Brain functional and effective connectivity based on electroencephalography recordings: A review

Jun Cao1 | Yifan Zhao1 | Xiaocai Shan1,2 | Huali-ling Wei3 | Yuzhu Guo4 | Liangyu Chen5 | John Ahmet Erkoyuncu1 | Ptolemaios Georgios Sarrigiannis6

1School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield
2Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China
3Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK
4School of Automation Science and Electrical Engineering, University of Beijing, Beijing, China
5Department of Neurosurgery, China Medical University, Shenyang, China
6Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

Correspondence
Yifan Zhao, School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield MK43 0AL, UK.
Email: yifan.zhao@cranfield.ac.uk

Funding Information
National Natural Science Foundation of China, Grant/Award Number: 61876015; Beijing Natural Science Foundation, China, Grant/Award Number: 4202040

Abstract
Functional connectivity and effective connectivity of the human brain, representing statistical dependence and directed information flow between cortical regions, significantly contribute to the study of the intrinsic brain network and its functional mechanism. Many recent studies on electroencephalography (EEG) have been focusing on modeling and estimating brain connectivity due to increasing evidence that it can help better understand various brain neurological conditions. However, there is a lack of a comprehensive updated review on studies of EEG-based brain connectivity, particularly on visualization options and associated machine learning applications, aiming to translate those techniques into useful clinical tools. This article reviews EEG-based functional and effective connectivity studies undertaken over the last few years, in terms of estimation, visualization, and applications associated with machine learning classifiers. Methods are explored and discussed from various dimensions, such as either linear or nonlinear, parametric or nonparametric, time-based, and frequency-based or time-frequency-based. Then it is followed by a novel review of brain connectivity visualization methods, grouped by Heat Map, data statistics, and Head Map, aiming to explore the variation of connectivity across different brain regions. Finally, the current challenges of related research and a roadmap for future related research are presented.

KEYWORDS
artificial intelligence, brain association, electroencephalogram, machine learning, survey

1 INTRODUCTION

In the past decades, plenty of prior researchers showed their interest in understanding brain activities through analyzing noninvasive brain signals. Scalp electroencephalography (EEG), has been yearly increasingly attractive resulting in a growing number of publications (Van Diessen et al., 2015). On one hand, to extract hidden information from EEG recordings, various traditional digital signal processing methods were employed, such as independent component analysis (ICA; van Mierlo et al., 2014b; Tafreshi, Daliri, & Ghodousi, 2019; van Mierlo et al., 2014a; Vecchio, Miraglia, Bramanti, & Rossini, 2014), power spectral density (PSD; Erra, Velazquez, & Rosenblum, 2017; Ko,
Figure 1: Number of publications per year from PubMed search using keywords “EEG and Machine learning” or “EEG and AI” (Trend1) and “brain connectivity and EEG” (Trend2) in the period 2005–2020.
applications, and particularly their associated ML approaches. Moreover, there is no review available on a variety of visualization methods to represent the variation of functional and effective connectivity.

Section 2 of this article systematically reviews brain functional and effective connectivity estimation and analysis methods in accordance with their properties and applied areas. Those popular methods to quantify brain connectivity are grouped as parametric and nonparametric techniques and their conceptual and mathematical details are represented in the Appendix. Various distinct approaches of connectivity visualization, grouped into three categories, are reviewed in Section 3. Section 4 provides a summary and critical evaluation of brain connectivity methods and their applications in the fields of neural disease and brain activity research. We then discuss the role of visualization for brain connectivity analysis in detail. Besides, a discussion on the significance of ML and its various effective combination with brain connectivity estimations are represented in Section 4. This section closes with the identified research gaps and future direction. The conclusions are given in Section 5.

2 | ESTIMATION

A review of high-frequently used methods to estimate brain connectivity based on EEG is given in Table 1 associated with their properties. These methods are either linear or nonlinear, parametric, or nonparametric, belong to functional or effective connectivity, in time, frequency or time-frequency domain. The conceptual and mathematical details of these approaches are represented individually in the Appendix. It can be observed from this table that the functional connectivity is usually estimated by nonparametric methods while the effective connectivity estimates are based on parametric techniques. In this review, parametric and nonparametric methods are grouped separately.

2.1 | Nonparametric methods

2.1.1 | Linear methods

Linear intercommunication of the brain has been studied for several decades and is relatively straightforward in terms of computation and interpretation (Van Diessen et al., 2015). Correlation measures are well-accepted to represent linear connections. Pearson correlation coefficient (PCC) and cross-correlation were applied in the time domain to estimate functional connectivity from multi-channel EEG recordings (Fagerholm, Moran, Violante, Leech, & Friston, 2020; Handojoso et al., 2013; Ibrahim et al., 2018; Lee & Hsieh, 2014; Tarokh, Carskadon, & Achermann, 2010; Vortmann, Kroll, &...
Putze, 2019). It should be noted that PCC is the cross-correlation value at the zero time lag. In other words, if there is any time delay between the two EEG signals, PCC cannot accurately represent the strength of linear correlation. The counterpart of the cross-correlation in the frequency domain is the coherence. Coherence is sensitive to both changes in power and phase relationships between EEG signals (Sakkalis, 2011). The magnitude squared coherence (MSC: Battaglia & Brovelli, 2019) and the imaginary part of coherence (Ewald, Marzetti, Zappasodi, Meinecke, & Nolte, 2012; Haufe & Ewald, 2019; Nolte et al., 2004; Stam, Nolte, & Daffertshofer, 2007) were widely utilized and the latter one is less sensitive to volume conduction (Nentwich et al., 2020; Nolte et al., 2004). The limitation of coherence and correlation is that only either spectral components or temporal information can be observed while no information about brain dynamics is provided. Time-frequency analysis methods are popular to study the changes in cortical connectivity, simultaneously extracting spectral and temporal information (Sankari, Adeli, & Adeli, 2012). The short-time Fourier coherence (STFC) (Chen, Ros, & Gruzelier, 2013; Wendling, Ansari-asl, Bartolomei, & Senhadji, 2009) and wavelet coherence (WC) (Ieracitano, Duun-Henriksen, Mammon, La Foresta, & Morabito, 2017; Lachaux et al., 2002; Qassim, Cutmore, & Rowlands, 2017; Sankari et al., 2012; Sankari & Adeli, 2011) are utilized by several studies to produce EEG functional connectivity in the time-frequency domain. STFC employs a fixed sliding window to achieve spectral analysis within the time window, while WC optimizes and adjusts the wavelet base for different signal frequencies to characterize time-varying coherence (Sakkalis, 2011).

### 2.2.2 Nonlinear methods

Phase synchronization is a category that focuses on the phase coupling of oscillation systems. The phase locking value (PLV; Bajo et al., 2015; Bedo, Ribary, & Ward, 2020; Delgado-Restituto, Romaine, & Rodríguez-Vázquez, 2019; Mheich, Hassan, Khalil, Berrou, & Wendling, 2015; Sadaghiani & Kleinschmidt, 2016) and the phase lag index (PLI; Chaturvedi et al., 2019; Fraga González et al., 2018; Liao, Zhou, & Luo, 2019; Stam et al., 2007) are high-frequently used to obtain the strength of phase synchronization. Information theory is regarded as another efficient method in the case of extracting nonlinear interactions among EEG signals. To be more specific, mutual information (MI; Melia et al., 2015; Meng, Yao, Sheng, Zhang, & Zhu, 2015; Piho & Tjahjadi, 2020; Rajendra Acharya et al., 2012; Yin et al., 2017) and synchronization likelihood (SL; Altenburg, Vermeulen, Strijers, Fetter, & Stam, 2003; Chriskos, Frantzidis, Gkivogkli, Bamidis, & Kourtidou-Papadeli, 2018; Muntzat, Ali, Yasin, & Malik, 2018; Pijnenburg et al., 2004; Yu et al., 2017) are used to estimate the undirected functional connectivity from EEG recordings, while the directed transfer of information can be quantified by transfer entropy (TE). Choosing SL as a connectivity measure might lead to biased results since SL is sensitive to volume conduction effects (Boersma et al., 2013). In contrast, TE has demonstrated its robustness against volume conduction. (Harmah et al., 2020; Huang, Pal, Chuang, & Lin, 2015; McBride et al., 2015; Yang, Le Bouquin Jeannes, Bellanger, & Shu, 2013; Yao & Wang, 2017). It should be noticed that information-based approaches are fully model-free. That is to say, relatively fewer assumptions tend to be imposed on the data, but it trades in the need for larger data sets (Seth et al., 2015; Zhao et al., 2020).

#### 2.2 Parametric methods

The parametric methods refer to a framework or procedure where there is a specification of a model that requires a set of fixed parameters to be fitted to the observed signals (Sakkalis, 2011; Salman, Grover, & Shankar, 2018; Zhao, Billings, Wei, & Sarrigiannis, 2012). Compared with model-free techniques, parametric-model-based methods are more commonly accepted to estimate the effective connectivity of multi-channel EEGs. In this case, parametric effective connectivity is based on theoretical models that describe how brain areas interact and influence each other (Sakkalis, 2011). Dynamic causal modeling (DCM; Brown & Friston, 2012; Herz et al., 2012; Lee, Yoon, & Lee, 2020; Van de Steen, Almgren, Razi, Friston, & Marinazzo, 2019) applies a Bayesian framework to assess model performance, and structural equation modeling (SEM; Astolfi et al., 2004; Babiloni et al., 2003; Sartori, Bertoldo, Zavaglia, Ursino, & Toffolo, 2012) is a generalized linear modeling framework combining factor analysis with path modeling (Kaur et al., 2019). DCM and SEM treat the brain as a deterministic nonlinear and linear system, respectively (Astolfi et al., 2004; David et al., 2006; Friston, Harrison, & Penny, 2003).

Unlike DCM and SEM that require a certain amount of a priori knowledge about connectivity, many researchers have developed various data-driven methods for quantifying effective connectivity based on Granger causality (Salman et al., 2018). Granger causality-related methods occupy a useful middle ground between fully model-free and highly model-dependent approaches (Seth et al., 2015). Most of the Granger causality measures are constructed based on autoregressive models, in which the present samples of the signals are predicted using a linear or nonlinear combination of the past samples (Omidvarnia, Mesbahi, Khilf, et al., 2011; Omidvarnia, Mesbah, O’Toole, et al., 2011; van Mierlo et al., 2014a; Zhao et al., 2020; Zhao, Billings, Wei, He, et al., 2013; Zhao, Billings, Wei, & Sarrigiannis, 2013). According to the considered coefficients of the autoregressive models, partial directed coherence (PDC; Ghumare, Schroten, Vandenbergh, & Dupont, 2018; He, Billings, Wei, & Sarrigiannis, 2014; Mazher, Abd Aziz, Malik, & Ullah Amin, 2017; Silferhuth, Hintsala, Kortelainen, & Seppänen, 2012; Varotto et al., 2012; Varotto et al., 2014) and directed transfer function (DTF; Haufe, Nikulin, & Nolte, 2011; Omidvarnia, Mesbahi, Khilf, et al., 2011; Omidvarnia, Mesbah, O’Toole, et al., 2011) were applied in the field of EEG-based neuroscience. PDC is computationally more efficient and more robust than DTF since it does not involve any matrix inversion (He, Billings, et al., 2014). Granger causality was originally developed in the context of linear autoregressive with
FIGURE 2  Legend on next page.
exogenous (ARX) input models (Geweke, 1982), while some researchers focused on nonlinear causality, which is generated by nonlinear ARX model in time and frequency domains (Chen et al., 2019; He et al., 2013; He, Wei, Billings, & Sarrièriani, 2014; Zhao, Billings, Wei, He, et al., 2013; Zhao, Billings, Wei, & Sarrièriani, 2013). Besides the traditional Granger method, the error reduction ratio-causality (ERR-causality) test was proposed to estimate the time-varying direction and strength of linear or nonlinear causality between two signals as well as their relative time shift (Sarrièriani et al., 2014; Zhao et al., 2012, 2020).

3 | VISUALISATION OF BRAIN CONNECTIVITY

To effectively interpret results obtained from EEG recordings, many researchers have developed or employed various distinguished visualization approaches. This section aims to review the typical visualization methods for functional and effective connectivity and provide corresponding critical comparisons. The visualization methods can be categorized into three groups: heat map, data statistics, and head map.

Heat map generally employs an adjacency matrix to represent the quantification of brain connectivity, which performs well in showing the overall relationship between all pairs of available EEG channels. Chu et al. (2015) found a significant similarity between functional and structural networks within dysplasia patients. Regions of Interests (ROIs) for structural network analysis were selected to overlap with ROIs used for constructing the functional networks. The functional connectivity was quantified by cross-correlation and coherence and visualized by a heat map, as shown in Figure 2a. This visualization method was also used in pediatric epilepsy studies (Sargolzaei et al., 2015) and the analysis of spontaneous blood oxygen level-dependent signals (Chang, Liu, Chen, Liu, & Duyn, 2013).

The data statistic method tends to quantitatively compare targeted pairs and visualize significant ones by setting a threshold. Blinowska et al. (2017) demonstrated a statistical difference among a series of directed pairs in the contribution of Nonnormalized directed transfer function pairs values (Figure 2b), indicating the most important connections for the discrimination between healthy elderly and Alzheimer’s disease individuals (Blinowska et al., 2017). Similarly, PLI was estimated and visualized by a statistic method for patients with Fragile X Syndrome (FXS; Van Der Molen, Stam, & Van Der Molen, 2014), where the healthy control group shows the brain connectivity difference in typical frequency ranges.

With a combination of brain connectivity and a head map, researchers were able to represent the distinction between health and disease and responses to external stimuli while demonstrating the importance of specific brain regions. Figures 2c,d visualize brain functional connectivity for epilepsy diagnosis (Sargolzaei et al., 2015) and effective connectivity for human emotional responses to various types of music (Shahabi & Moghimi, 2016) respectively. Besides, Zhao et al. (2020) proposed a revised circular graph to visualize the functional connectivity quantification for the classification of Alzheimer’s disease (Figure 2e), which demonstrates the potential contribution of specific regions to disease diagnosis.

Both heat map and head map methods can represent connectivity distributions reflecting an extent involvement of specific brain regions. The heat map employs an adjacency matrix to show interconnections between each electrode pair, while the head map helps isolate and visualize brain areas of interest that can be used in developing specific applications, such as the representation of connectivity differences for diagnostic purposes (i.e., brain network function in health and various neurological disorders). It is worth noting that the determination of an appropriated threshold is important for the implementation of an appropriate head map method, since too much or too little information may limit connectivity interpretation. Moreover, a heat map is usually used for visualizing functional connectivity but not effective connectivity while the head map can reveal both types of connectivity. Another advantage of heat map is that it can generate appropriate inputs for deep learning models. For instance, Chen et al. (2019) used MI to build the heat map, which is then used as the inputs to the CNNs. The data statistics approach, compared with the other two methods, focuses more on quantification. To be more specific, this method tends to provide numerical comparison and select the most valuable connections by setting thresholds. Therefore, it is more suitable for studies that aim to reflect on data difference among a limited number of electrode pairs. However, the overall topological connectivity characteristics cannot be represented, which limits the ability to infer the contribution of structural connectivity to the findings.

FIGURE 2 Examples of the visualization of brain connectivity. (a) Structural and functional networks are topologically similar. Examples of structural and functional adjacency matrices from one patient. Similarity between structural network architecture and cross-correlation and coherence functional networks is visually evident (Chu et al., 2015). (b) Contribution of non-normalized directed transfer function (NDTF) pairs variables to PC1 in terms of Principal component analysis (PCA) loadings. Only these NDTF pairs which showed statistical differences between Nold and AD groups on the level $p < 0.0005$ contributed. It provides information about the importance of a particular parameter in the classification procedures (Blinowska et al., 2017). (c) Constructed functional connectivity map (The threshold of is applied as the connection strength) for subject diagnosed with left frontal region epilepsy and subject diagnosed with generalized epilepsy (Sargolzaei, Cabrèrizo, Goryawala, Salah, & Adjouadi, 2015). (d) Effective brain networks (averaged over all participants) for responses to different emotional music (Shahabi & Moghimi, 2016). (e) A revised circular graph plot overlaid with EEG electrode locations to highlight the real electrode locations and their corresponding locations in the plot (Zhao et al., 2020)
4 | APPLICATIONS AND ML

4.1 | Practical application

Increasing evidence exists that brain connectivity analysis, quantified by statistical dependence (e.g., correlation and coherence), information theory, and parametric and nonparametric causality analysis, can reveal the changes of the brain signal behavior or patterns of neurological patients. Over the past few years, there was substantial developments and emergences of a great number of promising results for analyzing and understanding different types of neurological disorders such as Alzheimer’s disease (Bajo et al., 2015; delEtoile & Adeli, 2017; Engels et al., 2015; Sankari & Adeli, 2011), epilepsy (Clemens et al., 2013; Douw et al., 2010; Visani et al., 2010; Xie & Krishnan, 2013) and Parkinson’s disease (Chaturvedi et al., 2019; Evangelisti et al., 2019; Handojoseno et al., 2013; Yuvaraj et al., 2013) and Parkinson’s disease-related dementia and Alzheimer’s disease (Jeong, Do Kim, Song, Chung, & Jeong, 2016; Handojoseno et al., 2013; Yuvraj et al., 2016). In addition to neurological and neurophysiological studies and associated clinical applications, there is a wide range of applications where the research objectives are to assist human in understanding brain behavior, for example, emotion recognition (Khosrowabadi, Heijnen, Wahab, & Quek, 2010; Lee & Hsieh, 2014; Shahabi & Moghimi, 2016), object recognition (Kaur et al., 2019; Supp et al., 2007; Tafreshi et al., 2019), mental assessment (Al-Shargie et al., 2019; Cattai et al., 2018; Rathee, Cecotti, & Prasad, 2017), and biometric identification (Frascini, Pani, Didaci, & Marcialis, 2019; La Rocca et al., 2014; Wang et al., 2019).

In the meantime, the rapid progression of ML technology, applied in the field of brain connectivity has led to very significant developments, aiming to achieve a deeper and better understanding of brain network behaviors for the health group and patients with various neurological conditions. Table 2 shows a depiction of some recent applications of brain connectivity estimations supported by ML methods.

4.2 | Hand-crafted versus imaging features

On one hand, state-of-the-art methods for EEG-based applications mostly apply a procedure that comprises hand-crafted features and traditional ML classifiers. For example, “PCC, WC, MSC, PS, and MI + SVM” is applied to object recognition (Tafreshi et al., 2019), “WC + LDA” is used for the diagnosis of Parkinson’s disease-related dementia and Alzheimer’s disease (Jeong et al., 2016), “DTF + SVM” is used for the detection of brain responses to emotional music (Shahabi & Moghimi, 2016) and “SL + SVM, LR, and NB” is used for

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Recent applications combining brain connectivity estimations with machine learning methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applications</strong></td>
<td><strong>Estimation + ML method</strong></td>
</tr>
<tr>
<td>Object recognition (Tafreshi et al., 2019)</td>
<td>PCC, WC, MSC, PS, and MI + SVM</td>
</tr>
<tr>
<td>Diagnosis of Parkinson’s disease-related dementia and Alzheimer’s disease (Jeong, Do Kim, Song, Chung, &amp; Jeong, 2016)</td>
<td>WC + linear discriminant analysis (LDA)</td>
</tr>
<tr>
<td>Prediction of freezing of gait in Parkinson’s disease patients (Handojoseno et al., 2013)</td>
<td>PCC + multilayer perceptron neural network and k-nearest neighbor classifier</td>
</tr>
<tr>
<td>Emotion recognition (Piho &amp; Tjahjadi, 2020)</td>
<td>MI + SVM, naive Bayes (NB) classifier, and K-nearest neighbors (KNN)</td>
</tr>
<tr>
<td>Detection of brain responses to emotional music (Shahabi &amp; Moghimi, 2016)</td>
<td>DTF + SVM</td>
</tr>
<tr>
<td>Discrimination between Alzheimer’s patients and healthy individuals (Blinowska et al., 2017)</td>
<td>DTF + artificial neural networks (ANNs)</td>
</tr>
<tr>
<td>Depression diagnosis (Saeedi et al., 2020)</td>
<td>PDC and DTF + long short-term memory and convolutional neural networks (CNN)</td>
</tr>
<tr>
<td>Attention-deficit/hyperactivity disorder identification (Chen et al., 2019)</td>
<td>MI + CNN</td>
</tr>
<tr>
<td>Diagnosis of Alzheimer’s disease (Zhao et al., 2020)</td>
<td>ERR + KNN</td>
</tr>
<tr>
<td>Diagnosis of major depressive disorder (Mumtaz et al., 2018)</td>
<td>SL + SVM, logistic regression (LR) and NB</td>
</tr>
<tr>
<td>Classification of autism spectrum disorder (Jamil et al., 2014)</td>
<td>PS + LDA and SVM</td>
</tr>
<tr>
<td>Speech categorization decisions (Al-Fahad, Yasin, &amp; Bidelman, 2019)</td>
<td>PCC and graph network + SVM and LDA</td>
</tr>
<tr>
<td>Transcranial magnetic stimulation monitoring (Gupta, Du, Hong, &amp; Choa, 2019)</td>
<td>Coherence + principal component analysis (PCA) along sparse nonnegative matrix factorization (NMF)</td>
</tr>
<tr>
<td>Detecting disorders of consciousness (Wang, Tian, Zhang, &amp; Hu, 2020)</td>
<td>Ensemble of SVMs + power spectral density difference (PSDD) incorporating with a recursive cosine function</td>
</tr>
<tr>
<td>Sedation scale estimation (Sanz-García et al., 2019)</td>
<td>PS + SVM</td>
</tr>
<tr>
<td>Detecting psycho-physiological insomnia (Aydın, Tunga, &amp; Yetkin, 2015)</td>
<td>MI, PCC and MSC + NB, random forest, regression methods and nearest neighbor based methods</td>
</tr>
<tr>
<td>Investigation of the effect of Clozapine therapy (Ravan, Hasey, Reilly, MacCrimmon, &amp; Khodayari-Rostamabad, 2015)</td>
<td>Cross-power spectral density (CPSD) + fuzzy c-mean</td>
</tr>
<tr>
<td>Face perception tasks (Jamal, Das, Maharatna, Pan, &amp; Kuyucu, 2015)</td>
<td>PLV + LDA and KNN</td>
</tr>
</tbody>
</table>
the diagnosis of major depressive disorder (Mumtaz et al., 2018). On the other hand, various visualization methods provide robust features for deep learning in recent years. For instance, Saeedi et al. (2020) reconstructed a connectivity image using two connectivity measures (PDC and DTF) and eight frequency bands for the input of deep learning networks. The proposed achieved an accuracy of 99.24% in the case of major depressive disorder diagnosis.

5 | DISCUSSION

5.1 | Connectivity estimation

Brain connectivity has shown significant potential in assessing brain function in patients with different neurological disorders and tracking various cognitive and emotional brain states for healthy participants. It is crucial to select the most appropriate connectivity measure, as widely distributed complex brain networks generate diverse topological signal processing and intercommunications in distinct conditions. Therefore, many novel methods have been developed from manifold aspects. This includes consideration of linear or nonlinear behaviors, the information in time, frequency or time-frequency domain, parametric or nonparametric measures, directed or undirected information.

5.1.1 | Functional connectivity versus effective connectivity

In the studies of effective connectivity, parametric methods have been extensively used to quantify directional interconnections among separated brain regions. The main advantage of parametric modeling of effective analysis is that the model structure is usually transparent and compact, with lagged signals being the model variables. The estimation of parametric models usually does not need a large number of samples, therefore parametric modeling allows to perform a transient or time-dependent connectivity analysis (Li, Lei, Cui, Guo, & Wei, 2019; Zhao et al., 2012). On the other hand, functional connectivity captures statistical independence between distributed and spatially separated neuronal areas, using for example, a bi-variate measure (Wang et al., 2019). Functional connectivity is generally model-free, which decreases the need of setting strict assumptions on signals. However, the amount of data needs to be increased to meet the calculation demand of functional connectivity estimation.

5.1.2 | Phase-based connectivity versus amplitude-based connectivity

Brain connectivity can also be divided into phase-based connectivity and amplitude-based connectivity. Connectivity of the amplitude is usually estimated with correlation (Brookes et al., 2011; Chang et al., 2013; Hipp & Siegel, 2015; Wang et al., 2020), while phase coupling is generally estimated with coherence-based methods and phase synchronization (Bastos & Schoffelen, 2016; Chaturvedi et al., 2019; Fagerholm et al., 2020; Nolte et al., 2004; Stam et al., 2007). These aspects further capture different neural processes (He et al., 2019; Siems & Siegel, 2020). As a result, the phase- and amplitude-based approaches give partially overlapping, partially differing results. At the same time, it has been shown that in noisy signals, phase and amplitude dynamics influence each other and the reliability of phase estimation inherently depends on the signal-to-noise ratio (SNR) and may generally be more accurate in the presence of higher signal amplitudes (He et al., 2019).

5.1.3 | Precautions during analysis

Volume conduction, an important issue when analyzing EEG recordings, has been shown to influence sensor-space connectivity analyses (Haufe et al., 2013), for example, phase coupling (Palva et al., 2018), Granger causality (Haufe et al., 2013), and correlation methods (Hipp & Siegel, 2015). To avoid the effect of volume conduction, some methodshave been proposed, such as imaginary part of the coherency (van Mierlo et al., 2014a), DTF (Kamifiski & Bfinowska, 1991) and PDC (Baccalà & Sameshima, 2001). Besides, calculating source-level connectivity from sensor-level EEGs could get rid of the influence caused by volume conduction or field propagation (He et al., 2019). This process is the so-called inverse problem (Van Diessen et al., 2015). In the meantime, some researchers suggest the interpretation of connectivity measures from sensor-level EEG recordings is not straightforward. Instead, source-level EEG is believed to be a reliable tool for measuring connectivity and it can be reconstructed from the scalp EEG (Moezzi & Goldsworthy, 2018).

5.2 | Role of visualization

Visualization plays a crucial role in the research of EEG brain connectivity. Specifically, it establishes the base for further practical and clinical uses of novel methods. It could not only improve the efficiency for other researchers to understand and evaluate the proposed approaches and produced results, but also benefit experts in presenting and explaining the finding obtained from various advanced methods to people without basic knowledge of EEG signals and brain connectivity. The heat map, data statistics and head map generally conclude various popular visualization methods and meet the distinct requirements and aims of researchers.

Besides, EEG study could be combined with appropriate visualization methods to study potential functions, oscillations and intercommunications within a dynamic architecture of the human brain (Chen et al., 2013). The current medical instrument related to EEG focuses on oscillations and biomarkers within a signal channel (Ratti, Waninger, Berka, Ruffini, & Verma, 2017). Therefore, from the authors’ point of view, development and innovations of visualization
have the potential to promote the upgrading of EEG-related equipment, as well as explore more secrets within the human brain.

Another opportunity for brain connectivity visualization is the combination with deep learning methods, which has attracted increasing investigation for brain study. There is a significant amount of works to use a variety of deep learning methods in the understanding of EEG recording, but most of them use the raw data directly. The limitation of such approaches is the lack of transparency because they usually cannot reveal which brain region has the anomaly and how it leads to neurological disorders. There is increasing evidence that CNN has superior performance to classify images by learning the spatial pattern of raw images automatically, instead of using handcrafted features. Therefore, it is promising to use the brain connectivity visualization, in a form of images, to be fed into CNN in the classification. Essentially, in this approach, brain connectivity acts as handcrafted features, but in an image format instead of singular values. Chen et al. proposed a general framework for the attention-deficit/hyperactivity disorder (ADHD) identification problem by combining an EEG-based brain network, estimated by MI, with CNN techniques (2019). The proposed framework achieved a convincing performance with an accuracy of 94.67% with the heat map visualization. Besides functional connectivity, effective connectivity visualization can also provide valuable inputs for CNN. For example, Saeedi et al. (2020) constructed a connectivity image with two connectivity measures (PDC and DTF) and eight frequency bands as the input of deep learning networks. The experiments show that the CNN applied on the constructed image of effective connectivity achieves the best results with an accuracy of 99.24% in the case of major depressive disorder diagnosis.

5.3 Research gaps and future direction

Although significant outcomes in this area have been achieved, limitations and significant challenges remain. First, brain network interactions are dynamic and may be time-varying (the associated signals are nonstationary), as phase synchronization and phase scattering occur within the millisecond range (Varela, Lachaux, Rodriguez, & Martinerie, 2001). Transient associations, usually highly nonlinear, among different brain regions have been observed (Sarrigiannis et al., 2014, 2018). Even though recent years witnessed the development of brain connectivity techniques, there are limited methods that can analyze nonstationary and nonlinear behaviors of brain networks. Exploring the hidden information within EEG signals is far beyond the capabilities of commonly available methods. Second, simplifying the interpretation of data-driven methods is a challenge. While traditional parametric models extended from AR models, for example, autoregressive moving average with exogenous input (ARMAX) and nonlinear autoregressive moving average with exogenous input (NARMAX) models (Gu et al., 2021) and dynamic causal model (David et al., 2006) are parsimonious and transparent, whose individual model terms have a clear explanation and can be linked back to the original neural system, other types of parametric models for example, nonlinear kernel models (Shen, Baingana, & Giannakis, 2016) tend to be complicated and it becomes difficult to relate the models back to the underlying system due to the lack of transparency. Interpretation of the outcomes of nonparametric methods for example, neural networks (Abbasvandi & Nasrabad, 2019; Saeedi et al., 2020) require sufficient mathematical knowledge and probably some specific and professional modeling skills. Therefore, the overall performance of such methods heavily depends on the experience and knowledge of the end-users, typically clinicians or doctors if the tools are developed for clinical applications. Third, there is no real-time feedback of brain functional connectivity during data acquisition in current EEG systems. As a result, scientists (e.g., neurologists or neurophysiologists) cannot flexibly focus on the examination of specific network areas in real-time or determine where the most significant abnormalities state or behavior takes place. Some researchers have started to integrate EEG with augmented reality (Mercier-Ganady, Lotte, Loup-Escande, Marchal, & Lecuyer, 2014; Vortmann et al., 2019), but they only focused on the information obtained from a single channel. There is a lack of applications to track and visualize time-varying brain connectivity in real-time. Fourthly, from the authors’ point of view, there are a lot of visualization and imaging techniques in neuroscience and some of them used for other imaging modalities also have the potential to visualize EEG brain connectivity. For instance, BioImage Suite provides an attractive and interpretative visualization for fMRI functional connectivity (Finn et al., 2015; Shen et al., 2018).

<table>
<thead>
<tr>
<th>Future direction</th>
<th>Purpose or strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel estimation</td>
<td>Extract more valuable information from EEG signals by robust brain connectivity methods, especially nonstationary and nonlinear intercommunications.</td>
</tr>
<tr>
<td>Interpretability</td>
<td>Design appropriate visualization methods to reduce the difficulty of understanding the actual implication of brain connectivity estimation and its outcomes, such as disease diagnosis and brain activity analysis.</td>
</tr>
<tr>
<td>Universality</td>
<td>Build a large dataset covering people with different ages, genders and diseases conditions to develop and evaluate universal brain connectivity methods.</td>
</tr>
<tr>
<td>Real-time research</td>
<td>Establish a real-time sensor and monitoring system based on advanced brain connectivity estimation and visualization approaches, capturing dynamic neuro-connectivity and assisting observation.</td>
</tr>
<tr>
<td>Improved diagnosis</td>
<td>Using the visualization of estimated brain functional or effective connectivity as the input of deep learning methods to maintain the transparency and improve the classification accuracy.</td>
</tr>
<tr>
<td>Application extension</td>
<td>Pursue a deeper understanding of the brain network and explore potential fields where EEG brain connectivity can be used.</td>
</tr>
</tbody>
</table>
Therefore, even though brain connectivity techniques have been significantly developed over the past years, more validation and improvements/developments are still required. Table 3 suggests potential directions and possible strategies, which will help deepen our understanding of brain functional and effective connectivity from a variety of novel perspectives, as well as how to enhance their availability and reliability.

6 | CONCLUSIONS

This article reviewed recent studies in estimating brain functional and effective connectivity using EEG with associated visualization as well as their applications. It is observed that functional connectivity is usually measured using nonparametric methods while effective connectivity is measured by parametric methods. In comparison to parametric methods, nonparametric methods require relatively fewer assumptions, but they trade in the need for larger data sets. Nonlinear dynamics and the spatiotemporal characteristics of brain networks hidden within EEG recordings have the potential to be identified using appropriate functional and effective connectivity methods. Although progress has been made the potential of those methods is far from being fully explored. It is expected that future research in this area will focus on better integrating different methods in particular ML approaches to improve the accuracy of disease diagnosis meanwhile increasing transparency. Real-time visualization of brain connectivity during the data acquisition stage can be introduced to improve the data quality, and help scientists better identify which areas of the brain exhibit connectivity deficits under various neurological conditions.

ACKNOWLEDGMENT

Author Yuzhu Guo gratefully acknowledge the support from the Beijing Natural Science Foundation, China (Grant No. 4202040) and the National Natural Science Foundation of China (Grant No. 61876015).

Open access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this review article as no new data was created or analyzed in this study.

ORCID

Jun Cao  https://orcid.org/0000-0003-2121-7631
Yifan Zhao  https://orcid.org/0000-0003-2383-5724
Xiaocai Shan  https://orcid.org/0000-0003-0817-0357
Hua-liang Wei  https://orcid.org/0000-0002-4704-7346
Yuzhu Guo  https://orcid.org/0000-0002-8588-5172
Liangu Chen  https://orcid.org/0000-0002-6817-8644
John Ahmet Erkoyuncu  https://orcid.org/0000-0002-8046-9911
Ptolemaios Georgios Sarigiannis  https://orcid.org/0000-0002-8380-8755

REFERENCES


Sarrigiani, P. G., Zhao, Y., He, F., Wei, H., Billings, S. A., Lawrence, S., … Yanni, J. (2015). Direct functional connectivity between the thalamus (Vim) and the contralateral motor cortex: Just a single case observation or a common pathway in the human brain? Brain Stimulation, 8(6), 1230–1233. https://doi.org/10.1016/j.brs.2015.08.006


APPENDIX: THE CONCEPTUAL AND MATHEMATICAL DETAILS
FOR CALCULATING BRAIN CONNECTIVITY

Functional connectivity

Correlation

PCC is able to evaluate the linear interdependency between two signals in the time domain and it ranges from -1 to 1. The correlation coefficient between signal X and Y can be defined as follows (van Mierlo et al., 2014a):

$$\rho_{xy} = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sigma_x \sigma_y}$$

where $E$ is the expected value, $\mu_x$ and $\mu_y$ are the mean values and $\sigma_x$ and $\sigma_y$ are the standard deviations of X and Y time series.

Cross-correlation differs from PCC since it is a function with respect to time lag $\tau$, which can be expressed as (Ibrahim et al., 2018; van Mierlo et al., 2014a).

$$\rho_{xy}(\tau) = \frac{E[(x_n-\mu_x)(y_{n+\tau}-\mu_y)]}{\sigma_x \sigma_y}$$

Magnitude squared coherence

MSC is a linear method to estimate the interconnections between two signals in the frequency domain calculated by PSD. The MSC of signal X and Y can be written as (van Mierlo et al., 2014a, 2014b; Wendling et al., 2009)

$$\text{MSC}_{XY}(f) = C_{xy}^2 = \frac{S_{xy}(f)^2}{|S_{XX}(f)| \times |S_{YY}(f)|}$$

where $S_{XX}(f)$ and $S_{YY}(f)$ are the PSD of signal X and Y respectively, and $S_{xy}(f)$ is the cross PSD at frequency $f$.

To avoid the volume conduction effects, the imaginary part of the coherence is employed by a few studies instead of looking at the magnitude and phase of the coherence (Ewald et al., 2012; Van Diessen et al., 2015; van Mierlo et al., 2014a).

$$C_{xy} = RC_{xy} + IC_{xy}$$

Wavelet coherence

WC is generally acknowledged as a qualitative estimator that can represent the time-varying relations in the time-frequency domain between signals with a specifically determined scale (Labat, 2005; Tafreshi et al., 2019). To be more specific, the wavelet transforms of a signal $x$ is a function of both time and frequency, which is defined as the convolution of the input with a Wavelet family $\theta(u)$:

$$W_x(t,f) = \int_{-\infty}^{+\infty} x(u) \cdot \theta^*_f(u) du$$

Wavelet cross-spectrum around time $t$ and frequency $f$, given input signals $x$ and $y$, can be derived by the Wavelet transforms of $x$ and $y$:

$$CW_{xy}(t,f) = \int_{t-\delta/2}^{t+\delta/2} W_x(\tau,f) \cdot W^*_y(\tau,f) d\tau$$

where $^*$ defines the complex conjugate and $\delta$ is assumed as a frequency-depending scalar.

WC at the time $t$ and frequency $f$ is derived as:

$$WCo_{xy}(t,f) = \frac{|CW_{xy}(t,f)|}{|CW_{XX}(t,f) \times CW_{YY}(t,f)|^{1/2}}$$

Short-time Fourier coherence

Short-time Fourier transform (STFT) is also implemented in the time-frequency domain instead of the classical fast Fourier transform approach to calculate coherence, and coherence may be calculated around a number of time instants. This technique produces the so-called “coherogram,” which forms a three-dimensional matrix of time and frequency against coherence. However, stationarity is still required within each time interval for which coherence is calculated, meaning that in practice one should carefully decide on the optimal section length over which each coherence estimate is measured. Window length and overlapping within each coherence estimate affect the resolution of the measure (Delgado-Restituto et al., 2019; Sakkalis, 2011).

Phase synchronization

Phase synchronization (PS) assumes two oscillation systems without amplitude synchronization can have phase synchronization. The PLV and the PLI are high-frequently used to obtain the strength of phase synchronization (van Mierlo et al., 2014a). The instantaneous phase of a signal $x$ is given by:

$$\phi_x(t) = \text{arctan} \frac{\tilde{x}(t)}{x(t)}$$

where $\tilde{x}(t)$ is the Hilbert transform of $x(t)$ which is defined as:

$$\tilde{x}(t) = \frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau$$
where \( PV \) refers to the Cauchy principal value. The PLV for two signals is defined as:

\[
PLV = \left| \frac{1}{N} \sum_{n=0}^{N-1} \exp(i \theta_n) \right|
\]

where \( \Delta t \) defines the sampling period and \( N \) indicates the sample number of each signal (Mormann, Lehnertz, David, & Elger, 2000). The range of PLV is from 0 to 1, where 0 shows a lack of synchronization and 1 indicates strict phase synchronization.

The PLI captures the asymmetry of the distribution of phase differences between two signals and is calculated based on the relative phase difference between the two signals (Borrego-Salcido, Juárez-Del-Toro, & Cruz-Aké, 2019; Wendling et al., 2009).

\[
PLI = |E[\text{sign}(\Delta \theta(t))]|
\]

The resulting value lies in the interval \([0, 1]\), where a higher value indicates more phase synchrony.

### Synchronization likelihood

SL is a measure of multivariate synchronization to investigate the interactions between nonlinear dynamical systems without any knowledge about the governing equations (Stam & van Dijk, 2002). For an \( M \)-channel time series \( x_{k,n} \) (\( k \) denotes channel number \([k = 1, \ldots, M]\)), embedded vectors \( X_{k,n} \) are reconstructed with time-delay embedding:

\[
X_{k,n} = \{x_{kn}, x_{kn+1}, \ldots, x_{kn+(m-1)e}\}
\]

where \( m \) is the embedding dimension and \( e \) denotes time lag. The estimated probability that embedded vectors \( X_n \) are closer to each other than a distance \( \varepsilon \) is (Wendling et al., 2009):

\[
P_{e,k,n} = \frac{1}{Z_{w_2-w_1}} \sum_{j=1}^{N} \theta(\varepsilon - |x_{kj} - x_{kj}|) \quad w_1 < |i - j| < w_2
\]

where \( | \cdot | \) is the Euclidean distance; \( \theta \) stands for the Heaviside step function, \( w_1 \) is the Theiler correction and \( w_2 \) determines the length of the sliding window. Letting \( P_{e,k,n} = P_{e,ref} \), where \( P_{e,ref} < 1 \), the number of channels \( H_{ij} \) where the embedded vectors \( X_{ij} \) and \( X_{kj} \) will be closer together than this critical distance \( e_{ij} \):

\[
H_{ij} = \sum_{k=1}^{M} \theta(e_{ij} - |X_{ij} - X_{kj}|)
\]

SL at time \( n \) can be obtained by averaging over all values of \( j \)

\[
SL_n = \frac{1}{2P_{e,ref}(w_2 - w_1)} \sum_{j=1}^{N} \left( \frac{H_{ij} - 1}{M - 1} \right)
\]

where \( w_1 < |i - j| < w_2 \)

All aforementioned measures are normalized between 0 and 1; the 0 value means that the two signals are completely independent. On the opposite, the 1 value means that the two signals are completely synchronized (Wendling et al., 2009).

### Mutual information

According to information theory, MI of two random variables \( X \) and \( Y \) shows how a random variable is informative for the other one. Let, \( P(x) \) and \( P(y) \) be the probability distributions of random variables \( X \) and \( Y \), respectively. The entropy of \( X \) and \( Y \) is defined as (Tafreshi et al., 2019):

\[
H(X) = - \sum_{x \in X} P(x) \log_2 P(x)
\]

\[
H(Y) = - \sum_{y \in Y} P(y) \log_2 P(y)
\]

where \( n \) defines window length.

\( H(Y|X) \) and \( H(X,Y) \) are conditional entropy and joint entropy between \( X \) and \( Y \), defined respectively as:

\[
H(Y|X) = - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(y|x)
\]

\[
H(Y|X) = - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(y|x)
\]

where \( E \) is the expected value function.

MI of two random variables \( X \) and \( Y \) is computed as follows (Khosrowabadi et al., 2010):

\[
MI(X,Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X)
\]

\( MI(X,Y) = 0 \) if and only if random variables \( X \) and \( Y \) are statistically independent.

### Effective connectivity

#### Granger causality

Granger causality can estimate effective interactions from time-series data. Granger causality implements a statistical, predictive notion of causality whereby causes precede and help predict their effects (Seth et al., 2015). It is defined in both the time and frequency domains, and it allows for the conditioning out of common causal influences. One
time series is said to Granger cause a second one if the inclusion of the past values of the first into the modeling of the second significantly reduces the variance of the modeling error. According to Granger causality, if the past values of \( X \) contain information that helps to predict \( Y \) above and beyond the information contained in past values of \( Y \) alone then signal \( X \) “Granger-causes” signal \( Y \). The Granger causality from signal \( X \) to \( Y \) and the one from signal \( Y \) to \( X \) can be investigated separately. (van Mierlo et al., 2014a).

**Partial directed coherence**

PDC is a method that quantifies the relation between two among \( N \) signals, avoiding volume conduction by estimating influences of all other signals (Tafreshi et al., 2019). PDC improves the concept of Partial Coherence by estimating causal influences. This method is estimated on multivariate autoregressive (MVAR).

\( X(t) \) is a set of estimated signals from \( N \) recording channels:

\[
X = [x_1(t), x_2(t), \ldots x_N(t)]^T
\]  

(21)

The MVAR process is an expressive description of the data set \( X \):

\[
\sum_{\tau=0}^{P} A(\tau)X(t-\tau) = E(t)
\]  

(22)

In this model \( E(t) \) is a zero-mean multivariate uncorrelated white noise vector. \( A(\tau) \) is the autoregressive coefficients matrix and its elements \( a_{ij}(\tau) \) indicates the influence of \( X_j(t-\tau) \) on \( X_i(t) \) and \( p \) represents the border of the model.

PDC from the \( i \)th channel to the \( j \)th channel at frequency \( f \) is defined as follows:

\[
e_{ij}(f) = \frac{\mathcal{A}_{ij}(f)}{\sqrt{\sum_{m=1}^{N} \mathcal{A}_{jm}(f)\mathcal{A}_{jm}(f)}}
\]  

(23)

where \( \mathcal{A}_{ij}(f) \) is a frequency-domain description of \( a_{ij}(\tau) \) (Baccalà & Sameshima, 2001; Granger & Aug, 2007)

\[
\mathcal{A}_{ij}(f) = \begin{cases} 
1 - \sum_{\tau=1}^{P} a_{ij}(\tau)e^{-j2\pi\tau f}, & \text{if } i = j \\
-\sum_{\tau=1}^{P} a_{ij}(\tau)e^{-j2\pi\tau f}, & \text{otherwise}
\end{cases}
\]  

(24)

**Directed transfer function**

The directed transfer function (DTF) is also based on the concept of Granger causality estimated on MVAR, which models all signals simultaneously (Kamiliški & Bfinowska, 1991). \( X(t) \) is a set of estimated signals from \( N \) recording channels:

\[
X = [x_1(t), x_2(t), \ldots x_N(t)]^T
\]  

(25)

The MVAR process is an expressive description of the data set \( X \):

\[
\sum_{\tau=0}^{P} A(\tau)X(t-\tau) = E(t)
\]  

(26)

In this model \( E(t) \) is a zero-mean multivariate uncorrelated white noise vector with \( A(0) = 1 \). \( A(1), A(2), \ldots, A(p) \) is coefficients matrix and \( p \) represent the border of the model. Equation (20) can be transformed into the frequency domain, which is defined as:

\[
A(f)X(f) = E(f)
\]  

(27)

where

\[
A(f) = \sum_{\tau=0}^{P} A(\tau)e^{-j2\pi\tau f}
\]  

(28)

So, \( X(f) \) can be obtained by

\[
X(f) = A^{-1}(f)E(f) = H(f)E(f)
\]  

(29)

\( H(f) \) is the transfer function of the system and its elements \( h_{ij}(f) \) indicate the causal influence from the \( j \)th input to the \( i \)th output at frequency \( f \).

The DTF is defined as:

\[
\rho^2_{ij}(f) = |h_{ij}(f)|^2
\]  

(30)

Normalized DTF is defined as:

\[
\gamma^2_{ij}(f) = \frac{|h_{ij}(f)|^2}{\sum_{m=1}^{N} |H_{im}(f)|^2}
\]  

(31)

where \( \gamma^2_{ij}(f) \) describes the influence ratio of the \( j \)th channel-related cortical area on the \( i \)th channel-related cortical area with respect to the influence of all estimated cortical signals (Babiloni et al., 2005). This is an important difference between DTF and PDC since DTF is normalized for the structure that receives the signal, whereas PDC is normalized for the structure that sends the signal (Baccalà & Sameshima, 2001).

**Structural equation modeling**

A set of linear structural equations are employed by the structural equation modeling (SEM), exploring causal intercommunications among the observed variables and parameters. In terms of neural systems, a measure of covariance represents the degree to which the activities of two or more regions are related. The SEM technique estimates the parameters by minimizing the difference between the observed covariances and those implied by a structural or path model (Astolfi et al., 2004; Babiloni et al., 2003).
**Dynamic causal modeling**

The aim of DCM (Friston et al., 2003) is to make inferences about the coupling among brain regions or sources and how that coupling is influenced by experimental factors. DCM uses the notion of effective connectivity, defined as the influence one neuronal system exerts over another. The central idea behind DCM is to treat the brain as a deterministic nonlinear dynamical system that is subject to inputs and produces outputs. In neuroimaging, DCM starts with a reasonably realistic neuronal model of interacting cortical regions. This model is then supplemented with a forward model of how neuronal activity is transformed into measured responses. This enables effective connectivity of the neuronal model to be estimated from observed data (David et al., 2006; Yvert, Perrone-Bertolotti, Baciu, & David, 2012).

**Transfer entropy**

TE is an alternative measure of effective connectivity based on information theory. TE does not require a model of the interaction and is inherently nonlinear (Vicente, Wibral, & Lindner, 2011). TE for two observed time series \( x_t \) and \( y_t \) can be written as

\[
\text{TE}(X \rightarrow Y) = \sum_{y_t, u_t} p(y_t, x_t, x_{t-u}) \log \frac{p(y_t, x_t, x_{t-u})}{p(y_t|x_t) p(x_{t-u})} \tag{32}
\]

where \( t \) is a discrete-valued time-index and \( u \) denotes the prediction time, a discrete-valued time-interval. \( y_t^x \) and \( x_t^u \) are \( d_y \) - and \( d_x \) - dimensional delay vectors (Vicente et al., 2011).

**Error reduction ratio**

Error reduction ratio (ERR) is a parametric method in the time domain, which indicates how much of the change in the system response, in percentage, can be accounted for by including the relevant model terms. This capability allows the computing of the contribution of linear and nonlinear terms independently without fitting the full nonlinear model, where parameter estimation and model validation are required. Consider a linear regression

\[
Y = P\Theta
\]

where

\[
Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(M) \end{bmatrix}, \quad P = \begin{bmatrix} P^T(1) \\ \vdots \\ P^T(M) \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta(1) \\ \vdots \\ \theta(M) \end{bmatrix}
\]

and \( P^T(k) = (p_1(k), p_2(k), \ldots, p_M(k)) \). \( p_1, p_2, \ldots, p_M \) are the model terms selected from the candidate term set. Matrix \( P \) can be decomposed as \( P = W \times A \) where

\[
W = \begin{bmatrix} w_1(1) & w_2(1) & \ldots & w_M(1) \\ w_1(2) & w_2(2) & \ldots & w_M(2) \\ \vdots & \vdots & \ddots & \vdots \\ w_1(M) & w_2(M) & \ldots & w_M(M) \end{bmatrix}
\]

and \( A = \{a_{ij}\} \) is an upper triangular matrix with unity diagonal elements.

Equation (33) is then rewritten as

\[
Y = WG, \tag{36}
\]

where \( G = A\Theta = [g_1 g_2 \ldots g_M]^T \). Equation (36) is now ready to represent the relation between \( Y \) and \( G \).

The importance of each model term to the variation of the system output is then estimated. Initially, set values \( a_{ij} = 0 \) for \( i \neq j \) (A then becomes an identity matrix), so \( w_1(k) = p_1(k) \), and calculate \( g_1 \) as

\[
g_1 = \sum_{k=1}^M w_1(k) y(k) \sum_{k=1}^M w_1^2(k) \tag{37}
\]

For \( j = 2, 3, \ldots, M \), set \( a_{ij} = 1 \) and then calculate

\[
a_{ij} = \frac{\sum_{k=1}^M w_j(k) y(k)}{\sum_{k=1}^M w_j^2(k)}, \tag{38}
\]

where \( i = 1, 2, \ldots, j - 1 \). Next, the algorithm calculates

\[
w_j(k) = p_j(k) - \frac{1}{i-1} \sum_{i=1}^j a_{ij} w_i(k) \tag{39}
\]

and

\[
g_1 = \sum_{k=1}^M w_1(k) y(k) \sum_{k=1}^M y^2(k) \tag{40}
\]

The ERR value for each term \( p_i \) is finally defined as

\[
\text{ERR} = \sum_{k=1}^M y^2(k) \frac{\sum_{k=1}^M y^2(k)}{\sum_{k=1}^M y^2(k)} \tag{41}
\]

Values of ERR range from 0 to 100%. The larger the value of ERR, the higher the dependence between this term and the output. To stop the search procedure and determine the number of significant terms \( N \), a criterion called Penalised Error-to-Signal Ratio (PESR) is used (Zhao et al., 2012). It can be written as
This criterion is introduced to monitor the search procedure, where \( m \) denotes the index of the selected terms. The search procedure stops when \( PESR_m \) achieves a local minimum. In this article, the value of \( \lambda \) was chosen as 8.

To calculate the contribution of the input to the output, the sum of ERR of all selected terms, denoted by \( SERR \), is calculated by

\[
SERR = \sum_{i=1}^{N} ERR_i. \tag{43}
\]

Note \( N \) is the number of the selected terms, not the number of total candidate terms. The value of \( SERR \) (0 \( \leq \) SERR \( \leq \) 1) describes the percentage explained by the identified model to the system output. If the considered inputs can fully explain the variation of the system output, the value of \( SERR \) is equal to 100%.

The linearity of connectivity is represented by the sum of ERR of the linear terms, and it can be computed by

\[
ERR_L(y,u) = \frac{\sum_{b=1}^{N} (ERR_b | p_b \in C_1)}{SERR}. \tag{44}
\]

The nonlinearity of connectivity is represented by the sum of ERR of the nonlinear terms, and it can be computed by

\[
ERR_N(y,u) = \frac{\sum_{b=1}^{N} (ERR_b | p_b \notin C_1)}{SERR}. \tag{45}
\]