PAPER

A brain-age model for preterm infants based on functional connectivity

To cite this article: M Lavanga *et al* 2018 *Physiol. Meas.* **39** 044006

View the [article online](#) for updates and enhancements.

Related content

- [Interhemispheric connectivity estimated from EEG time-correlation analysis in preterm infants with normal follow-up at age of five](#)
E J Meijer, H J Niemarkt, I P P C Raaijmakers *et al.*
- [Comparison of connectivity analyses for resting state EEG data](#)
Elzbieta Olejarczyk, Laura Marzetti, Vittorio Pizzella *et al.*
- [Effect of intermittent hypoxic training on hypoxia tolerance based on brain functional connectivity](#)
Guang Li, Tinglin Zhang, Xiaojian Chen *et al.*

Recent citations

- [A perinatal stress calculator for the neonatal intensive care unit: an unobtrusive approach](#)
M Lavanga *et al*

IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,
biomedical engineering and related subjects.

Start exploring the collection—download the
first chapter of every title for free.

PAPER

A brain-age model for preterm infants based on functional connectivity

RECEIVED
20 December 2017REVISED
12 March 2018ACCEPTED FOR PUBLICATION
29 March 2018PUBLISHED
26 April 2018M Lavanga^{1,2} , O De Wel^{1,2}, A Caicedo^{1,2}, K Jansen^{3,4}, A Dereymaeker³, G Naulaers³ and S Van Huffel^{1,2}¹ Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Kasteelpark Arenberg 10, Box 2446, 3001, Leuven, Belgium² imec, Kapeldreef 75, 3001, Leuven, Belgium³ Department of Development and Regeneration, Neonatal Intensive Care Unit, UZ Leuven, UZ Herestraat 49, Box 7003 21, 3000 Leuven, Belgium⁴ Department of Development and Regeneration, Child Neurology, UZ Leuven, UZ Herestraat 49, box 7003 21, 3000 Leuven, BelgiumE-mail: mlavanga@esat.kuleuven.be**Keywords:** early brain development, functional connectivity, preterm neonatal EEG, preterm infant**Abstract**

Objective: In this study, the development of EEG functional connectivity during early development has been investigated in order to provide a predictive age model for premature infants. **Approach:** The functional connectivity has been assessed via the coherency function (its imaginary part (*ImCoh*) and its mean squared magnitude (*MSC*)), the phase locking value (*PLV*) and the Hilbert–Schmidt dependence (*HSD*) in a dataset of 30 patients, partially described and employed in previous studies (Koolen *et al* 2016 *Neuroscience* **322** 298–307; Lavanga *et al* 2017 *Complexity* **2017** 1–13). Infants' post-menstrual age (PMA) ranges from 27 to 42 weeks. The topology of the EEG couplings has been investigated via graph-theory indices. **Main results:** Results show a sharp decrease in *ImCoh* indices in θ , (4–8) Hz and α , (8–16) Hz bands and *MSC* in β , (16–32) Hz band with maturation, while a more modest positive correlation with PMA is found for *HSD*, *PLV* and *MSC* in δ_2 , θ , α bands. The best performances for the PMA prediction were *mean absolute error* equal to 1.51 weeks and adjusted coefficient of determination R^2_{adj} equal to 0.8. **Significance:** The reported findings suggest a segregation of the cortex connectivity, which favours a diffused tasks architecture on the brain scalp. In summary, the results indicate that the neonates' brain development can be described via lagged-interaction network features.

1. Introduction

Preterm infants represent 10% of the worldwide neonatal population and their survival rates have been characterized by a steady increase thanks to the most recent achievements in neonatal care (González *et al* 2011). Despite the improvements, these patients are at higher risk of poor neurodevelopmental outcome (Martin *et al* 2005, Grieve *et al* 2008). As Grieve pointed out, there is an extensive literature how the premature exposure to the extrauterine environment is linked to the disruption of the normal brain development. Due to the possible impact of prematurity on neurodevelopmental outcome, a need for maturation charts to detect neural disorders in advance has been discussed in the literature. In Franke *et al* (2012), Franke *et al* used fMRI to describe cerebral maturation.

The development of the infant brain during the perinatal period is characterized by an increased relevance of cortico-cortical connections and thalamo-cortical connections (Kostović and Judoš 2010). Specifically, the slow wave burst-modulation pattern is due to the cortical interactions, while subcortical patterns generate high-frequency oscillations (Witte *et al* 1997). The latter type of connections are dominant up to 32 weeks PMA, when the former starts prevailing (Kostović and Judoš 2010). Consequently, the anatomical development drives a change in neuroelectrical waveforms and one would expect a change in connectivity among the activity courses of the different neuron pools. It should not be a surprise that brain connectivity has been classified into two main types by Friston (1994). The first one is the description of the anatomical wiring, known as structural connectivity, and

usually described by imaging techniques. The second type is the functional connectivity (FC), which delves into the statistical relationships among the neural activities of different brain regions and provides insights into how the brain manages its neuronal resources to engage any type of tasks (Friston 1994). Clearly, the two architectures are not independent, but they influence each other (Goldenberg and Galván 2015), as discussed earlier. Therefore, the functional connectivity is a strong candidate to track the infant brain maturation. Franke *et al* (2012) has already described a brain maturation model via FC quantified using fMRI. However, compared to fMRI, the EEG can measure the functional relations among neuronal pools with a lesser intrusive and easier recording procedure (Meijer *et al* 2014). The literature provides an extensive range of linear and nonlinear methods to assess FC in EEG channels, such as the *mean squared coherence* (MSC) or *phase locking value* (PLV) (Pereda *et al* 2005).

In case of perinatal development, the common approach is to describe how the coupling of a selected pair of channels changes over age (González *et al* 2011, Meijer *et al* 2014, Myers *et al* 2015). However, there are two main drawbacks with this approach. First, the combination of channels to study their interactions increases exponentially with the number of channels. Unlike the application with fMRI (Batalle *et al* 2017), indices, which describe the topological distribution of EEG connectivity, are seldomly applied. The most recent approaches investigate multichannel EEG activity with network analysis, as proposed by Bullmore and Sporns (2009). As shown by the authors of this paper in a previous study (Lavanga *et al* 2017), graph indices, such as the path length or the clustering coefficient, derived from EEG effective connectivity, can be embodied into a brain-age model, similarly to Franke *et al* (2012).

The second drawback is the investigation of the connectivity at the sensors' level, which can generate spurious couplings due to volume conduction (Haufe *et al* 2013). Some authors suggested to retrieve sources to investigate the coupling. And yet, many doubts can be raised on this approach for two reasons. The first one is the requirement of a proper head-model for neonates to estimate the sources, which can turn into an expensive and unneeded process, as already mentioned in previous studies (Meijer *et al* 2014). The second reason is that the sources of spontaneous EEG are close to the surface of the scalp (Grieve *et al* 2008), therefore to the EEG recording sites, which questions the necessity to retrieve the sources to avoid the computation of spurious connectivity. Based on these considerations, FC can be described via different pairwise EEG connectivity methods, such as Nolte's imaginary part of coherency (*ImCoh*), which studies the lagged interactions among signals (Nolte *et al* 2004). Haufe *et al* (2013) reports that *ImCoh* highlights the strong and true interactions and disregards the weak interactions due to the volume conduction.

ImCoh is not new to perinatal development studies, such as (González *et al* 2011). However, previous research has not dealt with this method via growth charts as (Meijer *et al* 2014, Myers *et al* 2015) or Koolen *et al* (2016) proposed. Following the path by those authors, we decided to provide a systematic comparison of different EEG-based FC connectivity approaches in order to track the maturation of the infant brain. This study is an extension of previous research in our group (Lavanga *et al* 2016). *ImCoh* and the *Hilbert–Schmidt dependence* (HSD) (Szabó 2014) have been added to *MSC* and *PLV* to describe the evolution of the healthy preterm cerebral networks. There are two primary aims of this study: (1) to describe cortical connectivity as function of infant's age; (2) to provide a brain-age prediction model based on graph topological indices, similarly to Franke *et al* (2012) and O'Toole *et al* (2016).

2. Methods

2.1. Data

We used an EEG recordings dataset from 30 preterm neonates, who were recruited for a larger study to assess brain development (Koolen *et al* 2016, Lavanga *et al* 2017) between 2012 and 2014 at the same neonatal intensive care unit (NICU), in the University Hospitals Leuven, Belgium, with informed parental consent. Twenty-six neonates present normal neurodevelopment outcome at 2 years from the birth, while four subjects are declared normal at 9 months (Bayley Scales of Infant Development—II, mental and motor score >85). EEG measurements for each subject took place at least once during their stay at the unit and lasted at least 2 hours. The data were recorded with nine electrodes ($F_1, F_2, C_3, C_4, T_3, T_4, O_1, O_2$ and the reference C_z , which was disregarded in the study) according to the 10–20 international system monopolar set-up. The sampling frequency for the data collection was 250 Hz (BRAIN RT, OSG Equipment, Mechelen, Belgium). A total number of 103 recordings was obtained from 30 preterm patients, whose *post-menstrual age* (PMA) ranges from 27 weeks to 42 weeks. The two sleep stages, namely quiet sleep (QS) and non-quiet sleep (NQS, all other states), were annotated by two independent readers (KJ, AD). Each recording has at least one QS–NQS cycle. The average (standard deviation) duration for the EEG recording, for the QS and NQS are respectively 263.69(129.39) min, 73.58(33.01) min and 190.11(106.90) min, while the amount of multichannel EEG, QS and NQS data are respectively 452 h, 326 h, 126 h. Sleep states QS and NQS are respectively 29.46(9.83)% and 70.54(9.83)% of the trace.

2.2. Functional connectivity methods

In order to assess the functional connectivity (FC), the most common linear and nonlinear approaches to derive bidirectional interdependencies reported in Pereda *et al* (2005) have been applied. The first two measures are based on the definition of the coherency as

$$C_{xy}(f) = \frac{P_{xy}(f)}{(P_{xx}(f)P_{yy}(f))^{1/2}}, \quad (1)$$

which measures the linear relationship of two EEG channels x and y at a specific frequency and is in general a complex number. $P_{xy}(f)$ is the cross-spectrum between channels x and y , while $P_{xx}(f)$ and $P_{yy}(f)$ are autospectra of each channel, respectively. Consequently, the coherency is a normalized cross-spectrum. The most common way to describe the correlation of two signals in the frequency domain is the squared magnitude of the coherency

$$k_{xy}^2(f) = |C_{xy}(f)|^2, \quad (2)$$

which is also known as the *magnitude squared coherence* (MSC). However, Nolte *et al* (2004) pointed out that the true interactions can be measured via the imaginary part of the coherency (*ImCoh*)

$$ImCoh_{xy}(f) = \mathcal{I}(C_{xy}(f)), \quad (3)$$

which assesses only the lagged interactions. Those couplings emerge only if there is a phase shift among the time series. In this way, the index should be robust to spurious connectivity, which can arise from any linear mixing of independent sources due to volume conduction (Haufe *et al* 2013). According to Nolte *et al* (2004), if a linear superposition of K sources $s_k(f)$ is assumed for any channel x :

$$X(f) = \sum_{k=1}^K a_{xk}s_k(f), \quad (4)$$

the only distorted part of cross-spectrum $P_{xy}(f)$ will be the real part, as shown below:

$$P_{xy}(f) = \langle X(f), Y^*(f) \rangle = \sum_{k=1}^K a_{xk}a_{yk'} \langle s_k(f), s_k^*(f) \rangle > \quad (5)$$

$$= \sum_{k=1}^K a_{xk}a_{yk} \langle s_k(f), s_k^*(f) \rangle = \sum_{k=1}^K a_{xk}a_{yk} |s_k(f)|^2, \quad (6)$$

where $\langle X(f), Y^*(f) \rangle$ is the dot product between two signals. Consequently, this instantaneous mixing is only reflected in the real part of coherency and might only affect the MSC. For these reasons, if independent sources are mixed at the sensors' level, the imaginary coherency will vanish to zero since it measures only the time-shifted couplings. A similar FC index to *ImCoh* is the *phase locking value* (PLV), which also assesses the phase shift among two channels. It is defined as

$$PLV_{xy} = \frac{1}{N} \left| \sum_{t=0}^N e^{j\gamma_{xy}(t)} \right|, \quad (7)$$

where $\gamma_{xy}(t) = \varphi_x(t) - \varphi_y(t)$. $\varphi_x(t)$ and $\varphi_y(t)$ represent the phases of the time series x and y and are derived from the analytical signal $\tilde{x}(t) = x(t) + jx_H(t)$ (or $\tilde{y}(t)$), where $x_H(t)$ is the Hilbert transform of $x(t)$. Another general measure of interaction, which can capture both linear and nonlinear interrelationships, is the *Hilbert–Schmidt dependence* (HSD) defined as the Hilbert–Schmidt norm of following cross-covariance operator:

$$C_{xy} = \mathbb{E}([\phi(x) - \mu_x] \otimes [\psi(y) - \mu_y]), \quad (8)$$

where $\phi(x)$ and $\psi(y)$ are nonlinear feature maps, μ_x (μ_y) is the expected value $\mathbb{E}(\phi(x))$ ($\mathbb{E}(\psi(y))$) and $\otimes$ is the tensor product (Szabó 2014). In this study, the HSD analysis has been performed by means of the ITE toolbox, Matlab (Mathworks, Natick, Ma, USA) (Szabó 2014).

2.3. Graph theory

A graph is a diagram of points or nodes, whose relations are defined by the connecting lines among them. In order to give an idea of its possible applications, many authors conjure up the well-known picture of geographic maps, whose connections lines are the roadways among cities, or the social network, whose lines are the friendships among people. In case we are interested to assess not only the existence of a link, but also the strength of friendship in a social network or the amount of traffic to reach an urban center, a value could be assigned to each link and a weighted graph is obtained. In this study, the relationship among EEG channels can be described by a weighted network. Specifically, any FC method generates a coupling matrix $M \times M$, defined as follows:

$$\mathcal{A}_{ij} = \mathcal{C}_{x_i \leftrightarrow x_j}, \quad (9)$$

where $\mathcal{C}_{x_i \leftrightarrow x_j}$ is the general coupling intensity and M is the number of processes. As shown by Bullmore, the matrix $\mathbf{A}$ can be interpreted as an adjacency matrix of a brain scalp graph, with weighted edges (Bullmore and Sporns 2009). The FC generates undirected graphs, which result in a symmetric matrix $\mathbf{A}$, since the connectivity values are bidirectional (as highlighted by $x_i \leftrightarrow x_j$). In the literature, a list of topological and spectral indices have been thoroughly discussed to describe the architecture of networks (Bondy and Murty 1976, Bullmore and Sporns 2009). Among the topological measures, we recall the path length, the clustering coefficient and the diameter. All these indices assess the level of integration in the network.

Specifically, the path length is the average shortest path to reach any node in the network from any another node. The lower the value, the higher the integration in the network. The diameter is usually defined as the maximum eccentricity in the network, which also represents the maximum distance from one node to any other node. It is self-evident that those indices measure the distance among nodes similarly: the former is the mean of means, while the latter the maximum of the longest paths. The clustering coefficient is defined as the average of all triangles around it. This topological index mirrors the graph coupling density, which has also been investigated. The network density is defined as the sum or the average of part of the entries of $\mathbf{A}$. In particular, we define the density based on the different brain regions: the anterior and the posterior densities are computed as the average of the entries of $\mathbf{A}$ associated to the anterior/posterior electrodes; the left and right densities are computed using only the coupling values of the electrodes left/right hemisphere, while the interhemispheric density is the average of the symmetric channel combinations between hemispheres, which is the average of one of secondary diagonals. The total density is the average of all entries in $\mathbf{A}$. Besides these topological indices, the spectral indices such as the spectral radius, the spectral gap and the algebraic connectivity have been computed. The spectral radius and the spectral gap are the maximal absolute eigenvalue and the difference between the first and the second absolute eigenvalues of the adjacency matrix. The algebraic connectivity is the second smallest eigenvalue of the Laplacian matrix (Bondy and Murty 1976).

2.4. Algorithmic pipeline and statistical analysis

Each signal was resampled to 128 Hz and low-pass filtered at 32 Hz. Each FC method was computed in 1 min non-overlapping windows. *MSC* and *ImCoh* were computed with Welch's method using a 1 s subwindow and the overlap was 70%. Those two parameters were mainly chosen for a consistent and unbiased coherency estimation (Bortel and Sovka 2007). The *MSC* was then extracted as the average of $|C_{xy}(f)|^2$ in the following frequency bands: δ_1 (0–2) Hz, δ_2 (2–4) Hz, θ (4–8) Hz, α (8–16) Hz, β (16–32) Hz. The *ImCoh* was computed as the maximum of $\mathcal{I}(C_{xy}(f))$ in the same frequency bands. This choice is a consequence of the fact that the *ImCoh* is expected to be centered on one single frequency, as reported in Nolte et al (2004). In case of the *HSD* and *PLV*, the signal was split in the specified frequency bands via wavelet decomposition. The chosen basis was Daubechies 4, which is known to be the best wavelet for epileptic EEG: since neonatal EEG is also characterized by transient waveforms, the authors of this paper have also opted for this type of wavelet. The coefficients of each scale were used to compute the phase synchrony and the information entropy. In both types of segmentation, the mean was subtracted and the window was discarded wherever the maximum value was above 3000 μV . It is interesting to notice how shorter windows for *HSD* and *PLV* could have affected the significance of the coupling estimation. In case of *PLV*, enough basic periods are required to assess the oscillator interactions and the lowest frequency oscillations are just reported by the empirical threshold defined by Bezduschnko and Smirnov (2010). *MSC*, *PLV* and *MI* generate symmetric adjacency matrices, while the latter one is antisymmetric in case of *ImCoh*. In order to apply the undirected graph theory, we simply decided to ignore the sign of the entries and take $\mathcal{A}_{ij} = |A_{ImCoh_{ij}}|$. Consequently, each method generates a 8×8 matrix for each method and each window. In order to guarantee the reliability of each connectivity method, we applied a surrogate data test. Specifically, the data were shuffled via the *adjusted-amplitude Fourier transform algorithm* (AAFT) in each epoch, which rescales the signals with a Gaussian process and randomizes the phase (Schreiber and Schmitz 2000). Subsequently, we did not only compute the coupling between two generic channels X and Y, but also the coupling between X and 19 surrogates of Y in each epoch. As suggested by Pereda et al (2005), the coupling was significant if and only if it was higher than the surrogate-based 19 interaction intensities. The number 19 guarantees an $\alpha = 0.05$ significance level since $N_{surr} = \frac{1}{\alpha} - 1 = 19$. Each coupling value which didn't pass the surrogate test was set to zero. The AAFT surrogates were computed with the MATS toolbox (Kugiumtzis and Tsimpliris 2010). Since the label of sleep stage was known, the different FC adjacency matrices were averaged throughout the QS and NQS periods, such that two adjacency matrices $\mathbf{A}$ were obtained per recording.

The next step was the computation of the topological and spectral indices, whose output was arranged as a $103 \times 12 \times 5$ tensor for each method and sleep state. The number of recordings was 103. We obtained 12 network indices and we used five frequency bands. Due to the large number of features, we decided to report only the

path length as median(IQR), where IQR is *inter-quartile range*, in three PMA groups (≤ 31 Weeks, $\in (31 - 37)$ Weeks, ≥ 37 Weeks) to give a general overview of the predictive power for each method. Furthermore, the Pearson correlation coefficient between the path length and PMA on the entire dataset has been reported.

The topological features from each FC method were used to develop a linear mixed-effect (LME) model to predict the PMA in each sleep state. In the random-effects modeling, the subject or patient ID label was considered as a grouping variable. Since the topology features can be characterized by collinearity due to the similar definitions of graph indices, we decided to apply a least absolute shrinkage operator (LASSO) to select the features for the regression model. In particular, a LASSO algorithm was run 100 times on each dataset and only features selected more than 40% of times were included in the final model. The LME model with LASSO-filtered features was then estimated on 100 random splits of the dataset. In each split, the model was tuned on 70% of data, the training set, and tested on the remaining 30%, the test set. The model performance has been assessed via the *mean absolute error* (MAE) on the test set and the *adjusted coefficient of determination* R_{adj}^2 on the training set. The last step was to develop a model independently from the type of FC method or frequency band. Specifically, a subset of features with the highest predictive power to track PMA have been selected in this way: each attribute was used to predict age in a univariate LME fashion and ranked based on MAE. The 10 features with the lowest MAE were used to develop a final age regression model. In this phase, the FC coupling intensities together with the network indices were also used as candidate features. Eventually, the final model with 10 features was tested on 100 random splits of the dataset and evaluated using R_{adj}^2 and MAE.

3. Results

Table 1 gives a general overview of different methods to describe the connectivity over age by means of the path length. The different age groups in the two tables have respectively the following number of recordings: 24, 62, 17. The first remarkable aspect is a stronger correlation with PMA for *ImCoh* compared to other FC methods, in particular for the frequency bands δ_2 , θ and α , together with $k_{xy}^2(\theta, \beta)$ and *HSD* (θ, α). Secondly, PLV, *HSD* (θ, α) and $k_{xy}^2(\delta_1, \delta_2, \theta, \alpha)$ show a negative correlation between path length and PMA, while $k_{xy}^2(\beta)$ and *ImCoh* present a positive trend. Since the path length is a measure of distance, a decrease of this topological index relates with an increase of connectivity, while its increase is associated to a lesser degree of interaction. For the *ImCoh*(θ), LME-regression chart with the associated confidence intervals and data cloud of the path length and the clustering coefficient have been reported in figure 1. Each color in the plot represents a different patient. Figure 2 shows different density indices computed via *MSC* in β and θ bands. In particular, panel a shows posterior density for $k_{xy}^2(\beta)$ during QS, while panel c shows the anterior density for $k_{xy}^2(\theta)$ for QS. The anterior density $k_{xy}^2(\theta)$ increases with postnatal maturation, while we found a negative trend of $k_{xy}^2(\beta)$ with postnatal maturation. Table 2 reports the results for the prediction via a linear mixed-effects model. *ImCoh* clearly outperforms any other FC methods in PMA prediction both in terms of MAE and R_{adj}^2 . Considering the frequency bands, the θ band consistently outperforms the others in PMA prediction for the *ImCoh*, while *MSC* sees the β band as the most informative one, PLV the δ_1 and δ_2 band and *HSD* the θ and α band. These findings tend to be consistent in both sleep states. Table 2 also displays the result with the best selected features, which mainly come from *ImCoh* in θ and α bands. Consequently, the prediction power is similarly to that of the *ImCoh*. The topographic distribution of the coupling value for *ImCoh*(θ) is reported in figure 3.

4. Discussion

In this paper, we investigated recent approaches to describe EEG-based FC in order to develop a model to predict the age of preterm infants. Firstly, we compared different methodologies and frequency bands to describe the connectivity development; secondly, we merged different features to find the best PMA predictor. The first important aspect to highlight is that our results seem to confirm previous findings in the literature, since both Grieve and Meijer have shown an increase of the $k_{xy}^2(\theta)$ in the frontal area and a decrease of $k_{xy}^2(\beta)$ in the occipital area (at least, during gestational maturation) (Grieve et al 2008, Meijer et al 2014). Furthermore, Gonzalez has shown a decrease in *ImCoh*(δ_2) (González et al 2011). However, the neurophysiological implication of these results is more complicated. Meijer already described a segregation on the cortex due to the development of specialized tasks associated to the different brain areas, which finds support in the decrease of the *ImCoh* and in the hypothesis that sources are on the cerebral cortex of the infant (Meijer et al 2014). This result can be further confirmed by the decrease of connectivity in $k_{xy}^2(\beta)$. This interpretation can be contradicted by positive trends of *MSC*, *PLV* and *HSD* in the other bands. A possible explanation is that $k_{xy}^2 = |c_{xy}|^2$, which contains the real part of the coherency, and the *ImCoh* have different information. Grieve argued that the lower $k_{xy}^2(\theta)$ in preterm babies is a sign of higher reactivity to sensory inputs during sleep (Grieve et al 2008). An increase in *MSC* and *HSD* in

Table 1. The table displays the path length for the different FC methods during QS and NQS in the different age groups: PMA $\leq$ 31 Weeks, PMA $\in$ (31–37) Weeks, PMA $\geq$ 37 Weeks. Results are reported as median(IQR), where IQR stands for *interquartile range*. In the last column, the Pearson correlations of the path length with PMA is reported and the associated p_{value} is reported. The values $\times 10^3$ means that reported results are divided by $\times 10^3$. The path length is a measure of distance: a positive trend is a decrease of connectivity (segregation), while a negative one is an increase of connectivity (integration).

PMA	Path length—QS					Path length—NQS				
	$\leq$ 31 Weeks	$\in$ (31–37) Weeks	$\geq$ 37 Weeks	ρ (%)		$\leq$ 31 Weeks	$\in$ (31–37) Weeks	$\geq$ 37 Weeks	ρ %	
$ImCoh(\delta_1)$	14.54(5.32)	20.24(8.12)	24.81(16.75)	49.53 ^a		18.79(4.49)	21(6.11)	20.42(7.43)	–2.8% ^{n.s.}	
$ImCoh(\delta_2)$	8.97(3.63)	13.42(5.84)	20.17(5.01)	75.48 ^a		10.88(2.53)	14.73(3.87)	15.14(3.54)	40.89% ^a	
$ImCoh(\theta)$	6.11(1.85)	9.72(5.4)	17.71(6.19)	74.78 ^a		7.55(2.01)	12.44(5.69)	14.83(3.42)	59.12% ^a	
$ImCoh(\alpha)$	5.53(1.72)	7.2(3.21)	11.01(3.53)	74.49 ^a		6.68(1.31)	9.7(2.79)	10.72(1.39)	68.83% ^a	
$ImCoh(\beta)$	5.78(1.09)	6.03(1.31)	6.39(1.82)	17.63 ^{n.s.}		6.5(1.04)	6.6(0.98)	6.51(0.8)	1.95% ^{n.s.}	
$k_{xy}^2(\delta_1)$	6.18(1.54)	6.18(2.01)	5.53(2.09)	–9.61 ^{n.s.}		5.99(1.43)	5.82(1.7)	4.85(1.85)	–28.8% ^a	
$k_{xy}^2(\delta_2)$	6.17(2.36)	6.19(2.1)	5.52(1.77)	–18.89 ^{n.s.}		6.18(2.49)	5.45(1.88)	5.68(1.36)	–18.86% ^{n.s.}	
$k_{xy}^2(\theta)$	6.77(2.1)	6.67(1.92)	4.92(1.31)	–44.85 ^a		8.05(2.64)	6.19(2.29)	5.21(1.04)	–45.49% ^a	
$k_{xy}^2(\alpha)$	7.52(2.15)	8.04(2.83)	7.08(2.7)	–21.03 ^b		10.29(3.62)	9.12(3.4)	8.22(4.64)	–24.77% ^b	
$k_{xy}^2(\beta)$	6.45(2.67)	8.45(4.01)	10.1(6.25)	48.49 ^a		7.45(2.22)	11.59(6.1)	14.44(6.98)	56.21% ^a	
$PLV(\delta_1)$	3.86(1.78)	3.48(1.13)	3.66(1.35)	–11.41 ^{n.s.}		3.8(1.3)	4.01(1.22)	3.84(1.4)	–1.98% ^{n.s.}	
$PLV(\delta_2)$	3.02(0.68)	3.22(0.92)	3.09(0.67)	–9.91 ^{n.s.}		3.3(0.7)	3(0.79)	3.23(0.66)	–21.03% ^b	
$PLV(\theta)$	2.99(0.74)	3.17(0.8)	2.95(0.52)	–9.51 ^{n.s.}		3.25(0.91)	3.09(0.7)	2.97(0.56)	–24.43% ^b	
$PLV(\alpha)$	3.37(1.3)	3.54(0.91)	3.31(0.51)	–15.94 ^{n.s.}		3.82(0.82)	3.74(0.99)	3.65(0.57)	–20.84% ^b	
$PLV(\beta)$	3.71(1.59)	4.13(1.05)	4.57(1.15)	19.97 ^b		4.17(1.09)	4.55(1.18)	4.99(1.38)	32.21% ^a	
$HSD(\delta_1)$	154.57(65.05)	132.93(47.37)	130.55(68.01)	–20.45 ^b		140.41(51.88)	137.28(38.59)	127.26(43.46)	–15.78% ^{n.s.}	
$HSD(\delta_2)$	167.54(63.25)	155.04(55.88)	135.17(36.72)	–37.45 ^a		165.8(44.65)	133.66(44.25)	133.97(41.04)	–34.96% ^a	
$HSD(\theta)$	216.53(96.42)	172.87(52.74)	135.18(32.48)	–56.62 ^a		195.53(76.71)	152.3(57.33)	131.62(32.15)	–53.99% ^a	
$HSD(\alpha)$	237.61(78)	205.71(65.18)	182.09(31.04)	–46.08 ^a		268.11(46.97)	220.82(69.59)	193.22(37.79)	–44.76% ^a	
$HSD(\beta)$	236.16(100.17)	268.16(79.4)	274.02(98.75)	14 ^{n.s.}		266.93(77.01)	292.73(104.85)	355.53(170.72)	36.15% ^a	

^a Stands for $p \leq 0.01$.

^b Stands for $p \leq 0.05$, while ^{n.s.} means $p = \text{not significant}$.

Figure 1. The figure shows the linear mixed-effect regression between the clustering coefficient and the age (Panel (a)), and between path length and the age for the *ImCoh* in θ bands (Panel (b)) during QS. Specifically, the black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. Each color in the plot represents a different patient. In order to show the specific trend for a single patient, we highlighted the result for a chosen infant in thicker red line. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value} .

Figure 2. The figure shows the linear mixed-effect regression between different MSC density indices coefficient and the age. Panel (a) displays the density in posterior hemisphere in the β band during only QS, while panel (b) shows the density in anterior hemisphere in the θ band during QS. The black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. In this figure, we reported only the data cluster without highlighting each patient. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value} .

that band could mean a decrease in reactivity (higher instantaneous interaction), but a lesser degree of lagged interaction with a lower *ImCoh*. This opposite trend between the two indices was also reported by González et al (2011), who argued that stimuli and intervention in the postnatal maturation could influence those trends. Whether this hypothesis is true should be tested with a proper study. However, Myers et al (2015) has already shown that the skin-to-skin contact (which is a tactile stimulus) drives a sharper decrease rate in the MSC with postnatal maturation. This approach could be extended to *ImCoh* in future research.

In general, the decrease of cortical connectivity indices with *ImCoh* are in line with results obtained with effective connectivity in our previous study (Lavanga et al 2017), where we reported the cortical segregation as main reason for this decrease. This segregation is the development of the thalamo-cortical and long cortico-cortical connections, as reported in the introduction, while the short-range wiring weakens, as has been shown via fMRI (Batalle et al 2017). This process may trigger a separation and segregation of the different brain regions. The thalamo-cortical connections are dominant in the first part of the preterm development (up to 31 PMA Weeks), in order to develop a substrate for sensory-cortical activation. The cortico-cortical connections prevail in the second part of the development (after 31 PMA Weeks), in order to shape the folding of the different cortical areas (Kostović and Judoš 2010), which causes the vanishing discontinuity in the neonatal EEG (Witte et al 1997, Kostović and Judoš 2010). This could induce either a burst de-synchronization or instantaneous synchronization, which both will diminish the *ImCoh*. In this perspective, it is not surprising that the *ImCoh* is the best method to predict PMA. This segregation hypothesis has also been debated in Stevens (2009). The author reports

Table 2. The performance of the different linear mixed-effect models for the investigated FC methods is displayed. The table reports the adjusted explained variance R^2_{adj} on the training set and the MAE on the test set as median(IQR), where IQR stands for *interquartile range*. The results are reported for both sleep states, QS and NQS. The last row reports the result with selection of the best features via LASSO.

FC methods	QS		NQS	
	R^2_{Adj} [.]	MAE (Weeks _{PMA})	R^2_{Adj} [.]	MAE (Weeks _{PMA})
$ImCoh(\delta_1)$	0.37(0.1)	2.06(0.38)	0.09(0.06)	2.62(0.45)
$ImCoh(\delta_2)$	0.72(0.09)	1.61(0.27)	0.26(0.07)	2.31(0.34)
$ImCoh(\theta)$	0.81(0.08)	1.54(0.22)	0.62(0.06)	1.59(0.31)
$ImCoh(\alpha)$	0.66(0.14)	1.7(0.3)	0.61(0.08)	1.54(0.27)
$ImCoh(\beta)$	0.49(0.09)	2.08(0.29)	0.39(0.08)	2.21(0.42)
$k^2_{xy}(\delta_1)$	0.08(0.06)	2.64(0.36)	0.36(0.08)	2.07(0.39)
$k^2_{xy}(\delta_2)$	0.2(0.07)	2.46(0.4)	0.26(0.09)	2.47(0.45)
$k^2_{xy}(\theta)$	0.38(0.08)	2.03(0.25)	0.39(0.1)	2.05(0.38)
$k^2_{xy}(\alpha)$	0.38(0.09)	2.14(0.39)	0.28(0.09)	2.35(0.34)
$k^2_{xy}(\beta)$	0.41(0.12)	2.16(0.4)	0.47(0.08)	1.94(0.36)
$PLV(\delta_1)$	0.18(0.08)	2.47(0.36)	0.44(0.09)	1.96(0.36)
$PLV(\delta_2)$	0.27(0.07)	2.33(0.33)	0.32(0.11)	2.32(0.39)
$PLV(\theta)$	0.24(0.07)	2.31(0.33)	0.23(0.09)	2.44(0.42)
$PLV(\alpha)$	0.29(0.12)	2.26(0.36)	0.26(0.08)	2.4(0.35)
$PLV(\beta)$	0.15(0.07)	2.43(0.33)	0.34(0.14)	2.25(0.33)
$HSD(\delta_1)$	0.17(0.09)	2.66(0.37)	0.08(0.06)	2.67(0.36)
$HSD(\delta_2)$	0.54(0.1)	1.97(0.35)	0.26(0.07)	2.27(0.41)
$HSD(\theta)$	0.59(0.07)	1.82(0.33)	0.51(0.05)	1.87(0.33)
$HSD(\alpha)$	0.54(0.07)	1.92(0.34)	0.45(0.06)	1.96(0.34)
$HSD(\beta)$	0.28(0.08)	2.27(0.4)	0.28(0.09)	2.4(0.34)
Best features	0.8(0.11)	1.51(0.31)	0.65(0.07)	1.64(0.29)

the different studies which have investigated a decrease of FC in general brain development. A possible explanation is a shift from a more diffuse cortical organization to a focalized one, probably driven by the cortical folding.

5. Conclusions

This research provides a new FC-based PMA prediction model for healthy premature neonates. The main novelty lies on the extensive usage of graph features to describe EEG interdependencies, as well as the comparison of multiple regression models to quantify the predictive power of the different connectivity methods. The results showed that $ImCoh$ outperforms the other methods to predict the age of the patient. The highest predictive performances was obtained with $ImCoh$ in δ_2, θ, α bands. The performance parameters were $MAE = 1.51$ weeks and $R^2_{adj} = 0.80$. Based on FC analysis, the data shows a segregation of EEG nodes, at least in term of lagged interactions. This result could be explained by a more diffused architecture to perform different tasks on the cortex.

Acknowledgments

This research is supported by Bijzonder Onderzoeksfonds KU Leuven (BOF): The effect of perinatal stress on the later outcome in preterm babies (# C24/15/036); imec funds 2017; Belgian Foreign Affairs-Development Cooperation (VLIR UOS programs (2013–2019)); EU: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) / ERC Advanced Grant: BIOTENSORS (n° 339804); ERASMUS +: INGDIVS 2016-1-SE01-KA203-022114; EU H2020-FETOPEN 'AMPHORA' # 766456. This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information. ML is a SB PhD fellow at Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (FWO), supported by Flemish government.

ORCID iDs

M Lavanga <https://orcid.org/0000-0002-3615-033X>

References

- Batalle D *et al* 2017 NeuroImage early development of structural networks and the impact of prematurity on brain connectivity **149** 379–92
- Bezruschko B P and Smirnov D A 2010 *Extracting Knowledge from Time Series* (Berlin: Springer)
- Bondy J and Murty U 1976 *Graph Theory with Applications* (Berlin: Springer)
- Bortel R and Sovka P 2007 Approximation of statistical distribution of magnitude squared coherence estimated with segment overlapping *Signal Process.* **87** 1100–17
- Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nat. Rev. Neurosci.* **10** 186–98
- Franke K, Luders E, May A, Wilke M and Gaser C 2012 Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI *Neuroimage* **63** 1305–12
- Friston K J 1994 Functional and effective connectivity in neuroimaging: a synthesis *Hum. Brain Mapp.* **2** 56–78
- Goldenberg D and Galván A 2015 Developmental cognitive neuroscience the use of functional and effective connectivity techniques to understand the developing brain **12** 155–64
- González J J, Mañas S, De Vera L, Méndez L D, López S, Garrido J M and Pereda E 2011 Assessment of electroencephalographic functional connectivity in term and preterm neonates *Clin. Neurophysiol.* **122** 696–702
- Grieve P G, Isler J R, Izraelit A, Peterson B S, Fifer W P, Myers M M and Stark R I 2008 EEG functional connectivity in term age extremely low birth weight infants *Clin. Neurophysiol.* **119** 2712–20
- Haufe S, Nikulin V V, Müller K R and Nolte G 2013 A critical assessment of connectivity measures for EEG data: a simulation study *NeuroImage* **64** 120–33
- Koolen N, Dereymaeker A, Räsänen O, Jansen K, Vervisch J, Matic V, Naulaers G, De Vos M, Van Huffel S and Vanhatalo S 2016 Early development of synchrony in cortical activations in the human *Neuroscience* **322** 298–307
- Kostović I and Judoš M 2010 The development of the subplate and thalamocortical connections in the human foetal brain *Acta Paediatrica* **99** 1119–27
- Kugiumtzis D and Tsimpiris A 2010 Measures of analysis of time series (MATS): a MATLAB toolkit for computation of multiple measures on time series data bases *J. Stat. Softw.* **33** 1–30
- Lavanga M, De Wel O, Caicedo A, Jansen K, Dereymaeker A, Naulaers G and Van Huffel S 2017 Monitoring effective connectivity in the preterm brain: a graph approach to study maturation *Complexity* **2017** 1–13
- Lavanga M, De Wel O, Caicedo Dorado A, Jansen K, Dereymaeker A, Naulaers G and Van Huffel S 2016 Linear and nonlinear functional connectivity methods to predict brain maturation in preterm babies *8th Int. Workshop on Biosignal Interpretation*
- Martin J A, Kochanek K D, Strobino D M, Guyer B and MacDorman M F 2005 Annual summary of vital statistics—2003 *Pediatrics* **115** 619–34
- Meijer E, Hermans K H M, Zwanenburg A, Jennekens W, Niemarkt H J, Cluitmans P J M, Van Pul C, Wijn P F F and Andriessen P 2014 Functional connectivity in preterm infants derived from EEG coherence analysis *Eur. J. Paediatric Neurol.* **18** 780–9
- Myers M, Grieve P, Stark R, Isler J, Hofer M, Yang J, Ludwig R and Welch M 2015 Family nurture intervention in preterm infants alters frontal cortical functional connectivity assessed by EEG coherence *Acta Paediatrica* **104** 670–7
- Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S and Hallett M 2004 Identifying true brain interaction from EEG data using the imaginary part of coherency *Clin. Neurophysiol.* **115** 2292–307
- O'Toole J M, Boylan G B, Vanhatalo S and Stevenson N J 2016 Estimating functional brain maturity in very and extremely preterm neonates using automated analysis of the electroencephalogram *Clin. Neurophysiol.*
- Pereda E, Quiroga R Q and Bhattacharya J 2005 Nonlinear multivariate analysis of neurophysiological signals *Prog. Neurobiol.* **77** 1–37
- Schreiber T and Schmitz A 2000 Surrogate time series *Physica D* **142** 346–82
- Stevens M C 2009 The developmental cognitive neuroscience of functional connectivity *Brain Cogn.* **70** 1–12
- Szabó Z 2014 Information theoretical estimators toolbox *J. Mach. Learn. Res.* **15** 283–7
- Witte H, Putsche P, Eiselt M, Hoffmann K, Schack B, Arnold M and Jäger H 1997 Analysis of the interrelations between a low-frequency and a high-frequency signal component in human neonatal EEG during quiet sleep *Neurosci. Lett.* **236** 175–9