq i < j \leq n} d(i, j)}{n(n-1)}$$

Where  $n$  denotes the number of nodes in the graph. If two nodes are not connected, then we set their shortest path length to the maximum shortest path length between connected pairs of nodes in the graph.

## 2.7 Small-world network evaluation

The small-worldness of a graph is defined using the average clustering coefficient and the characteristic path length of the graph divided by the equivalent measures in a random graph of the same density. The comparison to the random graph is defined by the following two parameters:

$$\gamma = \frac{CC}{CC_r} \quad , \quad \lambda = \frac{SP}{SP_r}$$

Where,  $CC$  and  $SP$  respectively denote the average clustering coefficient and the characteristic path length of the network and  $CC_r$  and  $SP_r$  are the same quantities for the random graph with the same density. The small-worldness is defined as the following ratio:

$$\rho = \frac{\gamma}{\lambda}$$

Initially, Watts and Strogatz proposed to use Erdős–Rényi random graphs as a base random model (Watts and Strogatz 1998). In the neuroscience literature, however, it is suggested to compare random graphs with the same degree distribution to have a fair definition of small-worldness (Sporns and Zwi 2004; Maslov et al. 2002). Therefore, we constructed 10 random surrogate networks with the same degree distribution as the initial network. The parameters  $CC_r$  and  $SP_r$  are the average over the set of 10 random graphs. Similar to other small-world

studies in neuroscience (Fornito, Zalesky, and Bullmore 2016; van den Heuvel et al. 2008; Eguíluz et al. 2005), each random surrogate network was built by the following procedure:

We started with the initial network and performed a random rewiring step on edges which preserves the degree of nodes. Next, we chose a set of 4 nodes at random  $\{i, j, k, l\}$ , such that the edges  $i \rightarrow j$  and  $k \rightarrow l$  exist but the edges  $i \rightarrow l$  and  $k \rightarrow j$  do not exist. Note that if the network is very sparse or densely connected, it is not possible to find a quadruple of nodes with the above property. Since we had networks out of these two extreme cases, we could always find a proper set of four nodes. Then we removed the two existing edges  $i \rightarrow j$  and  $k \rightarrow l$  and formed two new edges  $i \rightarrow l$  and  $k \rightarrow j$ . In this way the connections are randomized but the degree of involved nodes is not changed. The rewiring was done 1000 times with random quadruple of nodes. Considering the number of nodes (90), this number of rewiring steps was enough to produce a random graph.

## 2.8 Scale-free network evaluation

If  $p(d)$  represents the empirical probability of having a node with degree  $d$ , then the degree distribution of a scale free network should follow a power-law, meaning that  $p(d) = \beta d^{-\alpha}$ . In this formulation,  $\alpha$  is a constant parameter which we refer to as the exponent, and  $\beta$  is a constant to make sure the function represents a probability mass function. In scale-free networks the exponent range should be  $2 < \alpha < 3$  (Girvan et al. 2007; Choromański, Matuszak, and Miękiś 2013).

The typical approach to investigate the presence of power-law degree distribution is through the log-log plot representation of the degree distribution. A linear log-log plot indicates a power-law distribution. The R-squared of the linear regression can represent how much of the variance can be expressed by a linear model. If the R-squared is close to one, we can conclude

a power-law degree distribution, and thus the presence of scale-free network. Moreover, the slope of the log-log plot represents the exponent parameter  $\alpha$ , which needs to be in the desirable range ( $2 < \alpha < 3$ ) for scale-free networks.

One can obtain the log-log degree distribution plot for each subject and calculate the R-squared value and the slope of the regression curve. However, we intended to also illustrate the average distributions over subjects. Therefore, we presented the results of the scale-free evaluation in two parts. First, we illustrated the average (over subjects) results, then we reported the estimated values of parameters for the whole sample.

Here we explain the procedure to compute the average degree distribution over the subjects: Let assume we estimate the parameters  $\alpha$  and  $\beta$  by a linear regressor, for each subject. We denote the estimation of parameters for the subject  $i$  by  $\hat{\alpha}_i$  and  $\hat{\beta}_i$ , which are respectively the slope and the intercept of the estimated linear regressor for the log-log degree distribution plot of subject  $i$ . The two parameters can vary significantly among subjects. Some studies compute the average distribution over all subjects in order to estimate a final  $\alpha$  (van den Heuvel, Stam, Boersma, & Hulshoff Pol, 2008). In an arithmetic mean over the degree distributions, the subjects who have higher  $\beta$  values have more contribution to the final mean. In this way the final estimation of  $\alpha$  can be biased. Here we propose a slightly different approach to overcome this issue. We do not compute the average over distributions, but instead, we compute the average over logarithm value of distributions. Note that by taking the average over  $\log p_i(d)$  values, the final plot will represent:  $\frac{1}{n} \sum_{i=1}^n \log p_i(d) = \frac{1}{n} \sum_{i=1}^n \log \beta_i + \frac{1}{n} \sum_{i=1}^n \alpha_i \log d$ . The slope of this plot is the average over the slope values of subjects,  $\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n \alpha_i$ .

### 3. Results

In this section, we reported the results of two graph construction methods: density threshold and kNN. The comparison between the two construction methods was done based on two criteria. The first criterion was small-worldness, as efficient networks show small-world properties (Hilgetag and Goulas 2016; Bassett and Bullmore 2006). Secondly, we checked whether the two graph construction methods can produce scale-free networks which is related to resilience of the network (Cohen et al. 2000; Albert, Jeong, and Barabási 2000; Albert and Barabási 2002).

### 3.1. Small-world comparison

First, we reported the two parameters, clustering coefficient ratio  $\gamma$  and characteristic path length ratio  $\lambda$  (Fig 1) for both graph construction methods. The average value of the parameters was calculated over all subjects and the standard deviation was depicted by means of error bars. Note that in each vertical point of the plot the parameter  $k$  and the density, chosen for the two methods, lead to graphs with the same number of edges, hence the comparison between the two graph construction methods can be considered fair. Moreover,  $\gamma$  and  $\lambda$  were calculated using random graph with the same density.

The clustering coefficient ratio  $\gamma$  (Fig 1, left) was significantly higher than 1 for the kNN-built graphs (in blue) across all thresholds, meaning more clusters than those for random graphs, especially for  $k$  values between 8 and 30. While the density threshold graphs presented clustering coefficient ratio close to 1 and below those for the kNN method. This suggests having more well-clustered graphs with the kNN method as compared to density threshold. The characteristic path length ratio  $\lambda$  (Fig 1, right) for the kNN-built graphs showed longer path lengths than random graphs for  $k < 23$ , however it showed nearly equal characteristic

path lengths to the random graphs for  $k > 23$ . Density graphs produced longer characteristic path length than kNN-built ones. Both parameters suggest that the kNN method will produce networks which have higher small-worldness.

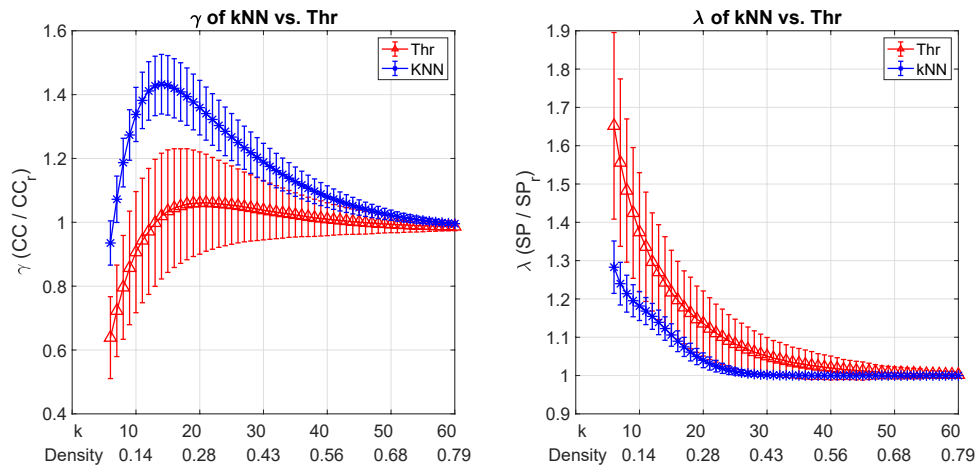


Fig. 1. Comparison of parameters ( $\gamma$ ) and ( $\lambda$ ) on the left and right respectively, for two graph construction methods: density threshold (Thr in red) and the kNN (in blue). The x-axis of the plots denotes the parameter  $k$  and density for the kNN and density threshold methods respectively. The errorbars show the standard deviation over subjects. The clustering coefficient ratio ( $\gamma$ ) is greater for the kNN graph, while the characteristic path ratio ( $\lambda$ ) is smaller.

As next step, we calculated the small-worldness ( $\rho$ ) (Fig 2). The kNN-built graphs showed on average higher small-worldness than density threshold ones, especially for  $k < 30$ . The most efficient networks were constructed with  $k \approx 18$ , showing the highest  $\rho$ . Moreover, the error-



bars indicate less variability in the kNN method, that is, a consistent behavior of the kNN method over the set of subjects.

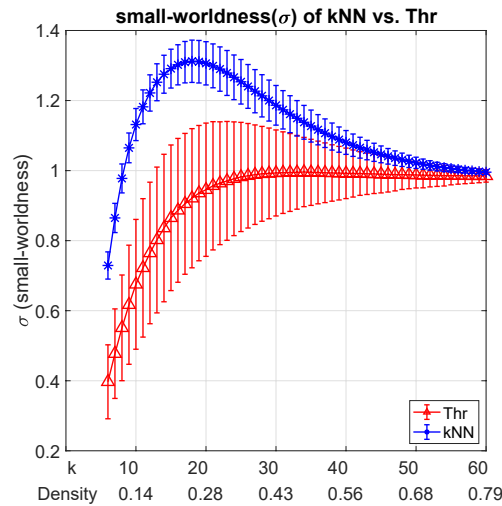


Fig. 2. The comparison of small-worldness ( $\rho$ ) for two graph construction methods, density threshold (in red) and the kNN (in blue). The parameters  $k$  and density are shown in the x-axis of the plots. The error-bars denote the standard deviation of the small-worldness with different subjects. The kNN graphs showed significantly higher small-worldness, in particular for  $\approx 18$ , in comparison to density-thresholded graphs.

### 3.2. Scale-free comparison

We plotted the average (over subjects) logarithm of probability mass functions versus the logarithm of degree, for  $k$  values by increasing in steps of 10 (Fig. 3). The top row shows log-log plots of the kNN methods, while the bottom row depicts the same plots for the density threshold.

In all log-log plots for the kNN method a linear fit was very close to the actual points. Moreover, the R-squared values (denoted at the top of each plot) were close to one. This result suggests that the degree distribution is a power-law  $p(d) = \beta d^{-\alpha}$ . In addition, except for very

sparse graphs ( $k = 6$ ), the whole range of kNN graphs had estimated exponents in the range of  $2 < \alpha < 3$  that characterizes scale-free networks. In comparison, the density threshold method did not show a linear behavior in the log-log plots (Fig. 3 - bottom row), particularly with increasing threshold values. Moreover, the estimated exponent  $\alpha$  failed to fall within the desirable range of scale-free networks.

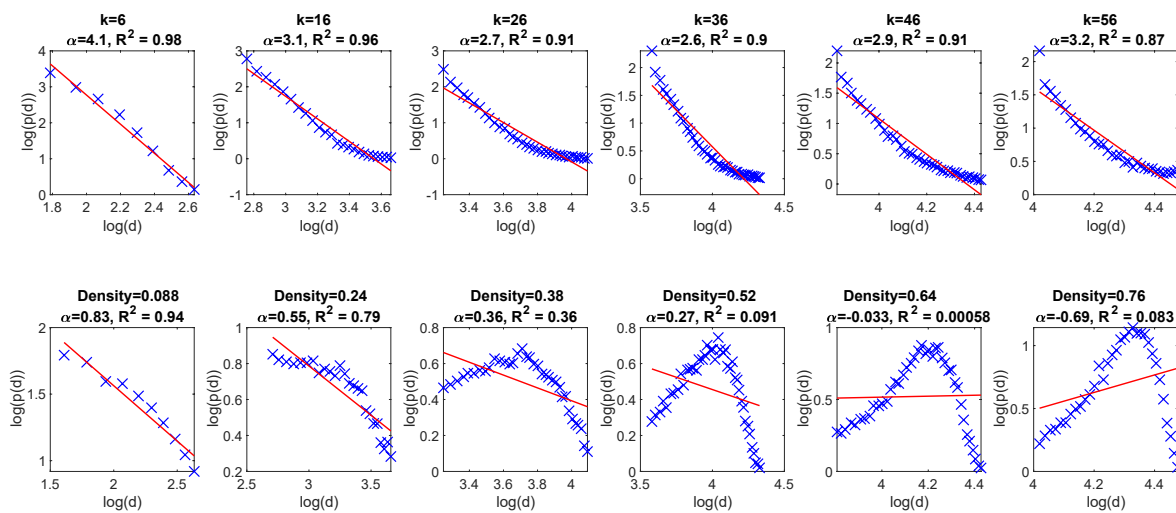


Fig. 3. The log-log plots of the average distributions over the subjects. The x-axis is the logarithm of the degree, while the y-axis is the logarithm of the empirical probability of having a node with this degree. The first row shows the results of kNN graph construction method, while the second row shows the results of density threshold graphs. The parameter  $k$  and the density are denoted at the top of each plot. In each column, the graphs have the same number of edges. The R-squared value of the regression and the estimated slope  $\alpha$  is also shown at the top of each plot. The kNN graphs (top row) fit to a linear estimation, moreover, they have exponents ( $\alpha$ ) in the range of scale-free networks. The density threshold graphs are not linear for most thresholds.

We also reported the summary statistics of the scale-free test for the whole sample. We computed the R-squared of the linear regression for each subject with the fixed graph construction parameters  $k$  and density. Moreover, the exponent  $\alpha$  was estimated based on a linear regression. Fig. 4 (top row) shows the population of R-squared values for the kNN and density threshold methods. We showed see that the kNN method, especially in case of sparse graphs ( $k < 20$ ) presented high R-squared values, which suggests a proper linear fit to the plot. Conversely, the R-squared values were mostly lower than 0.5 for the density threshold method, which means that a linear regressor cannot express the function depicted in the log-log plot. The bottom row of Fig. 4 depicts the population of exponent  $\alpha$  for the two methods. Again, for the kNN method most of the subjects had exponents in the desirable range while the density threshold method failed to do so.

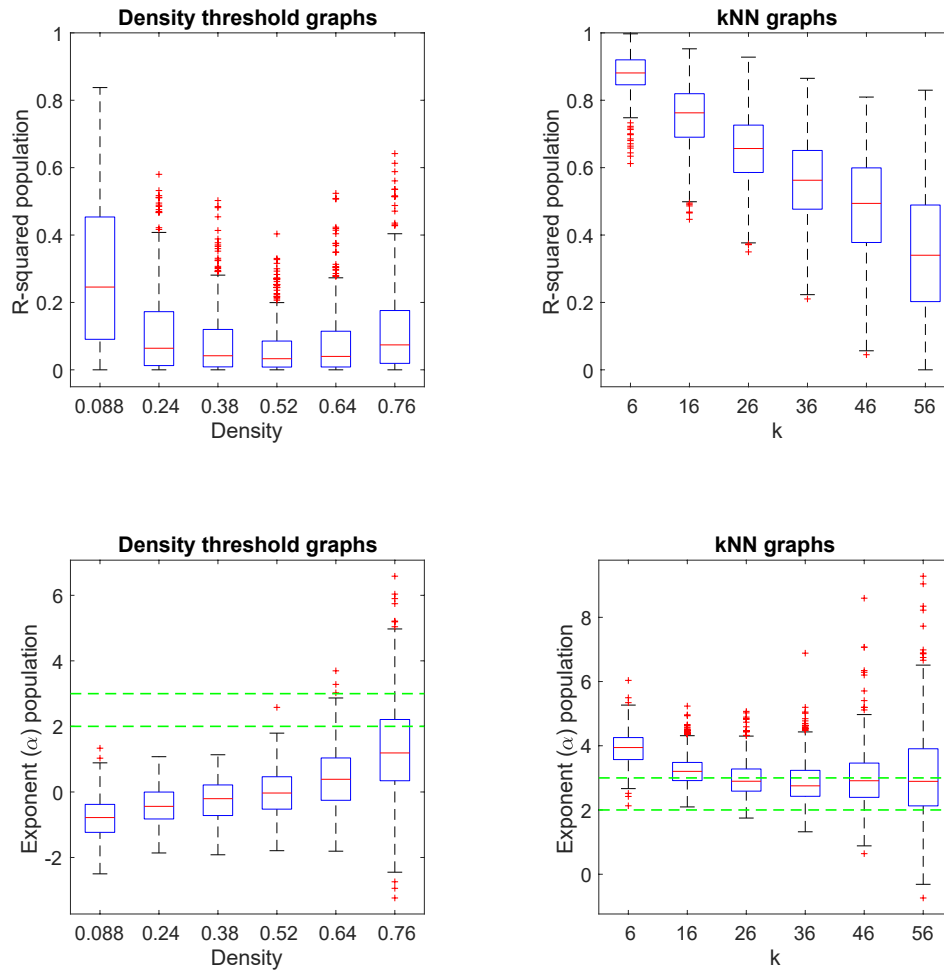


Fig. 4. The box-plot of the regression output values, R-squared and exponent  $\alpha$ . Each box represents the population of R-squared or the estimated parameter  $\alpha$  for the regression of log-log plot. The central mark in the box is the median, the edges of the box are the 25th and 75th percentiles, while whiskers goes to the maximum and minimum values. Outliers are depicted as red plus signs. The R-squared values are significantly higher for the kNN graphs, suggesting a better linear fit. In addition, the exponent ( $\alpha$ ) is in the range of scale-free network for kNN graphs, with  $k > 20$ , while this is never the case for density threshold method.

### 3.3. Comparison of networks constructed using KNN and density

Here we present an illustration of the two graph construction methods by plotting the final brain network of a single subject when KNN is applied in contrast to density threshold method (Fig. 5). The parameter  $k$ , was chosen to be  $k = 16$  for the KNN graph. The corresponding density threshold, for the density threshold graphs, was chosen such that both methods produced the same number of edges (same wiring budget). The choice of  $k = 16$  was based on the highest small-worldness (Fig. 2).

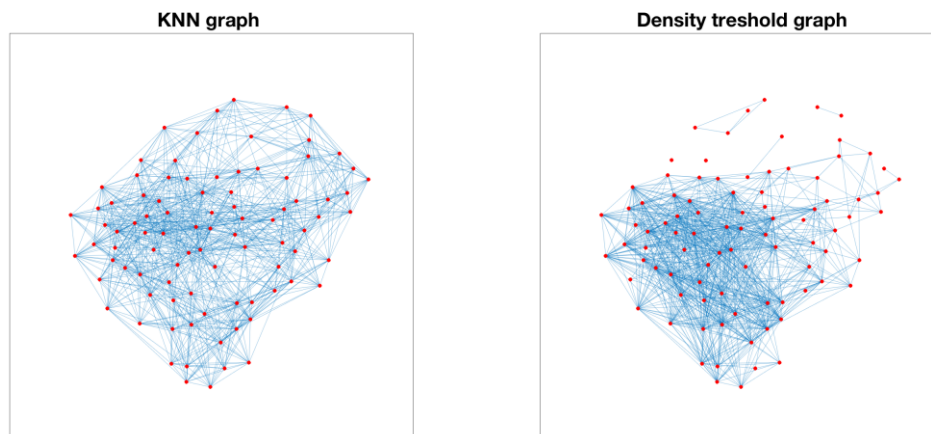


Fig. 5. The KNN (left) and density threshold (right) graphs for a single subject.

The main difference observed in the two network construction methods was related to the total interconnectivity of the brain areas (nodes). The graph built using the density threshold method (Fig. 5 - right) is not connected, thus nodes without any edge connection were observed. On the other hand, the KNN graph (Fig. 5 - left) is connected.

The same interconnectivity pattern were seen for other subjects and for the whole range of  $k$  parameter values. The reason for such pattern relies on the fact that the density threshold graph is intrinsically prone to produce unconnected components. In brief, the density threshold method uses a global threshold to make edges. In this way, if some of the nodes have systematically weaker connections they cannot form any edge. In other words, if the scale of connectivity values differs among the different regions of the brain (nodes), the density threshold method gives no opportunity to some areas of the brain that have weaker connection, to form an edge. A non-connected graph might not be a favorable structure for the human brain, as some areas of the brain would be isolated and therefore not taken part in the global information processing. In contrast, the KNN method ensures that every node is connected to  $k$  neighbors. This condition leads to connected graphs for values of  $k$  which are not extremely small. For a detailed discussion please see (Luxburg 2007).

#### **4. Discussion**

Our brain is a complex organ with different levels of interaction and connections, from local cellular action potentials to information travelling across the body in our nervous system. The brain network is formed by interconnected areas called nodes and their connections are known as links. Due to the nature of the interactions in the brain, the optimal network should be able to process locally specialized information as well as to connect localized information centers between one another, so that information can travel globally and fast in the brain. A network with such features is said to present small-world characteristics (Watts and Strogatz 1998). Small-world networks work with maximal efficiency and therewith low energetic cost (Bassett and Bullmore 2006). These optimal networks are reflected topologically by having high clustering and short paths connecting the clusters. Beside the efficient organization of networks based on small-world properties, the resilience of the brain network is also desirable. A resilient

network is less vulnerable to random perturbation or loss of connections. Such networks exhibit power-law degree distribution and are called scale-free networks.

In practice, brain networks can be inferred using multiple neuroimaging measures such as fMRI, EEG, fiber tracts in DTI, etc. Here, we focus on rsfMRI. In rsfMRI we infer a functional brain network at rest, in which, nodes are voxels or anatomical brain areas and links are the connectivity between the nodes, resulting in a fully connected matrix. In network analysis, connectivity is interpreted as a distance function between nodes: higher connectivity represents short distances, that is, a closer relationship implies strong connectivity. Nevertheless, there is no consensus on how to treat these connectivity values: a fully connected graph may contain spurious connections, e.g. due to noise in data acquisition (Fornito, Zalesky, and Bullmore 2016). However, which connections should be included? This question leads to a major problem in brain network analysis, the network construction.

It is common practice in neuroscience to construct a network, by using global thresholds, such as absolute and density thresholds, the latter being the most popular. When applying a density threshold, the number of links (and consequently the density) is fixed and the stronger connections are kept until the desired density is achieved and the remaining links are set to zero. Considering groups of subjects, the number of connections is the same across subjects which is an advantage when comparing topological properties between groups, however, the final network after thresholding will depend on average strength, for example, subjects with lower average fully connected matrix will have low-weight connectivity links, whereas, for subjects with high average fully connected matrix, only high-weight links will be included. Therefore, when applying global thresholds where nodes are treated the same, as the case for the density threshold, if the connectivity values of specific brain areas are systematically

weaker than those for other areas, global thresholds give no opportunity to that particular nodes to make connections in the final graph.

Thresholding in brain network construction is addressed in detail in (van Wijk, Stam, and Daffertshofer 2010). As a final consideration, there is no consensus on network construction methods as there is no unbiased thresholding. Therefore the thresholding method should be chosen carefully by the researcher and should reflect the network features one desires to stress (van Wijk, Stam, and Daffertshofer 2010).

Here we proposed an alternative thresholding method, called kNN (Luxburg 2007; Daitch, Kelner, and Spielman 2009). kNN is used in machine learning, where the problem of graph construction also emerges. kNN connects items that belong to the same neighborhood by allowing the k nearest neighbors of each node. The remaining links, that are not neighbors, are set to 0. In neuroimaging, the connectivity strength can be also thought as the distance function between brain areas. In rsfMRI more specifically, it represents a functional distance and, therefore, is related to the exchange of information in the brain: regions exchanging information have high connectivity values. In turn, spurious connections would present lower connectivity since there is no real relationship between regions. Thus, brain areas can assume the role of the items that must be grouped based on proximity defined by exchange of information.

In this paper, we applied kNN and density thresholding methods to functional connectivity matrices from rsfMRI of a big cohort of healthy subjects. We showed that networks built using kNN method presented small-world and scale-free properties, that is efficient and resilient networks. Our interpretation is that kNN method can preserve key connections and yet reduce



the average shortest path length. Shortest paths are essential to the efficient functionality of any network promoting fast communication among specialized centers of information processing.

Regarding efficiency and resilience in the brain, it has been argued that it is plausible that the brain would evolve to minimize energetic costs of information processing (Hilgetag and Goulas 2016; Bassett and Bullmore 2006). In addition, a brain working in small-world scale-free regime would have advantages as efficiency, rapid adaptive reconfiguration and resilience which are essential features for plasticity and cognition (Baronchelli et al. 2013) as well as neurological diseases as stroke and dementia. These advantages would be possible because a failure in a network with such an architecture would most likely affect a peripheral node rather than a hub, as a consequence neither the mean shortest path is strongly increased nor the clustering coefficient decreases. In short, in case of a random failure in the system, e.g. a removal of a node or a malfunctioning given an external or internal perturbation, the information can still reach its final destination in a fast way through other paths. A potential weakness of small-world scale-free networks is the sensitivity in case of targeted attacks, e.g. a hub in the network.

## **5. Conclusion**

We proposed a new method to construct brain networks based on k-nearest neighbors (kNN). We showed that kNN-built networks presented small-world and scale-free properties, while the most frequently used method based on density thresholding, did not show them to a significant level: decreased small-worldness in all threshold range as compared to kNN and no scale-free. Considering the small-worldness and scale-free experiments in a large cohort with healthy subjects,  $k \sim 18$  is advisable to construct networks with maximal efficiency.

## **6. Acknowledgements**

The authors thank Ulrike von Luxburg and Debarghya Goshdastidar for the discussions and comments.

## 7. Funding

This work has been supported by the German Research Foundation DFG (SFB 936/ Z3 to S.H., SFB 936/C7 to C.G.F. and S.K. and DFG KU 3322/1-1 to S.K.), the European Union (ERC-2016-StG-Self-Control-677804 to S.K.), and a Fellowship from the Jacobs Foundation (JRF 2016-2018 to S.K.)

The Authors have declared that there are no conflicts of interest.

## 8. References

Albert, Réka, and Albert-László Barabási. 2002. “Statistical Mechanics of Complex Networks.”

*Reviews of Modern Physics* 74 (1): 47–97. <https://doi.org/10.1103/RevModPhys.74.47>.

Albert, Réka, Hawoong Jeong, and Albert-László Barabási. 2000. “Error and Attack Tolerance of Complex Networks. *Nature*, 406: 378–482, 2000.” *Nature* 406 (6794): 378–82.

<https://doi.org/10.1038/35019019>.

Alexander-Bloch, Aaron F., Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois

Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T. Bullmore. 2010. “Disrupted Modularity and Local Connectivity of Brain Functional Networks in Childhood-Onset Schizophrenia.”

*Frontiers in Systems Neuroscience* 4 (October): 1–16. <https://doi.org/10.3389/fnsys.2010.00147>.

Amaral, L A N, A Scala, M Barthelemy, and H E Stanley. 2000. “Classes of Small-World Networks.”

*Proceedings of the National Academy of Sciences USA* 97 (21): 11149–52.

<https://doi.org/10.1073/pnas.200327197>.

Arbabshirani, Mohammad R., Kent A. Kiehl, Godfrey D. Pearlson, and Vince D. Calhoun. 2013.

“Classification of Schizophrenia Patients Based on Resting-State Functional Network Connectivity.” *Frontiers in Neuroscience* 7 (7 JUL): 1–16.

<https://doi.org/10.3389/fnins.2013.00133>.

Baronchelli, Andrea, Ramon Ferrer-i-Cancho, Romualdo Pastor-Satorras, Nick Chater, and Morten H.

Christiansen. 2013. “Networks in Cognitive Science.” *Trends in Cognitive Sciences* 17 (7): 348–60. <https://doi.org/10.1016/j.tics.2013.04.010>.

Bassett, Danielle Smith, and Ed Bullmore. 2006. “Small-World Brain Networks.” *The Neuroscientist*

12 (6): 512–23. <https://doi.org/10.1177/1073858406293182>.

Bertram, Lars, Anke Böckenhoff, Ilja Demuth, Sandra Düzel, Rahel Eckardt, Shu-Chen Li, Ulman

Lindenberger, et al. 2014. “Cohort Profile: The Berlin Aging Study II (BASE-II)†.”

*International Journal of Epidemiology* 43 (3): 703–12. <https://doi.org/10.1093/ije/dyt018>.

Bullmore, Ed, and Olaf Sporns. 2009. “Complex Brain Networks: Graph Theoretical Analysis of

Structural and Functional Systems.” *Nature Reviews Neuroscience* 10 (3): 186–98.

<https://doi.org/10.1038/nrn2575>.

Choromański, Krzysztof, Michał Matuszak, and Jacek Miękiś. 2013. “Scale-Free Graph with

Preferential Attachment and Evolving Internal Vertex Structure.” *Journal of Statistical Physics*

151 (6): 1175–83. <https://doi.org/10.1007/s10955-013-0749-1>.

Cohen, Reuven, Keren Erez, Daniel Ben-Avraham, and Shlomo Havlin. 2000. “Resilience of the

Internet to Random Breakdowns.” *Physical Review Letters* 85 (21): 4626–28.

<https://doi.org/10.1103/PhysRevLett.85.4626>.

Daitch, Samuel I., Jonathan A. Kelner, and Daniel A. Spielman. 2009. “Fitting a Graph to Vector

Data.” *Proceedings of the 26th Annual International Conference on Machine Learning - ICML*

'09, 1–8. <https://doi.org/10.1145/1553374.1553400>.

Eguíluz, Victor M., Dante R. Chialvo, Guillermo A. Cecchi, Marwan Baliki, and A. Vania Apkarian.

2005. “Scale-Free Brain Functional Networks.” *Physical Review Letters* 94 (1): 1–4.

<https://doi.org/10.1103/PhysRevLett.94.018102>.

Fornito, Alex, Andrew Zalesky, and Edward T Bullmore. 2016. “Fundamentals of Brain Network Analysis - 1st Edition.” 2016.

Gang, Zhang, Zhou Dan, Huang Ying, Huang Xiaobo, Zhang Yong, Luo Weishi, and Zhang Yang.

2013. “An Unsupervised Method for Brain MRI Segmentation.” *International Journal of Emerging Technology and Advanced Engineering* 3 (10): 8–13.

Girvan, M., MEJ. Newman, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási.

2007. “Community Structure in Social and Biological Networks.” *PNAS* 99 (12): 7821–26.

<https://doi.org/10.1073/pnas.122653799>.

Heuvel, M. P. van den, C. J. Stam, M. Boersma, and H. E. Hulshoff Pol. 2008. “Small-World and Scale-Free Organization of Voxel-Based Resting-State Functional Connectivity in the Human Brain.” *NeuroImage* 43 (3): 528–39. <https://doi.org/10.1016/j.neuroimage.2008.08.010>.

Hilgetag, Claus C, and Alexandros Goulas. 2016. “Is the Brain Really a Small-World Network?”

*Brain Structure and Function* 221 (4): 2361–66. <https://doi.org/10.1007/s00429-015-1035-6>.

Kortelainen, Jukka, Eero Väyrynen, and Tapio Seppänen. 2011. “Isomap Approach to EEG-Based Assessment of Neurophysiological Changes During Anesthesia.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19 (2): 113–20.

<https://doi.org/10.1109/TNSRE.2010.2098420>.

Luxburg, Ulrike Von. 2007. “A Tutorial on Spectral Clustering A Tutorial on Spectral Clustering.”

*Statistics and Computing* 17 (March): 395–416.

Maslov, S., Kim Sneppen, B. Charloreaux, M. Galli, S. J. Pevzner, M. Tasan, Y.-Y. Ahn, et al. 2002.

“Specificity and Stability in Topology of Protein Networks.” *Science* 296 (5569): 910–13.

<https://doi.org/10.1126/science.1065103>.

Ng, Andrew Y., Michael I. Jordan, and Yair Weiss. 2002. “On Spectral Clustering: Analysis and an

Algorithm.” *Advances in Neural Information Processing Systems* 14, 849–56.

<https://doi.org/10.1.1.19.8100>.

Power, Jonathan D, Kelly A Barnes, Abraham Z Snyder, Bradley L Schlaggar, and Steven E Petersen.

2012. “Spurious but Systematic Correlations in Functional Connectivity MRI Networks Arise from Subject Motion.” *NeuroImage* 59 (3): 2142–54.

<https://doi.org/10.1016/j.neuroimage.2011.10.018>.

Rubinov, Mikail, and Olaf Sporns. 2010. “Complex Network Measures of Brain Connectivity: Uses and Interpretations.” *NeuroImage* 52 (3): 1059–69.

<https://doi.org/10.1016/j.neuroimage.2009.10.003>.

Song, Xiao-Wei, Zhang-Ye Dong, Xiang-Yu Long, Su-Fang Li, Xi-Nian Zuo, Chao-Zhe Zhu, Yong

He, Chao-Gan Yan, and Yu-Feng Zang. 2011. “REST: A Toolkit for Resting-State Functional Magnetic Resonance Imaging Data Processing.” *PloS One* 6 (9): e25031.

<https://doi.org/10.1371/journal.pone.0025031>.

Sporns, O, and JD Zwi. 2004. “The Small World of the Cerebral Cortex.” *Neuroinformatics* 2.

Suárez Sánchez, A., F. J. Iglesias-Rodríguez, P. Riesgo Fernández, and F. J. de Cos Juez. 2014.

“Applying the K-Nearest Neighbor Technique to the Classification of Workers According to Their Risk of Suffering Musculoskeletal Disorders.” *International Journal of Industrial*

*Ergonomics*. <https://doi.org/10.1016/j.ergon.2015.09.012>.

Telesford, Qawi K., Karen E. Joyce, Satoru Hayasaka, Jonathan H. Burdette, and Paul J. Laurienti.

2011. "The Ubiquity of Small-World Networks." *Brain Connectivity* 1 (5): 367–75.

<https://doi.org/10.1089/brain.2011.0038>.

Tzourio-Mazoyer, N., B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. 2002. "Automated Anatomical Labeling of Activations in SPM Using a

Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain." *NeuroImage* 15

(1): 273–89. <https://doi.org/10.1006/nimg.2001.0978>.

Watts, D J J, and S H H Strogatz. 1998. "Collective Dynamics of 'small-World' Networks." *Nature*

393 (6684): 440–42. <https://doi.org/10.1038/30918>.

Wijk, Bernadette C.M. van, Cornelis J. Stam, and Andreas Daffertshofer. 2010. "Comparing Brain

Networks of Different Size and Connectivity Density Using Graph Theory." *PLoS ONE* 5 (10).

<https://doi.org/10.1371/journal.pone.0013701>.

Zhou, Dengyong, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf.

2004. "Learning with Local and Global Consistency." *Nips* 1: 8. <https://doi.org/10.1.1.115.3219>.

Zhu, Mu, Wenhong Chen, John P. Hirdes, and Paul Stolee. 2007. "The K-Nearest Neighbor

Algorithm Predicted Rehabilitation Potential Better than Current Clinical Assessment Protocol."

*Journal of Clinical Epidemiology* 60 (10): 1015–21.

<https://doi.org/10.1016/j.jclinepi.2007.06.001>.